

---

# ENVIRONMENTAL TRIGGERS FOR GEOSMIN AND 2-MIB PRODUCTION IN DRINKING WATER RESERVOIRS

---

Thesis submitted for the degree of

*Doctor of Philosophy*

Cardiff University



Annalise Sara Hooper BSc RSci

Supervised by: Rupert G. Perkins & Peter Kille

SCHOOL OF EARTH AND ENVIRONMENTAL SCIENCES

APRIL 2023



Dŵr Cymru  
Welsh Water

“I am enough of the artist to draw  
freely upon my imagination.  
Imagination is more important than  
knowledge. Knowledge is limited.  
Imagination encircles the world.”

- Albert Einstein (1929)

## SUMMARY

The presence of taste and odour compounds (T&O) in drinking water lead to numerous complaints to water companies worldwide. Geosmin and 2-MIB are common T&O compounds, with *Cyanobacteria* being the primary biological source in drinking water reservoirs. Both compounds have low odour thresholds in humans and require expensive additional treatment. This thesis used molecular and statistical analysis of water from Welsh Water reservoirs, to provide a framework for predicting and monitoring T&O events and understanding their causes.

Elevated T&O concentrations were confined to warmer months, except for a one geosmin event in winter 2019. There was no correlation between cyanobacterial abundance and T&O concentrations, but qPCR analysis based on eDNA sampling demonstrated that geosmin synthase (*geoA*) was a suitable proxy for predicting geosmin concentrations. Abundances of *geoA* and 2-MIB cyclase (*mic*) were significantly non-linearly associated with high ammonium-to-nitrate ratios, identifying thresholds for heightened T&O risk. The ratio of total inorganic nitrogen to total phosphorous was significantly non-linearly associated with increases in *geoA*. Increased *geoA* was also significantly negatively associated with temperature and dissolved reactive silicate in all reservoirs. Next-generation sequencing of bacterial and algal communities showed that community compositions clustered according to T&O concentrations. Bacterial and algal co-occurrence networks uncovered significant positive and negative associations, highlighting cyanospheres in all reservoirs. Random Forest models were developed for geosmin (Alaw) and 2-MIB (Pentwyn) using significantly co-occurring taxa exposing indicative T&O taxa and the probable *Cyanobacteria* causing the T&O. *Cyanobacteria* had more negative than positive associations in their cyanospheres.

This thesis illustrates the importance of nutrient ratios in triggering potential geosmin and 2-MIB events. It also indicates that *Cyanobacteria* subjected to environmental stress (negative biotic interactions and low temperatures) increase their T&O-production. These findings provide a useful framework for water monitoring to enable the prediction and possible prevention of T&O events.

## DEDICATION

To my beautiful grandmother, Jean Furlong, it was your words of wisdom to continue learning and to bury my head in a book that I have achieved and accomplished what I have today. As you always said, wealth is knowledge. Without you here, it has been hard to find the inspiration, but dedicating this thesis to you has kept me focussed. Without you, I would not be here, and I will always thank you for making my father the man he is today, my rock and my hero. Words cannot express how grateful I am to have my father and all the support he has given me throughout my life. I will never be able to thank my father, Paul Hooper, enough for the countless times he has lifted me when life has been challenging. Especially keeping me on track in my studies and ensuring I am happy and well. I would also like to dedicate this thesis to my not-so-baby brother, Liam Hooper. My brother, my best friend, and my cheerleader. One of the only people who can cheer me up and make me laugh when I can't find the strength. I would also like to thank all four of my incredible sisters for their support throughout my studies.

My friends are my family, and I would not be here today if it wasn't for my ultimate best friend, Hannah Evans. My oldest and closest friend, I will forever be in debt for all her emotional support. You always kept me going when no one else knew how to. I would also like to thank my best friend, Rebecca Hooper, especially for her support during my studies when I didn't think I could achieve a PhD. Such a kind individual whom I am blessed to have in my life. Nadine Beavis, writing a dedication section without getting your name in my "book" would not be acceptable. You're one of the most caring people I know, and you have kept me grounded for as long as I can remember. Sarah Morgan, thank you for being my rocker; life is better with you in it. Hannah Clee/ Gibbins, you are also one person I would love to thank for always making me laugh.

## ACKNOWLEDGEMENTS

This dream would never have become a reality if it wasn't for the chance my supervisor, Dr Rupert Perkins, took on me. It has been his enthusiasm for the work that has kept my imagination going. I also will never be able to thank Rupert enough for all his fantastic support when things haven't gone to plan. In addition to one excellent supervisor, I am eternally grateful for the support, knowledge, and wisdom from my secondary supervisor, Prof Peter Kille. I have been so lucky to be trained by you both and have gained extensive knowledge that I will always value. Similarly, I have also been blessed with a fantastic academic team at Cardiff University. I would like to take this opportunity to thank Dr Sarah Christofides for all her statistical support throughout my PhD and the weekly coffee meetings. My work would also not be complete if it wasn't for the excellent technical assistance I received through the genome hub and from Inge Elfferich. I want to thank Angela Marchbank, Trudy Workman, and Catherine Bresner for all their hard work. My thesis would also not be complete if it wasn't for all the edits Dr Sophie Watson has done for me; she is a great friend and a great work colleague. Also, I would like to mention and thank Dr Amie Parris and Dr Jordan Cuff for all their outstanding support and help along the way.

I also want to thank the catchment, inorganic chemistry, and organic chemistry teams in Dŵr Cymru Welsh Water. Without the assistance of Gemma Godwin and Greg Bullock from the catchment team, I would not have been able to get the samples I needed. I also would not have been able to of achieved this PhD without the support and guidance of Matthew Jones. As a previous employee of Dŵr Cymru Welsh Water, it has been amazing staying connected and working with such a great team,

To say this PhD has been easy would be a lie. Possibly one of the most challenging parts of my life. Although, I would never change my decision to do this PhD and would do it again if I had a second chance. It was from the support I received during my PhD from the people that I have mentioned that I have been able to complete this thesis and PhD.

This research was supported by funding from NERC Centre for Doctoral Training in Freshwater Biosciences and Sustainability (GW4 FRESH CDT) from the Natural Environment Research Council [NE/R0115241] in collaboration with Dŵr Cymru, Welsh Water (DCWW).

## Contents

|                        |       |
|------------------------|-------|
| SUMMARY .....          | ii    |
| DEDICATION.....        | iii   |
| ACKNOWLEDGEMENTS ..... | iv    |
| Contents.....          | vi    |
| List of Figures .....  | xiv   |
| List of Tables.....    | xxxii |

### **Chapter 1: General Introduction**

|  |    |
|--|----|
| 1.1 – General introduction .....   | 3  |
| 1.2 – Geosmin and 2-methylisoborneol.....  | 4  |
| 1.2.1 – Metabolic pathways of geosmin and 2-MIB.....                             | 6  |
| 1.2.2 – Geosmin synthase ( <i>geoA</i> ) and 2-MIB cyalse ( <i>mic</i> ) .....   | 7  |
| 1.3 – Adaptations of <i>Cyanobacteria</i> .....                                  | 9  |
| 1.3.1 – Cyanobacterial T&O productivity driven rather than biomass-related ..... | 10 |
| 1.4 – Environmental triggers .....   | 11 |
| 1.4.1 – Abiotic triggers .....   | 11 |
| 1.4.1.1 – Nutrients .....  | 11 |
| 1.4.1.2 – Light and temperature.....   | 13 |
| 1.4.1.3 – Hydrology and stratification.....                                      | 15 |
| 1.4.2 – Biotic triggers .....  | 16 |
| 1.4.2.1 – The cyanobacterial phycosphere – the “cyanosphere” .....               | 16 |
| 1.4.2.2 – Geosmin and 2-MIB as allelopathic tools.....                           | 17 |
| 1.4.2.3 – Geosmin and 2-MIB degraders.....                                       | 18 |
| 1.5 – Using eDNA to help predict T&O events.....                                 | 19 |
| 1.5.1 – Modelling T&O events.....  | 22 |
| 1.6 – Treating T&O compounds.....  | 22 |

|  |    |
|--|----|
| 1.7 – Research aims.....   | 24 |
| 1.8 – Research hypotheses.....   | 25 |
| <b><u>Chapter 2: Materials and Methods</u></b>                                       |    |
| 2.1 – Sampling methods and sample site details .....                                 | 29 |
| 2.1.1 – Llyn Alaw reservoir, Anglesey .....  | 32 |
| 2.1.2 – Llyn Cefni reservoir, Anglesey .....   | 33 |
| 2.1.3 – Dolwen and Plas Uchaf Reservoirs, Conwy .....                                | 35 |
| 2.1.4 – Alwen Reservoir, Conwy .....   | 37 |
| 2.1.5 – Llwyn On Reservoir, Taf Fawr Valley .....                                    | 38 |
| 2.1.6 – Pentwyn and Ponsticill Reservoirs, Taf Fechan .....                          | 39 |
| 2.1.7 – Llandegfedd Reservoir, Newport.....  | 41 |
| 2.1.8 – COVID-19 sampling disruption statement .....                                 | 42 |
| 2.2 – Defining a Taste and Odour event .....   | 42 |
| 2.3 – Water chemistry and physical analysis.....                                     | 42 |
| 2.3.1 – Determination of geosmin and 2-MIB concentrations.....                       | 43 |
| 2.3.2 – Nutrient analysis .....  | 43 |
| 2.3.3 – Metals .....   | 44 |
| 2.3.4 – Chlorophyll <i>a</i> (Chl <i>a</i> ) .....                                   | 44 |
| 2.3.5 – Total Organic Carbon (TOC).....  | 44 |
| 2.3.6 – pH, conductivity, and Turbidity .....  | 45 |
| 2.4 – Water filtering and Extraction of total genomic environmental DNA (eDNA) ..... | 45 |
| 2.4.1 – Primary Illumina-Nextera tag PCR for 16S rRNA and <i>rbcL</i> genes .....    | 46 |
| 2.4.2 – 16S rRNA and <i>rbcL</i> amplicon clean-up .....                             | 47 |
| 2.4.3 – Secondary Illumina-Nextera indexing adapter PCR.....                         | 48 |
| 2.4.4 – SequelPrep Normalisation Plates .....  | 50 |
| 2.4.6 – Databases.....   | 50 |
| 2.4.7 – Bioinformatic analysis using QIIME2.....                                     | 51 |



**Chapter 3: Monthly changes in T&O concentrations reveal ammonium and phosphorous to be key triggers**

|  |    |
|--|----|
| 3.1 Introduction .....   | 55 |
| 3.2 – Materials and Methods.....   | 59 |
| 3.2.1 – Sample collection .....  | 59 |
| 3.2.2 – Physical and chemical water quality testing .....  | 59 |
| 3.2.3 – Assignment of Trophic State Indices and overall Carlson Trophic State Index by season and year .....   | 59 |
| 3.2.4 – Defining a Taste and Odour event .....   | 61 |
| 3.2.5 – Principal Component Analysis of chemical and physical parameters .....                                 | 62 |
| 3.2.6 – Monthly changes in T&O concentrations and nutrients .....  | 62 |
| 3.2.7 – Monthly changes in nutrients and T&O concentrations correlogram's .....                                | 65 |
| 3.3 – Results.....   | 66 |
| 3.3.1 – Concentrations of geosmin and 2-MIB in studied reservoir and tributary locations.....                  | 66 |
| 3.3.2 – Seasonality of geosmin and 2-MIB concentrations in reservoir locations for each studied reservoir..... | 68 |
| 3.3.3 – CTSI and TSI indicators per season and year for each reservoir.....                                    | 72 |
| 3.3.3.1 – Alaw .....   | 72 |
| 3.3.3.2 – Alwen .....  | 74 |
| 3.3.3.3 – Cefni .....  | 76 |
| 3.3.3.4 – Dolwen.....  | 78 |
| 3.3.3.5 – Llandegfedd .....  | 80 |
| 3.3.3.6 – Llwyn On.....  | 82 |
| 3.3.3.7 – Pentwyn.....   | 84 |
| 3.3.3.8 – Plas Uchaf .....   | 86 |
| 3.3.3.9 – Pontsticill.....   | 88 |

|  |     |
|--|-----|
| 3.3.4 – Significant associations between physical and chemical water parameters and T&O levels using PCA.....            | 90  |
| 3.3.4.1 – Alaw – geosmin and 2-MIB concentrations.....   | 90  |
| 3.3.4.2 – Alwen – 2-MIB concentrations .....   | 92  |
| 3.3.4.3 – Cefni – geosmin concentrations.....  | 93  |
| 3.3.4.4 – Dolwen – geosmin concentrations .....  | 94  |
| 3.3.4.5 – Llandegfedd – 2-MIB concentrations .....   | 95  |
| 3.3.4.6 – Llwyn On – geosmin concentrations .....  | 96  |
| 3.3.4.7 – Pentwyn – 2-MIB concentrations .....   | 97  |
| 3.3.4.8 – Plas Uchaf – geosmin concentrations.....   | 98  |
| 3.3.4.9 – Pontsticill – geosmin and 2-MIB concentrations .....   | 99  |
| 3.3.4 – Monthly changes in nutrients associated with monthly changes in T&O concentrations.....                          | 101 |
| 3.3.4.1 – Alaw – geosmin .....   | 101 |
| 3.3.4.2 – Cefni – geosmin.....   | 104 |
| 3.3.4.3 – Dolwen – geosmin .....   | 107 |
| 3.3.4.4 – Llwyn On – geosmin .....   | 110 |
| 3.3.4.5 – Plas Uchaf – geosmin.....  | 113 |
| 3.3.4.6 – Pontsticill – geosmin .....  | 116 |
| 3.3.4.7 – Pentwyn – 2-MIB .....  | 119 |
| 3.3.4.8 – Pontsticill – 2-MIB .....  | 121 |
| 3.3.5 – Correlation matrices of monthly changes in nutrients associated with monthly changes in T&O concentrations ..... | 123 |
| 3.4 – Discussion.....  | 127 |
| 3.4.1 – Seasonality of T&O compounds in relation to CTSI and TSI indicators ..   | 127 |
| 3.4.2 – Indirect influences of changes in nutrients on the change in T&O concentrations.....                             | 129 |

|   |     |
|---|-----|
| 3.4.3 – Monthly changes in N and P fractions associated with monthly changes in T&O concentrations .....  | 130 |
| 3.4.4 – Time lagged monthly changes in geosmin concentrations in a reservoir experiencing severe geosmin events with monthly changes in N and P fractions ..... | 131 |
| 3.4.5 – Monthly changes in sulphate, dissolved iron and dissolved reactive silicate and implications on monthly changes in T&O concentrations .....             | 132 |
| 3.4.6 – Limitations of the study .....  | 134 |
| 3.5 – Conclusions .....   | 135 |

**Chapter 4: Negatively co-occurring taxa with Cyanobacteria inducing T&O events: modelling T&O using indicative taxa**

|  |     |
|--|-----|
| 4.1 – Introduction .....   | 139 |
| 4.2 – Materials and Methods .....  | 142 |
| 4.2.1 – Sample collection .....  | 142 |
| 4.2.2 – eDNA extraction and community analysis using the 16S rRNA and <i>rbcL</i> genes .....  | 142 |
| 4.2.3 – Water chemistry and physical analysis .....  | 143 |
| 4.2.4 – Categorising T&O concentration levels .....  | 143 |
| 4.2.5 – Spatial and temporal changes in 16S rRNA phyla community composition .....   | 143 |
| 4.2.6 – Co-occurrence of 16S rRNA and <i>rbcL</i> genera and network visualisation .....   | 144 |
| 4.2.7 – Non-metric Multi-dimensional Scaling of 16S rRNA and <i>rbcL</i> communities .....   | 146 |
| 4.2.8 – Random Forest to identify indicator genera from significant co-occurring organisms in reservoirs experiencing T&O “events” ..... | 147 |
| 4.3 – Results .....  | 148 |
| 4.3.1 – 16S rRNA seasonal succession of phyla abundance between years .....  | 148 |
| 4.3.2 – Binomial co-occurrence analysis of 16S rRNA and <i>rbcL</i> communities .....  | 150 |

|  |     |
|--|-----|
| 4.3.2.1 – Alaw .....   | 151 |
| 4.3.2.2 – Alwen .....  | 160 |
| 4.3.2.3 – Cefni .....  | 163 |
| 4.3.2.4 – Dolwen.....  | 181 |
| 4.3.2.5 – Llandegfedd .....  | 184 |
| 4.3.2.6 – Llwyn On.....  | 189 |
| 4.3.2.7 – Pentwyn.....   | 192 |
| 4.3.2.8 – Plas Uchaf .....   | 197 |
| 4.3.2.9 – Pontsticill.....   | 201 |
| 4.3.3 – NMDS analysis for 16S rRNA genera .....  | 205 |
| 4.3.3.1 – Alaw .....   | 205 |
| 4.3.3.2 – Cefni .....  | 207 |
| 4.3.3.3 – Dolwen.....  | 208 |
| 4.3.3.4 – Llwyn On.....  | 210 |
| 4.3.3.5 – Pentwyn.....   | 212 |
| 4.3.3.6 – Plas Uchaf .....   | 214 |
| 4.3.3.7 – Pontsticill.....   | 216 |
| 4.3.4 – NMDS analysis for rbcL genera .....  | 218 |
| 4.3.4.1 – Alaw .....   | 218 |
| 4.3.4.2 – Cefni .....  | 220 |
| 4.3.4.3 – Dolwen.....  | 222 |
| 4.3.4.4 – Llwyn On.....  | 224 |
| 4.3.4.5 – Pentwyn.....   | 226 |
| 4.3.4.6 – Plas Uchaf .....   | 228 |
| 4.3.4.7 – Pontsticill.....   | 230 |
| 4.3.5 – Indicative genera of a T&O ‘event’ in Alaw, Llwyn On and Pentwyn<br>reservoirs experiencing high T&O concentrations..... | 232 |

|   |     |
|---|-----|
| 4.3.5.1 – Alaw .....  | 236 |
| 4.3.5.2 – Llwyn On.....   | 242 |
| 4.3.5.3 – Pentwyn.....  | 243 |
| 4.4 – Discussion.....   | 247 |
| 4.4.1 – Productivity driven cyanobacterial T&O events.....  | 248 |
| 4.4.2 – The “cyanosphere” in relation to T&O events.....  | 249 |
| 4.4.3 – Signature 16S rRNA and <i>rbcL</i> communities reflecting T&O concentrations<br>in reservoirs experiencing high T&O concentrations.....   | 257 |
| 4.4.4 – Indicative taxa for T&O events .....  | 258 |
| 4.5 – Conclusions.....  | 262 |
| <br>  |     |
| <b><u>Chapter 5: The importance of nutrient ratios in determining elevations in<br/>geosmin synthase (<i>geoA</i>) and 2-MIB cyclase (<i>mic</i>) resulting in Taste and<br/>Odour events</u></b> |     |
| 5.1 – Introduction.....   | 268 |
| 5.2 – Materials and methods.....  | 271 |
| 5.2.1 – Defining a T&O event.....   | 271 |
| 5.2.2 – Sample locations.....   | 271 |
| 5.2.3 – Genomic eDNA extraction.....  | 271 |
| 5.2.4 – Putative classification of standards used for 16S rRNA, <i>geoA</i> and <i>mic</i><br>standard curves .....   | 271 |
| 5.2.5 – qPCR 16S rRNA, <i>geoA</i> and <i>mic</i> standards preparation.....  | 274 |
| 5.2.6 – Reverse primer design for <i>mic</i> gene .....   | 276 |
| 5.2.7 – qPCR reactions for 16S rRNA, <i>geoA</i> and <i>mic</i> genes .....   | 278 |
| 5.2.8 – qPCR set-up on the Eppendorf epMotion robot .....   | 279 |
| 5.2.9 – qPCR reactions.....   | 280 |
| 5.2.10 – Cross validation of 16S rRNA and <i>geoA</i> standards.....  | 283 |
| 5.3 – Results.....  | 284 |

|  |     |
|--|-----|
| 5.3.1 – Temporal changes in gene copy numbers and T&O concentrations by year and month.....          | 284 |
| 5.3.2 – Temporal changes in gene copy numbers and T&O concentrations by season and year .....        | 286 |
| 5.3.3 – Modelling a T&O event.....   | 288 |
| 5.3.3.1 - <i>geoA</i> :16S abundance in a reservoir with major geosmin events (Alaw) .....           | 289 |
| 5.3.3.2 - <i>geoA</i> :16S abundance in reservoirs with minor geosmin concentrations .....           | 292 |
| 5.3.3.3 – <i>mic</i> :16S abundance in all reservoirs .....  | 295 |
| 5.4 – Discussion.....  | 298 |
| 5.4.1 – Triggers of geosmin “events” .....   | 299 |
| 5.4.2 – <i>mic</i> and 2-MIB concentrations.....   | 301 |
| 5.4.3 – qPCR primer specificity for <i>mic</i> and <i>geoA</i> .....                                 | 302 |
| 5.4.4 – Seasonal influence of T&O compounds .....  | 303 |
| 5.5 – Conclusions .....  | 303 |
| <b><u>Chapter 6: General Discussion</u></b>  |     |
| 6.1 – T&O events are productivity driven and not always biomass-related.....                         | 307 |
| 6.3 –T&O triggers and indicators .....   | 309 |
| 6.1 – Seasonal occurrences of T&O concentrations .....   | 316 |
| 6.6 – Potential biological and ecological functions of T&O.....                                      | 318 |
| 6.7 – Cyanobacterial response to environmental stress – the cyanosphere T&O stable state theory..... | 320 |
| Bibliography.....  | 329 |
| Appendix .....   | 394 |

## List of Figures

|  |    |
|--|----|
| <i>Figure 1.1: Simplified biosynthetic pathways for the formation of geosmin and 2-MIB.</i>  | 7  |
| <i>Figure 1.2: An annotated Anabaena ucrainica CHAB2155 geosmin synthesis operon (HQ404996) and the associated biosynthetic pathway for geosmin production.</i>  | 8  |
| <i>Figure 1.3: An annotated Pseudanabaena sp. Dqp15 (HQ830028.1) 2-MIB synthesis operon and the associated biosynthetic pathway for 2-MIB production.</i>  | 9  |
| <i>Figure 1.4: Hypervariable regions in the 16S rRNA gene in Pseudomonas. The plotted line reflects fluctuations in variability amongst aligned 16S rRNA gene sequences of 79 known Pseudomonas strains. Adapted from Bodilis et al. (2012).</i>   | 20 |
| <i>Figure 1.5: The location of the rbcL gene and coding sequence within Nitzschia palea NIES-2729 chloroplast.</i>   | 21 |
| <i>Figure 2.1: An overview of reservoir locations in Wales, U.K., used within this study; 1 = Alaw. 2 = Cefni. 3 = Dolwen. 4 = Plas Uchaf. 5 = Alwen. 6 = Llwyn On. 7 = Pentwyn. 8 = Pontsticill. 9 = Llandegfedd. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).</i> | 31 |
| <i>Figure 2.2: Llyn Alaw reservoir and sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1 – R5 = reservoir locations. T1 – T4 = tributary locations. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).</i>             | 32 |
| <i>Figure 2.3: Close up of T2 (tributary two) and R1 (reservoir one) locations in Alaw reservoir with an inset map of the position of these sampling points within Alaw reservoir. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).</i>                                 | 33 |
| <i>Figure 2.4: Llyn Cefni reservoir and sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1 – R5 = reservoir locations. T1 – T3 = tributary locations. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).</i>            | 34 |
| <i>Figure 2.5: Close up of R3 and R4 (reservoir 3 and 4) locations in Cefni reservoir with an inset map of the position of these sampling points within Cefni reservoir. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).</i>   | 35 |
| <i>Figure 2.6: Plas Uchaf reservoir (top) and Dolwen reservoir (bottom) sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1</i>   |    |

– R4 = Plas Uchaf reservoir locations, T1 = Plas Uchaf tributary location located in the Afon Aled. R1 = Dolwen reservoir location, T1 – T3 = Dolwen tributary locations. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017). ..... 36

Figure 2.7: Alwen reservoir and sample point locations with an inset map of the reservoir’s location in Wales U.K., bottom right. R1 – R3 = Reservoir locations. T1 – T4 = Tributary locations. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017). ..... 38

Figure 2. 8: Llwyn On reservoir and sample point locations with an inset map of the reservoir’s location in Wales U.K., bottom right. R1 – R2 = Reservoir locations. T1 – T5 = Tributary locations. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017). ..... 39

Figure 2. 9: Pentwyn reservoir (top) and Pontsticill reservoir (bottom) sample point locations with an inset map of the reservoir’s location in Wales U.K., bottom right. R1 – R2 = Pentwyn reservoir locations, T1 = Pentwyn tributary location. R1 – R3 = Pontsticill reservoir location, T1 = Pontsticill tributary locations. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017). ..... 40

Figure 2.10: Llandegfedd reservoir and sample point locations with an inset map of the reservoir’s location in Wales U.K., bottom right. R1 – R4 = Reservoir locations. T1 – T2 = Tributary locations. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017)..... 41

Figure 2.11: Illumina-Nextera secondary PCR set up: columns 1-12 in the 96 well plate containing cleaned amplicons have the addition of a unique index 1 S5XX adapters (orange caps) and rows A-H have the addition of unique index 2 (N7XX) adapters (white caps)..... 49

Figure 3.1: Box and whisker plot showing  $\log_{10}$  geosmin concentrations + 1 ( $\text{ng L}^{-1}$ ) from the reservoir (Res in red) and tributary (Trib in blue) locations for all nine reservoirs over the total sampling period (16/7/2019 – 30/11/2020). Raw data is indicated by the coloured dots grouped by reservoir and tributary locations. The length of the box indicates the interquartile range, extending from the 25<sup>th</sup> to the 75<sup>th</sup> percentile and the horizontal bar within the box denotes the median value. The whiskers display the range. .... 67

Figure 3.2: Box and whisker plot showing  $\log_{10}$  2-MIB concentrations + 1 ( $\text{ng L}^{-1}$ ) from the reservoir (Res in red) and tributary (Trib in blue) locations for all nine reservoirs over total sampling period (16/7/2019 – 30/11/2020). Raw data is



indicated by the coloured dots grouped by reservoir and tributary locations. The length of the box indicates the interquartile range, extending from the 25<sup>th</sup> to the 75<sup>th</sup> percentile and the horizontal bar within the box denotes the median value. The whiskers display the range. .... 68

Figure 3.3: Log<sub>10</sub> + 1 geosmin concentrations (ng L<sup>-1</sup>) measured in water samples from the nine Welsh Water reservoirs between 16/7/2019 – 30/11/2020, coloured by season and faceted by individual reservoirs. The dotted lines indicate COVID-19 sample disruption. .... 70

Figure 3.4: Log<sub>10</sub> + 1 2-MIB concentrations (ng L<sup>-1</sup>) measured in water samples from nine Welsh reservoirs between 16/7/2019 – 30/11/2020, coloured by season and faceted by individual reservoirs. The dotted lines indicate COVID-19 sample disruption..... 71

Figure 3.5: Four circular bar plots representing means and standard deviation error bars for the Alaw reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (µg L<sup>-1</sup>) per season and year, C) average Chl a (µg L<sup>-1</sup>) per season and year and D) average CTSI measurements per season and year..... 73

Figure 3.6: Four circular bar plots representing means and standard deviation error bars for the Alwen reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (µg L<sup>-1</sup>) per season and year, C) average Chl a (µg L<sup>-1</sup>) per season and year and D) average CTSI measurements per season and year..... 75

Figure 3.7: Four circular bar plots representing means and standard deviation error bars for the Cefni reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (µg L<sup>-1</sup>) per season and year, C) average Chl a (µg L<sup>-1</sup>) per season and year and D) average CTSI measurements per season and year..... 77

Figure 3.8: Four circular bar plots representing means and standard deviation error bars for the Dolwen reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (µg L<sup>-1</sup>) per season and year, C) average Chl a (µg L<sup>-1</sup>) per season and year and D) average CTSI measurements per season and year..... 79

Figure 3.9: Four circular bar plots representing means and standard deviation error bars for the Llandegfedd reservoir for A) average Secchi depth measurement (m) per

|  |           |
|--|-----------|
| season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.....  | 81        |
| <i>Figure 3.10: Four circular bar plots representing means and standard deviation error bars for the Llwyn On reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (<math>\mu\text{g L}^{-1}</math>) per season and year, C) average Chl a (<math>\mu\text{g L}^{-1}</math>) per season and year and D) average CTSI measurements per season and year.....</i>    | <i>83</i> |
| <i>Figure 3.11: Four circular bar plots representing means and standard deviation error bars for the Pentwyn reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (<math>\mu\text{g L}^{-1}</math>) per season and year, C) average Chl a (<math>\mu\text{g L}^{-1}</math>) per season and year and D) average CTSI measurements per season and year.....</i>     | <i>85</i> |
| <i>Figure 3.12: Four circular bar plots representing means and standard deviation error bars for the Plas Uchaf reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (<math>\mu\text{g L}^{-1}</math>) per season and year, C) average Chl a (<math>\mu\text{g L}^{-1}</math>) per season and year and D) average CTSI measurements per season and year.....</i>  | <i>87</i> |
| <i>Figure 3.13: Four circular bar plots representing means and standard deviation error bars for the Pontsticill reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous (<math>\mu\text{g L}^{-1}</math>) per season and year, C) average Chl a (<math>\mu\text{g L}^{-1}</math>) per season and year and D) average CTSI measurements per season and year.....</i> | <i>89</i> |
| <i>Figure 3.14: PCA biplot of Alaw reservoir and tributary locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned T&amp;O level.....</i>   | <i>91</i> |
| <i>Figure 3.15: PCA biplot of Alwen reservoir's tributary locations using component 1 (PC1) and component 2 (PC2) with individual points coloured according to the assigned 2-MIB level.....</i>   | <i>92</i> |
| <i>Figure 3.16: PCA biplot of Cefni reservoir's reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level.....</i>  | <i>93</i> |

|  |            |
|--|------------|
| <i>Figure 3.17: PCA biplot of Dolwen reservoir’s reservoir locations using component 1 (PC1) and component 2 (PC2) with individual points coloured according to the assigned geosmin level. ....</i>   | <i>94</i>  |
| <i>Figure 3.18: PCA biplot of Llandegfedd reservoir’s tributary locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned 2-MIB level. ....</i>   | <i>95</i>  |
| <i>Figure 3.19: PCA biplot of Llwyn On reservoir’s reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level. ....</i>  | <i>96</i>  |
| <i>Figure 3.20: PCA biplot of Pentwyn reservoir’s reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned 2-MIB level. ....</i>   | <i>97</i>  |
| <i>Figure 3.21: PCA biplot of Plas Uchaf reservoir’s reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level. ....</i>  | <i>98</i>  |
| <i>Figure 3.22: PCA biplot of Pontsticill reservoir’s reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level. ....</i>   | <i>100</i> |
| <i>Figure 3.23: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in NH<sub>4</sub><sup>+</sup> concentrations (mg L<sup>-1</sup>) (blue dots) and B – monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (mg L<sup>-1</sup>) (red dots) given from each consecutive sampled month of the study period in Alaw reservoir locations. ....</i>  | <i>102</i> |
| <i>Figure 3.24: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) and B – monthly changes in orthophosphate concentrations (mg L<sup>-1</sup>) (orange dots) given from each consecutive sampled month of the study period in Alaw reservoir locations. ....</i>             | <i>103</i> |
| <i>Figure 3.25: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in NH<sub>4</sub><sup>+</sup> concentrations (mg L<sup>-1</sup>) (blue dots) and B – monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (mg L<sup>-1</sup>) (red dots) given from each consecutive sampled month of the study period in Cefni reservoir locations. ....</i> | <i>105</i> |
| <i>Figure 3.26: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) and B – monthly changes in orthophosphate concentrations (mg L<sup>-1</sup>) (orange dots)</i>   |            |

|  |     |
|--|-----|
| given from each consecutive sampled month of the study period in Cefni reservoir locations. ....   | 106 |
| Figure 3.27: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in $\text{NH}_4^+$ concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in $\text{NO}_3^-$ concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Dolwen reservoir locations.....               | 108 |
| Figure 3.28: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) and B – monthly changes in orthophosphate concentrations ( $\text{mg L}^{-1}$ ) (orange dots) given from each consecutive sampled month of the study period in Dolwen reservoir locations. ....   | 109 |
| Figure 3.29: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in $\text{NH}_4^+$ concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in $\text{NO}_3^-$ concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Llwyn On reservoir locations.....             | 111 |
| Figure 3.30: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) and B – monthly changes in orthophosphate concentrations ( $\text{mg L}^{-1}$ ) (orange dots) given from each consecutive sampled month of the study period in Llwyn On reservoir locations. .... | 112 |
| Figure 3. 31: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in $\text{NH}_4^+$ concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in $\text{NO}_3^-$ concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Plas Uchaf reservoir locations. ....         | 114 |
| Figure 3.32: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) given from each consecutive sampled month of the study period in Plas Uchaf reservoir locations. ....   | 115 |
| Figure 3. 33: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in $\text{NH}_4^+$ concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in $\text{NO}_3^-$ concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations. ....        | 117 |
| Figure 3.34: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots)   |     |

|   |     |
|---|-----|
| given from each consecutive sampled month of the study period in Pontsticill reservoir locations. ....  | 118 |
| Figure 3.35: Monthly changes in 2-MIB concentrations ( $\text{ng L}^{-1}$ ) (dark blue bars) and A – monthly changes in $\text{NH}_4^+$ concentrations ( $\text{mg L}^{-1}$ ) (light blue dots) and B – monthly changes in $\text{NO}_3^-$ concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Pentwyn reservoir locations. ...                                       | 120 |
| Figure 3.36: Monthly changes in 2-MIB concentrations ( $\text{ng L}^{-1}$ ) (dark blue bars) and monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) given from each consecutive sampled month of the study period in Pentwyn reservoir locations. ....   | 121 |
| Figure 3.37: Monthly changes in 2-MIB concentrations ( $\text{ng L}^{-1}$ ) (dark blue bars) and A – monthly changes in $\text{NH}_4^+$ concentrations ( $\text{mg L}^{-1}$ ) (light blue dots) and B – monthly changes in $\text{NO}_3^-$ concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations. ...                                   | 122 |
| Figure 3.38: Monthly changes in 2-MIB concentrations ( $\text{ng L}^{-1}$ ) (dark blue bars) and monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations. ....   | 123 |
| Figure 3.39: Correlogram of monthly changes in nutrients and monthly changes in geosmin concentrations with a one-month lag for all reservoirs. Colour-coded by the individual reservoir. Pearson correlation is displayed on the right-hand side of the figure, and variable distribution is available on the diagonal. Correlation significance codes: ‘.’ $p < 0.1$ , $p < 0.05$ ‘*’, ‘**’ $p < 0.01$ , ‘***’ $p < 0.001$ . .... | 125 |
| Figure 3.40: Correlogram of monthly changes in nutrients and monthly changes in geosmin concentrations for all reservoirs. Colour-coded by the individual reservoir. Pearson correlation is displayed on the right-hand side of the figure, and variable distribution is available on the diagonal. Correlation significance codes: ‘.’ $p < 0.1$ , $p < 0.05$ ‘*’, ‘**’ $p < 0.01$ , ‘***’ $p < 0.001$ . ....                      | 126 |
| Figure 4.1: Alaw top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location. ....  | 149 |
| Figure 4.2: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Aphanizomenon NIES81 node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the   |     |

nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978472>.

..... 152

Figure 4.3: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Cyanobium PCC-6307 node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link:

<https://rpubs.com/ASHooper/978472>. ..... 154

Figure 4.4: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Snowella OTU37S04 node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link:

<https://rpubs.com/ASHooper/978472>. ..... 156

Figure 4.5: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Nostocaceae family node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978472>.

..... 159

Figure 4.6: Subset of the significant ( $< 0.05$ ) co-occurring taxa associated with the Aphanizomenon NIES81 node during the entirety of this study in the Alwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link:

<https://rpubs.com/ASHooper/978480>. ..... 160

Figure 4.7: Subset of the significant ( $< 0.05$ ) co-occurring taxa associated with the Dolichospermum NIES41 node during the entirety of this study in the Alwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978480>.

Figure 4.8: Subset of the significant ( $< 0.05$ ) co-occurring taxa associated with the Nostocaceae family node during the entirety of this study in the Alwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978480>. ..... 161

Figure 4.9: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Aphanizomenon MDT14a node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>. ..... 164

Figure 4.10: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Pseudanabaena PCC-7429 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>. ..... 165

Figure 4.11: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Aphanizomenon NIES81 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>. ..... 168

Figure 4.12: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Cyanobium PCC-6307 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>. ..... 169

Figure 4.13: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Gloeotrichia PYH6 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>. ..... 171

Figure 4.14: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Microcystis* PCC-7914 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link:

<https://rpubs.com/ASHooper/978493>. ..... 173

Figure 4.15: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the Nostocaceae family node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each genus. Link:

<https://rpubs.com/ASHooper/978493>. ..... 175

Figure 4.16: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Geitlerinema* LD9 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.

..... 179

Figure 4.17: Subset of the significant ( $< 0.1$ ) co-occurring taxa associated with the *Aphanizomenon* MDT14a and Nostocaceae family nodes during the entirety of this study in the Dolwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978497>..... 181

Figure 4.18: Subset of the significant ( $< 0.1$ ) co-occurring taxa associated with the *Dolichospermum* NIES41 node during the entirety of this study in the Dolwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978497>. ..... 182

Figure 4.19: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Aphanizomenon* MDT14a node during the entirety of this study in the Llandegfedd reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial



nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link:

<https://rpubs.com/ASHooper/978520>. ..... 185

Figure 4.20: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Llandegfedd reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978520>. ..... 186

Figure 4.21: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Aphanizomenon* NIES81 node during the entirety of this study in the Llwyn On reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978527>. ..... 189

Figure 4.22: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Llwyn On reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978527>. ..... 190

Figure 4.23: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Pentwyn reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978532>. ..... 193

Figure 4.24: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Dolichospermum* NIES41 node during the entirety of this study in the Pentwyn reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978532>. ..... 195

Figure 4.25: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Nostocaceae* family node during the entirety of this study in the Plas Uchaf reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each genus. Link:

<https://rpubs.com/ASHooper/978536>. ..... 198

Figure 4.26: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Microcystis* PCC-7914 node during the entirety of this study in the Plas Uchaf reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978536>. ..... 199

Figure 4.27: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Pontsticill reservoir. Cyanobacterial nodes are illustrated in green. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978544>. ..... 202

Figure 4.28: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Dolichospermum* NIES41 node during the entirety of this study in the Pontsticill reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978544>. ..... 203

Figure 4.29: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Snowella* out37S04 node during the entirety of this study in the Pontsticill reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978544>. ..... 203

Figure 4.30: Three 16S rRNA NMDS ordination of samples from Alaw ( $n = 34$ ) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring Cyanobacteria from co-occurrence

|  |            |
|--|------------|
| <i>analysis B) significant co-occurring T&amp;O degrading bacteria from co-occurrence analysis C) environmental variables of interest. ....</i>  | <i>206</i> |
| <i>Figure 4.31: Three 16S rRNA NMDS ordination of samples from Cefni (n = 33) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring Cyanobacteria from co-occurrence analysis B) significant co-occurring T&amp;O degrading bacteria from co-occurrence analysis C) environmental variables of interest. ....</i>  |            |
| <i>Figure 4.32: Three 16S rRNA NMDS ordination of samples from Dolwen (n = 6) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring Cyanobacteria from co-occurrence analysis B) significant co-occurring T&amp;O degrading bacteria from co-occurrence analysis C) environmental variables of interest. ....</i>  | <i>207</i> |
| <i>Figure 4.33: Three 16S rRNA NMDS ordination of samples from Llwyn On (n = 20) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring Cyanobacteria from co-occurrence analysis B) significant 16S rRNA genera identified by NMDS analysis (<math>p \leq 0.001</math>) C) environmental variables of interest. ....</i>   |            |
| <i>Figure 4.34: Three 16S rRNA NMDS ordination of samples from Pentwyn (n = 31) reservoir coloured by 2-MIB level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring Cyanobacteria from co-occurrence analysis B) significant co-occurring T&amp;O degrading bacteria from co-occurrence analysis C) environmental variables of interest. ....</i>  | <i>213</i> |
| <i>Figure 4.35: Three 16S rRNA NMDS ordination of samples from Plas Uchaf (n = 22) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring Cyanobacteria from co-occurrence analysis B) significant co-occurring T&amp;O degrading bacteria from co-occurrence analysis C) environmental variables of interest. ....</i>   |            |
| <i>Figure 4.36: Six 16S rRNA NMDS ordination of samples from Pontsticill (n = 24) reservoir coloured by geosmin (left – A, C and E) and 2-MIB level (right – B, D and F) and shaped by sample point locations, arrows represent vectors of A+B) significant co-occurring Cyanobacteria from co-occurrence analysis C+D) significant co-occurring T&amp;O degrading bacteria from co-occurrence analysis E+F) environmental variables of interest. ....</i> | <i>217</i> |

*Figure 4.37: Three rbcL NMDS ordination of samples from Alaw (n = 34) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant rbcL genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring rbcL genera with Cyanobacteria from co-occurrence analysis C) environmental variables of interest. .... 219*

*Figure 4.38: Three rbcL NMDS ordination of samples from Cefni (n = 32) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant rbcL genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring rbcL genera with cyanobacteria from co-occurrence analysis C) environmental variables of interest. .... 221*

*Figure 4.39: Three rbcL NMDS ordination of samples from Dolwen (n = 6) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant rbcL genera governed by NMDS analysis ( $p \leq 0.05$ ) B) significant co-occurring rbcL genera with Cyanobacteria from co-occurrence analysis C) environmental variables of interest. .... 223*

*Figure 4.40: Three rbcL NMDS ordination of samples from Llwyn On (n = 21) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant rbcL genera governed by NMDS analysis ( $p \leq 0.005$ ) B) environmental variables of interest. .... 225*

*Figure 4.41: Three rbcL NMDS ordination of samples from Pentwyn (n = 31) reservoir coloured by 2-MIB level and shaped by sample point locations, arrows represent vectors of A) significant rbcL genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring rbcL genera with Cyanobacteria from co-occurrence analysis C) environmental variables of interest. .... 227*

*Figure 4.42: Three rbcL NMDS ordination of samples from Plas Uchaf (n = 22) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant rbcL genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring rbcL genera with Cyanobacteria from co-occurrence analysis C) environmental variables of interest. .... 229*

*Figure 4.43: Six rbcL NMDS ordination of samples from Pontsticill (n = 25) reservoir coloured by geosmin (left – A, C and E) and 2-MIB level (right – B, D and F) and shaped by sample point locations, arrows represent vectors of A+B) significant rbcL genera governed by NMDS analysis ( $p \leq 0.001$ ) C+D) significant co-occurring rbcL*

|   |            |
|---|------------|
| <i>genera with Cyanobacteria from co-occurrence analysis E+F) environmental variables of interest. ....</i>   | <i>231</i> |
| <i>Figure 4.44: Local variable importance plot displaying the Mean Decrease in Accuracy by each taxon (x-axis) out-of-bag cross-validated prediction (left), and the mean decrease in Gini coefficient measuring how each taxon contributes to the homogeneity of the nodes and leaves in Alaw reservoirs random forest.....</i>      | <i>238</i> |
| <i>Figure 4.45: Multi-dimensional Scaling Plot of Proximity matrix from the Random Forest output for Alaw reservoir grouped by geosmin levels. Ellipses represent a 95% confidence interval. ....</i>   | <i>239</i> |
| <i>Figure 4.46: Grouped bar plots of A) Mean Decrease in Accuracy values and B) the relative abundance (%) of each proposed “true indicator taxa” coloured by geosmin concentration level classification in Alaw reservoir.....</i>   | <i>240</i> |
| <i>Figure 4.47: Grouped bar plots of A) Mean Decrease in Accuracy values and B) the relative abundance (%) of each proposed indicator taxa coloured by geosmin concentration level classification in Alaw reservoir.....</i>  | <i>241</i> |
| <i>Figure 4.48: Local variable importance plot displaying the Mean Decrease in Accuracy by each taxon (x-axis) out-of-bag cross-validated prediction (left), and the mean decrease in Gini coefficient measuring how each taxon contributes to the homogeneity of the nodes and leaves in Llwyn On reservoirs random forest. ....</i> | <i>243</i> |
| <i>Figure 4.49: Local variable importance plot displaying the Mean Decrease in Accuracy by each taxon (x-axis) out-of-bag cross-validated prediction (left), and the mean decrease in Gini coefficient measuring how each taxon contributes to the homogeneity of the nodes and leaves in Pentwyn reservoirs random forest.....</i>   | <i>245</i> |
| <i>Figure 4.50: Multi-dimensional Scaling Plot of Proximity matrix from the Random Forest output for Pentwyn reservoir grouped by 2-MIB levels. Ellipses represent a 95% confidence interval. ....</i>  | <i>246</i> |
| <i>Figure 4. 51: Grouped bar plots of A) Mean Decrease in Accuracy values and B) the relative abundance (%) for the genera with the highest informative contribution to the random forest according to the overall MDA, coloured by 2-MIB concentration level classification in Pentwyn reservoir.....</i>                            | <i>247</i> |
| <i>Figure 5.1: Neighbour-joining (NJ) phylogenetic tree of 2-MIB associated genes using Tamura-Nei genetic distance model, with Streptomyces anulatus as an outgroup. Accession numbers: Plankothrix sp. PD12 (MT515744.1) Pseudanabaena sp. PD8 (MT3602661.1) Pseudanabaena yagii NIES-4238 (LC507459.1)</i>                         |            |

*Pseudanabaena* sp. 1 (LC507462.1) *Pseudanabaena* *yagii* NIES-4237 (LC507458.1)  
*Pseudanabaena* sp. NIVA-CYA 1119 (HQ630887.1) *Pseudanabaena* sp. 2  
(HQ830028.1) *Pseudanabaena* *galeata* NIES-512 (AB826230.1) *Pseudanabaena*  
*galeata* (MN167115.1) *Pseudanabaena* *limnetica* (HQ630883.1) *Pseudanabaena*  
*cinerea* (LC507461.1) *Pseudanabaena* *cinerea* NIES-4062 (LC507460.1)  
*Microcoleus* *pseudautumnalis* (LC486303.1) *Oscillatoria* *limosa* (HQ630885.1)  
*Leptolyngbya* sp. (KP013063.1) *Oscillatoria* *prolifera* (MK759878.1) *Planktothrix* sp.  
328 (KJ658378.1) *Oscillatoria* *prolifera* (MK759878.1) *Planktothricoides* *raciborskii* 1  
(LC512931.1) *Planktothricoides* *raciborskii* 2 (LC157992.1) *Planktothricoides*  
*raciborskii* 3 (LC157991.1) *Planktothricoides* *raciborskii* 4 (LC157990.1)  
*Planktothricoides* *raciborskii* 5 (LC157989.1) *Planktothricoides* *raciborskii* 6  
(LC157988.1) *Planktothricoides* *raciborskii* 7 (LC157987.1) *Planktothricoides*  
*raciborskii* 8 (LC157986.1) *Planktothricoides* *raciborskii* 9 (HQ830029.1). Genbank  
Accession numbers are provided with associated version identifiers. .... 277  
Figure 5.2: Sequence view of the newly developed primer MIBHR (24 bp). ..... 278  
Figure 5.3: An overview of plate setups for the Eppendorf epMotion® 5075 robot.  
One 96-well plate setup for diluted samples (1:100 16S rRNA, 1:10 geoA and mic)  
and another 96-well plate for the reagent stock pates and standards. A general map  
for the qPCR reaction in a 384-well plate. .... 279  
Figure 5.4: Primer positions and alignment of the amplicon on *Anabaena* *ucrainica*  
CHAB2155 geosmin synthesis associated operon, complete sequence (HQ404996).  
..... 281  
Figure 5.5: Primer positions and alignment of the amplicon on *Pseudoanabaena* sp.  
*dqp15* 2-methylisoborneol (2-MIB) synthesis associated operon, complete sequence  
(HQ830028.1). .... 281  
Figure 5.6: Histograms of (A) the frequency of 16S rRNA copy numbers from this  
study (red – AH.Frequency) and the frequency of copy numbers of 16S rRNA  
processed previously (blue – IE.Frequency) (B) frequency of 16S rRNA copy  
numbers from this study (red – AH.Frequency) and the frequency of copy numbers  
of 16S rRNA processed previously (blue – IE.Frequency) after the calibration  
coefficient had been added. .... 284  
Figure 5.7: Scatterplots of log<sub>10</sub> concentrations of geosmin (ng L<sup>-1</sup>) and geoA:16S  
(copy numbers mL<sup>-1</sup>) from all reservoirs. Individual points are coloured  
corresponding to the reservoir and facet wrapped according to sampling month and

year. Each scatterplot includes a linear regression line of best fit with the  $R^2$  value and associated significance assigned by  $p$  values. .... 285

Figure 5.8: Scatterplots of  $\log_{10}$  concentrations of 2-MIB ( $\text{ng L}^{-1}$ ) and mic:16S (copy numbers  $\text{mL}^{-1}$ ) from all reservoirs. Individual points are coloured corresponding to the reservoir and facet wrapped according to sampling month and year. Each scatterplot includes a linear regression line of best fit with the  $R^2$  value and associated significance assigned by  $p$  values. .... 286

Figure 5.9: Scatterplots of  $\log_{10}$  concentrations of geosmin ( $\text{ng L}^{-1}$ ) and geoA:16S (copy numbers  $\text{mL}^{-1}$ ) from all reservoirs. Individual points are coloured corresponding to the reservoir and facet wrapped according to sampling season and year. Each scatterplot includes a linear regression line of best fit with the  $R^2$  result and associated significance assigned by  $p$  values. .... 287

Figure 5.10: Scatterplots of  $\log_{10}$  concentrations of 2-MIB ( $\text{ng L}^{-1}$ ) and mic:16S (copy numbers  $\text{mL}^{-1}$ ) from all reservoirs. Individual points are coloured corresponding to the reservoir and facet wrapped according to sampling season and year. Each scatterplot includes a linear regression line of best fit with the  $R^2$  result and associated significance assigned by  $p$  values. .... 288

Figure 5.11: Smooth function plots for predictor variables in the Alaw geoA:16S copy numbers  $\text{mL}^{-1}$  GAM. Estimated smooth functions (solid lines) with 95% confidence intervals (grey shaded area) are shown for each smoothed predictor: (top) Geosmin concentrations  $\text{ng L}^{-1}$  (middle)  $\text{NO}_3^-:\text{NH}_4^+$  ratio and (bottom) TIN:TP ratio. The y-axis denotes the partial effect size, the comb on the x-axis shows where the value of predictor data points lie and the points are the residuals. The horizontal red line at  $y = 0$  intercept indicates the overall mean of the response (geoA abundance). .... 291

Figure 5.12: Smooth function plots for predictor variables in the control geoA:16S copy numbers  $\text{mL}^{-1}$  GAM. Estimated smooth functions (solid lines) with 95% confidence intervals (grey shaded area) are shown for each smoothed predictor: (top) Geosmin concentrations  $\text{ng L}^{-1}$  (middle)  $\text{NO}_3^-:\text{NH}_4^+$  ratio and (bottom) TIN:TP ratio. The y-axis denotes the partial effect size, the comb on the x-axis shows where the value of predictor data points lie and the points are the residuals. The horizontal red line at  $y = 0$  intercept indicates the overall mean of the response (geoA abundance). .... 294

Figure 5.13: Smooth function plots for predictor variables in the mic:16S copy numbers  $\text{mL}^{-1}$  GAM. Estimated smooth functions (solid lines) with 95% confidence

*intervals (grey shaded area) are shown for each smoothed predictor: (top) 2-MIB concentrations ng L<sup>-1</sup> (middle) NO<sub>3</sub>:NH<sub>4</sub><sup>+</sup> ratio (bottom) TIN:TP ratio and (C). The y-axis denotes the partial effect size, the comb on the x-axis shows where the value of predictor data points lie and the points are the residuals. The horizontal red line at y = 0 intercept indicates the overall mean of the response (mic abundance)..... 297*

*Figure 6.1: Illustrations of linear regression models conducted on the combined relative abundances of significantly co-occurring T&O-producing Cyanobacteria against geosmin concentrations in A) the Alaw reservoir and 2-MIB concentrations in B) the Pentwyn reservoir. Associated p and r<sup>2</sup> values are located in the top left corner..... 309*

*Figure 6.2: Example of ways a cyanobacterial equilibrium state can vary with environmental stressors that create suboptimal conditions for Cyanobacteria and lead to T&O events. In a) and b), only one equilibrium exists for each environmental stressor. However, if the equilibrium curve is folded backwards (c), three equilibria can exist for a given environmental stressor. It can be seen from the red arrows indicating the direction of change that, in this case, equilibria on the dashed middle section (between green balls) are unstable and represent the transitional state between the two states..... 322*

*Figure 6.3: Diagram illustrating the “cyanosphere stable state theory”, displaying A) the stable cyanosphere state where geoA abundance remains at a steady level with stable cyanobacterial productivity and hence low geosmin production and B) the unstable cyanosphere state where geoA abundance increases directly proportional to increases in cyanobacterial productivity leading to increased geosmin production. The size of the green balls is representative of geoA abundance..... 324*



## List of Tables

|   |            |
|---|------------|
| <i>Table 1.1: Chemical structure, molecular weight, and the molecular formulas of geosmin and 2-MIB. ....</i>   | <i>5</i>   |
| <i>Table 2.1: Sampled months calculated from the start date (7/2019) to the end date (9/2020) of this study period. ....</i>  | <i>30</i>  |
| <i>Table 2.2: Primer sets used. The primer sets were analysed using the integrated DNA technologies OligoAnalyzer tool (Owczarzy et al., 2008). ....</i>                    | <i>46</i>  |
| <i>Table 2.3: PCR reaction composition for 16S rRNA and rbcL gene. ....</i>   | <i>47</i>  |
| <i>Table 2.4: PCR conditions for 16S rRNA and rbcL gene. ....</i>   | <i>47</i>  |
| <i>Table 2.5: PCR conditions used for Illumina PCR. ....</i>  | <i>49</i>  |
| <i>Table 2.6: Parameters used for all samples processed through the DADA2 bioinformatics pipeline for both 16S rRNA and rbcL genes. ....</i>                                | <i>52</i>  |
| <i>Table 3.1: Trophic state classification scheme based on Carlson's (1977) trophic state index. ....</i>   | <i>61</i>  |
| <i>Table 3.2: CTSI values assigned to trophic status of the reservoir and relevant attributes characterising the trophic state (Carlson and Simpson, 1996). ....</i>        | <i>61</i>  |
| <i>Table 3.3: Sampled months used to calculate the monthly changes in nutrients and T&amp;O concentrations for each reservoir. ....</i>                                     | <i>64</i>  |
| <i>Table 4.1: Total 16S rRNA and rbcL taxa present at <math>\geq 1\%</math> proportion. ....</i>  | <i>145</i> |
| <i>Table 4.2: p-values used for filtering significant positive and negative co-occurring 16S rRNA and rbcL genera and the number of samples used for analysis. ....</i>     | <i>145</i> |
| <i>Table 4.3: Online binomial 16S rRNA and rbcL networks for each reservoir and the p-value used to determine negative and positive relationships between genera. ....</i>  | <i>150</i> |
| <i>Table 4.4: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Alaw reservoir. ....</i>   | <i>153</i> |
| <i>Table 4.5: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Alwen reservoir. ....</i>  | <i>162</i> |
| <i>Table 4.6: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir. ....</i>  | <i>166</i> |
| <i>Table 4.7: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Dolwen reservoir. ....</i> | <i>183</i> |

|  |     |
|--|-----|
| <i>Table 4.8: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Llandegfedd reservoir.</i>  | 187 |
| <i>Table 4.9: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within Llwyn On reservoir.</i>   | 191 |
| <i>Table 4.10: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Pentwyn reservoir.</i>   | 194 |
| <i>Table 4.11: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Plas Uchaf reservoir.</i>  | 200 |
| <i>Table 4.12: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Pontsticill reservoir.</i>   | 204 |
| <i>Table 4.13: Results from random forest classification of T&amp;O status from significantly co-occurring taxa from reservoirs experiencing high T&amp;O concentrations.</i>  | 233 |
| <i>Table 4.14: An overview of indicative taxa and their function-specific traits for the production of geosmin as determined by Random Forest analysis in the Alaw reservoir.</i>  | 264 |
| <i>Table 4.15: Additional indicator taxa taken from co-occurrence analysis for geosmin production based on shared function-specific traits with indicator taxa identified by Random Forest analysis in the Alaw reservoir.</i> | 265 |
| <i>Table 5.1: List of strains used for standard curves in all qPCR reactions with their origin and the putative 16S rRNA classification after genomic classification with corresponding percentage identity.</i>               | 273 |
| <i>Table 5.2: Isolated bacterial strains used for the qPCR template preparation for 16S rRNA, geoA and mic used in the standard curve set-up.</i>  | 274 |
| <i>Table 5.3: Primer pairs used to amplify 16S rRNA, geoA and mic gene templates during the initial PCR when generating qPCR standards.</i>  | 274 |
| <i>Table 5.4: PCR master mix volumes for a 50 µL reaction volume for primer pairs 27F and 1492R, geo78F and geo982R and MIBS02F and MIBS02R respectively.</i>  | 275 |
| <i>Table 5.5: PCR conditions used for 16S rRNA primer pair 27F and 1492R.</i>  | 275 |
| <i>Table 5.6: The PCR conditions used for geoA primer pair 78F and 982R.</i>   | 275 |
| <i>Table 5.7: The PCR conditions used for mic primer pair MIBS02F and MIBS02R.</i>   | 276 |
| <i>Table 5.8: Triplicate Qubit readings for both 16S rRNA, geoA and mic concentrations and an average in ng µL<sup>-1</sup>.</i>   | 278 |

|   |            |
|---|------------|
| <i>Table 5.9: Primer pairs for 16S rRNA, geoA and mic used for the qPCR.....</i>  | <i>280</i> |
| <i>Table 5.10: qPCR master mix volumes for a 10 µL reaction volume for primer pairs 534F, 907R and geo799F geo982R and MIBS02F and MIBAHR respectively. ....</i>  | <i>282</i> |
| <i>Table 5.11: The qPCR conditions used for 16S rRNA primer pair 534F and 907R.</i>   | <i>282</i> |
| <i>Table 5.12: The qPCR conditions used for geoA primer pair geo799F and geo982R. ....</i>  | <i>282</i> |
| <i>Table 5.13: The qPCR conditions used for mic primer pair MIBS02F and MIBAHR. ....</i>  | <i>282</i> |
| <i>Table 5.14: GAM model results for Alaw reservoir with geoA:16S copy numbers mL<sup>-1</sup> as the response variable. Using summer 2019 as the reference level for seasonal comparison.....</i>  | <i>290</i> |
| <i>Table 5.15: GAM model results for all reservoirs except Alaw, with geoA:16S copy numbers mL<sup>-1</sup> as the response variable. Using winter 2019 as the reference level for seasonal comparison and Llandegfedd as the reference level for reservoir comparison.....</i> | <i>293</i> |
| <i>Table 5.16: GAM model results with mic:16S copy numbers mL<sup>-1</sup> as the response variable. Using winter 2019 as the reference level for seasonal comparison, and Llandegfedd as the reference level for reservoir comparison.....</i>                                 | <i>296</i> |
| <i>Table 6.1: Buffer mechanisms maintaining stable and unstable states through bottom-up and top-down approaches involved in the cyanosphere stable state theory. ....</i>  | <i>325</i> |

# Chapter 1: General Introduction



## 1.1 – General introduction

The aesthetic properties of drinking water, such as Taste and Odour (T&O), are used by customers to assess the quality of their drinking water (Zhou et al., 2017). T&O associated with drinking water is often perceived to be linked to contamination and is a common cause of complaints to water companies worldwide (Krasner et al., 1983; Webber et al., 2015). T&O associated with drinking water is becoming an increasing problem for water companies as T&O problems in drinking water reservoirs are increasing in frequency and magnitude globally (Winter et al., 2011). Most reported cases of T&O events are said to be biological in origin (Watson, 2003). Although, tracing the sources of biological T&O has proven to be challenging, requiring expert knowledge of both abiotic and biotic factors contributing to these events. Therefore, it is vital to evaluate the environmental triggers that can influence T&O events and use this information to aid water companies in structuring their water monitoring and implementing measures to help prevent T&O events.

T&O imparted on source waters is suspected to be caused by one or more of the many volatile organic compounds often present at any given time, yet most odour events are associated with a small number of volatile compounds (Watson, 2010). Generally, the T&O compounds brought to attention are earthy-musty odorants (Gerber and Brunswick, 1965; Olsen et al., 2016). The odour threshold concentrations of these compounds can be exceptionally low ( $\leq 10 \text{ ng L}^{-1}$ ; Ömür-Özbek et al., 2007); as a result, they are easily detected in drinking water, even in small concentrations. Consequently, this leads to customer distrust in the water they consume, culminating in customer complaints to the water industry.

A range of algae and bacteria produces T&O compounds; however, *Cyanobacteria* are widely considered the most common source of the earthy-musty T&O compounds (Otten et al., 2016). The molecules trans-1,10-dimethyl-trans-9-decalol (geosmin) and 2-methylisoborneol (2-MIB), produced by *Cyanobacteria*, impart an earthy/musty/mouldy T&O in the water column (Izaguirre et al., 1982) and are currently at the forefront of attention concerning T&O events. Large proportions of geosmin and 2-MIB get retained in the cells of *Cyanobacteria*, but the liberation of the molecules can be initiated through cell rupture by physical processes such as

grazing or treatment (Liato and Aïder, 2017). Depending on the growth phase and environmental factors, *Cyanobacteria* store or release geosmin and 2-MIB (Watson et al., 2016). For example, *Dolichospermum circinalis* can release geosmin during the stationary phase due to cell lysis, with intracellular geosmin at its greatest during the exponential growth phase (Rosen et al., 1992). Previous hypotheses suggest that the release of intracellular pools of geosmin/2-MIB occurs during the stationary phase or at the onset of senescence (Watson, 2003). Lysis of cyanobacterial cells releases any remaining intracellular geosmin/2-MIB into the water column (Lee et al., 2017; Watson et al., 2008), which explains why heightened T&O concentrations can be associated occasionally with high biomass.

## 1.2 – Geosmin and 2-methylisoborneol

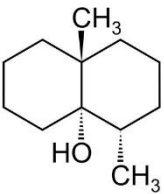
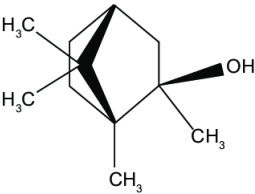
The two secondary metabolites, geosmin and 2-MIB, are produced by *Cyanobacteria* and a range of specific microorganisms, including *Actinomycetes*, *Myxobacteria* and fungi (Jüttner and Watson, 2007), but are mainly associated with *Cyanobacteria* in waters where photosynthesis can occur. Occurrences of earthy-musty odours in source waters are historically reported – having first been described in the 1850s (Wood et al., 2001). T&O compounds, geosmin and 2-MIB, responsible for T&O imparted on drinking water were described in the 1960s, with geosmin identified in 1965 (Gerber and Brunswick, 1965) and 2-MIB in 1969 (Gerber, 1969).

Geosmin and 2-MIB are highly potent, with human odour threshold concentrations being detectable between 4 and 10 ng L<sup>-1</sup> (Chong et al., 2018). Geosmin and 2-MIB are tertiary alcohols that exist as (+) and (-) enantiomers (Table 1.1: Jüttner and Watson, 2007). Geosmin and 2-MIB outbreaks originate from the (-) enantiomer and are 10 times more potent than the (+) forms (Jüttner and Watson, 2007). The two secondary terpenoids are ubiquitous and most problematic in freshwater ecosystems, and they are exceedingly resistant to natural degradation, boiling, and conventional treatment processes (Kim et al., 2016).

There is currently no known function of both T&O compounds, although terpenoid compounds such as geosmin primarily function as chemicals that exert a beneficial or detrimental physiological effect on individuals of another species, such as in plants (Kiran et al., 2007). T&O compounds appear to be precursors for

important cellular compounds such as sterols and pigments (Bentley and Meganathan, 1981; Watson, 2004). Demonstrations of geosmin derived from *Cyanobacteria* linked to essential functions exist, for example, geosmin suppressing the growth of the green alga *Chlorella pyrenoidosa* (Ikawa et al., 2001). In addition, the geosmin gene cluster has two global transcription regulator genes known to modify cellular signals associated with responses to environmental stress (Zhou et al., 2012). Thus, the highly conserved arrangement of the cluster in *Cyanobacteria* could suggest its importance in adapting to environmental stresses, indicating that the synthesis of geosmin and other volatile organic compounds, like 2-MIB, could be related to defence/offence methods towards other microorganisms (Asquith et al., 2013; Wang et al., 2013). Paerl and Millie (1996) and Watson (2003) suggested that secondary metabolites, like geosmin and 2-MIB, are released to remove excess metabolites during periods of environmental stress.

Table 1.1: Chemical structure, molecular weight, and the molecular formulas of geosmin and 2-MIB.

| Compound Name                                 | Compound Structure  | Molecular weight | Molecular formula                 |
|---|---|------------------|-----------------------------------|
| Trans-1,10-dimethyl-trans-9-decalol (Geosmin) |  | 182.3            | C <sub>12</sub> H <sub>22</sub> O |
| 2-Methylisoborneol (2-MIB)                    |  | 168.3            | C <sub>11</sub> H <sub>20</sub> O |



### 1.2.1 – Metabolic pathways of geosmin and 2-MIB

Bentley and Meganathan (1981) were the first to investigate the biosynthetic pathways of geosmin and 2-MIB metabolism in the genus *Streptomyces* using radio-gas chromatography. Bentley and Meganathan (1981) hypothesised that geosmin and 2-MIB get synthesised through an isoprenoid pathway (the terpenoid or mevalonate pathway). In the isoprenoid pathway, a monoterpene precursor, geranyl diphosphate (GPP), existed for 2-MIB, and geosmin had a sesquiterpene precursor, farnesyl diphosphate (FPP). Subsequent studies also confirmed the same pathway for *Cyanobacteria* (Naes et al., 1989; Zimba, et al., 1999). In *Cyanobacteria*, synthesis of geosmin and 2-MIB originate from the methylerythritol-phosphate (MEP), the mevalonate (MVA), and the Leucine pathways along with other terpenoids that are vital for the growth and survival of photosynthetic organisms (Jüttner and Watson, 2007; Pattanaik and Lindberg, 2015). The MEP pathway is considered an alternate or concomitant pathway to the MVA and Leucine pathways, all found in *Cyanobacteria* (Perkins et al., 2019) (Figure 1.1). The MEP pathway is known for synthesising carotenoids, chlorophyll and other terpenoid products (Pattanaik and Lindberg, 2015).

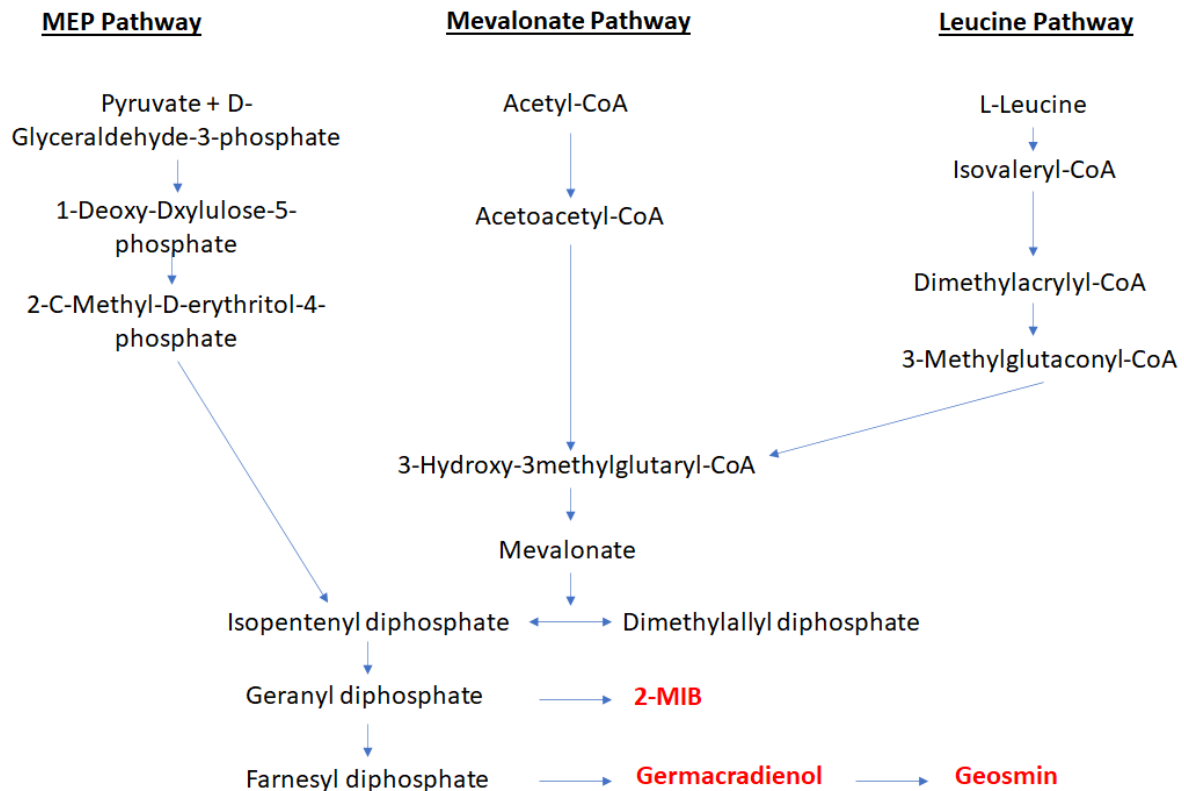


Figure 1.1: Simplified biosynthetic pathways for the formation of geosmin and 2-MIB.

### 1.2.2 – Geosmin synthase (*geoA*) and 2-MIB cyclase (*mic*)

The molecular foundation of geosmin and 2-MIB production comes from geosmin synthase (*geoA*) and 2-MIB cyclase (*mic*) genes that encode for enzymes responsible for the conversion of the intermediate precursors, as shown in Figures 1.2 – 1.3. Churro et al. (2020) found that *Cyanobacteria* have a very conserved *geoA* and share similar gene architecture with *Deltaproteobacteria*, with a three-gene operon assembly. Theories suggest that due to the distribution of *geoA* not being heterogeneous within *Cyanobacteria*, *geoA* is transferred through horizontal gene transfer (Wang et al., 2019). From the eight orders Churro et al. (2020) described, *Oscillatoriales*, *Nostocales*, and *Synechococcales* possess the majority of known geosmin producers. The organisation of the gene cluster in *Cyanobacteria* is highly conserved, except for *Phormidium* sp. (Ludwig et al., 2007). The *geoA* cluster has two global transcription regulator genes, previously described as modulating cellular signals associated with responses to environmental stress (Zhou et al., 2012).

Geosmin is a sesquiterpene formed from converting the precursor farnesyl diphosphate (FPP) through two steps catalysed by a bi-functional geosmin synthase encoded by the *geoA* gene in the presence of  $Mg^{2+}$  (Jiang et al., 2007). The enzyme first catalyses the conversion of FPP to germacradienol by its N-terminal domain, and finally to geosmin by its C-terminal domain (Jiang et al., 2007).

Wang et al. (2011) demonstrated that the cyanobacterial production of 2-MIB originates from two adjacent genes, the SAM-dependent methyltransferase (*mtf*) gene and the *mic* gene. The *mtf* and *mic* genes are accountable for geranyl diphosphate (GPP) methylation and consequent cyclization to 2-MIB. These two genes located between two homologous cyclic nucleotide-binding protein genes collectively form a putative operon. The cyclic nucleotide-binding protein genes were predicted members of the Crp-Fnr family; universal regulators of photosynthesis and nitrogen fixation in many microorganisms (Körner et al., 2003). 2-MIB synthesis consists of two crucial reactions: SAM-dependent 2-C-methyltransferase (*mtf*) catalysed methylation of GPP producing 2-methyl-GPP, and finally, 2-MIB cyclase (*mic*) catalysed cyclization of the 2-methyl-GPP to 2-MIB (Giglio et al., 2011).

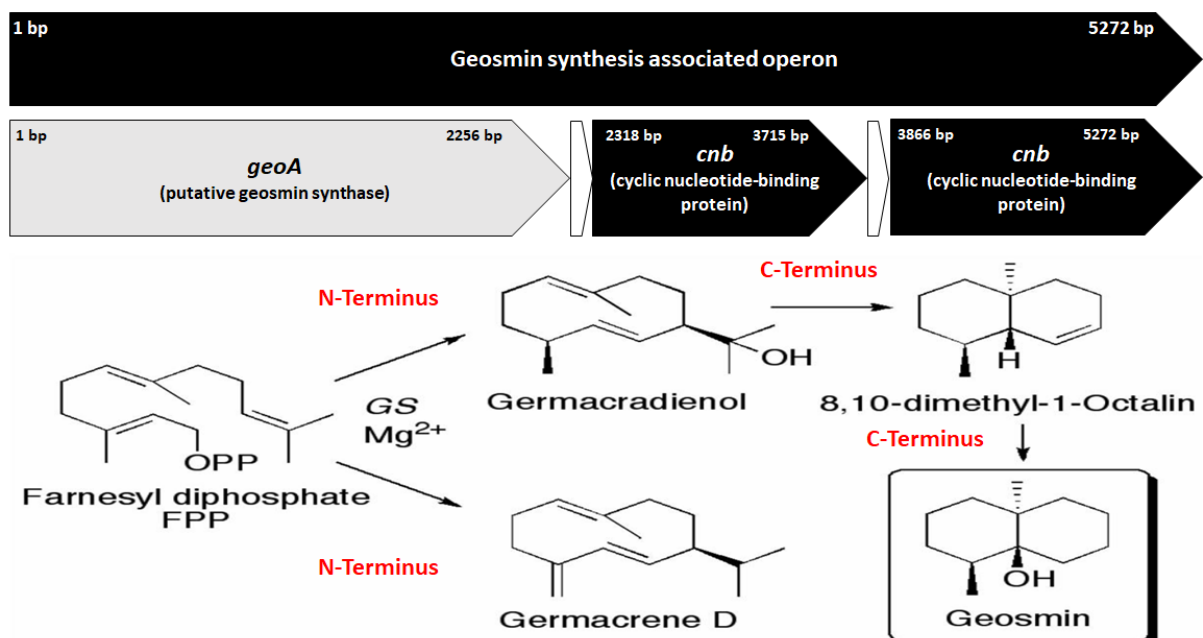


Figure 1.2: An annotated *Anabaena ucrainica* CHAB2155 geosmin synthesis operon (HQ404996) and the associated biosynthetic pathway for geosmin production.

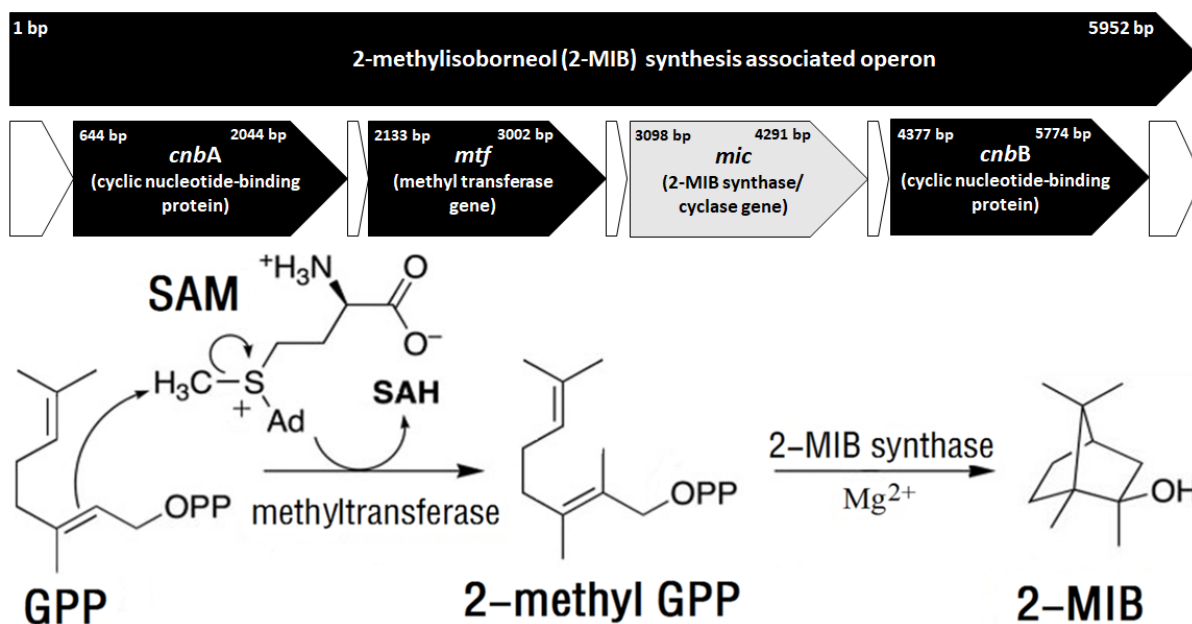


Figure 1.3: An annotated *Pseudanabaena sp. Dqp15* (HQ830028.1) 2-MIB synthesis operon and the associated biosynthetic pathway for 2-MIB production.

### 1.3 – Adaptations of *Cyanobacteria*

*Cyanobacteria* produce oxygen, reduce carbon dioxide (CO<sub>2</sub>) and fix nitrogen (N<sub>2</sub>) in aerobic conditions, and therefore play a significant role in the nitrogen and carbon cycles (Parmar et al., 2011). *Cyanobacteria* have diversified throughout their evolution, producing both unicellular and multicellular forms containing specialised cells (e.g., heterocysts), enabling the fixation of atmospheric N<sub>2</sub> (Flores and Herrero, 2010). Specialised cells, like heterocysts, provide *Cyanobacteria* with a competitive advantage when the ratio of nitrogen to phosphorous is low, and nitrogen availability is limiting growth rates (Tilman et al., 1982). Phosphorous is vital in regulating primary productivity in *Cyanobacteria* (Lu et al., 2019). The dissolved organic phosphate pool consists of diverse compounds, including phosphoesters and phosphonates (Kathuria and Martiny, 2011). The enzyme alkaline phosphatase (APase) produced by *Cyanobacteria* readily hydrolyse Phosphoesters (Štrojsová et al., 2003). *Cyanobacteria* also possess a high affinity for, and ability to store, phosphorous using polyphosphate synthesis is known as “luxury uptake”, whereby under conditions of phosphorous in excess, *Cyanobacteria* take up phosphorous

exceeding growth demand to accumulate polyphosphate as a phosphorous reserve (Li and Dittrich, 2018). Bar-Yosef et al. (2010) reported *Cyanobacteria* exploiting the APase pathway in green algae, showing phosphatase expression in response to a cyanotoxin (cylindrospermopsin) signal from the *Cyanobacterium*, *Aphanizomenon ovalisporum*. While *A. ovalisporum* can produce extracellular APase, its first response is to induce cells of other phytoplankton to produce extracellular APase for *A. ovalisporum* to use (Raven, 2010). More specifically, cylindrospermopsin initiates the production of APase in other phytoplankton species to obtain inorganic phosphate (Raven, 2010).

In addition to mechanisms *Cyanobacteria* utilise to attain nutrients, many planktonic *Cyanobacteria* can produce gas vesicles and/ or alter cell density to optimise their position in the water column through buoyancy hence optimising light and nutrient availability (Walsby, 1994). Other cell adaptations for filamentous forms of *Cyanobacteria*, including the genera *Oscillatoria* and *Dolichospermum*, include akinetes, a resistant spore tolerant to cold and desiccation, aiding their survival in adverse conditions and cold winters. Germination of akinetes occurs when the environmental conditions are favourable for the growth of a vegetative filament (Soule and Garcia-Pichel, 2019).

### 1.3.1 – Cyanobacterial T&O productivity driven rather than biomass-related

Correlation-based analysis of T&O concentrations and cyanobacterial abundance is well studied, focusing mainly on T&O originating from cyanobacterial blooms (Kim et al., 2015; Wang et al., 2016). Although, some studies fail to find a correlation between geosmin/2-MIB concentrations and cyanobacterial biomass over time (Graham et al., 2010; Seto et al., 1996; Wang et al., 2016; Watson et al., 2008; Xuwei et al., 2019). However, periods of high productivity have been associated with high geosmin and 2-MIB production in *Cyanobacteria* (Zimmerman et al., 1995). Previous evidence indicates that productivity-driven T&O events would explain why heightened 2-MIB concentrations occur at times of high ATP synthesis (Behr et al., 2014). Consequently, caution must be applied when using water quality models, such as Chong et al. (2018), which are based on cyanobacterial abundance, to predict the on-set of T&O events.

## 1.4 – Environmental triggers

A range of abiotic environmental factors such as nutrients (Amé and Wunderlin, 2005; Schindler, 2012; Van de Waal et al., 2014), light (Alghanmi et al., 2018; Wang and Li, 2015; Zhang et al., 2009), temperature (Shen et al., 2022), hydrology (Havens et al., 2019) and stratification (Merino-Ibarra et al., 2021) along with biotic interactions such as predation (Zuo, 2019), commensalism (Zhang et al., 2021), mutualism (Ramanan et al., 2016), resource competition (Holland and Kinnear, 2013; Lewington-Pearce et al., 2019), and parasitism (Ramanan et al., 2016) shown to influence cyanobacterial growth and productivity, and hence T&O outcome. However, the results of previous findings differ in conclusions, and there is no explicit understanding of how these environmental factors trigger T&O events. The uncertainty of influential environmental triggers for T&O production highlights the requirement for further investigation in determining the specific abiotic and biotic environmental triggers underpinning cyanobacterial T&O production.

### 1.4.1 – Abiotic triggers

#### 1.4.1.1 – Nutrients

T&O production is linked with the rapid growth phases of *Cyanobacteria* when subjected to optimal growth conditions when productivity is high (Alghanmi et al., 2018; Perkins et al., 2019). Initially, phosphorous was considered the primary nutrient limiting the growth of *Cyanobacteria* (Schindler, 1974, 1977), despite extensive evidence suggesting that nitrogen limitation is a common, widespread phenomenon in reservoirs (Elser et al., 2007). Nutrient enrichment of water columns is attributable to the increased application of nitrogen fertilisers (Balls et al., 1995) and slurry (Chapuis-Lardy et al., 2003), all of which can alter nitrogen to phosphorous loading in already nutrient-rich waters. Generally, cyanobacterial growth partly depends on the TN:TP ratio, with cyanobacterial dominance often intensifying when the ratio is low (<29:1 by mass; Smith, 1983). Additionally, relative increases in phosphorous limitation leading to the subsequent stimulation of geosmin production have previously shown negative correlations with orthophosphate (Dae-

Kyun et al., 2001). Orthophosphate is the most readily utilisable form of phosphorous available to *Cyanobacteria*, with evidence of significant correlations existing between orthophosphate and geosmin concentrations (Perkins et al., 2019). Orthophosphate concentrations, in turn, affect the TN:TP ratio as orthophosphates constitute the organic phosphorous fraction. However, in eutrophic systems where nutrients are not limited, the TN:TP ratio appears less predictive of cyanobacterial dominance (Paerl and Fulton, 2006). Harris et al. (2016) found that simultaneous decreases in TN:TP and nitrate to ammonia ( $\text{NO}_3^-:\text{NH}_3^+$ ) ratios can potentially create prime conditions favouring T&O events. Harris et al. (2016) proposed that decreases in cellular energetics when switching from oxidized ( $\text{NO}_3^-$ ) to reduced forms of nitrogen ( $\text{NH}_3^+$ ) might allow reallocation of energy from reducing oxidized forms of nitrogen to build carbon-rich secondary metabolites (like geosmin and 2-MIB). Thus, ammonia may pose a key trigger in initiating T&O events. Likewise, Perkins et al. (2019) identified ammonium, the most reduced form of nitrogen in aquatic systems, as a critical trigger for T&O production in *Cyanobacteria*. Similarly, Howard (2020) found that low TN:TP ratios favoured cyanobacterial dominance and growth, while low  $\text{NO}_3^-:\text{NH}_3^+$  ratios were associated with increased T&O production. The ratio of oxidized nitrogen (nitrate and nitrite) to reduced nitrogen (ammonia and ammonium) has significant consequences for the growth rate, mineral composition, and production of carbon-rich organic compounds like geosmin and 2-MIB (Domingues et al., 2011). Furthermore, Domingues et al. (2011) demonstrated that increases in TP or  $\text{NH}_3^+$  (or both), causing decreases in TN:TP and  $\text{NO}_3^-:\text{NH}_3^+$  ratios, respectively, may initiate the metabolic capability in *Cyanobacteria* to produce T&O compounds.

In addition to nitrogen and phosphorous, Sinang et al. (2015) found that cyanobacterial biomass and dominance correlated with iron concentrations. Iron is a vital nutrient for nitrogen-fixing *Cyanobacteria* (Larson et al., 2018). Molot et al. (2014) proposed a novel model for cyanobacterial dominance, linking anoxia, phosphorous, nitrogen, iron, and sulphate to the formation of harmful cyanobacterial blooms. This model suggests that phosphorous regulates cyanobacterial biomass and productivity until excessive phosphorous renders the water column either nitrogen or light limited. However, the availability of ferrous ions ( $\text{Fe}^{2+}$ ) regulates the cyanobacterial ability to outcompete its eukaryotic counterparts. The scarcity of  $\text{Fe}^{2+}$  in phosphorous-limited oxygenated waters can severely limit the growth of

*Cyanobacteria* if not supplemented by migrating down to anoxic Fe<sup>2+</sup>-rich waters (Molot et al., 2014). Some *Cyanobacteria* can also produce siderophores to enable iron acquisition under iron-limited conditions (Hopkinson and Morel, 2009). Siderophores are absent in picocyanobacteria, yet some *Synechococcus* can produce siderophores (Kranzler et al., 2013). In parallel with Fe<sup>2+</sup>, sulphate reduction to sulphide has proven to limit Fe<sup>2+</sup> diffusion rates from anoxic sediments because of the insoluble iron sulphide formation (Carignan and Tessier, 1988).

#### 1.4.1.2 – Light and temperature

Light is harvested in *Cyanobacteria* by chlorophyll-containing photosystems immersed in the thylakoid membranes and phycobilisomes, creating the photosystem-associated light-harvesting antennae (Montgomery, 2014). *Cyanobacteria* have evolved many photoreceptor-controlled mechanisms, particularly in their photosynthetic activities, that enable *Cyanobacteria* to sense and react to environmental light conditions (Allahverdiyeva et al., 2015; Gutu and Kehoe, 2012; Ho et al., 2017). For example, complementary chromatic adaptation supports *Cyanobacteria* with the ability to tailor their light-harvesting antennae to the spectral distribution of ambient light (Gutu and Kehoe, 2012). Complementary chromatic adaptation also enforces light-dependent alterations of cellular or filament morphology (Bennett and Bogorad, 1973; Bordowitz and Montgomery, 2008). Morphological phenotypes can shift between spherical and rod-shaped cells under changes in prevalent wavelengths and intensity of light-induced complementary chromatic adaptation (Pattanaik et al., 2012; Walters et al., 2013). Thus, complementary chromatic adaptation increases the fitness of *Cyanobacteria*, allowing them to acclimatise to low light conditions, which could explain T&O events in winter months (Parinet et al., 2010; Wang et al., 2014).

Wang et al. (2011) elucidated the genetic background, light regulation, and biochemical mechanisms of 2-MIB synthesis in *Cyanobacteria* using genome walking and Polymerase Chain Reaction (PCR) methods. Using light-induced reactions, Wang et al. (2011) identified the inhibition of expressed 2-MIB-associated genes by low light and activated by intense light. These findings were concurrent with previous geosmin synthesis studies described by Bowmer et al. (1992) and Zhang et al. (2009). Increases in the geranyl diphosphate pool during periods of



reduced photosynthesis in low-light environments could activate the transcription of 2-MIB synthesis genes, while geranyl diphosphate depletion under intense light may turn transcription off (Wang et al., 2011). In addition, Wang et al. (2011) reported that methyltransferase (*mtf*) and *mic* genes were still active in dark conditions, indicating that other internal or external factors may activate and regulate the synthesis of 2-MIB in *Cyanobacteria*. Ludwig et al. (2007) demonstrated a lack of expression of *geoA* after 24 hours in the dark; it is unclear whether the differences in results from these studies reflect different regulatory mechanisms for geosmin and 2-MIB expression.

Environmental variables such as light and temperature can influence *Cyanobacteria*'s production of T&O compounds (Tung, 2004). Although, a study conducted by Alghanmi et al. (2018) revealed no correlation between temperature and light intensity with geosmin and 2-MIB production. However, the optimal temperature and light intensity for the growth and production of these compounds were 25 °C and 17  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  (Alghanmi et al., 2018). Extra- and intracellular geosmin and 2-MIB was affected by altering the light intensity and temperatures in this study; this study's heightened concentrations of extracellular geosmin arose during the stationary and death phases. Intracellular geosmin and 2-MIB concentrations were high during the late exponential phase and decreased with the growth decline. According to Shen et al. (2022), the expression levels of *geoA* and *mic* genes at 15 °C were significantly higher than those at 25 and 35 °C. Although, the transcription of *mic* genes in *Pseudanabaena foetida* var. *intermedia* was higher at 35 °C than at 25 °C (Shen et al., 2022). These results highlight that unfavourable temperatures can increase the potential of geosmin and 2-MIB synthesis from the gene expression level in *Cyanobacteria*.

Global temperatures may increase by over two degrees celsius by 2100 (Shen and Guoya, 2013). Reservoirs subjected to augmented atmospheric temperature increases will experience increased temperatures; this is predicted to result in more intense and extended periods of thermal stratification and altered hydrology (Carey et al., 2011). Heightened temperatures also increase the stability of the water column, thus reducing vertical turbulent mixing and shifting the competitive balance in favour of buoyant *Cyanobacteria* (Jöhnk et al., 2008). It is well known that photosynthetic capacity, respiration rate and growth rate are temperature dependent,

with the optimum temperature at 25 °C or greater (Robarts and Zohary, 1987). Nevertheless, direct temperature effects most likely act synergistically with other factors like mixing and nutrient availability. Elevated biovolumes of *Cyanobacteria* are seasonal, occurring mainly in the warmer months (summer and autumn) when geosmin and 2-MIB are said to be at their highest concentrations (Kong et al., 2019). In accordance, connections with warming are propounded to encourage cyanobacterial growth as they have a higher optimal growth temperature than their eukaryotic counterparts (Zhang et al., 2017). In contrast, evidence demonstrates that enhanced cyanobacterial T&O concentrations can present at lower temperatures, with reports of winter T&O events in the absence of cyanobacterial blooms (Parinet et al., 2010).

#### 1.4.1.3 – Hydrology and stratification

Hydrological processes control the delivery and retention of nutrients and suspended sediments to lakes and reservoirs, influencing the composition of phytoplankton communities (Kimmel and Groeger, 2009). Seasonal alterations and interannual differences in the hydrological regimes can influence the temporal variations in *Cyanobacteria* abundance (Halac et al., 2019). Warming of the surface waters promotes the vertical stratification of stable water columns, favouring buoyant cyanobacterial growth in the epilimnion; an advantage for low light acclimated *Cyanobacteria*, as more light is available for photosynthesis (Dantas et al., 2011). Thermal stratification in the summer months can lead to lower dissolved oxygen concentrations near and in the sediments, promoting the reduction of  $\text{Fe}^{3+}$  to  $\text{Fe}^{2+}$  and releasing iron bound phosphorous into the labile pool (Wang et al., 2019). The liberation of labile phosphate into the water column through diffusion or turbulent mixing can alter the TN:TP ratio and increase the phosphorous supply. Stratified and eutrophic reservoirs have a low redox potential at the sediment-water interface (Søndergaard, 2009); a combination of strongly reducing conditions (due to anoxia) and high temperatures (20°C) leading to significant amounts of nutrient release in the overlying water (Dadi et al., 2020).

## 1.4.2 – Biotic triggers

Food chain manipulation can alter the food web structure in water bodies enabling the control of cyanobacterial T&O compounds (Fink, 2007). However, It is vital to understand better the biological functions of geosmin and 2-MIB in nature, e.g., their potential roles in the attraction of herbivores or the defence from predators (Fink et al., 2006a, 2006b; Watson, 2003).

### 1.4.2.1 – The cyanobacterial phycosphere – the “cyanosphere”

*Cyanobacteria* can offer niche microenvironments that facilitate interactions with heterotrophic bacteria (Adam et al., 2016; Dziallas and Grossart, 2011; Salomon et al., 2003; Woodhouse et al., 2018). This microenvironment is called the ‘phycosphere’ (Bell and Mitchell, 1972). The cyanobacterial phycosphere is equivalent to the rhizosphere in soils and sediments, providing a reservoir of exudates rich in a diverse range of organic molecules and even bioactive compounds such as toxins (Sivonen and Börner, 2008). Associated bacteria can be attached to the cyanobacterial cells (Hmelo, et al., 2012; Ploug et al., 2011), whereas others grow on extracellular mucus or develop free-living populations (Brauer et al., 2014). There is substantial evidence showing that heterotrophic bacteria coexist with *Cyanobacteria* in freshwater environments, and substantial evidence suggests that bacteria influence cyanobacterial productivity (Salomon et al., 2003). Bacterial remineralisation of organic matter can benefit *Cyanobacteria* from vitamins and growth factors produced by the associated microflora in the phycosphere (Simon et al., 2002). Thus, cyanobacterial compounds can support the growth of heterotrophic bacteria and shape the structure of the bacterioplankton community (Eiler and Bertilsson, 2004; Louati et al., 2015). A recent metagenomic study by Louati et al. (2015) revealed that marked shifts in the phycosphere community accompany changes in the cyanobacterial species composition.

Louati et al. (2015) have shown that the genus, and potentially the species of *Cyanobacteria* and its metabolic capacities, select for the bacterial community in the phycosphere. The availability of dissolved organic matter provided by *Cyanobacteria* underpins the foundation of cyanobacterial-bacterial interactions. *Cyanobacteria* can

generate a wealth of dissolved organic carbon substances to benefit nearby heterotrophic bacteria that can subsequently reciprocate the favour by removing reactive oxygen species, CO<sub>2</sub> generation, and nutrient recycling (Dziallas and Grossart, 2011; Paerl and Otten, 2013; Steffen et al., 2012). For example, *Microcystis* alters the ambient environmental conditions by decreasing the availability of light and oxygen concentrations (Paerl and Otten, 2016), as well as modifying CO<sub>2</sub> and pH levels (Havens, 2008), conditions likely to affect nearby bacteria. One hypothesis for such a close-knit relationship is that *Cyanobacteria* have small genomes compared to eukaryotic organisms (Herdman et al., 1979; Humbert et al., 2013). While this benefits *Cyanobacteria* for rapid reproduction and evolution, it is not automatically constructive for cyanobacterial bloom formation as some functions are lost (Giovannoni et al., 2014). Genome reduction can be advantageous if some lost functions are regained through “public goods”, as Morris et al. (2012) mentioned in the Black Queen Hypothesis. In addition to bacterial interactions, *Cyanobacteria* complement algae in marine environments by excreting large amounts of exometabolites (growth factors and biosynthetic precursors) as well as processing toxins (Morris et al., 2011; Pérez et al., 2016; Wienhausen et al., 2017).

#### 1.4.2.2 – Geosmin and 2-MIB as allelopathic tools

Allelopathy is the common biological phenomenon by which an organism (*Cyanobacteria*) produces biochemicals (e.g. geosmin and 2-MIB) that influence the growth, survival, development, and reproduction of another organism (Cheng and Cheng, 2015). These biochemicals are known as allelochemicals and can have either a beneficial or detrimental effect on target organisms. The synthesis of T&O compounds is an energy-consuming process; thus, bacteria must use these compounds as adaptive responses to environmental change (Koksharova, 2020). These compounds, therefore, have the potential use as “infochemicals” and/ or “allelochemicals”; these metabolites might be functioning in a variety of chemical interactions such as attractants, deterrents and protectants (Watson et al., 2007; Zuo, 2019).

Biosynthesis and release of allelopathic compounds can interfere with the settlement and growth of competitors (Li and Li, 2012). Geosmin can inhibit competitors such as *Chlorella pyrenoidosa* (Ikawa et al., 2001). T&O can also attract

insects such as the mosquito *Aedes aegypti* (Melo et al., 2019), hypothesised to serve as a nutritional source. Although geosmin-producing *Cyanobacteria* attract mosquitoes, other cyanobacterial strains show strong larvicidal effects when producing geosmin (Koksharova, 2020). Therefore, Geosmin can benefit producers' growth in competition for available nutrients (Ikawa et al., 2001). *Cyanobacteria* could be using these compounds as allelochemicals in their defence against herbivores; since the biosynthesis of T&O compounds is energetically costly, producing the metabolites upon stress is more beneficial.

#### 1.4.2.3 – Geosmin and 2-MIB degraders

Previous reports demonstrate that several gram-positive and gram-negative bacteria can degrade geosmin, including gram-positive strains of *Bacillus cereus* (Narayan and Nunez, 1974), *Bacillus subtilis* (Narayan and Nunez, 1974; Yagi et al., 1988), *Rhodococcus* genus (Guttman and van Rijn, 2012) and gram-negative strains of *Alphaproteobacteria* and *Sphingopyxis alaskensis* (Hoefel et al., 2009) and a *Bacteroidetes* clade (Hania et al., 2017). Four species of *Flavobacterium* have been noted to degrade 2-MIB (Clercín, 2019), along with two strains of bacteria belonging to *Pseudomonas* sp. and *Enterobacter* sp. (Tanaka et al., 1996). However, some reports of T&O degradation are associated with consortiums of bacteria (e.g., *Sphingopyxis alaskensis*, *Novosphingobium stygiae* and *Pseudomonas veronii*), with no biodegradation evident when one of the bacteria was absent (Hoefel et al., 2006). Only a few studies have tried to identify potential degraders from natural habitats such as rivers (Du et al., 2017) and reservoirs (Clercín et al., 2021; Westerhoff et al., 2005). One study found the dynamics of T&O degrading bacteria to mimic the seasonal peaks of geosmin and 2-MIB (Clercín et al., 2021). Elucidating the interactions between T&O producers and degraders may give insight into the proliferation of T&O concentrations, and the identification of degraders may give rise to indicative taxa.

Members of the *Bacteroidetes* phylum are known for their well-known capabilities to degrade polymeric compounds such as peptidases, glycoside hydrolases, glycosyl transferases and adhesion proteins (Fernández-Gómez et al., 2013); therefore, they may play a crucial role in the degradation and cycling of T&O compounds. Hania et al. (2017) demonstrated that specific clades of *Bacteroidetes*

scavenge *Cyanobacteria* and decay their biomass. Other studies have also reported a high abundance of *Bacteroidetes* present during cyanobacterial blooms (Berg et al., 2009; Moradinejad et al., 2020; Woodhouse et al., 2016; Zheng et al., 2020). Given the ability of *Bacteroidetes* to degrade polymeric compounds and the fact that they scavenge *Cyanobacteria* and are almost always present during cyanobacterial blooms, they likely play a role in the degradation of T&O compounds. Several members of the *Proteobacteria* are known to attach to *Cyanobacteria* and to actively lyse the cells (Manage et al., 2000; Maruyama et al., 2003; Yamamoto et al., 1993).

## 1.5 – Using eDNA to help predict T&O events

Previous identification of cyanobacterial strains has been objectively based on morphological characteristics such as trichome width, cell size, division planes, shape and arrangement, pigmentation, and gas vacuoles and sheath (Baker, 1992, 1991; Komárek and Anagnostidis, 1989). Despite considerable expertise required to identify cyanobacterial strains through microscopy, subjective judgement can lead to errors resulting in the incorrect assignment of isolates. Komárek and Anagnostidis (1989) predicted that over 50% of the strains in culture collections were wrongly classified. More reliable testing has promoted molecular approaches in cyanobacterial taxonomy. Methods often incorporate environmental DNA (eDNA), which can be harvested from environmental samples and preserved, extracted, amplified, sequenced and assigned taxonomy (Ruppert et al., 2019). A widely used method in eDNA molecular detection of phytoplankton and bacterioplankton communities is using methods involving DNA barcodes to differentiate samples within a pooled DNA library (Curren et al., 2019; Santi et al., 2021; Santoferrara, 2019).

For metagenetics, target genes must be variable enough to enable the distinction between closely related organisms, yet conservation of the flanking regions must be high to reduce variability within a single taxon (Mallott et al., 2018). Additionally, specific primer sets are required to readily amplify the target genes for a range of species within a large taxonomic group, and the target region must be long enough to generate the required taxonomic resolution. For bacterial identification, the 16S rRNA gene is commonly used (Bukin et al., 2019). The 16S rRNA gene is

approximately 1600 bp long and homes nine hypervariable regions of varying conservation (V1 – V9), as depicted in Figure 1.4 (Tringe and Hugenholtz, 2008). For eukaryotic species, the Consortium for the Barcode of Life (CBOL) Plant Working group recommends the plastid DNA genes *rbcL* and *matK* as standard metagenetic target genes (CBOL Plant Working Group, 2009). The *rbcL* gene is a part of the DNA sequence located within the chloroplast DNA, as illustrated in Figure 1.5. The *rbcL* gene is approximately 1400 bp long and has a low level of mutation compared with other target genes in the chloroplast DNA (Nurhasanah and Papuangan, 2019). Isolated eDNA can remain viable for long periods, allowing for applications in molecular biology, ecology and environmental sciences, especially when coupled with Next-Generation Sequencing (NGS) (Thomsen and Willerslev, 2015).

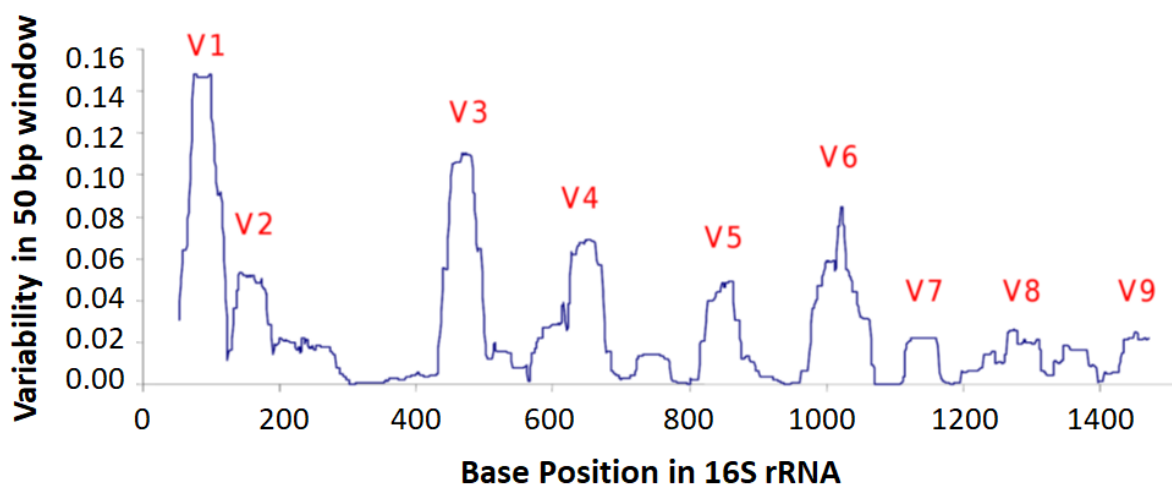


Figure 1.4: Hypervariable regions in the 16S rRNA gene in *Pseudomonas*. The plotted line reflects fluctuations in variability amongst aligned 16S rRNA gene sequences of 79 known *Pseudomonas* strains. Adapted from Bodilis et al. (2012).

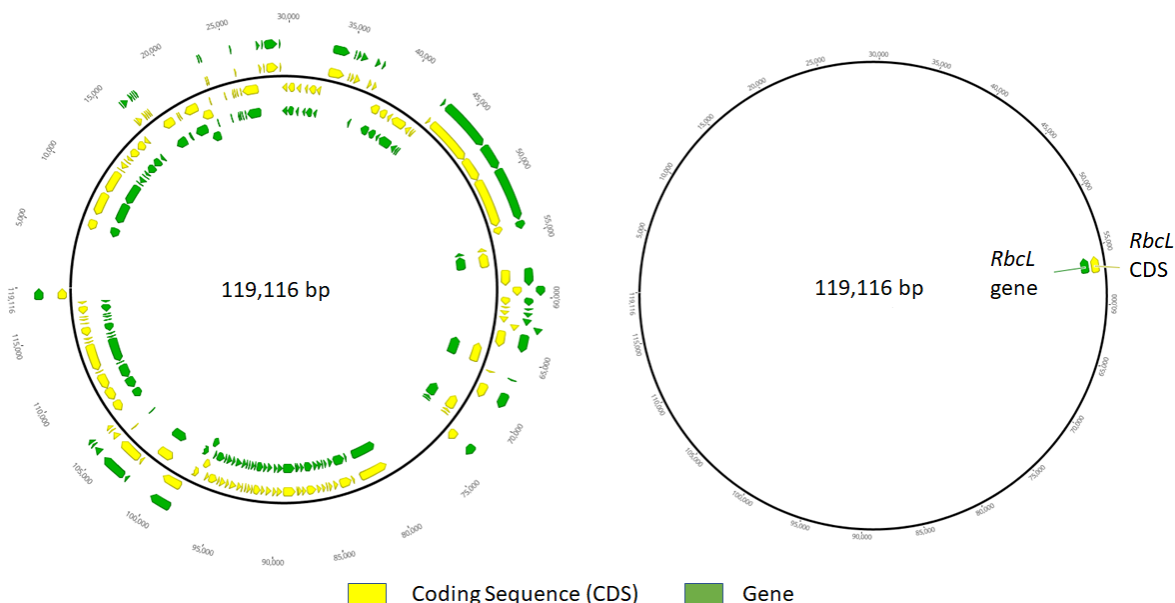


Figure 1.5: The location of the *rbcL* gene and coding sequence within *Nitzschia palea* NIES-2729 chloroplast.

Constant advancements in analytical chemistry have provided ultrasensitive detection methods for T&O compounds, like solid phase microextraction (SPME), coupled with gas chromatography-mass spectrometry (GC-MS) (Pochiraju et al., 2021). Advanced chemical detection of T&O compounds in parallel with the discovery of biosynthetic pathways associated with the production of T&O has lured scientists towards the development of biomolecular approaches to predict and forecast T&O events (Lee et al., 2017; John et al., 2018; Lu et al., 2019; Lukassen et al., 2019; Lukassen et al., 2019; Wang et al., 2019). One promising tool for early on-site detection of geosmin and 2-MIB is quantitative PCR (qPCR), through amplifying implicated genes, *geoA* and *mic* (Cane and Watt, 2003; Dickschat et al., 2007; Giglio et al., 2008; S Giglio et al., 2011; Steven Giglio et al., 2011; Gust et al., 2003; Komatsu et al., 2008; Wang and Cane, 2008). These methodologies have gained considerable attention due to their specificity, source identification, low limit of detection, and handling of multiple samples in a single run. These benefits allow them to be used for rapid on-site monitoring of T&O events (Chiu et al., 2017; Lee et al., 2017; John et al., 2018; Lu et al., 2019). The *geoA* gene is the main target for applying qPCR primers specific for the cyanobacterial production of geosmin (Devi et al., 2021). Most designed qPCR primers target the *mic* gene encoding for 2-MIB



cyclase of the 2-MIB operon (Figure 1.3). A total of 132 strains across 21 genera for geosmin production and 72 cyanobacterial strains across 13 genera for 2-MIB production have been reported (detailed in Devi et al., 2021). John et al. (2018) developed a novel, small amplicon PCR primer set (geo799F and geo927R) capable of distinguishing all geosmin-producing *Cyanobacteria* using a single protocol. However, no known universal PCR primers have been described for *mic* to capture all 2-MIB-producing *Cyanobacteria*; this highlights the need to revise currently developed *mic* primers and to create new universal primers.

### 1.5.1 – Modelling T&O events

Predicting the occurrence of geosmin and 2-MIB in drinking water reservoirs before reaching the treatment works is necessary to ease customer complaints and the cost of treatment (Kakimoto et al., 2014). In recent years, the application of predictive models to forecast geosmin and 2-MIB events has been proposed (Bertone and O'halloran, 2016; Dzialowski et al., 2009; Paerl et al., 2022). These predictive models mainly focus on regression analysis to develop empirical models that relate geosmin and 2-MIB concentrations using key water quality variables strongly correlate with T&O concentrations to predict T&O events. Although a lot of these models have also focussed on correlating cyanobacterial biomass (Christensen et al., 2006; Mau et al., 2004; Smith et al., 2002; Sugiura et al., 2004), and previous studies have shown that a correlation is not always established between T&O concentrations and cyanobacterial biomass (Graham et al., 2010; Xuwei et al., 2019). When modelling T&O events, caution is needed when using cyanobacterial biomass; for example, Chong et al. (2018) found an overestimation of geosmin concentrations when *Microcystis* sp. was blooming when taking biomass into account.

## 1.6 – Treating T&O compounds

Due to the unfavourable T&O that geosmin and 2-MIB bestow on drinking water leading to customer complaints, it has become increasingly important to remove these compounds (Jung et al., 2004; Kim et al., 2014a; Srinivasan and Sorial, 2011; WHO, 2016; Yang et al., 2010; Zamyadi et al., 2016). The tertiary

alcohol structure of geosmin and 2-MIB makes them highly resistant to oxidation processes commonly used in water purification (Clerc and Druschel, 2019). Previous studies have shown that geosmin is exceptionally resistant to treatment in conventional water treatment procedures like coagulation, sedimentation and filtration (Cook et al., 2001; Srinivasan and Sorial, 2011). Adsorption by activated carbon, either granular activated carbon (GAC) or powdered activated carbon (PAC), is thought to be the most effective technology for removal, and it is often used for trace organic contaminants removal from surface water (Srinivasan and Sorial, 2011). Yang et al. (2010) evaluated the cost-effectiveness of process A - conventional treatment followed by ozonation followed by biological activated carbon (O<sub>3</sub>/BAC); process B – process A followed by ultrafiltration (UF) – nanofiltration (NF); and process C – conventional treatment followed by GAC UF-NF. Process A had an annual cost of \$2,574,405, process B cost of \$7,525,581, and process C cost of \$7,059,465 (Yang et al., 2010). Yang et al. (2010) found that the GAC or O<sub>3</sub>/BAC processes effectively removed odorous compounds, but prices remained highly costly. Due to the costs of treating taste and odour compounds and filtration processes being clogged by *Cyanobacteria*, prevention is the preferred approach.

Attention has come from the biodegradation of T&O compounds from biologically active carbon (Clerc, 2019; Yagi et al., 1988; Zhou et al., 2011) and sand filters (Ho et al., 2007). A table of genera responsible for the degradation of geosmin and 2-MIB is displayed in Appendix 1.1. Currently, it is difficult to remove both compounds by conventional water treatment processes; biodegradation with microorganisms (Ishida and Miyaji, 1992) is effective for this purpose (Ho et al., 2007; Hsieh et al., 2010; Xue et al., 2011; Yuan et al., 2012). In sand filters, the primary cause of the removal of geosmin and 2-MIB was predominantly through biodegradation (Ho et al., 2007); a pseudo-first-order reaction caused by a *Pseudomonas* sp., *Alphaproteobacterium*, *Sphingomonas* sp. and an *Acidobacteriaceae* member. With the knowledge of T&O degraders, specific biofilters have been cultivated to degrade geosmin and 2-MIB from sand filters from drinking water treatment works (McDowall et al., 2007); sand filters with a well-established biofilm were capable of removing geosmin and 2-MIB to below detection limits after 11 days of 15 minutes contact with the sand.

## 1.7 – Research aims

Environmental triggers for geosmin and 2-MIB production have been widely reported (Journey et al., 2013; Oh et al., 2017; Sinang et al., 2015), but a lot remains unclear about the dynamics controlling T&O events, and no studies to date have identified biological indicators for T&O events. This thesis works in collaboration with Dwr Cymru Welsh Water and reports on research conducted with the incentive to accomplish the following aims by analysing nine reservoirs in Wales:

- ❖ Comparing tributary and reservoir locations for each site for geosmin and 2-MIB concentrations to locate T&O events. Assessing the reservoir's trophic state by season and year to better understand the seasonal patterns and driving factors of geosmin and 2-MIB events. Guided by literature, named variables  $\text{NH}_4^+$ ,  $\text{NO}_3^-$ , TP, Orthophosphate,  $\text{Fe}^{2+}$ , sulphate, and dissolved reactive silicate were selected to evaluate the effects of monthly changes in these nutrients on T&O outcome. Effects of monthly changes in nutrient concentrations and monthly lagged changes in T&O compounds were also assessed to evaluate any time-lapsed influences on T&O production. This research will provide the chemical mechanisms underpinning nutrient relationships in conjunction with geosmin and 2-MIB production, revealing important triggers for T&O events – Chapter 3.
- ❖ Employing NGS using the 16S rRNA and *rbcL* genes to explore the community compositions of the bacterial and algal communities from each studied reservoir. Using the community compositions as an informative response to nutrient influxes and environmental variables that had previously occurred in the water bodies. Exploring temporal (seasonal) and spatial differences in bacterial phyla over the study period to reveal the window for cyanobacterial growth. Identify significantly co-occurring bacterial and algal genera that could be associated in the cyanobacterial phycosphere (cyanosphere) potentially attributing to T&O. Examine bacterial and algal communities to see if low, medium and high T&O concentrations influence the composition of these communities. Establish a model that pinpoints indicative taxa associated with T&O concentration levels – Chapter 4.

- ❖ Using biosynthetic genes *geoA* and *mic* to assess the association between seasonality, temperature and nutrients and the production of geosmin and 2-MIB in reservoir drinking water; utilising qPCR to quantify the gene abundance of *geoA* and *mic* whilst creating a new primer for the *mic* gene to capture a broader range of 2-MIB-producing *Cyanobacteria*. Combine the molecular dataset with the corresponding water chemistry data to conduct predictive analysis to reveal linear variable triggers and nutrient ratio thresholds for heightened *geoA* and *mic* abundance that can result in geosmin and 2-MIB events – Chapter 5.

## 1.8 – Research hypotheses

- ❖ The seasonality of T&O events will not be localised only to the summer and autumn months. Carlson's Trophic State Index (CTSI) may provide a basis for understanding each reservoir's biological productivity sustained by season and year. Understanding the biological productivity capacity of the reservoirs by season and year could indicate nutrient enrichment which could lead to T&O events. Enhanced ammonium supply will stimulate rapid cyanobacterial productivity increases, stimulating the synthesis of T&O compounds, geosmin and 2-MIB. Increases in ammonium relative to decreases in nitrate concentrations will increase cyanobacterial productivity, although there will be a delay in T&O outcome due to time owed to the implicated biosynthetic pathways. Other vital nutrients involved in T&O production will be direct influences of orthophosphate (this form of phosphate is the most readily available form of P and hence no time delay) and a time delay for total phosphorous. Increasing orthophosphate and total phosphorous will alter the TN:TP ratio, favouring diazotrophic cyanobacterial productivity and resulting in time-lagged T&O events. Dissolved reactive silicate is expected to directly influence T&O production as this can be used as a proxy for diatom formation, revealing potential relationships between *Cyanobacteria* and diatoms – Chapter 3.
- ❖ T&O signature communities will be identified according to T&O concentration levels. Investigating bacteria and algae community compositions will reveal them to co-exist in complex synergistic networks, with taxa forming positive

and negative relationships. Within the architecture of the community profiles, cyanospheres will be exposed, and positive relationships will influence T&O events by creating positive feedback loops and recycling nutrients within the cyanospheres. Negatively associated taxa with geosmin and 2-MIB-producing *Cyanobacteria* will induce T&O events; competition for resources in limited supply will stimulate cyanobacterial productivity and hence T&O events. Using the biology consortium (16S rRNA and *rbcL* community compositions) as a response to environmental triggers for T&O production in model analysis, indicator taxa will be identified for reservoirs experiencing extreme T&O events – Chapter 4.

- ❖ *geoA* will provide a suitable direct proxy for geosmin concentrations when geosmin concentrations are elevated. Due to the lack of extreme 2-MIB events recorded in these data and sample type (open-water, not sediment samples), the *mic* gene will not correlate with 2-MIB concentrations. Modelling *geoA* copy numbers in response to environmental variables will reveal nutrient ratios ( $\text{NH}_4^+:\text{NO}_3^-$  and TIN:TP) significantly influential in heightening *geoA* abundance. Direct significant negative linear relationships will exist between mean temperature (highlighting geosmin production is connected to unfavourable conditions) and dissolved reactive silicate (indicating a potential link to diatom formation). Comparison of biosynthetic genes, *geoA* and *mic*, by year and season will expose significant differences between the same seasons of different years – Chapter 5.

# Chapter 2: Materials and Methods



DŴI MRU.  
GYMRU.  
FOR WALES

## 2.1 – Sampling methods and sample site details

Surface water (500 mL) was abstracted by Welsh Water, Dŵr Cymru, between July 2019 – September 2020 (sampled months 1 – 15 see Table 2.1 for sampled month assignment to date). Sampling included 52 sites across nine reservoirs in Wales, United Kingdom: 27 reservoir sites and 25 tributary sites. Locations of the nine reservoirs are displayed in Figure 2.1; five reservoirs were situated in North Wales (Dolwen, Plas Uchaf, Cefni, Alwen and Alaw reservoirs), and four were in South Wales (Llandegfedd, Pentwyn, Pontsticill and Llwyn Onn reservoirs). Detailed descriptions of each sampling location are provided in Sections 2.1.1 – 2.1.7. Water samples were collected using a 500 mL vessel attached to an extendable pole. The vessel was rinsed with reservoir water at each sample point before collecting a water sample from about 20 – 50 cm below the reservoir surface. Water samples were then transferred into a sterile 500 mL container, labelled with the unique reservoir sample point identifiers, and stored at four degrees celsius until further analysis.

Sampled months were calculated using the first month of the sampling period (7/2019) as sampled month one. A summary of the sampled months is displayed in Table 2.1.



Table 2.1: Sampled months calculated from the start date (7/2019) to the end date (9/2020) of this study period.

| Mon | Year | Date    | Sampled Months |
|-----|------|---------|----------------|
| 7   | 2019 | 7/2019  | 1              |
| 8   | 2019 | 8/2019  | 2              |
| 9   | 2019 | 9/2019  | 3              |
| 10  | 2019 | 10/2019 | 4              |
| 11  | 2019 | 11/2019 | 5              |
| 12  | 2019 | 12/2019 | 6              |
| 1   | 2020 | 1/2020  | 7              |
| 2   | 2020 | 2/2020  | 8              |
| 5   | 2020 | 5/2020  | 11             |
| 6   | 2020 | 6/2020  | 12             |
| 7   | 2020 | 7/2020  | 13             |
| 8   | 2020 | 8/2020  | 14             |
| 9   | 2020 | 9/2020  | 15             |
| 10  | 2020 | 10/2020 | 16             |
| 11  | 2020 | 11/2020 | 17             |

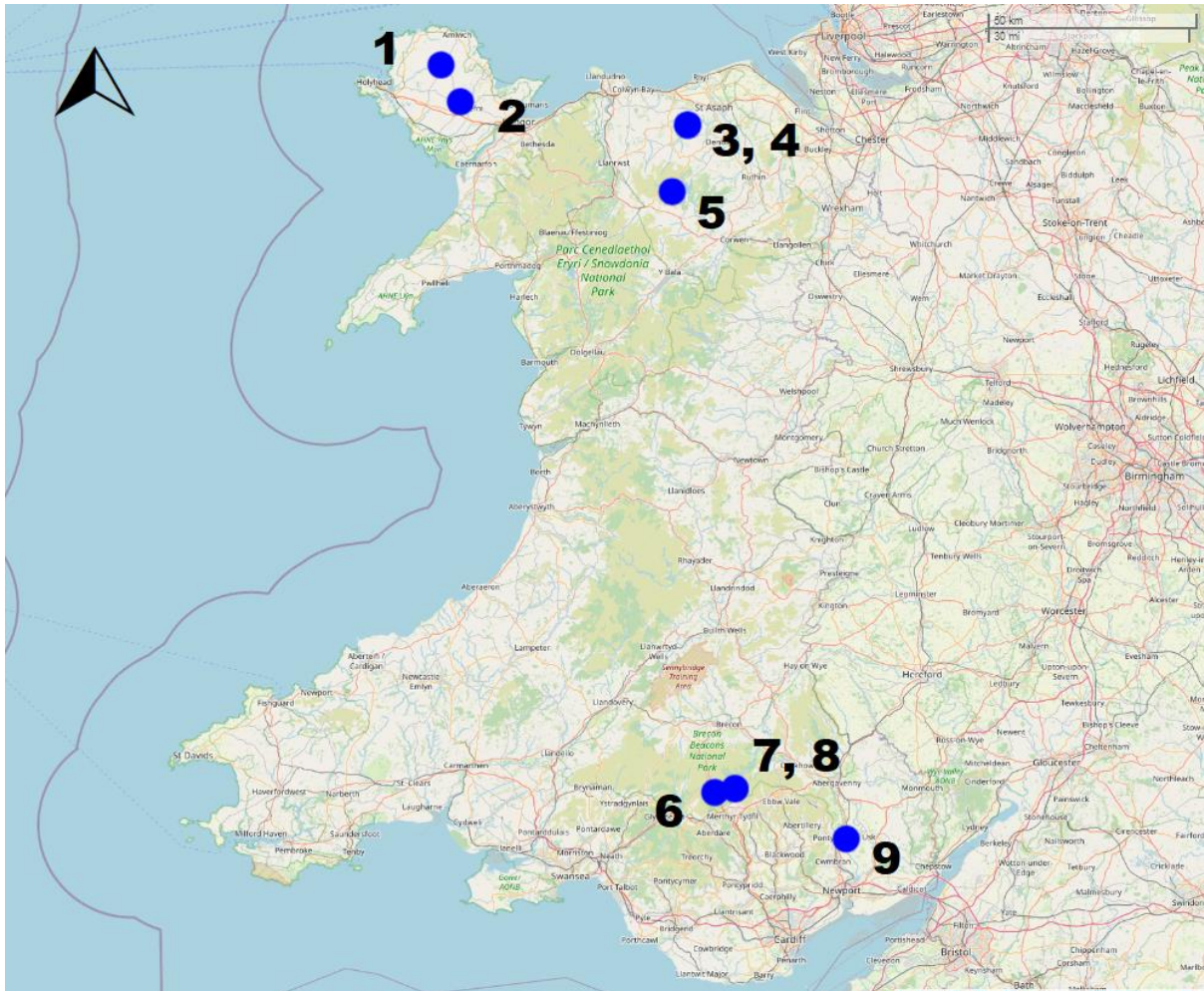


Figure 2.1: An overview of reservoir locations in Wales, U.K., used within this study; 1 = Alaw. 2 = Cefni. 3 = Dolwen. 4 = Plas Uchaf. 5 = Alwen. 6 = Llwyn On. 7 = Pentwyn. 8 = Pontsticill. 9 = Llandegfedd. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).

### 2.1.1 – Llyn Alaw reservoir, Anglesey

Llyn Alaw is an artificially constructed reservoir built in the 1960s to supply drinking water for the northern half of Anglesey. It covers an area of approximately 315 ha and is 4.8 km long. It is moderately alkaline and shallow (5.2 m at its deepest). The reservoir was constructed on a lowland fen, thus making the water naturally peaty (Hatton-ellis, 2016): due to the influence of peat, total phosphorous levels may be naturally higher than other reservoirs. In 2016 Natural Resources Wales noted nitrogen levels in the Alaw reservoir to be slightly elevated, with total nitrogen exceeding the  $1 \text{ mg L}^{-1}$  threshold with low confidence (total nitrogen =  $1.03 \pm 0.07 \text{ mg L}^{-1}$  (95% CI)). The catchment of this reservoir consists mainly of intensive agricultural land (91%) (Edwards, 2016).

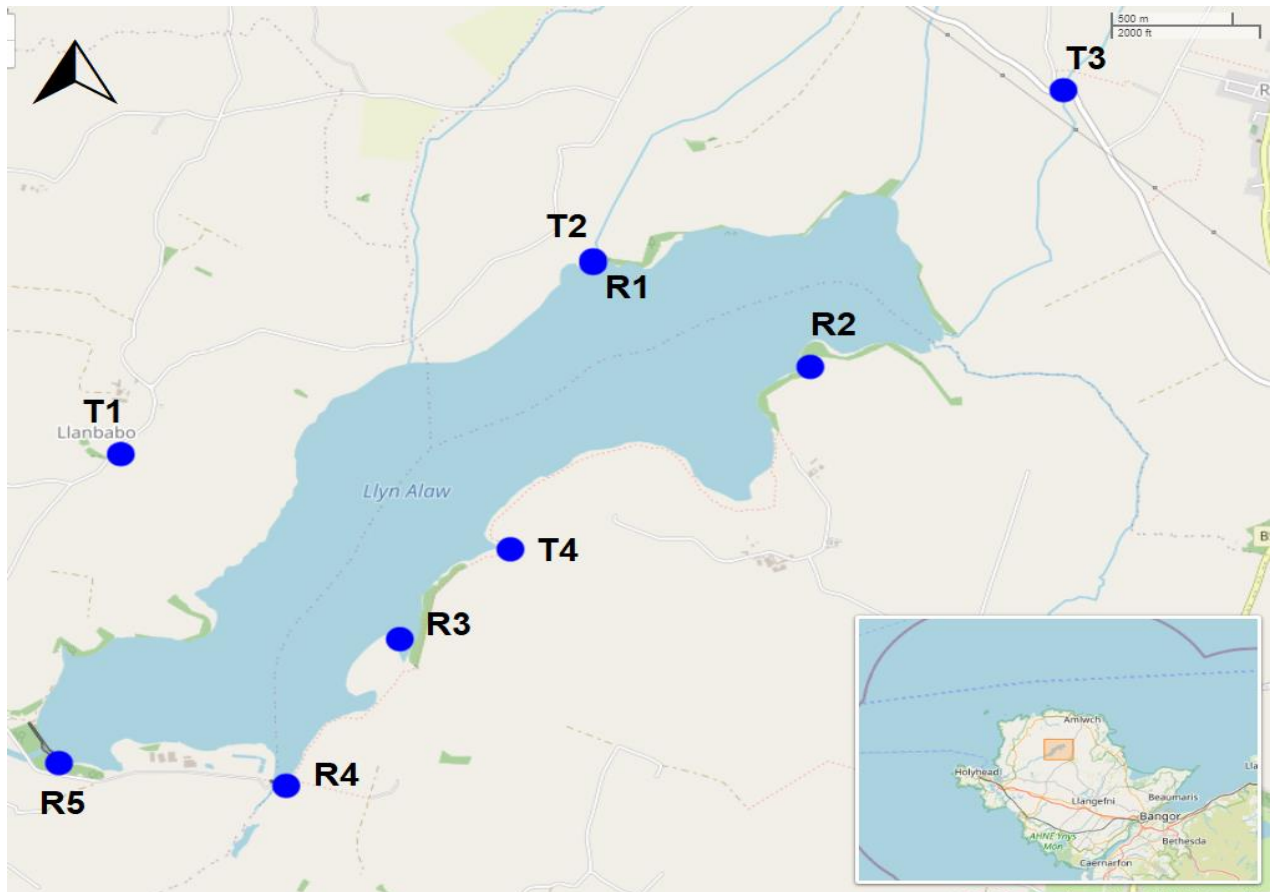


Figure 2.2: Llyn Alaw reservoir and sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1 – R5 = reservoir locations. T1 – T4 = tributary locations. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).

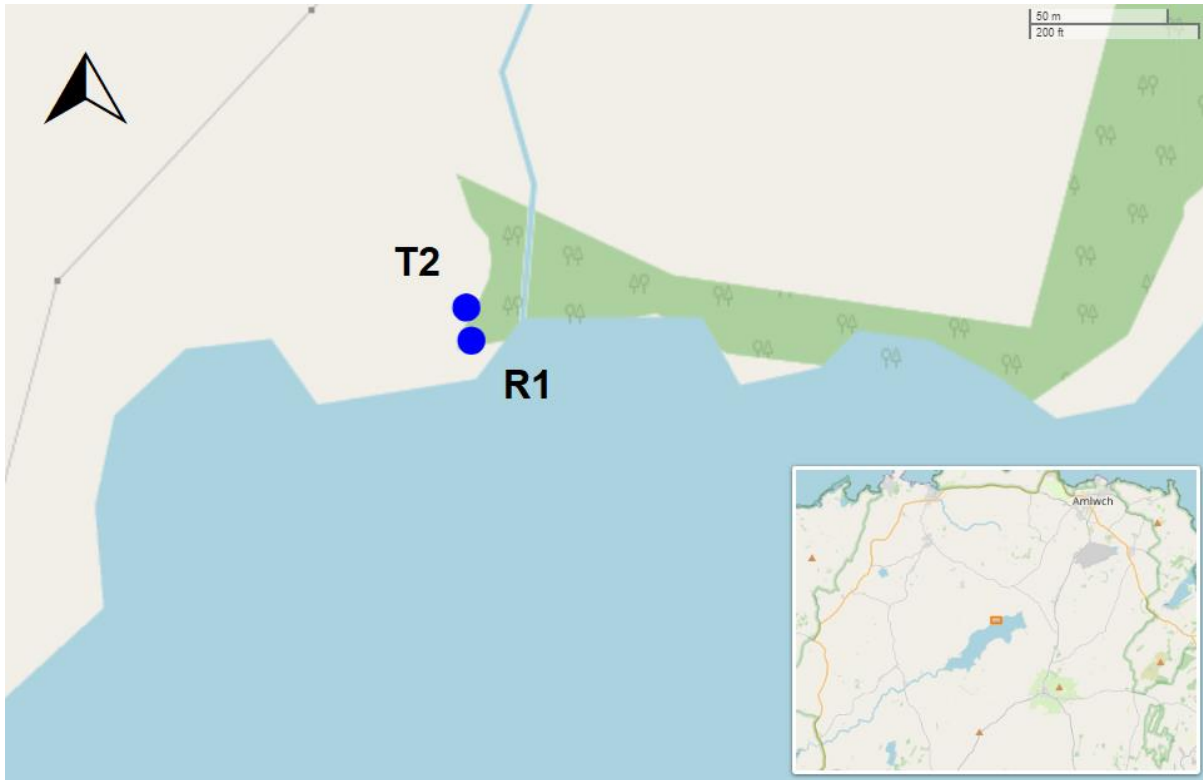


Figure 2.3: Close up of T2 (tributary two) and R1 (reservoir one) locations in Alaw reservoir with an inset map of the position of these sampling points within Alaw reservoir. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017).

### 2.1.2 – Llyn Cefni reservoir, Anglesey

Cefni reservoir is a shallow, alkaline artificial water body with a catchment area of over 4068 ha. It is one of the primary drinking water supplies for Ynys Mon and is fed by the water from the catchment area. More than 85% of the catchment comprises intensive agricultural land (Hatton-ellis, 2016). In 2016, Natural Resources Wales reported water quality data from the Cefni reservoir from 2010 to 2014 to be elevated in nitrogen and phosphorous, exceeding the  $1 \text{ mg L}^{-1}$  thresholds (Hatton-ellis, 2016). Although levels constituted a risk factor for eutrophication, this was not reflected in the biological data, with both phytoplankton and macrophytes showing “good status”.

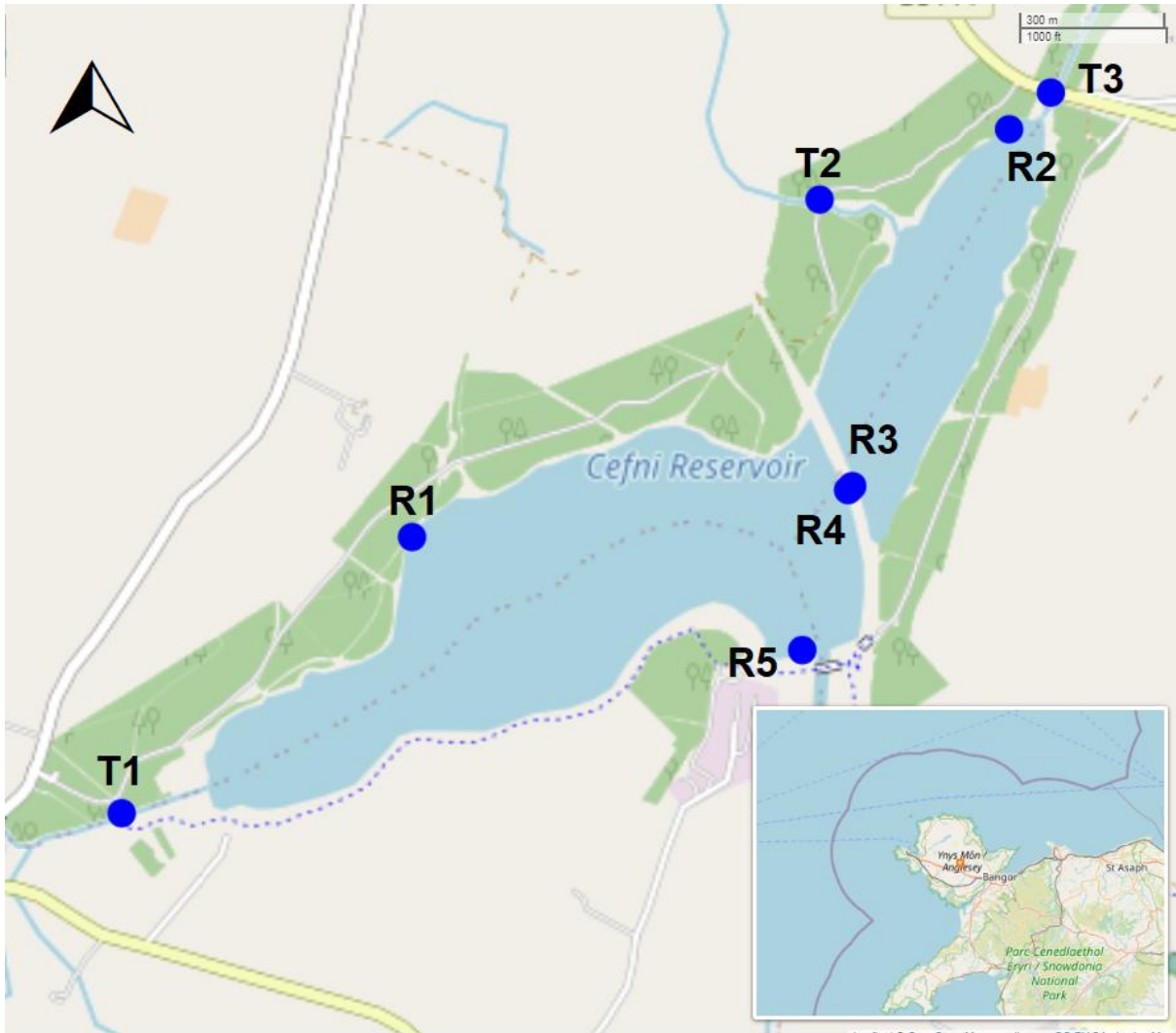


Figure 2.4: Llyn Cefni reservoir and sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1 – R5 = reservoir locations. T1 – T3 = tributary locations. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).

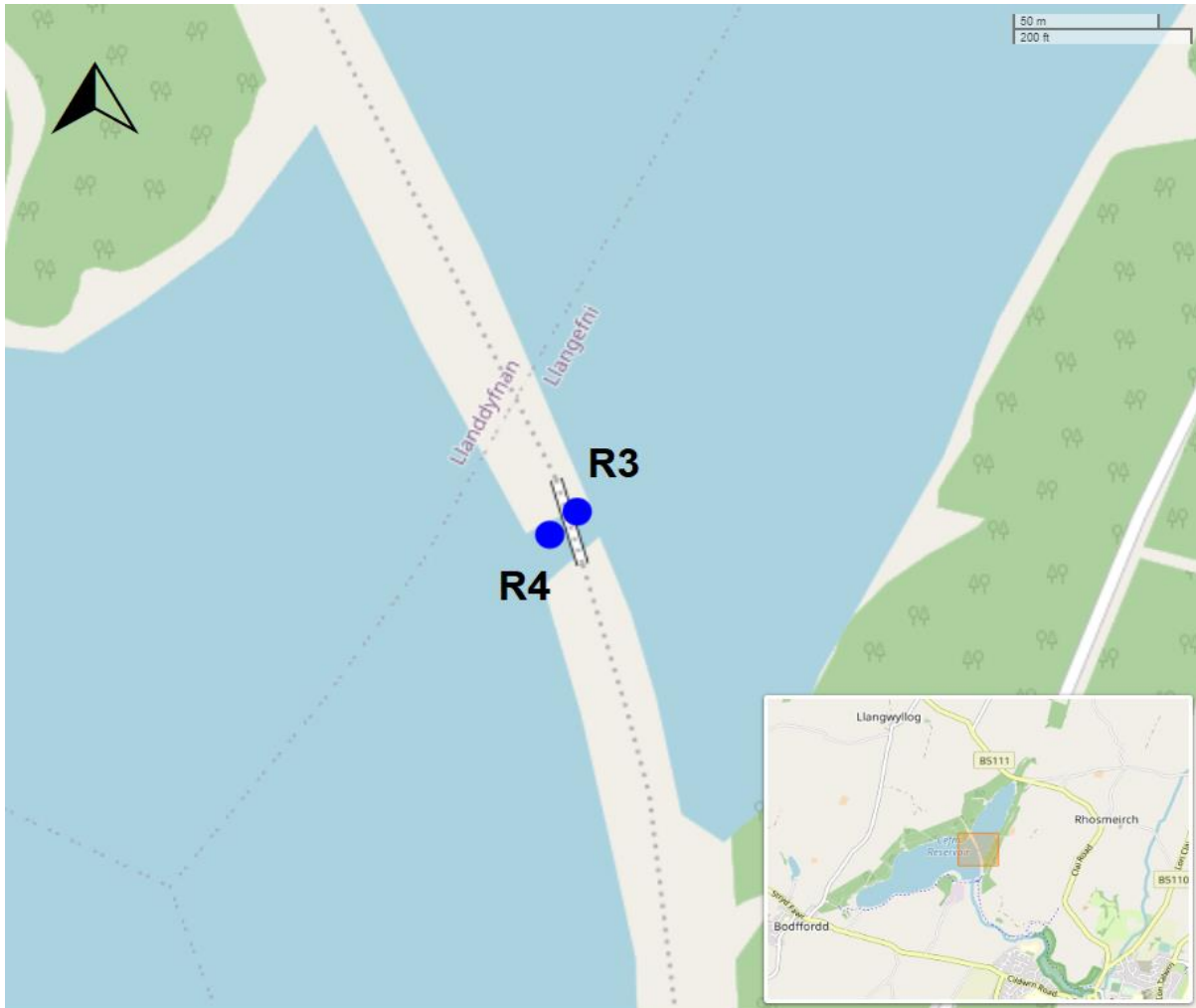


Figure 2.5: Close up of R3 and R4 (reservoir 3 and 4) locations in Cefni reservoir with an inset map of the position of these sampling points within Cefni reservoir. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).

### 2.1.3 – Dolwen and Plas Uchaf Reservoirs, Conwy

Plas Uchaf Reservoir and Dolwen Reservoir are artificial reservoirs in Denbighshire that span a catchment area of 226 ha. They impound a small tributary of the Afon Elwy and are designated drinking water supplies contributing to the main water supply for the Vale of Clwyd. Both reservoirs are shallow and alkaline, and the immediate catchment of the lakes consists of 85% intensive agricultural land use (Hatton-ellis, 2016). Dolwen reservoir is situated approximately 1 km to the south of Plas Uchaf reservoir (Perkins et al., 2019). Plas Uchaf reservoir is partially fed by an abstraction from the Afon Aled (T1: Figure 2.6), aided by releases from the Aled reservoir. Dolwen reservoir is fed by surface water drainage. It is estimated that the

% nitrogen contribution from intensive agriculture is around 86.8% (Edwards, 2016). Natural Resources Wales water quality data from 2010 – 2014 indicated that both total oxidised nitrogen and total nitrogen levels exceeded  $2 \text{ mg L}^{-1}$  with high confidence; phosphorous was deemed to be moderate, although with low confidence (Hatton-ellis, 2016).

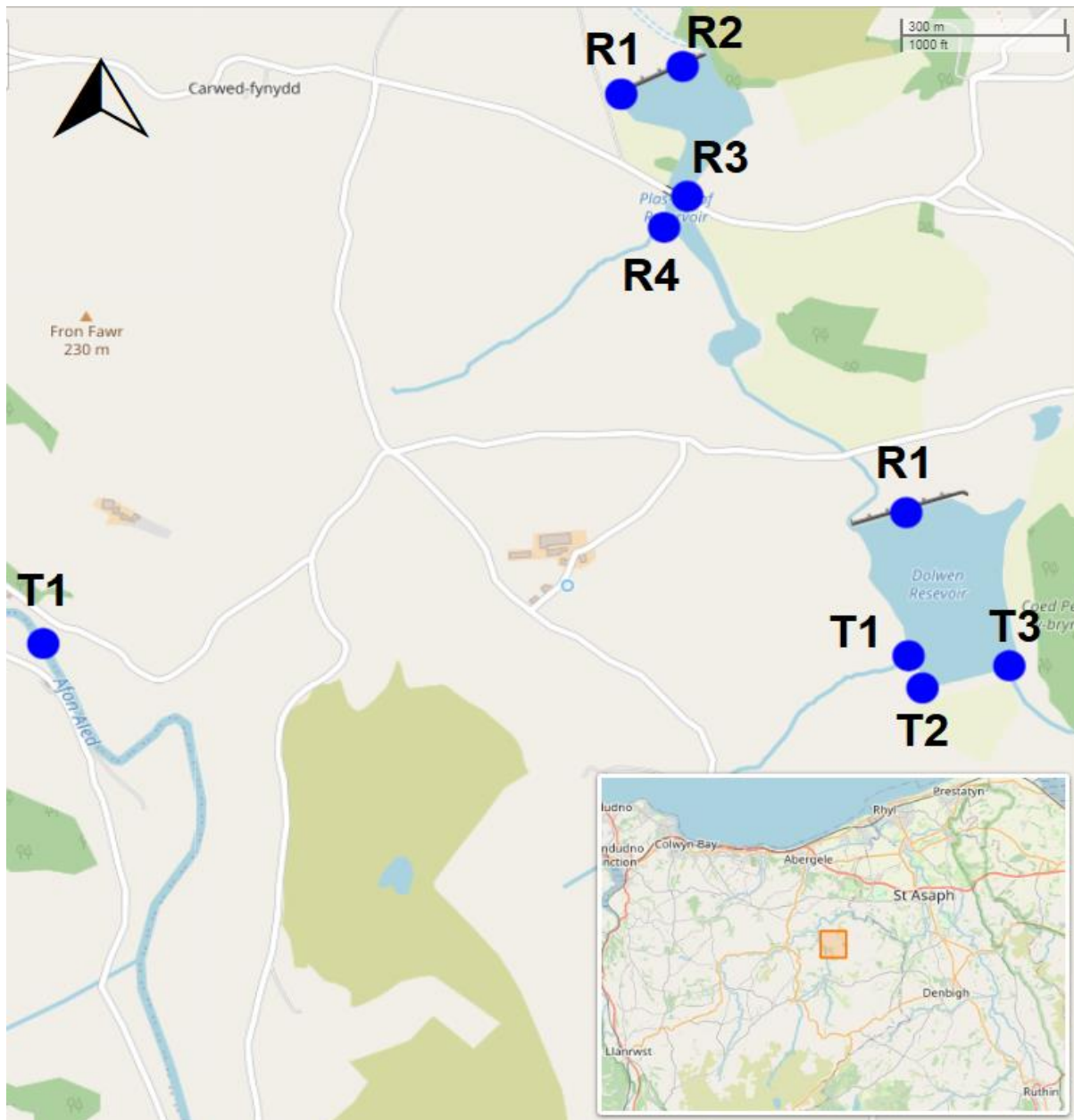


Figure 2. 6: Plas Uchaf reservoir (top) and Dolwen reservoir (bottom) sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1 – R4 = Plas Uchaf reservoir locations, T1 = Plas Uchaf tributary location located in the Afon Aled. R1 = Dolwen reservoir location, T1 – T3 = Dolwen tributary locations. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).

## 2.1.4 – Alwen Reservoir, Conwy

Alwen reservoir and dam were constructed in the 1920s to provide water for the port of Birkenhead, covering over 900 acres with a capacity of  $15 \times 10^6 \text{ m}^3$  (Gurnell et al., 1994). The primary source of drinking water supply for north-east Wales comes from Alwen reservoir and river Dee via an abstraction point at Poulton (Morrison, 1997). The reservoir is a tributary of the river Dee, joining it at Corwen. The upper river Dee and river Alwen catchments are underlain almost entirely by fine-grained sedimentary mudstones and siltstones (Natural Resources Wales, 2014). The Alwen dam impounds the river Alwen and is 9 km downstream from Alwen reservoir, a natural reservoir at the head of the catchment. The dam (located at R1: Figure 2.7) comprises two parts; the initial section consists of a gravity dam constructed in mass concrete faced with concrete blocks, and a subsidiary earthen embankment extending to the hillside (Hopkins, 2018). The dam is 27 m high and can store 14,240 Million Litres of water when full.



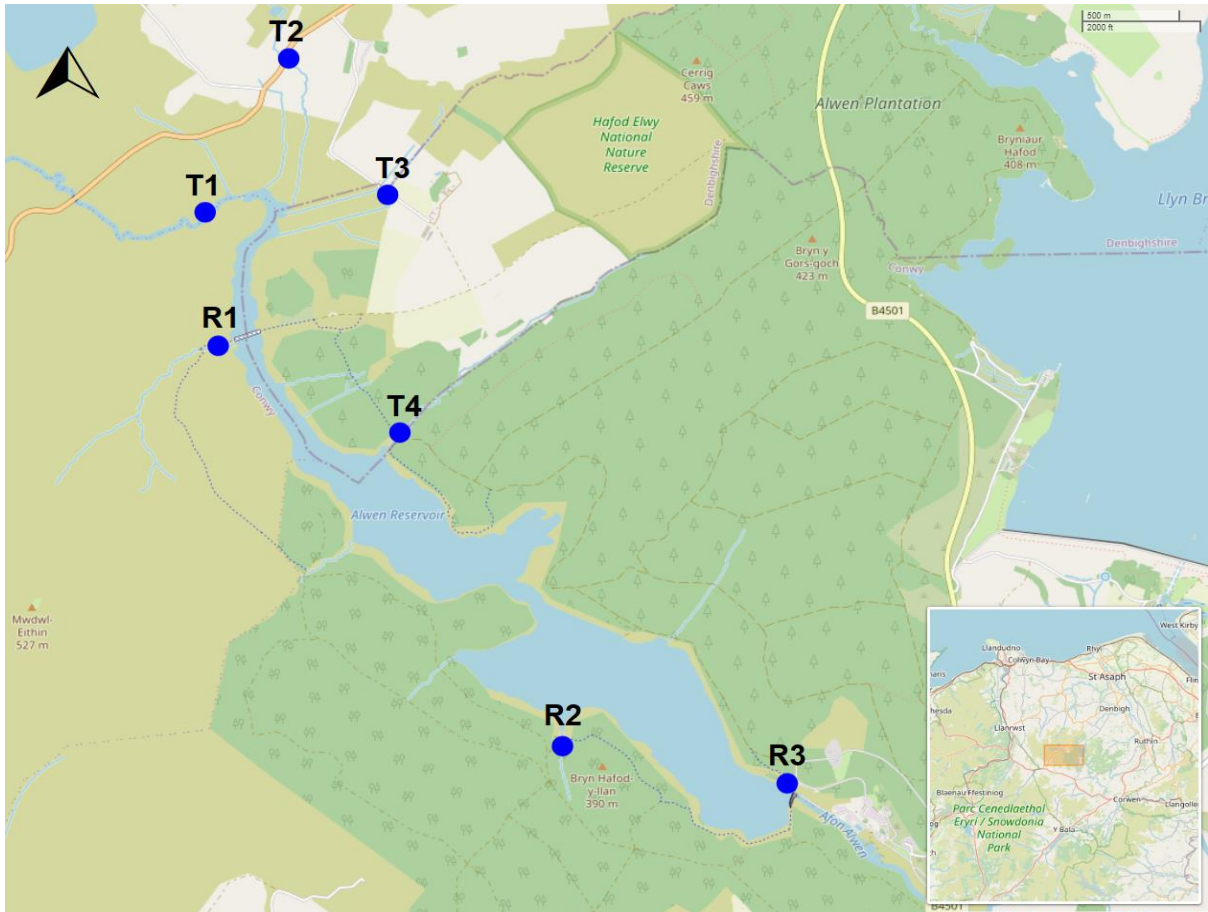


Figure 2.7: Alwen reservoir and sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1 – R3 = Reservoir locations. T1 – T4 = Tributary locations. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).

### 2.1.5 – Llwyn On Reservoir, Taf Fawr Valley

Llwyn On reservoir is an artificially made reservoir of just over 50 ha in Merthyr Tydfil and is moderately alkaline. The reservoir was constructed by damming the upper Afon Taf Fawr. The water is considered relatively peaty and is classified as shallow, but the reservoir stratifies in the summer. Intensive agriculture only accounts for 2% of land use (Edwards, 2016). From 2009 to 2014, nitrogen concentrations were deemed low (total oxidised nitrogen =  $0.18 \text{ mg L}^{-1}$ , total nitrogen =  $0.51 \pm 0.11$  (95% CI)  $\text{mg L}^{-1}$ ), and phosphorous concentrations were around  $16.3 \text{ } \mu\text{g L}^{-1}$  (Hatton-ellis, 2016). Due to the peaty nature of the water, this typically results in higher concentrations of phosphorous, although this was not accounted for in the

water quality data reported by Natural Resources Wales, hence confidence in phosphorous classification was “uncertain” (Hatton-ellis, 2016).

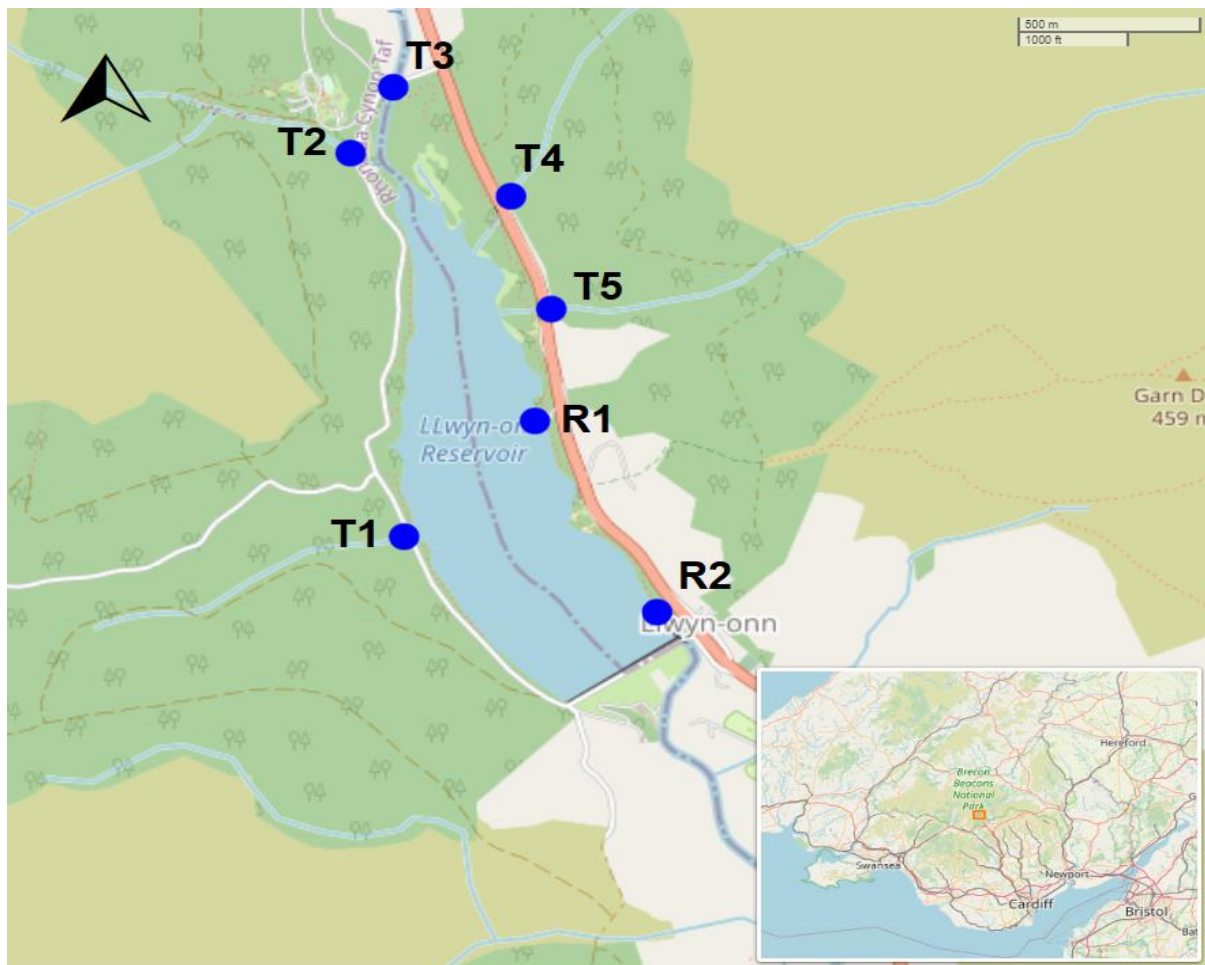


Figure 2. 8: Llwyn Onn reservoir and sample point locations with an inset map of the reservoir’s location in Wales U.K., bottom right. R1 – R2 = Reservoir locations. T1 – T5 = Tributary locations. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017).

### 2.1.6 – Pentwyn and Pontsticill Reservoirs, Taf Fechan

Pentwyn and Pontsticill reservoir lie within a catchment comprised predominantly of grasslands (56%) and uplands (22%), with sectors of woodland and commercial coniferous forestry (14%) (Barry et al., 2021). Pontsticill is a large reservoir on the Taf Fechan, lying partly in the county of Powys and partly within the county borough of Merthyr Tydfil. Pontsticill is an artificial reservoir with 110 ft high embankments completed in 1927 and holds approximately 13.6 giga litres of water and is a crucial element in the supply of drinking water to 106,000 customers in this

region (Gurney Environmental, 2016). The more recent reservoir incorporates the earlier Pentwyn (also known as Dolygaer) reservoir (1859 – 1863), which suffered significant water losses after completion due to the presence of severe fractures in the bedrock beneath its dam (Bowtell and Hill, 2006). Pentwyn is the smaller reservoir out of the two, with a capacity of 346 million gallons of water.

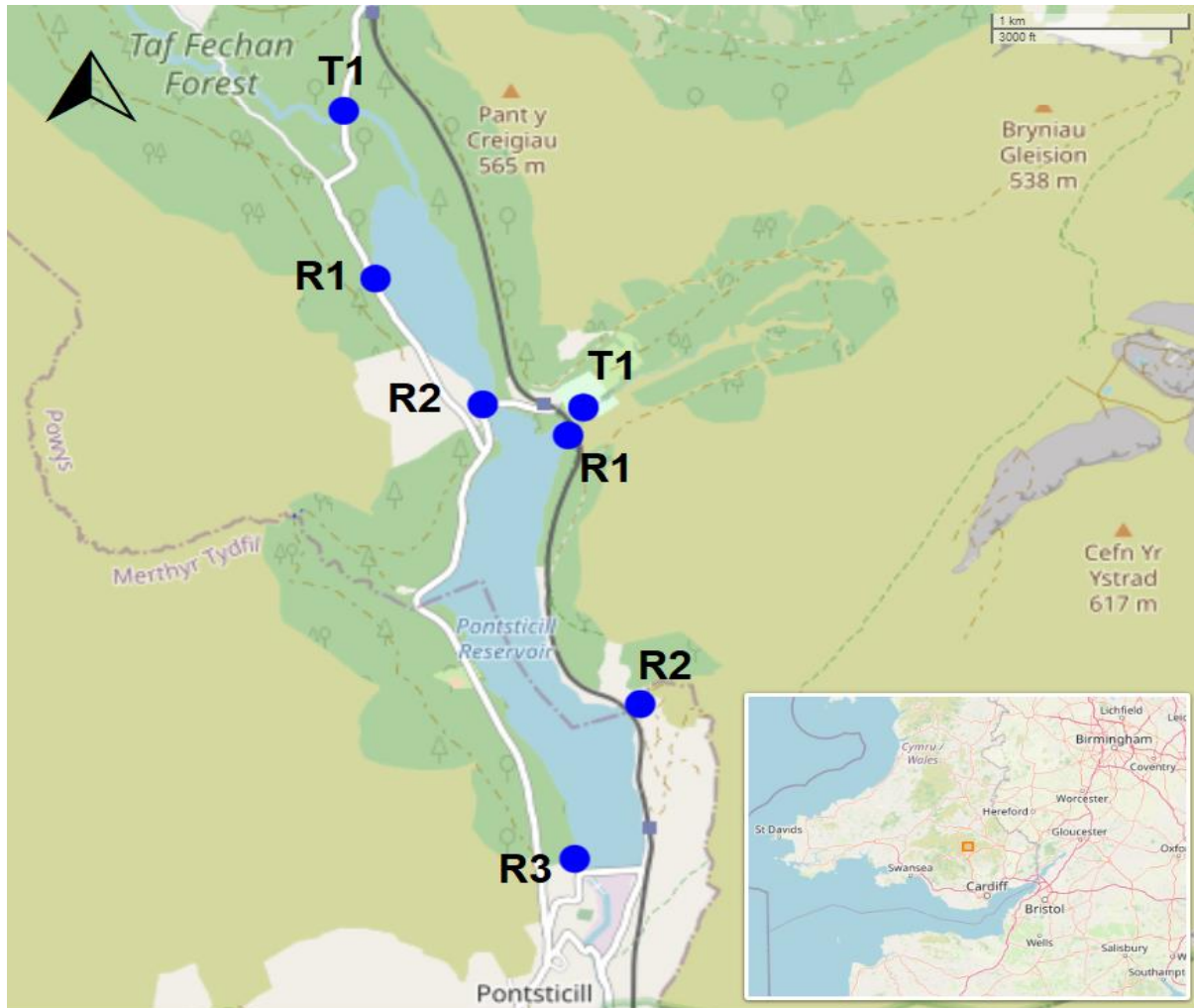


Figure 2. 9: Pentwyn reservoir (top) and Pontsticill reservoir (bottom) sample point locations with an inset map of the reservoir’s location in Wales U.K., bottom right. R1 – R2 = Pentwyn reservoir locations, T1 = Pentwyn tributary location. R1 – R3 = Pontsticill reservoir location, T1 = Pontsticill tributary locations. Map created using R 4.1.0 and package ‘leaflet’ (Cheng et al., 2017).

### 2.1.7 – Llandegfedd Reservoir, Newport

Llandegfedd reservoir is an artificially constructed drinking water supply situated between Usk and Cwmbrân in south Wales. It is a shallow, very alkaline reservoir that stratifies over the summer and spans over 176 ha (Edwards, 2016). It is a pumped storage reservoir for the river Usk and Wye abstractions and forms part of the south-east Wales strategic water supply network. Water quality data from 2009 – 2014 showed some evidence of eutrophication; phosphorous concentrations were consistent with moderate ecological status (governed by the Water Framework Directive), and total nitrogen and winter nitrate exceeded the 1 mg L<sup>-1</sup> threshold (Hatton-ellis, 2016).

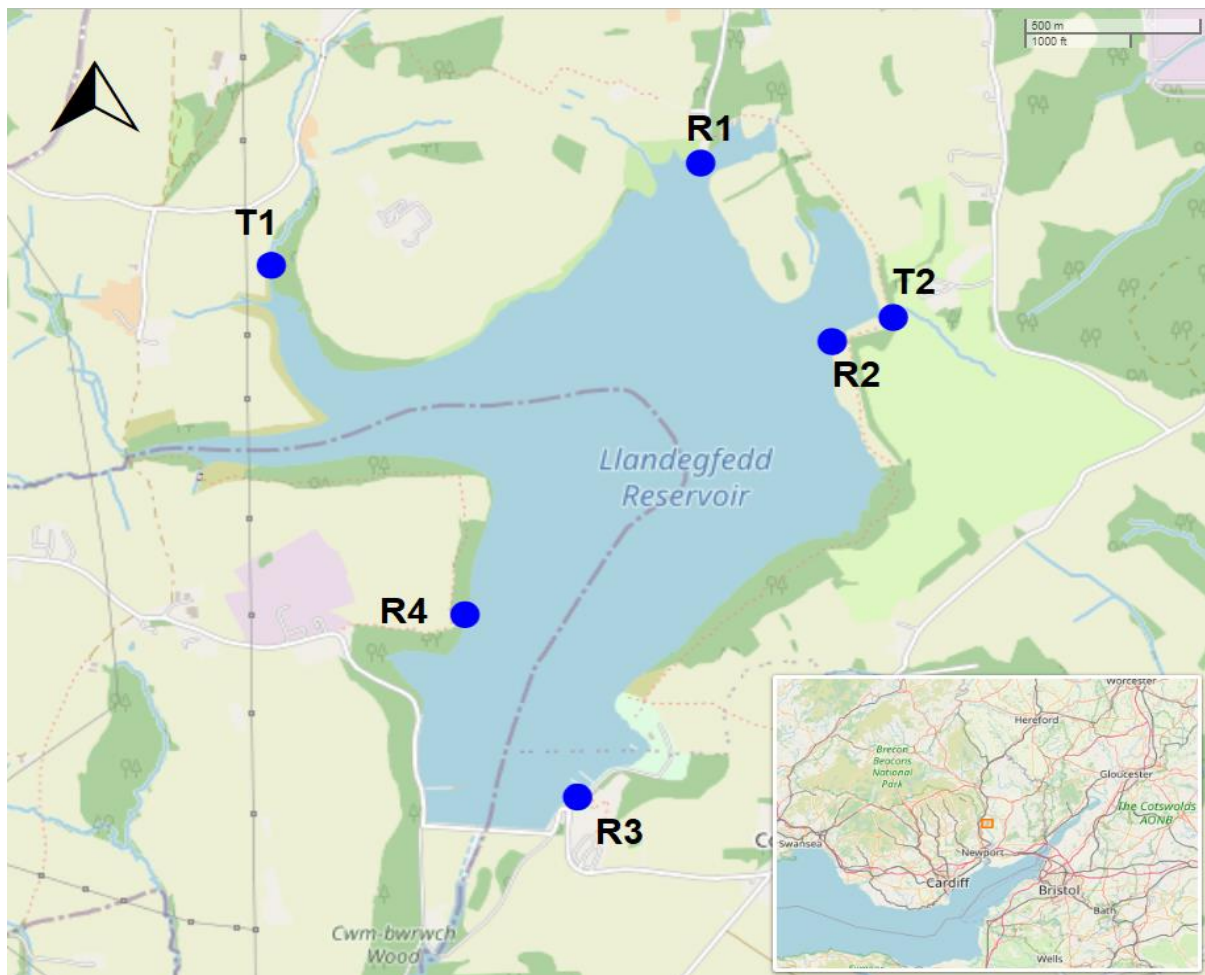


Figure 2.10: Llandegfedd reservoir and sample point locations with an inset map of the reservoir's location in Wales U.K., bottom right. R1 – R4 = Reservoir locations. T1 – T2 = Tributary locations. Map created using R 4.1.0 and package 'leaflet' (Cheng et al., 2017).

## 2.1.8 – COVID-19 sampling disruption statement

Due to COVID-19, sampling generally ceased between March – April 2020 (sampling months 9 – 10), with restrictions prolonging the sampling of some reservoirs. Affecting Chapters 3 – 5 and reservoirs described in Sections 2.1.1 – 2.1.7. The study period was extended from July 2020 to September 2020 to give additional data and the benefit of highlighting any reoccurring seasonal trends. Data were generally recorded from July 2019 – September 2020 for each reservoir, except for missing data points due to the pandemic.

## 2.2 – Defining a Taste and Odour event

For geosmin, customer complaints start at 7 ng L<sup>-1</sup>, and customer complaints start at 12 ng L<sup>-1</sup> for 2-MIB (Simpson and MacLeod, 1991). According to DCWW, increased sampling frequency on reservoir water commences when geosmin is >5 ng L<sup>-1</sup> and 2-MIB is >2.5 ng L<sup>-1</sup>. When levels are >10 ng L<sup>-1</sup> for geosmin and >5 ng L<sup>-1</sup> for 2-MIB, Granular Activated Carbon (GAC) filters are turned on if the drinking water treatment works have the facilities. If no GAC is present, initiation of Powdered Activated Carbon (PAC) dosing at low levels commences, and customer complaints are checked daily. Returning to regular operation is only authorised once at least 3 samples are below 10 ng L<sup>-1</sup> for geosmin and 5 ng L<sup>-1</sup> for 2-MIB.

This study defines a T&O event as a concentration measurement >10.0 ng L<sup>-1</sup> for geosmin and >5.0 ng L<sup>-1</sup> for 2-MIB. Classification of a low geosmin event is <5.0 ng L<sup>-1</sup>, a medium event is between 5.0 – 20.0 ng L<sup>-1</sup>, and a high event is any concentration above 20.0 ng L<sup>-1</sup>. Classification of a low 2-MIB event is <2.5 ng L<sup>-1</sup>, a medium event is between 2.5 – 10 ng L<sup>-1</sup>, and a high event is any concentration above 10.0 ng L<sup>-1</sup>.

## 2.3 – Water chemistry and physical analysis

Dŵr Cymru Welsh Water conducted all chemical analyses in Glaslyn laboratories in Newport, South Wales, U.K, adhering to ISO/IEC 17025:2019. Mean

temperature (°C) for all reservoir locations (Figure 2.1) was obtained through the Meteorological Office (Met Office).

### 2.3.1 – Determination of geosmin and 2-MIB concentrations

Each 500 mL water sample was extracted using a solid phase extraction (SPE) technique. The analytes were eluted from the SPE cartridge with Dichloromethane. An internal deuterated standard was added to the extract before being transferred into a labelled 2 mL autosampler vial to assess the efficiency of the run. The extract was then injected into a gas chromatograph through a multi-mode inlet in solvent vent mode. This removed most of the extraction solvent before the analytes of interest entered the column. A helium flow carried the analytes onto the column, where they were separated by boiling point. The separated analytes were then detected by a triple Quadrupole Mass Selective detector operating in multiple reaction monitoring mode.

### 2.3.2 – Nutrient analysis

Nutrients, ammonium ( $\text{NH}_4^+$ ) Total Oxidized Nitrogen (TON), nitrate ( $\text{NO}_3^-$ ), nitrite ( $\text{NO}_2^-$ ), orthophosphate, total phosphate (TP), sulphate ( $\text{SO}_4^{2-}$ ), and dissolved reactive silicate were analysed using a discrete analyser (Thermo Scientific Aquakem 600). Nutrient concentrations were measured colourimetrically or turbidimetrically. Turbidimetrically – The analytical method of measurement is dependent upon the reaction between the analyte within the sample and the reagents; sulphates react with the reagent to produce an insoluble precipitate which is measured turbidimetrically. Colourimetrically – all other determinants produce a coloured complex when reacted with the reagents; at a predetermined wavelength, the intensity of the coloured or turbid sample solution is proportional to the concentration of the analyte within the sample. The monochromatic light is focused upon the sample (in the form of a homogeneous medium), and the sample absorbs an amount of the focused light; the remaining light (not absorbed) is then transmitted and measured by a detector within the instrument. The ratio of the incident light to the transmitted light is called absorbance. Per Beer-Lambert's Law, concentration is

proportional to absorbance, so we can use calibration standards to calibrate the instrument and accurately quantify sample concentrations.

### 2.3.3 – Metals

Phosphorous and dissolved iron in samples were brought into solution by overnight digestion in 1% hydrochloric acid in an oven set at  $85 \pm 5^\circ\text{C}$ . Acidified samples were introduced to the Agilent 7700 Inductively Coupled Plasma Mass Spectrometer (ICP-MS) with a Cetac ASX-500 Autosampler (Agilent Technologies, USA) where the sample was sprayed as an aerosol into the plasma where it was desolvated, atomised and ionised. The sample ions were then extracted from the plasma into a quadrupole mass spectrometer, where the ions were separated according to their mass-to-charge ratio ( $m/z$ ) and detected using an electron multiplier. The concentration at a specific mass-to-charge ratio was measured against a calibration curve.

### 2.3.4 – Chlorophyll *a* (Chl *a*)

For Chlorophyll *a*, any plant material such as plankton was obtained from 500 mL of a water sample using an initial filtration step through a 70 cm Whatman Glass microfibre filter (GF/C) with a pore size of  $1.2 \mu\text{m}$ . This filtration step was followed by solvent extraction of Chlorophyll *a* using methanol. Once the sample had been extracted, Chlorophyll *a* concentration was determined by spectrophotometric evaluation of the extract. This was done by carrying out absorbance measurements on a Spectroquant® Pharo 300 spectrophotometer (MilliporeSigma, USA) at two wavelengths: 665 nm (the absorption of Chlorophyll A) and 750 nm (to compensate for background turbidity).

### 2.3.5 – Total Organic Carbon (TOC)

TOC is measured using a combustion-based method whereby each sample was acidified with phosphoric acid and sparged by the Shimadzu TOC Analyser (TOC-Vwp) (Shimadzu, Japan) to remove any inorganic carbon as carbon dioxide ( $\text{CO}_2$ ). Each sample was then passed through the reaction vessel within the instrument, where irradiation with UV light at  $80^\circ\text{C}$  in the presence of persulphate

converted all the carbon to CO<sub>2</sub>. This CO<sub>2</sub> was then transported using nitrogen gas to an Infrared detector, where the concentration of CO<sub>2</sub> was determined.

### 2.3.6 – pH, conductivity, and Turbidity

A water sample was taken up by the sample probes into the Skalar SP2000 robot (Skalar Analytical B.V., Netherlands) and then analysed using pH, conductivity, and turbidity meters. Conductivity is temperature dependent, and measurements were electronically compensated to a fixed temperature of 20°C for reporting purposes. Turbidity is an expression of the optical property of a liquid that causes light to be scattered and absorbed rather than transmitted in straight lines through the sample. The Turbidity module is based on Nephelometric measurements. The standard units are Nephelometric Turbidity Units (NTU).

## 2.4 – Water filtering and Extraction of total genomic environmental DNA (eDNA)

Reservoir water (500 mL) from each sampling point was filtered through a Sterivex filter (0.2 µM) using a diaphragm vacuum pump (Fisherbrand™, FB65453). 1 mL of ATL buffer (QIAGEN, Germany) was added to preserve the samples, and the filter was capped off with Luer locks before being frozen at -20°C for later use.

After each sample was defrosted, 100 µL of proteinase K (10 mg mL<sup>-1</sup>) was added to each filter before being put on a turntable for 2 hours. Sample (100 µL) was mixed with 200 µL of extraction buffer (5M NaCl, 30 mM NaEDTA and 70 mM tris pH 7.0), 25 µL of 10% DTAB and 200 µL of chloroform and added to a tube containing 0.1 mm glass beads following methods described by Fawley and Fawley (2004). The sample was agitated using an MP Biomedical, FastPrep-24™ at 5 ms<sup>-1</sup> for 30 seconds twice (with a 5 min incubation at room temperature). The sample was then centrifuged (Eppendorf, 5417C) at 14000 rpm for 2 minutes. The top phase of the sample was transferred to a sterile collection tube (200 µL), and 200 µL of buffer AL (QIAGEN, Germany) was added along with 200 µL of ethanol before being vortexed for 30 seconds. DNA was then further purified using the procedure and reagents provided within the DNAeasy® Blood & Tissue Kit (QIAGEN, Germany). This was



achieved by transferring the sample into a spin column and centrifuging at 14000 rpm for 1 minute. The collection tube was emptied, and 500 µL of AW1 buffer was added before centrifugation at 14000 rpm for 1 minute. The collection tube was again emptied, and 500 µL of AW2 buffer was added before another centrifugation at 14000 rpm for 3 minutes. The column was then inserted into a sterile Eppendorf tube, and 50 µL of nuclease-free water was added and left for 5 minutes (to maximise DNA recovery) before being centrifuged at 14000 rpm for 2 minutes.

#### 2.4.1 – Primary Illumina-Nextera tag PCR for 16S rRNA and *rbcL* genes

Two sets of primers were used, each having been tailored to include an Illumina-Nextera tag incorporated on the 5' end (Table 2.2). The composition of the PCR reaction (Table 2.3) and PCR conditions (Table 2.4) were used for all reactions, which were performed in triplicate to account for any PCR bias. All PCRs were carried out on a SimpliAmp™ Thermal Cycler (Thermo Fisher Scientific, USA).

Table 2.2: Primer sets used. The primer sets were analysed using the integrated DNA technologies OligoAnalyzer tool (Owczarzy et al., 2008).

| Name             | Use         | Sequence (5'-3')                   | Length (bases) | Amplicon size | Reference                       |
|------------------|-------------|------------------------------------|----------------|---------------|---------------------------------|
| Forward overhang | Nextera tag | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG  | 33             | -             | -                               |
| Reverse overhang | Nextera tag | GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG | 34             | -             | -                               |
| <i>rbcL</i> 646F | Algae       | ATGCGTTGGAGAGARCGTTTC              | 21             | 419           | (Glover, 2019)                  |
| <i>rbcL</i> 998R | Algae       | GATCACCTTCTAATTTACWACA ACTG        | 27             | 419           | (Glover, 2019)                  |
| 515F             | Bacteria    | GTGCCAGCMGCCGCGGTAA                | 19             | 358           | (Caporaso <i>et al.</i> , 2011) |
| 806R             | Bacteria    | GACTACHVGGGTWTCTAAT                | 20             | 358           | (Caporaso <i>et al.</i> , 2011) |

Table 2.3: PCR reaction composition for 16S rRNA and *rbcL* gene.

|                                | Single reaction / $\mu\text{L}$ |
|--------------------------------|---------------------------------|
| Nuclease free H <sub>2</sub> O | 13                              |
| AllTaq Mastermix (4x)          | 5                               |
| Forward primer                 | 0.5                             |
| Reverse primer                 | 0.5                             |
|                                |                                 |
| Volume of template DNA         | 1                               |
| Final volume of PCR reaction   | 20                              |
| Final primer concentration     | 0.30 pmol / $\mu\text{L}$       |

Table 2.4: PCR conditions for 16S rRNA and *rbcL* gene.

|           | PCR stages           | Time (minutes) | Temperature ( $^{\circ}\text{C}$ ) |
|-----------|----------------------|----------------|------------------------------------|
|           | Initial denaturation | 2:00           | 95                                 |
| 40 Cycles | Denaturation         | 0:05           | 95                                 |
|           | Annealing            | 0:15           | 55                                 |
|           | Extension            | 0:10           | 72                                 |
|           | Final Extension      | 5:00           | 72                                 |
|           | Hold                 | $\infty$       | 4                                  |

PCR success was assessed using the QIAxcel (QIAGEN, Germany) to ensure the correct product size and concentrations were present. The QIAxcel DNA High-Resolution Kit was employed, enabling the detection of DNA fragments between 15 base pair (bp) and 5 kilobase (kb) in size. Successful amplifications were stored at  $-20^{\circ}\text{C}$ , and any weak amplifications were reamplified to improve product generation. The successful triplicates were subsequently pooled together before the clean stage (detailed in Section 2.4.2).

#### 2.4.2 – 16S rRNA and *rbcL* amplicon clean-up

Successfully amplified samples were cleaned using a Zymo Research 96 DNA clean-up kit (Zymo Research, Cambridge) as per the manufacturer's instructions. In brief, 300  $\mu\text{L}$  of DNA binding buffer was added to each pooled triplicated sample contained in each well in the 96 well plates. The plate was sealed

and vortexed for 1 minute and subsequently centrifuged at 3000 x g for 5 minutes. The samples were then added to the wells of a Silicon-A plate mounted onto a collection plate. This plate was then centrifuged at 3000 x g for 5 minutes until sample mixtures were filtered entirely. 300 µL of wash buffer was then added to each well of the silicon-A plate and centrifuged at 3000 x g for 5 minutes; this step was repeated once. DNA was eluted by adding 30 µL of nuclease-free water into the silicon-A plate mounted onto an elution plate and centrifuged at 3000 x g for 3 minutes.

### 2.4.3 – Secondary Illumina-Nextera indexing adapter PCR

Cleaned amplicon samples underwent a secondary amplification with Illumina-Nextera index adapters. The addition of the Nextera index adapter to the primary nextera tagged primer skips the tagmentation of adapters in the standard Nextera XT protocol where PCR amplification bias can occur. In a 96-well plate format, 2.5 µL of a unique Nextera index adapter 1 (N7XX) for each column was added to each well going down rows A-H (see Figure 2.11). Additionally, 2.5 µL of a unique Nextera index adapter 2 (S5XX) for each row was added across each columns 1-12 (see Figure 2.11). KAPA HiFi HotStart ReadyMix (12.5 µL) and 5 µL of nuclease-free water were added to each well of the 96 well plates. Amplified DNA (2.5 µL) was then transferred into the corresponding wells of the new plate. The plate was sealed and vortexed to ensure the mixture was homogenous and then centrifuged for 1 minute at 14,000 rpm. This plate then underwent a PCR amplification using the conditions listed below in Table 2.5 on a MultiGene™ OptiMax Thermal Cycler (Labnet International).

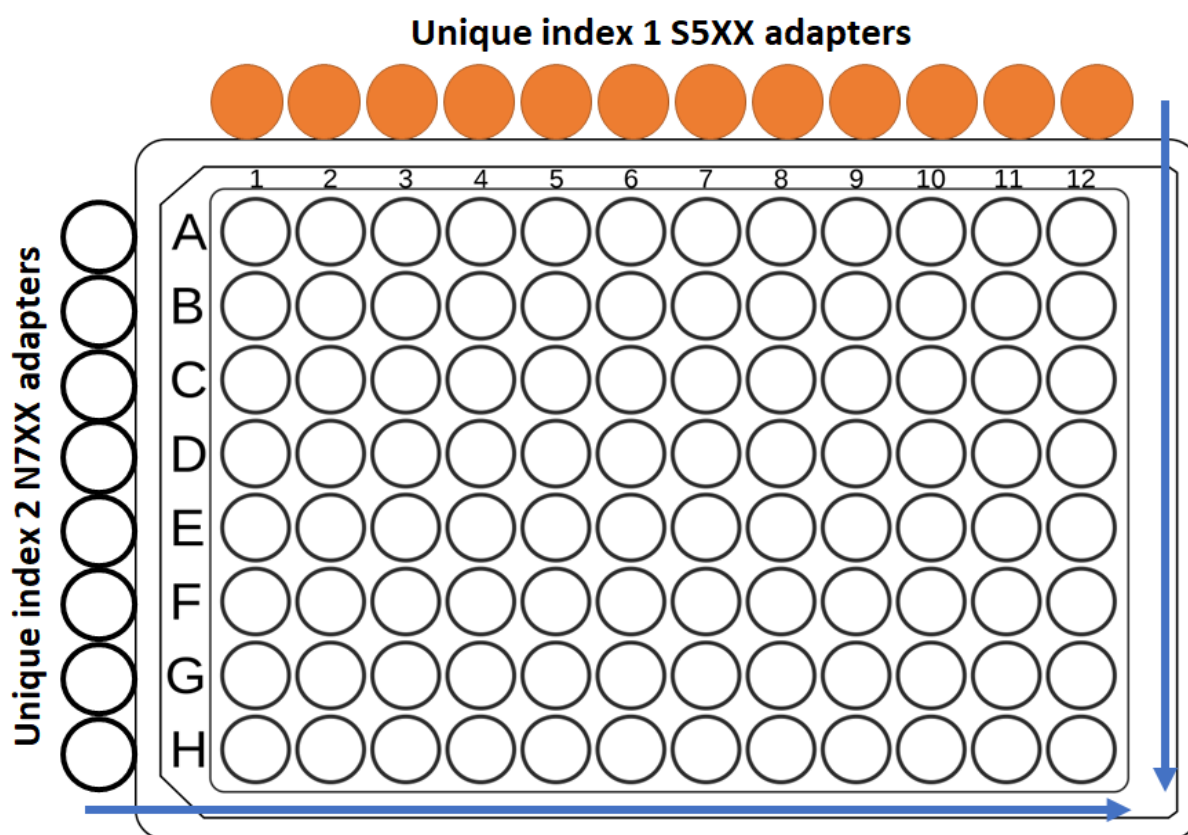


Figure 2.11: Illumina-Nextera secondary PCR set up: columns 1-12 in the 96 well plate containing cleaned amplicons have the addition of a unique index 1 S5XX adapters (orange caps) and rows A-H have the addition of unique index 2 (N7XX) adapters (white caps).

Table 2.5: PCR conditions used for Illumina PCR.

| PCR stages           | Time<br>(seconds) | Temp<br>°C |
|----------------------|-------------------|------------|
| Initial denaturation | 3 minutes         | 95         |
| Cycle number         | 8                 |            |
| Denaturation         | 30                | 95         |
| Annealing            | 30                | 55         |
| Extension            | 30                | 72         |
| Final extension      | 5 minutes         | 72         |
| Hold                 | ∞                 | 4          |

Random samples were assessed via a 4200 TapeStation System (Agilent) to determine the success of the secondary PCR using a High Sensitivity D1000

ScreenTape. All sample concentrations were measured through a QIAxcel measurement which ensured that a band was present for 16S rRNA and *rbcL* gene samples; if there were less than 90% of samples with a band, these samples were repeated.

#### 2.4.4 – SequelPrep Normalisation Plates

SequelPrep normalisation plates (Invitrogen, USA) were used to normalise the concentration of DNA in a 96-well plate format. The desired PCR product (20 µL) was transferred from the PCR plate into the well of the SequelPrep normalisation plate. An additional 20 µL of normalisation buffer was then added to the SequelPrep plate. The plate was then sealed with foil tape and subsequently vortexed for 1 minute to mix the contents. The plate was then incubated at room temperature for 1 hour. Without touching the plate sides, all liquid was aspirated from the wells of the SequelPrep plate. SequelPrep normalisation wash buffer (50 µL) was then added to the wells and mixed up and down through pipetting to improve the removal of contaminants. The wash buffer was then aspirated and discarded from the wells. SequelPrep normalisation elution buffer (20 µL) was then added to each well of the plate, and the plate was sealed and vortexed for 1 minute and incubated at room temperature for 5 minutes. The purified DNA was then pooled together for the next stage.

#### 2.4.5 – Next Generation Sequencing on the Illumina-MiSeq platform

Samples were sequenced for 16S rRNA and *rbcL* genes on four separate runs with either an Illumina MiSeq-nano or v2 cartridge for both genes using 2 x 250 bp paired-end reads. There were seven 96-well plates with an additional pilot set containing 18 samples.

#### 2.4.6 – Databases

A representative sequence encompassing 16S rRNA variable regions V4 & V5 (515 to 806 bp) and associated taxonomic classification extracted from the silva database v132 (Callahan, 2018) was downloaded from the qiime2 repository on 31 August 2020. A Bayesian classifier was generated using qiime2 (version 2021.2 - <https://view.qiime2.org/>) using the fit-classifier-naive-bayes function.

Diatom (*Bacillariophyceae*) RuBisCO large (*rbcL*) subunit sequences and associated taxonomy were downloaded from DIAT.BARCODE [Diat.barcode, an open-access curated barcode library for diatoms | Scientific Reports (nature.com)] curated database (25 October 2021 - version 9.2 of the database). In parallel, all *rbcL* entries were downloaded from GenBank (25 Oct 2021). Any GenBank accession appearing DIAT was removed from the GenBank download. Subsequently, GenBank accessions were used to derive NCBI taxonomic identifiers using `nucl_gb.accession2taxid` table, this allowed complete taxonomic classification to be extracted from `rankedlineage.dmp`, both of these files were downloaded from NCBI on 25th of October 2021. The sequences and taxonomy were then formatted as described in the qiime2 (v. 2021.2) instructions, and DIAT and GenBank libraries were concatenated. The concatenated sequences and taxonomy were imported into qiime2, and the `extract-reads` function was used to filter out sequences containing the forward and reverse primers (the default 0.8 identify threshold was applied). The extracted and trimmed reads were then combined with the taxonomy using a bayesian classifier approach qiime2 (version 2021.2 - <https://view.qiime2.org/>) using the `fit-classifier-naive-bayes` function to generate a diatom *rbcL* classifier.

#### 2.4.7 – Bioinformatic analysis using QIIME2

The bioinformatics programme QIIME2 was used to analyse the sequence data (Bolyen et al., 2019) (<https://view.qiime2.org/>). First, all raw sequence fastq files for replicate samples were concatenated before the bioinformatic pipeline. Paired-end sequences were then imported and demultiplexed before being processed using the DADA2 bioinformatics pipeline (Callahan et al., 2016) to produce merged, denoised, chimera-free, inferred sample sequences. Parameters used in the DADA2 analysis are listed in Table 2.6; quality scores for the forward and reverse reads for all MiSeq runs for both 16S rRNA, and *rbcL* genes (Appendix 2.1 – 2.8) were used to assess truncation. Samples from each run were merged using the amplicon sequence variant (ASV) tables produced from the DADA2 outputs. ASVs were subjected to the taxonomic assignment using the qiime `feature-classifier classify-sklearn` function against the Silva database 138 (Callahan, 2018) for 16S rRNA samples, and the curated *rbcL* database as mentioned in Section 2.4.6 for *rbcL*

samples. Samples then underwent proportional filtering per taxa using the qiime feature-table filter-features-conditionally function. The range was set to 0.001 for a 0.1% removal of the minimum relative abundance for a feature to be retained using --p-abundance. Then the minimum portion of samples that a feature must have a relative abundance of at least 'abundance' to be retained was set to 0.001 (0.1 %) for the --p-prevalence function.

Table 2.6: Parameters used for all samples processed through the DADA2 bioinformatics pipeline for both 16S rRNA and *rbcL* genes.

| Parameter       | 16S rRNA |           |           |           | <i>rbcL</i> |           |           |           |
|-----------------|----------|-----------|-----------|-----------|-------------|-----------|-----------|-----------|
|                 | Pilot    | Plate 1-2 | Plate 3+7 | Plate 4-6 | Pilot       | Plate 1-2 | Plate 3+7 | Plate 4-6 |
| --p-trim-left-f | 0        | 0         | 0         | 0         | 0           | 0         | 0         | 0         |
| --p-trim-left-r | 0        | 25        | 40        | 0         | 0           | 10        | 0         | 0         |
| --p-trunc-len-f | 250      | 250       | 250       | 250       | 250         | 250       | 250       | 250       |
| --p-trunc-len-r | 250      | 250       | 250       | 250       | 250         | 250       | 250       | 250       |
| --p-max-ee-r    | 10       | 10        | 10        | 10        | 10          | 10        | 10        | 10        |
| --p-max-ee-f    | 5        | 5         | 5         | 5         | 5           | 5         | 5         | 5         |
| --p-n-threads   | 32       | 32        | 32        | 32        | 32          | 32        | 32        | 32        |

# Chapter 3: Monthly changes in T&O concentrations reveal ammonium and phosphorous to be key triggers





### 3.1 Introduction

With the pressure of T&O events culminating in numerous customer complaints to water companies, it is pertinent for water companies to find preventative measures. Current treatments include coagulation, flocculation, adsorption, precipitation, reverse osmosis, membrane bioreactor, nanofiltration, and ozonation electro dialysis (Mustapha et al., 2021). However, installing technology in water treatment works exclusively for treating T&O compounds is not always practical or economically beneficial (Srinivasan and Sorial, 2011). Advanced water treatment processes, such as ozonation and activated carbon adsorption, are costly in capital and operational costs, so combining methods would exacerbate overall costs. Costly treatment processes have led to water companies exploring integrated methods for more effective removal of these T&O compounds (Mustapha et al., 2021). These methods aim to implement preventative measures and formulate predictive models, although, first, more must be known about the triggers that influence *Cyanobacteria* to produce geosmin and 2-MIB.

Production of T&O compounds by *Cyanobacteria* is species and strain-specific – this, combined with environmental parameters, represents a complex system composed of multiple interactions and not necessarily linear behaviours, which can be difficult to predict (Bruder et al., 2014). Although geosmin and 2-MIB are severe problems (Yang et al., 2010), especially in drinking water, few studies have focused on potential causes and solutions (Abd El-Hack et al., 2022). One well-known cause is that *Cyanobacteria* thrive in environments where nitrogen-to-phosphorous is low and when ammonium-to-nitrate ratios are high (Howard, 2020). Additionally, seasonal variations of *Cyanobacteria* are said to be linked closely to fluctuations in water quality (Watson et al., 2007), like nutrients (T. Zhang et al., 2017). Predictive tools for monitoring T&O “events” based on certain water quality parameters like trophic status of the water body would be ideal for cost effective treatment for drinking water treatment.

The distribution of geosmin and 2-MIB in drinking water reservoirs differs with different trophic statuses (Jüttner and Watson, 2007). Tucker (2000) reported that due to the light-limited nature of eutrophic ecosystems, planktonic organisms present

in the upper levels of the water column have a competitive advantage compared with benthic or substrate-attached organisms. Studies conducted by Smith et al. (2002) and Downing et al. (2001) revealed the trophic state of the reservoir, especially chlorophyll *a* (chl *a*) levels, to be a good indicator of T&O compounds. Both studies reported *Cyanobacteria* and T&O concentrations to be directly proportional to the chl *a* concentrations in the water body. In addition, some state regulatory agencies have adopted a chl *a* criterion to maintain water quality in reservoirs by preventing the formation of T&O compounds (Oklahoma Water Resources Board, 2005). However, studies by Dzialowski et al. (2009) using a series of predictive models to relate reservoir geosmin concentrations to water quality variables found that trophic state alone could not be used to predict T&O “events”. Total Phosphorous (TP), chl *a*, Total Nitrogen (TN) and Secchi disk (SD) measurements are frequently used as Trophic State Indicators (TSIs) in water quality and trophic evaluation models (Páez et al., 2001). Carlson first proposed the Carlson Trophic State Index (CTSI) using TSIs to determine the eutrophication levels of lakes (Carlson, 1977). Subsequently, CTSI has been widely used in lakes and reservoirs for assessing the water quality of the water body (An and Park, 2003; Jarosiewicz et al., 2012; Matthews et al., 2002; Sechi and Sulis, 2007; Sharma et al., 2010; Sheela et al., 2011). Since then, CTSI has been widely used for determining the trophic state of aquatic ecosystems and has been utilised to estimate the abundance or potential abundance of phytoplankton, including T&O-producing *Cyanobacteria*, which are becoming an increasing problem to surface drinking water utilities worldwide.

Nutrient concentrations in conjunction with cyanobacterial abundance have been well studied (Davis et al., 2009; Domingues et al., 2011; James et al., 2009; Shen et al., 2021), with the primary focus being on the role of phosphorous and nitrogen. Generally, cyanobacterial dominance is in part dependent on the total nitrogen to total phosphorous ratio (TN:TP), with *Cyanobacteria* at their highest growth rates during phases when TN:TP ratio is low (<29:1 by mass; Smith, 1983). Heightened cyanobacterial growth rates at lower TN:TP ratios is thought to be due to some species having the ability to form differentiated cells, and heterocysts, enabling these *Cyanobacteria* to fix atmospheric nitrogen (N<sub>2</sub>) (Shatwell and Köhler, 2019). Heterocysts give *Cyanobacteria* a competitive advantage when the ratio of nitrogen to phosphorous is low and the availability of nitrogen is limiting (Grover et al., 2019).

Recent research has shown that the type of nitrogen available also plays a large role in cyanobacterial abundance (Flores and Herrero, 2005). Oxidised inorganic nitrogen ( $\text{NO}_3^-$ ), compared to reduced inorganic nitrogen ( $\text{NH}_4^+$ ), possesses different energetic costs and assimilation rates for different phytoplankton taxa, including *Cyanobacteria*, which prefer reduced forms of nitrogen (Flores and Herrero, 2005). A more recent study has focussed on the TN:TP ratio alongside  $\text{NO}_3^-:\text{NH}_4^+$  with the production of T&O compounds (Howard, 2020), revealing that whilst low TN:TP ratios favoured cyanobacterial dominance and growth, low  $\text{NO}_3^-:\text{NH}_4^+$  ratios were associated with a heightened production of T&O compounds. T&O production during low  $\text{NO}_3^-:\text{NH}_4^+$  ratios is consistent with proceedings from Perkins et al. (2019), which found associations between low  $\text{NO}_3^-:\text{NH}_4^+$  ratios and heightened geosmin and 2-MIB production with  $\text{NH}_4^+$  identified as a critical trigger. In addition, the availability of ferrous iron ( $\text{Fe}^{2+}$ ) has been proposed to be important in regulating the ability of *Cyanobacteria* to compete with eukaryotic competitors (Molot et al., 2010). Molot et al. (2014) suggested that the scarcity of  $\text{Fe}^{2+}$  in P-limited oxygenated waters severely limits *Cyanobacteria* productivity unless supplemented by migrating down into  $\text{Fe}^{2+}$ -rich anoxic water enriched by internal loading at the anoxic sediment interface. Although  $\text{Fe}^{2+}$  can be sequestered in ferrous phosphate or ferrous sulphides as sulphate reduction to sulphide can limit  $\text{Fe}^{2+}$  diffusion rates from anoxic sediments because of insoluble iron sulphide formation (Carignan and Tessier, 1988). Hence, the availability of  $\text{Fe}^{2+}$  in relation to sulphate could help explain T&O events in mesotrophic water bodies (Durrer et al., 1999). Although nutrients are associated with increases in cyanobacterial productivity promoting T&O compounds, some studies have shown direct correlations between nutrients and T&O compounds to be weak (Bai et al., 2017). Interestingly, no studies to date have evaluated nutrient changes and how this influences the rate of change in T&O production.

The abundance of T&O compounds in a water body can vary greatly seasonally and across years (Stumpf et al., 2012); T&O compounds tend to increase during the warmer summer months when elevated temperatures and prolonged light irradiance stimulate cyanobacterial productivity and hence T&O production (Jöhnk et al., 2008; Watson et al., 2008). Westerhoff et al. (2005) found 2-MIB concentrations to increase from spring to late summer in three water supply reservoirs in America. Similarly, an investigation into T&O compounds in tap water during 1994 – 1997

found that geosmin and 2-MIB episodes occurred during warmer seasons, specifically from late spring to early autumn (Bruchet, 1999). In accordance, Yagi (2005) reviewed T&O-related problems in the southern basin of Lake Biwa from 1970 – 2005 and found that geosmin and/ or 2-MIB were consistently present from July to October. However, there is emerging evidence that T&O problems can occur during the colder months and are not necessarily confined to warmer seasons (Li et al., 2016). Jiang et al. (2016) revealed winter peaks in geosmin concentrations in the surface water of Lake Chaohu. In addition, Dzialowski et al. (2009) confirmed geosmin to be present at elevated concentrations during the winter. Thus, seasonal variations need to be considered in monitoring geosmin and 2-MIB concentrations. Further research is needed to elucidate mechanisms underpinning the production of T&O compounds during winter months.

To better understand the seasonal patterns and driving factors of geosmin and 2-MIB, water chemistry and physical data processed by Welsh Water Dŵr Cymru (DCWW) was analysed from nine reservoirs in Wales (see Chapter 2 – Materials and Methods) to determine triggers of T&O events. The trophic status of each reservoir was determined to see if the CTSI and/ or TSIs could be used as an appropriate indicator for T&O production. In addition, monthly changes in nutrients, namely,  $\text{NH}_4^+$ ,  $\text{NO}_3^-$ , TP, Orthophosphate,  $\text{Fe}^{2+}$ , sulphate, and dissolved reactive silicate, were established to see the effects of the monthly changes in T&O concentrations. Effects of monthly changes in nutrient concentrations and monthly lagged changes in T&O compounds were also assessed to evaluate any time-lapsed influences on T&O production. It is expected that general CTSI will not necessarily indicate T&O events, although TSIs may show some relation to T&O outcome mainly when analysed by season and year. Due to recorded monthly data, the monthly changes in nutrients and T&O concentrations will show better correlations leading to T&O events with  $\text{NH}_4^+$ ,  $\text{NO}_3^-$ , TP and Orthophosphate showing significant influences. This research will provide the chemical mechanisms underpinning nutrient relationships in conjunction with geosmin and 2-MIB production, revealing important triggers for T&O events.

## 3.2 – Materials and Methods

### 3.2.1 – Sample collection

Samples were collected according to Chapter 2 – Materials and methods 2.1, from sites detailed in Sections 2.1.1 – 2.1.7 with the sampling exception mentioned in Section 2.1.8. Sampled months were calculated according to Chapter 2 – material and methods 2.1. Seasons were assigned as spring; March 1<sup>st</sup> – May 31<sup>st</sup>, summer; June 1<sup>st</sup> – August 31<sup>st</sup>, autumn; September 1<sup>st</sup> – November 30<sup>th</sup>, and winter; December 1<sup>st</sup> – February 28<sup>th</sup>.

### 3.2.2 – Physical and chemical water quality testing

All chemical and physical analyses of water samples were conducted by Dŵr Cymru, Welsh Water in Glaslyn, Newport. All tests were carried out adhering to ISO/IEC 17025:2019. pH and turbidity were measured according to Chapter 2, Materials and Methods Section 2.3.6. Chlorophyll *a* (Chl *a*) was measured according to Chapter 2, Materials and Methods Section 2.3.4. Nutrients ammonium (NH<sub>4</sub><sup>+</sup>), Total Oxidized Nitrogen (TON), nitrate (NO<sub>3</sub><sup>-</sup>), nitrite (NO<sub>2</sub><sup>-</sup>), orthophosphate, total phosphate (TP), sulphate (SO<sub>4</sub><sup>2-</sup>), and dissolved reactive silicate were analysed using materials and methods from Chapter 2 – Materials and Methods, Section 2.3.2. Concentrations of T&O compounds, geosmin and 2-MIB were determined using materials and methods detailed in Chapter 2 – Materials and Methods, Section 2.3.1.

### 3.2.3 – Assignment of Trophic State Indices and overall Carlson Trophic State Index by season and year

Carlson's Trophic State Index (CTSI) is a classic method for characterising a reservoir's trophic state or overall health. The trophic state is defined as the total weight of the biomass in a water body at a specific location and time (Prasad and Siddaraju, 2012), and is the biological response to nutrient additions to the water bodies. Nutrient effects may be modified by specific parameters, for example; seasonal variations, grazing of phytoplankton/ zooplankton, mixing depth of the water. CTSI uses three main variables, namely Chl *a*, Secchi disc depth (SD) and

total phosphorous (TP) to calculate three Trophic State Indices (TSI) to calculate CTSI.

SD depth was calculated from turbidity (T) measured in NTU, using an inverse power function conversion factor (Equation 1) as described by Çako et al. (2013) with a high correlation coefficient ( $R^2 = 0.85$ ).

$$SD = 3.22T^{-0.222} \quad \text{Equation 1}$$

The three TSI ( $TSI_{SD}$ ,  $TSI_{TP}$  and  $TSI_{Chl a}$ ) were calculated using the following equations using the trophic state indicators (SD, TP, Chl a) listed below (Equation 2 – 4):

$$TSI(SD) = 60 - 14.41 \ln(SD) \quad \text{Equation 2}$$

$$TSI(TP) = 14.42 \ln(TP) + 4.15 \quad \text{Equation 3}$$

$$TSI(Chl a) = 9.81 \ln(Chl a) + 30.6 \quad \text{Equation 4}$$

TSI assigned to each reservoir using all reservoir locations (excluding tributary locations) were grouped by season and year, and an average and standard deviation was obtained. The overall CTSI for each reservoir by season and year was then calculated by averaging the TSI values obtained from the three trophic state indicators listed above. The CTSI and TSI indicators assigned to each reservoir were then compared with the criteria displayed in Table 3.1: revealing the reservoir's overall trophic status and trophic status of individual TSI indicators during the recorded season and year. Attributing characteristics of the trophic status are displayed in Table 3.2 based on the CTSI values.

Table 3.1: Trophic state classification scheme based on Carlson's (1977) trophic state index

| CTSI       | Trophic Status | Values before TSI calculations |   |  |
|------------|----------------|--------------------------------|---|--|
|            |                | Secchi Depth (SD) m            | Total Phosphorous (TP) $\mu\text{g L}^{-1}$ | Chlorophyll-a (chl <sub>a</sub> ) $\mu\text{g L}^{-1}$ |
| 0 - 40     | Oligotrophic   | > 8 – 4                        | 0 – 12                                      | 0 – 2.6  |
| 40 - 50    | Mesotrophic    | 4 – 2                          | 12 – 24                                     | 2.6 – 7.3  |
| 50 - 70    | Eutrophic      | 2 – 0.5                        | 24 – 96                                     | 7.3 – 56   |
| 70 - 100 + | Hypereutrophic | 0.5 – < 0.25                   | 96 – 384 +                                  | 56 – 155 +   |

Table 3.2: CTSI values assigned to trophic status of the reservoir and relevant attributes characterising the trophic state (Carlson and Simpson, 1996).

| CTSI values | Trophic Status | Attributes   |
|-------------|----------------|--|
| < 30        | Oligotrophic   | Clearwater and oxygen throughout the year in the hypolimnion                               |
| 30 - 40     | Oligotrophic   | Oligotrophic, although shallow reservoirs may become anoxic during the summer              |
| 40 - 50     | Mesotrophic    | Water is moderately clear, but increasing probability of anoxia during the summer          |
| 50 - 60     | Eutrophic      | The lower boundary of classical eutrophy: decreased transparency                           |
| 60 - 70     | Eutrophic      | The dominance of <i>Cyanobacteria</i> , algal scum probable, extensive macrophyte problems |
| 70 - 80     | Hypereutrophic | Heavy algal blooms are possible throughout the summer                                      |
| > 80        | Hypereutrophic | Algal scum, summer fish kills, few macrophytes   |

### 3.2.4 – Defining a Taste and Odour event

Classification of a defined low, medium and, high geosmin and 2-MIB “event” is detailed in Chapter 2 – Materials and Methods, Section 2.2.



### 3.2.5 – Principal Component Analysis of chemical and physical parameters

PCA analysis was conducted using R studio (version 4.1.0) with called variables shifted to be zero-centred and scaled to have unit variance. Graphical representations were completed by colouring individual points with geosmin or 2-MIB levels to identify variables associated with 1.Low, 2.Medium, and 3.High geosmin and 2-MIB levels.

PCAs were employed on reservoir and tributary locations that experienced maximum geosmin and 2-MIB concentrations from the entirety of this study exceeding the “event” level classification. Variables used for PCA analysis included pH, conductivity (CON) ( $\mu\text{S}/\text{cm}$  at  $20^\circ\text{C}$ ), turbidity (Turb) (NTU), total organic carbon (TOC) ( $\text{mg L}^{-1}$ ), ammonium ( $\text{NH}_4$ ) ( $\text{mg L}^{-1}$ ), nitrate ( $\text{NO}_3$ ) ( $\text{mg L}^{-1}$ ), nitrite ( $\text{NO}_2$ ) ( $\text{mg L}^{-1}$ ), Orthophosphate (Ortho) ( $\text{mg L}^{-1}$ ), dissolved reactive silicate (Sil) ( $\text{mg L}^{-1}$ ), sulphate ( $\text{SO}_4$ ) ( $\text{mg L}^{-1}$ ), total phosphorous (TP) ( $\text{mg L}^{-1}$ ), dissolved organic carbon (DOC) ( $\text{mg L}^{-1}$ ), dissolved magnesium (DMn) ( $\text{mg L}^{-1}$ ) and dissolved iron (DFe) ( $\text{mg L}^{-1}$ ). Variables for reservoir and tributary locations with a constant value (thus no standard deviation) were removed from the dataset before PCA analysis.

### 3.2.6 – Monthly changes in T&O concentrations and nutrients

The R package “data.table” (Dowle et al., 2019) was utilised to calculate the monthly changes in geosmin and 2-MIB concentrations, along with monthly changes in nutrients: ammonium ( $\text{NH}_4^+$ ), total phosphorous (TP), nitrate ( $\text{NO}_3^-$ ), sulphate ( $\text{SO}_4^{2-}$ ), dissolved iron ( $\text{Fe}^{2+}$ ) and dissolved reactive silicate (DSil) between each consecutive sampled month. Samples were generally taken monthly, but in the case of locations that had more than one set of data present for the same sampled month, an average was taken. Only reservoir locations from reservoirs that had maximum T&O concentrations from the entirety of the study period above the stated “event” thresholds were analysed – Alaw (geosmin:  $520 \text{ ng L}^{-1}$ ), Cefni (geosmin:  $13 \text{ ng L}^{-1}$ ), Dolwen (geosmin:  $29 \text{ ng L}^{-1}$ ), Llwyn On (geosmin:  $25 \text{ ng L}^{-1}$ ), Plas Uchaf (geosmin:  $11 \text{ ng L}^{-1}$ ), Pentwyn (2-MIB:  $58 \text{ ng L}^{-1}$ ), and Pontsticill (geosmin:  $15 \text{ ng L}^{-1}$  and 2-MIB:  $7.90 \text{ ng L}^{-1}$ ). Tributary data were excluded from this analysis due to the higher

flushing rates experienced in tributaries; monthly monitoring would not reveal any significant relationships.

Initial reservoir data frames containing raw data were first converted into `data.table`'s for compatibility with the R package "data.table". Each reservoir's `data.table` was then used to calculate the monthly changes in geosmin, 2-MIB,  $\text{NH}_4^+$ , TP,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{Fe}^{2+}$  and DSil grouped according to reservoir location and sampled month, using the `shift()` function. The `shift()` function was implemented to lag the sampled months by one consecutive month, giving the monthly change as:

$$\text{current value} - \text{previous value}$$

For missing monthly data, data available from the previous month was used to calculate the monthly change in mentioned variables for the next sampled month. Monthly changes determined when monthly gaps in the data were apparent need to be interpreted with caution. Reservoir location Res-1 was removed from Alaw, and reservoir locations Res-1 and Res-2 were removed from Cefni due to limited data present.

Table 3.3: Sampled months used to calculate the monthly changes in nutrients and T&O concentrations for each reservoir.

| Reservoir        | Location | Month in plot | Sampled Months used for change | Reservoir             | Location | Month in plot | Sampled Months used for change | Reservoir                  | Location | Month in plot | Sampled Months used for change | Reservoir             | Location | Month in plot | Sampled Months used for change | Reservoir                                 | Location | Month in plot | Sampled Months used for change | Reservoir   | Location | Month in plot | Sampled Months used for change | Reservoir                       | Location | Month in plot | Sampled Months used for change |    |         |    |         |    |         |    |         |   |       |   |       |   |       |
|------------------|----------|---------------|--------------------------------|-----------------------|----------|---------------|--------------------------------|----------------------------|----------|---------------|--------------------------------|-----------------------|----------|---------------|--------------------------------|---|----------|---------------|--------------------------------|---|----------|---------------|--------------------------------|---------------------------------|----------|---------------|--------------------------------|----|---------|----|---------|----|---------|----|---------|---|-------|---|-------|---|-------|
| A<br>L<br>A<br>W | Res - 1  | 1             | N/A                            | C<br>E<br>F<br>N<br>I | Res-3    | 1             | N/A                            | D<br>O<br>L<br>W<br>E<br>N | Res-1    | 1             | N/A                            | L<br>L<br>W<br>Y<br>N | Res-1    | 1             | N/A                            | P<br>L<br>A<br>S<br>U<br>C<br>H<br>A<br>F | Res-1    | 1             | N/A                            | P<br>O<br>N<br>T<br>S<br>T<br>I<br>C<br>I<br>L<br>L | Res-1    | 1             | N/A                            | P<br>E<br>N<br>T<br>W<br>Y<br>N | Res-1    | 1             | N/A                            |    |         |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 2             | 2 - 1                          |                       |          | 2             | 2 - 1                          |                            |          | 2             | 2 - 1                          |                       |          | 2             | 2 - 1                          |   |          | 2             | 2 - 1                          |   |          | 2             | 2 - 1                          |                                 |          | 2             | 2 - 1                          | 2  | 2 - 1   | 2  | 2 - 1   | 2  | 2 - 1   | 2  | 2 - 1   | 2 | 2 - 1 |   |       |   |       |
|                  |          | 3             | 3 - 2                          |                       |          | 3             | 3 - 2                          |                            |          | 3             | 3 - 2                          |                       |          | 3             | 3 - 2                          |   |          | 3             | 3 - 2                          |   |          | 3             | 3 - 2                          |                                 |          | 3             | 3 - 2                          | 3  | 3 - 2   | 3  | 3 - 2   | 3  | 3 - 2   | 3  | 3 - 2   | 3 | 3 - 2 | 3 | 3 - 2 |   |       |
|                  |          | 4             | 4 - 3                          |                       |          | 4             | 4 - 3                          |                            |          | 4             | 4 - 3                          |                       |          | 4             | 4 - 3                          |   |          | 4             | 4 - 3                          |   |          | 4             | 4 - 3                          |                                 |          | 4             | 4 - 3                          | 4  | 4 - 3   | 4  | 4 - 3   | 4  | 4 - 3   | 4  | 4 - 3   | 4 | 4 - 3 | 4 | 4 - 3 |   |       |
|                  |          | 5             | 5 - 4                          |                       |          | 5             | 5 - 4                          |                            |          | 5             | 5 - 4                          |                       |          | 5             | 5 - 4                          |   |          | 5             | 5 - 4                          |   |          | 5             | 5 - 4                          |                                 |          | 5             | 5 - 4                          | 5  | 5 - 4   | 5  | 5 - 4   | 5  | 5 - 4   | 5  | 5 - 4   | 5 | 5 - 4 | 5 | 5 - 4 | 5 | 5 - 4 |
|                  |          | 6             | 6 - 5                          |                       |          | 6             | 6 - 5                          |                            |          | 6             | 6 - 5                          |                       |          | 6             | 6 - 5                          |   |          | 6             | 6 - 5                          |   |          | 6             | 6 - 5                          |                                 |          | 6             | 6 - 5                          | 6  | 6 - 5   | 6  | 6 - 5   | 6  | 6 - 5   | 6  | 6 - 5   | 6 | 6 - 5 | 6 | 6 - 5 | 6 | 6 - 5 |
|                  |          | 7             | 7 - 6                          |                       |          | 7             | 7 - 6                          |                            |          | 7             | 7 - 6                          |                       |          | 7             | 7 - 6                          |   |          | 7             | 7 - 6                          |   |          | 7             | 7 - 6                          |                                 |          | 7             | 7 - 6                          | 7  | 7 - 6   | 7  | 7 - 6   | 7  | 7 - 6   | 7  | 7 - 6   | 7 | 7 - 6 | 7 | 7 - 6 | 7 | 7 - 6 |
|                  |          | 8             | 8 - 7                          |                       |          | 8             | 8 - 7                          |                            |          | 8             | 8 - 7                          |                       |          | 8             | 8 - 7                          |   |          | 8             | 8 - 7                          |   |          | 8             | 8 - 7                          |                                 |          | 8             | 8 - 7                          | 8  | 8 - 7   | 8  | 8 - 7   | 8  | 8 - 7   | 8  | 8 - 7   | 8 | 8 - 7 | 8 | 8 - 7 | 8 | 8 - 7 |
|                  | Res - 2  | 1             | N/A                            |                       | 9        | 9 - 8         | 9                              |                            | 9 - 8    | 9             | 9 - 8                          |                       | 9        | 9 - 8         | 9                              |   | 9 - 8    | 9             | 9 - 8                          |   | 9        | 9 - 8         | 9                              |                                 | 9 - 8    | 9             | 9 - 8                          | 9  | 9 - 8   | 9  | 9 - 8   | 9  | 9 - 8   | 9  | 9 - 8   |   |       |   |       |   |       |
|                  |          | 2             | 2 - 1                          |                       | 13       | 13 - 9        | 13                             |                            | 13 - 9   | 13            | 13 - 9                         |                       | 13       | 13 - 9        | 13                             |   | 13 - 9   | 13            | 13 - 9                         |   | 13       | 13 - 9        | 13                             |                                 | 13 - 9   | 13            | 13 - 9                         | 13 | 13 - 9  | 13 | 13 - 9  | 13 | 13 - 9  | 13 | 13 - 9  |   |       |   |       |   |       |
|                  |          | 3             | 3 - 2                          |                       | 14       | 14 - 13       | 14                             |                            | 14 - 13  | 14            | 14 - 13                        |                       | 14       | 14 - 13       | 14                             |   | 14 - 13  | 14            | 14 - 13                        |   | 14       | 14 - 13       | 14                             |                                 | 14 - 13  | 14            | 14 - 13                        | 14 | 14 - 13 | 14 | 14 - 13 | 14 | 14 - 13 | 14 | 14 - 13 |   |       |   |       |   |       |
|                  |          | 7             | 7 - 3                          |                       | 1        | N/A           | 1                              |                            | N/A      | 1             | N/A                            |                       | 1        | N/A           | 1                              |   | N/A      | 1             | N/A                            |   | 1        | N/A           | 1                              |                                 | N/A      | 1             | N/A                            | 1  | N/A     | 1  | N/A     | 1  | N/A     | 1  | N/A     |   |       |   |       |   |       |
|                  |          | 8             | 8 - 7                          |                       | 2        | 2 - 1         | 2                              |                            | 2 - 1    | 2             | 2 - 1                          |                       | 2        | 2 - 1         | 2                              |   | 2 - 1    | 2             | 2 - 1                          |   | 2        | 2 - 1         | 2                              |                                 | 2 - 1    | 2             | 2 - 1                          | 2  | 2 - 1   | 2  | 2 - 1   | 2  | 2 - 1   | 2  | 2 - 1   |   |       |   |       |   |       |
|                  |          | 13            | 13 - 8                         |                       | 3        | 3 - 2         | 3                              |                            | 3 - 2    | 3             | 3 - 2                          |                       | 3        | 3 - 2         | 3                              |   | 3 - 2    | 3             | 3 - 2                          |   | 3        | 3 - 2         | 3                              |                                 | 3 - 2    | 3             | 3 - 2                          | 3  | 3 - 2   | 3  | 3 - 2   | 3  | 3 - 2   | 3  | 3 - 2   |   |       |   |       |   |       |
|                  |          | 14            | 14 - 13                        |                       | 4        | 4 - 3         | 4                              |                            | 4 - 3    | 4             | 4 - 3                          |                       | 4        | 4 - 3         | 4                              |   | 4 - 3    | 4             | 4 - 3                          |   | 4        | 4 - 3         | 4                              |                                 | 4 - 3    | 4             | 4 - 3                          | 4  | 4 - 3   | 4  | 4 - 3   | 4  | 4 - 3   | 4  | 4 - 3   |   |       |   |       |   |       |
|                  |          | 17            | 17 - 14                        |                       | 5        | 5 - 4         | 5                              |                            | 5 - 4    | 5             | 5 - 4                          |                       | 5        | 5 - 4         | 5                              |   | 5 - 4    | 5             | 5 - 4                          |   | 5        | 5 - 4         | 5                              |                                 | 5 - 4    | 5             | 5 - 4                          | 5  | 5 - 4   | 5  | 5 - 4   | 5  | 5 - 4   | 5  | 5 - 4   |   |       |   |       |   |       |
|                  | Res - 4  | 1             | N/A                            |                       | 6        | 6 - 5         | 6                              |                            | 6 - 5    | 6             | 6 - 5                          |                       | 6        | 6 - 5         | 6                              |   | 6 - 5    | 6             | 6 - 5                          |   | 6        | 6 - 5         | 6                              |                                 | 6 - 5    | 6             | 6 - 5                          | 6  | 6 - 5   | 6  | 6 - 5   | 6  | 6 - 5   | 6  | 6 - 5   |   |       |   |       |   |       |
|                  |          | 2             | 2 - 1                          | 7                     | 7 - 6    | 7             | 7 - 6                          | 7                          | 7 - 6    | 7             | 7 - 6                          | 7                     | 7 - 6    | 7             | 7 - 6                          | 7   | 7 - 6    | 7             | 7 - 6                          | 7   | 7 - 6    | 7             | 7 - 6                          | 7                               | 7 - 6    | 7             | 7 - 6                          | 7  | 7 - 6   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 3             | 3 - 2                          | 8                     | 8 - 7    | 8             | 8 - 7                          | 8                          | 8 - 7    | 8             | 8 - 7                          | 8                     | 8 - 7    | 8             | 8 - 7                          | 8   | 8 - 7    | 8             | 8 - 7                          | 8   | 8 - 7    | 8             | 8 - 7                          | 8                               | 8 - 7    | 8             | 8 - 7                          | 8  | 8 - 7   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 5             | 5 - 3                          | 9                     | 9 - 8    | 9             | 9 - 8                          | 9                          | 9 - 8    | 9             | 9 - 8                          | 9                     | 9 - 8    | 9             | 9 - 8                          | 9   | 9 - 8    | 9             | 9 - 8                          | 9   | 9 - 8    | 9             | 9 - 8                          | 9                               | 9 - 8    | 9             | 9 - 8                          | 9  | 9 - 8   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 6             | 6 - 5                          | 13                    | 13 - 9   | 13            | 13 - 9                         | 13                         | 13 - 9   | 13            | 13 - 9                         | 13                    | 13 - 9   | 13            | 13 - 9                         | 13  | 13 - 9   | 13            | 13 - 9                         | 13  | 13 - 9   | 13            | 13 - 9                         | 13                              | 13 - 9   | 13            | 13 - 9                         | 13 | 13 - 9  |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 7             | 7 - 6                          | 14                    | 14 - 13  | 14            | 14 - 13                        | 14                         | 14 - 13  | 14            | 14 - 13                        | 14                    | 14 - 13  | 14            | 14 - 13                        | 14  | 14 - 13  | 14            | 14 - 13                        | 14  | 14 - 13  | 14            | 14 - 13                        | 14                              | 14 - 13  | 14            | 14 - 13                        | 14 | 14 - 13 |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 8             | 8 - 7                          | 16                    | 16 - 14  | 16            | 16 - 14                        | 16                         | 16 - 14  | 16            | 16 - 14                        | 16                    | 16 - 14  | 16            | 16 - 14                        | 16  | 16 - 14  | 16            | 16 - 14                        | 16  | 16 - 14  | 16            | 16 - 14                        | 16                              | 16 - 14  | 16            | 16 - 14                        | 16 | 16 - 14 |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 13            | 13 - 8                         | 17                    | 17 - 16  | 17            | 17 - 16                        | 17                         | 17 - 16  | 17            | 17 - 16                        | 17                    | 17 - 16  | 17            | 17 - 16                        | 17  | 17 - 16  | 17            | 17 - 16                        | 17  | 17 - 16  | 17            | 17 - 16                        | 17                              | 17 - 16  | 17            | 17 - 16                        | 17 | 17 - 16 |    |         |    |         |    |         |   |       |   |       |   |       |
|                  | Res - 5  | 1             | N/A                            | 1                     | N/A      | 1             | N/A                            | 1                          | N/A      | 1             | N/A                            | 1                     | N/A      | 1             | N/A                            | 1   | N/A      | 1             | N/A                            | 1   | N/A      | 1             | N/A                            | 1                               | N/A      | 1             | N/A                            | 1  | N/A     |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 2             | 2 - 1                          | 2                     | 2 - 1    | 2             | 2 - 1                          | 2                          | 2 - 1    | 2             | 2 - 1                          | 2                     | 2 - 1    | 2             | 2 - 1                          | 2   | 2 - 1    | 2             | 2 - 1                          | 2   | 2 - 1    | 2             | 2 - 1                          | 2                               | 2 - 1    | 2             | 2 - 1                          | 2  | 2 - 1   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 3             | 3 - 2                          | 3                     | 3 - 2    | 3             | 3 - 2                          | 3                          | 3 - 2    | 3             | 3 - 2                          | 3                     | 3 - 2    | 3             | 3 - 2                          | 3   | 3 - 2    | 3             | 3 - 2                          | 3   | 3 - 2    | 3             | 3 - 2                          | 3                               | 3 - 2    | 3             | 3 - 2                          | 3  | 3 - 2   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 5             | 5 - 3                          | 4                     | 4 - 3    | 4             | 4 - 3                          | 4                          | 4 - 3    | 4             | 4 - 3                          | 4                     | 4 - 3    | 4             | 4 - 3                          | 4   | 4 - 3    | 4             | 4 - 3                          | 4   | 4 - 3    | 4             | 4 - 3                          | 4                               | 4 - 3    | 4             | 4 - 3                          | 4  | 4 - 3   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 6             | 6 - 5                          | 5                     | 5 - 4    | 5             | 5 - 4                          | 5                          | 5 - 4    | 5             | 5 - 4                          | 5                     | 5 - 4    | 5             | 5 - 4                          | 5   | 5 - 4    | 5             | 5 - 4                          | 5   | 5 - 4    | 5             | 5 - 4                          | 5                               | 5 - 4    | 5             | 5 - 4                          | 5  | 5 - 4   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 7             | 7 - 6                          | 6                     | 6 - 5    | 6             | 6 - 5                          | 6                          | 6 - 5    | 6             | 6 - 5                          | 6                     | 6 - 5    | 6             | 6 - 5                          | 6   | 6 - 5    | 6             | 6 - 5                          | 6   | 6 - 5    | 6             | 6 - 5                          | 6                               | 6 - 5    | 6             | 6 - 5                          | 6  | 6 - 5   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 8             | 8 - 7                          | 7                     | 7 - 6    | 7             | 7 - 6                          | 7                          | 7 - 6    | 7             | 7 - 6                          | 7                     | 7 - 6    | 7             | 7 - 6                          | 7   | 7 - 6    | 7             | 7 - 6                          | 7   | 7 - 6    | 7             | 7 - 6                          | 7                               | 7 - 6    | 7             | 7 - 6                          | 7  | 7 - 6   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 13            | 13 - 8                         | 8                     | 8 - 7    | 8             | 8 - 7                          | 8                          | 8 - 7    | 8             | 8 - 7                          | 8                     | 8 - 7    | 8             | 8 - 7                          | 8   | 8 - 7    | 8             | 8 - 7                          | 8   | 8 - 7    | 8             | 8 - 7                          | 8                               | 8 - 7    | 8             | 8 - 7                          | 8  | 8 - 7   |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 14            | 14 - 13                        | 13                    | 13 - 9   | 13            | 13 - 9                         | 13                         | 13 - 9   | 13            | 13 - 9                         | 13                    | 13 - 9   | 13            | 13 - 9                         | 13  | 13 - 9   | 13            | 13 - 9                         | 13  | 13 - 9   | 13            | 13 - 9                         | 13                              | 13 - 9   | 13            | 13 - 9                         | 13 | 13 - 9  |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          | 17            | 17 - 14                        | 14                    | 14 - 13  | 14            | 14 - 13                        | 14                         | 14 - 13  | 14            | 14 - 13                        | 14                    | 14 - 13  | 14            | 14 - 13                        | 14  | 14 - 13  | 14            | 14 - 13                        | 14  | 14 - 13  | 14            | 14 - 13                        | 14                              | 14 - 13  | 14            | 14 - 13                        | 14 | 14 - 13 |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          |               |                                | 16                    | 16 - 14  | 16            | 16 - 14                        | 16                         | 16 - 14  | 16            | 16 - 14                        | 16                    | 16 - 14  | 16            | 16 - 14                        | 16  | 16 - 14  | 16            | 16 - 14                        | 16  | 16 - 14  | 16            | 16 - 14                        | 16                              | 16 - 14  | 16            | 16 - 14                        | 16 | 16 - 14 |    |         |    |         |    |         |   |       |   |       |   |       |
|                  |          |               |                                | 17                    | 17 - 16  | 17            | 17 - 16                        | 17                         | 17 - 16  | 17            | 17 - 16                        | 17                    | 17 - 16  | 17            | 17 - 16                        | 17  | 17 - 16  | 17            | 17 - 16                        | 17  | 17 - 16  | 17            | 17 - 16                        | 17                              | 17 - 16  | 17            | 17 - 16                        | 17 | 17 - 16 |    |         |    |         |    |         |   |       |   |       |   |       |

### 3.2.7 – Monthly changes in nutrients and T&O concentrations correlogram's

Correlograms were established using the R statistical package 'GGally' (Emerson et al., 2013). Correlograms enabled graphical illustrations of paired variables (monthly changes in  $\text{NH}_4^+$ , TP,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{Fe}^{2+}$ , DSil and monthly changes in geosmin and 2-MIB concentrations) drawn using scatterplots on the left part of the correlogram. Pearson correlation was used to determine significant relationships from the paired variables displayed on the right of the correlogram. Pearson correlation was used instead of Spearman's rank correlation as the data was assumed to be normally distributed (see Appendix 3.11); Pearson correlation evaluates the linear relationship between two continuous variables. Assigned significance of the paired variables as governed by Pearson correlation was established through the R statistical package 'Hmisc' (Tikhonov et al., 2019). Categorical geosmin and 2-MIB levels were added to the monthly change data by merging data frames for each reservoir based on the unique identifiers: reservoir location and sampled month.

To better understand variable monthly changes influencing geosmin events, monthly changes in geosmin were lagged by one consecutive sampled month and used for comparison. Two correlograms were created, one using standard monthly changes in geosmin concentrations and the second using lagged monthly changes in geosmin concentrations. For these correlograms, data were filtered to ensure no monthly gaps were present; this excluded Llwyn On reservoir from the analysis. No correlogram was implemented for 2-MIB concentrations due to limited data with missing monthly data.

## 3.3 – Results

### 3.3.1 – Concentrations of geosmin and 2-MIB in studied reservoir and tributary locations

For geosmin concentrations, sampling locations within each reservoir at six reservoirs had maximum concentrations above “event” level classification (See Section 3.2.6): Alaw (520 ng L<sup>-1</sup>), Cefni (13 ng L<sup>-1</sup>), Dolwen (29 ng L<sup>-1</sup>), Llwyn On (25 ng L<sup>-1</sup>), Plas Uchaf (11 ng L<sup>-1</sup>) and Pontsticill (15 ng L<sup>-1</sup>). Geosmin concentrations were generally lower in tributary locations within each reservoir’s catchment area (Figure 3.1), with only Alaw having a maximum concentration (16 ng L<sup>-1</sup>) above the “event” threshold. Although reservoir locations from six reservoirs experienced maximum concentrations exceeding geosmin “event” status, mean concentrations were generally low and below the “event” threshold: Cefni (3.46 ± 2.72 ng L<sup>-1</sup>), Dolwen (8.89 ± 9.95 ng L<sup>-1</sup>), Llwyn On (5.35 ± 5.56 ng L<sup>-1</sup>), Plas Uchaf (4.38 ± 2.46 ng L<sup>-1</sup>) and Pontsticill (2.66 ± 3.78 ng L<sup>-1</sup>). Conversely, Alaw’s reservoir locations had a high mean concentration with a large standard deviation (87.26 ± 121.63 ng L<sup>-1</sup>). Three reservoirs out of nine investigated in this study did not experience any geosmin concentrations above the “event” threshold; these reservoirs were Alwen (1.39 ± 0.62 ng L<sup>-1</sup>), Llandegfedd (1.80 ± 0.42 ng L<sup>-1</sup>) and Pentwyn (2.24 ± 0.72 ng L<sup>-1</sup>).

“Event” classification for 2-MIB concentrations was only assigned to maximum concentrations of reservoir locations in two reservoirs, Pentwyn (58 ng L<sup>-1</sup>) and Pontsticill (7.90 ng L<sup>-1</sup>). However, these two reservoirs had mean 2-MIB concentrations that did not exceed the “event” threshold: Pentwyn (4.44 ± 11.59 ng L<sup>-1</sup>) and Pontsticill (1.75 ± 2.14 ng L<sup>-1</sup>). Concentrations of 2-MIB generally appeared to be higher in reservoir locations compared to tributary locations (Figure 3.2), except for maximum concentrations found in Alaw (11 ng L<sup>-1</sup>), Alwen (11 ng L<sup>-1</sup>) and Llandegfedd (5.70 ng L<sup>-1</sup>) tributaries. These maximum concentrations had a minimal effect on mean 2-MIB concentrations for Alaw (1.60 ± 1.95 ng L<sup>-1</sup>), Alwen (0.99 ± 1.58 ng L<sup>-1</sup>) and Llandegfedd (1.14 ± 1.61 ng L<sup>-1</sup>) tributary locations, all mean concentrations being below the “event” concentration threshold.

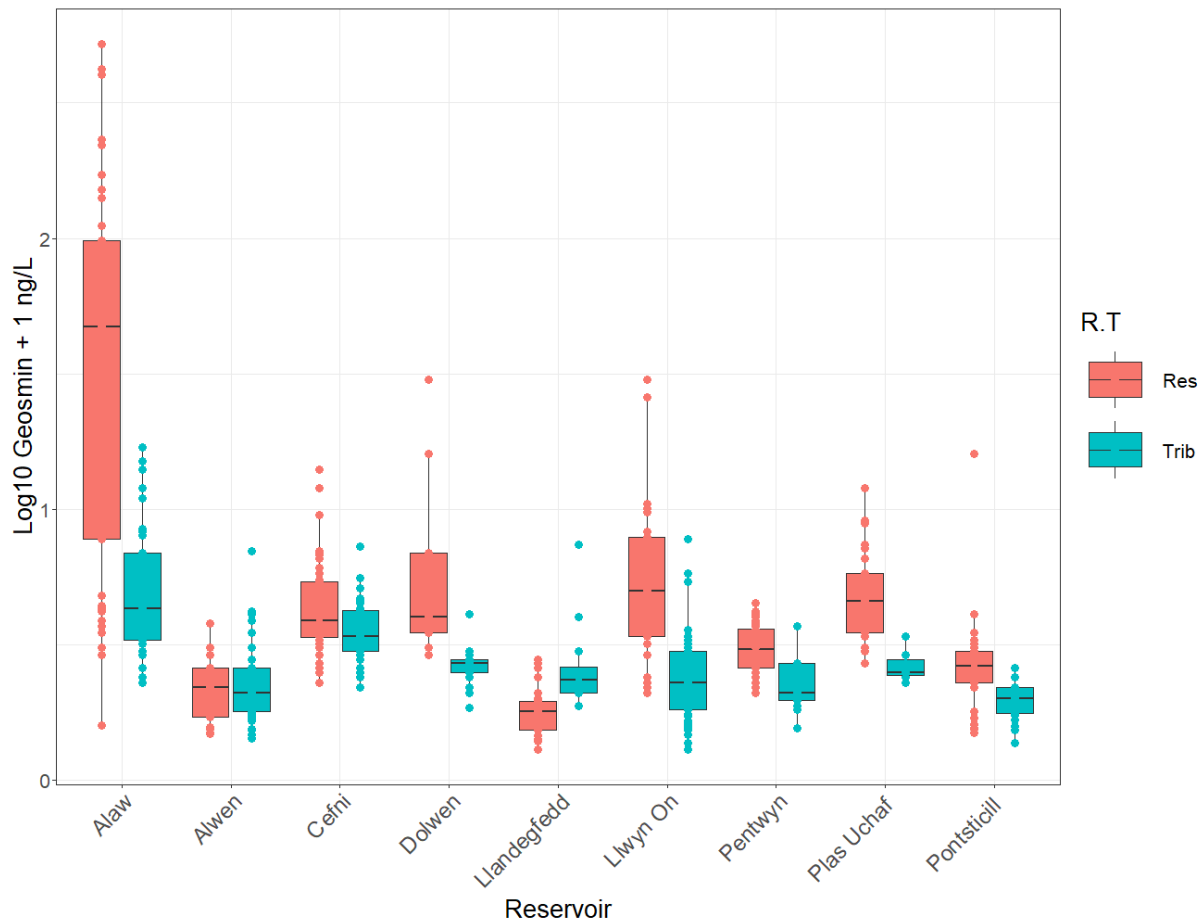


Figure 3.1: Box and whisker plot showing  $\log_{10}$  geosmin concentrations + 1 ( $\text{ng L}^{-1}$ ) from the reservoir (Res in red) and tributary (Trib in blue) locations for all nine reservoirs over the total sampling period (16/7/2019 – 30/11/2020). Raw data is indicated by the coloured dots grouped by reservoir and tributary locations. The length of the box indicates the interquartile range, extending from the 25<sup>th</sup> to the 75<sup>th</sup> percentile and the horizontal bar within the box denotes the median value. The whiskers display the range.

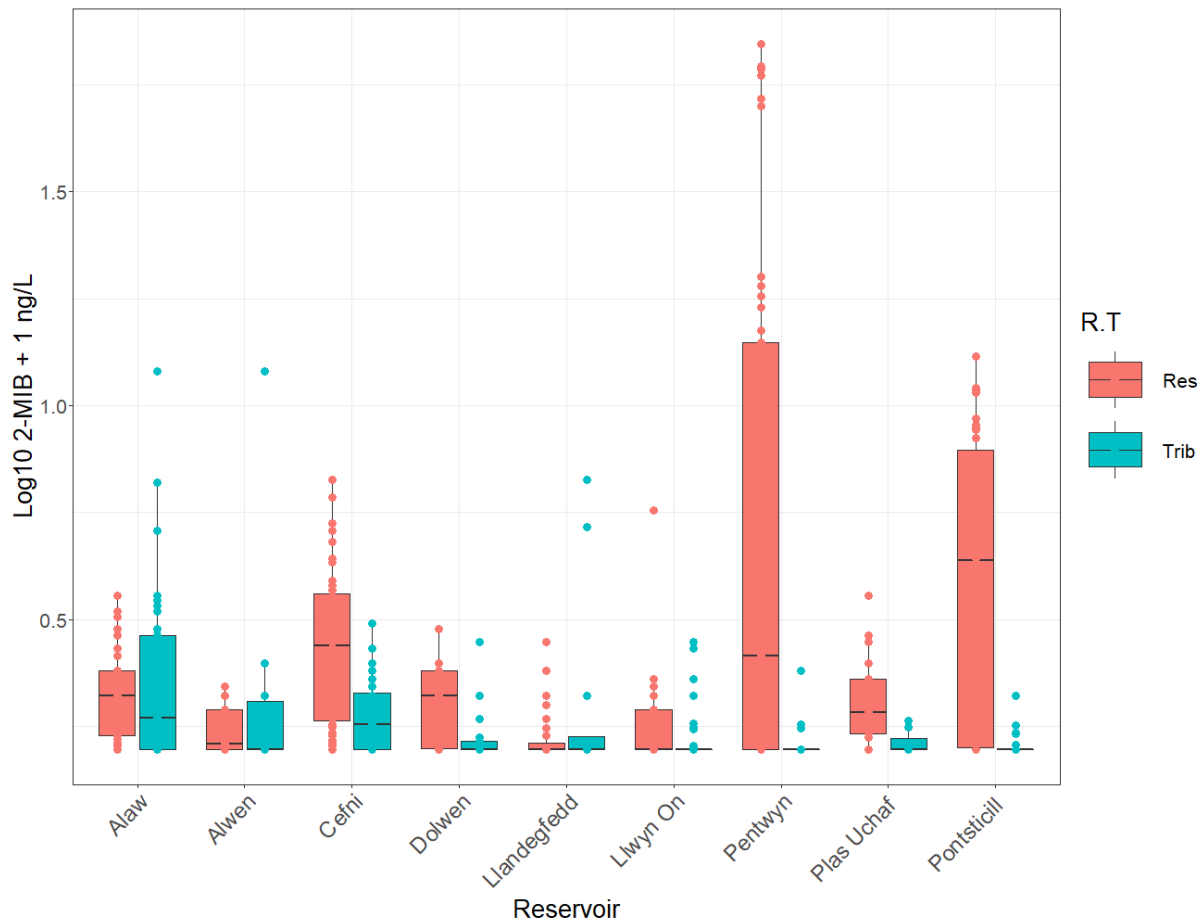


Figure 3.2: Box and whisker plot showing  $\log_{10}$  2-MIB concentrations + 1 ( $\text{ng L}^{-1}$ ) from the reservoir (Res in red) and tributary (Trib in blue) locations for all nine reservoirs over total sampling period (16/7/2019 – 30/11/2020). Raw data is indicated by the coloured dots grouped by reservoir and tributary locations. The length of the box indicates the interquartile range, extending from the 25<sup>th</sup> to the 75<sup>th</sup> percentile and the horizontal bar within the box denotes the median value. The whiskers display the range.

### 3.3.2 – Seasonality of geosmin and 2-MIB concentrations in reservoir locations for each studied reservoir

In a reservoir experiencing severe geosmin “events”, as seen in Alaw, elevated geosmin concentrations were predominantly confined to the summer and autumn months (Figure 3.3). Geosmin concentrations in the Alaw reservoir decreased from summer to autumn in both studied years. The lowest geosmin

concentrations were witnessed in winter, although there was evidence of one geosmin event ( $400 \text{ ng L}^{-1}$ ) during winter 2019, which exceeded geosmin concentrations recorded in summer 2019. All other reservoirs apart from Alwen and Llandegfedd showed a similar trend to Alaw, with heightened geosmin concentrations confined to the summer and autumn months. There was no evidence of a winter geosmin event in other reservoirs; the second-highest recorded winter geosmin concentration was  $7.90 \text{ ng L}^{-1}$  in the Plas Uchaf reservoir in 2019. Trends in geosmin concentrations were not as evident for both Alwen and Llandegfedd reservoirs, which was likely owed to such low concentrations present in these reservoirs throughout the study period, with maximum concentrations of  $6 \text{ ng L}^{-1}$  and  $1 \text{ ng L}^{-1}$ , respectively.

The seasonality of 2-MIB concentrations followed a similar pattern to geosmin concentrations (Figure 3.4), with heightened 2-MIB concentrations confined to summer and autumn. Only Pentwyn and Pontsticill reservoirs experienced 2-MIB concentrations over the event threshold ( $58$  and  $7.90 \text{ ng L}^{-1}$ , respectively).



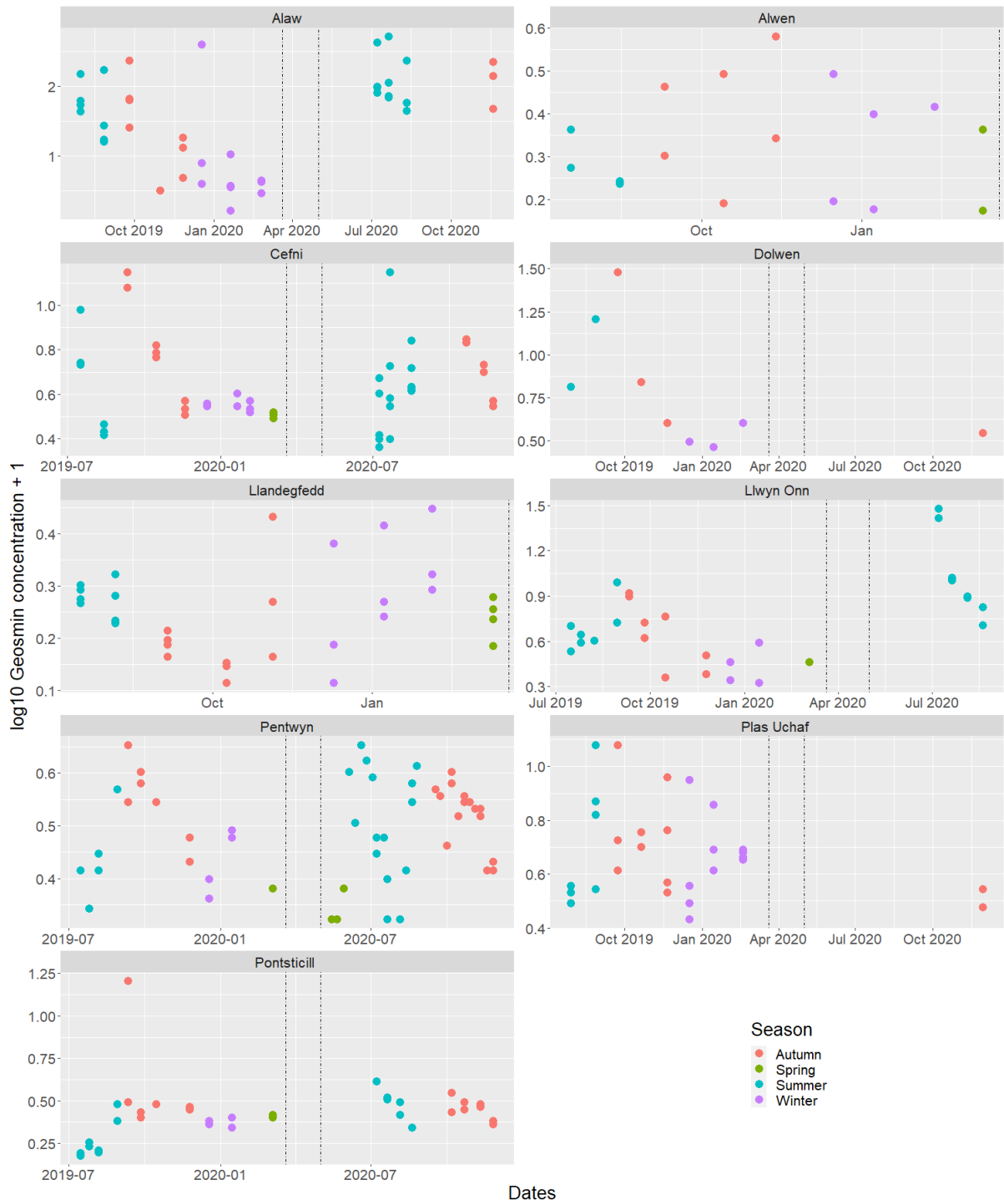


Figure 3.3: Log<sub>10</sub> + 1 geosmin concentrations (ng L<sup>-1</sup>) measured in water samples from the nine Welsh Water reservoirs between 16/7/2019 – 30/11/2020, coloured by season and faceted by individual reservoirs. The dotted lines indicate COVID-19 sample disruption.

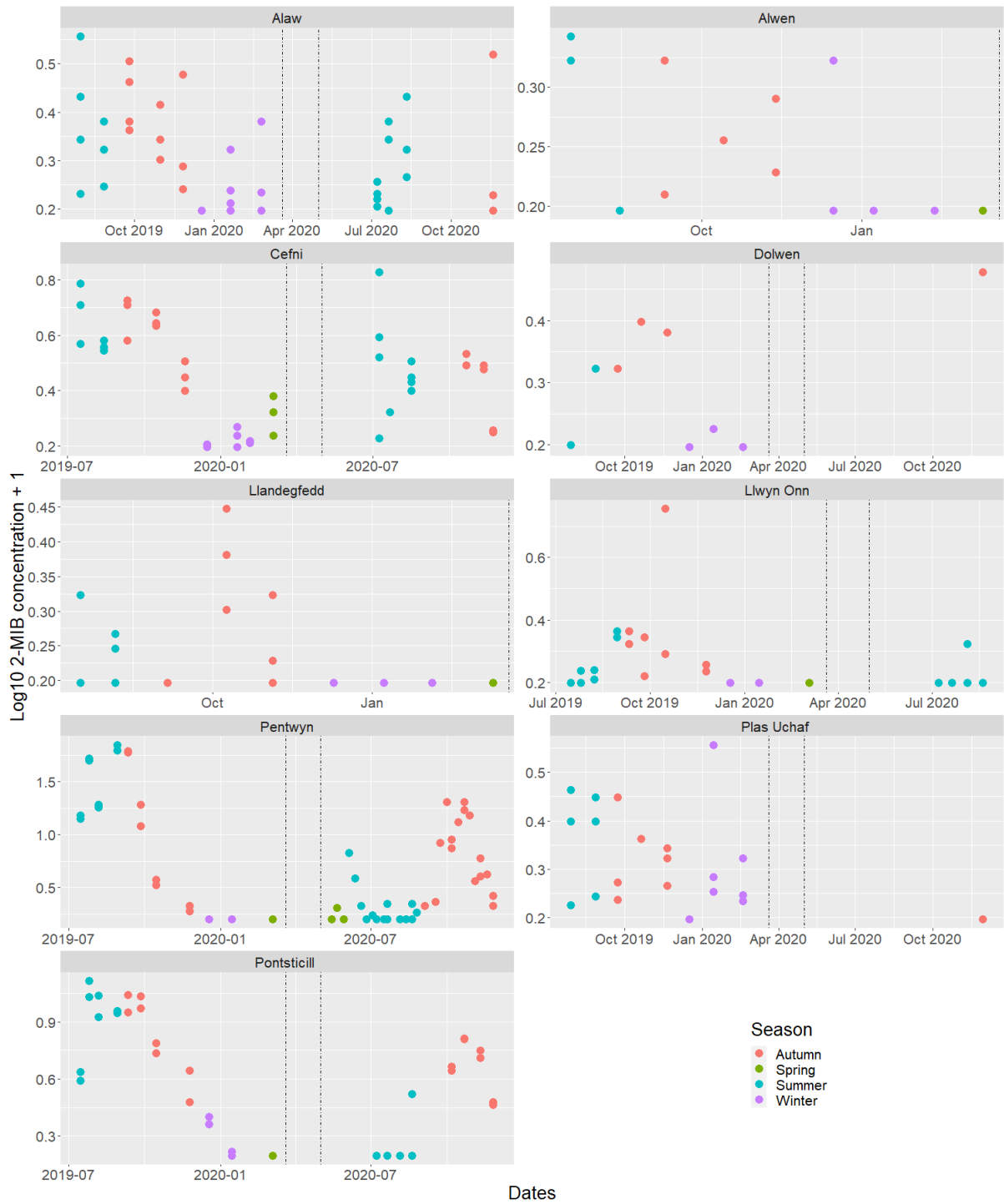


Figure 3.4: Log<sub>10</sub> + 1 2-MIB concentrations (ng L<sup>-1</sup>) measured in water samples from nine Welsh reservoirs between 16/7/2019 – 30/11/2020, coloured by season and faceted by individual reservoirs. The dotted lines indicate COVID-19 sample disruption.

### 3.3.3 – CTSI and TSI indicators per season and year for each reservoir

#### 3.3.3.1 – Alaw

CTSI for all seasons and years during the study period for the Alaw reservoir were categorised as eutrophic (53.07 – 56.93) with minor variations in these calculations ( $\pm 2.44$  –  $\pm 3.07$ ) (Figure 3.5D). TSI indicator SD also reflected a eutrophic status of the reservoir (1.79 – 2.55 m) with slight variations around the mean ( $\pm 0.23$  – 0.35 m) (Figure 3.5A). TSI indicator TP, had the most considerable variation in reflecting the trophic status of the reservoir (46.25 – 96.30  $\mu\text{g L}^{-1}$ ) with the highest standard deviations ( $\pm 7.70$  –  $\pm 82.96$   $\mu\text{g L}^{-1}$ ) (Figure 3.5B); shifting from a eutrophic TP status to a hypereutrophic status during Autumn 2019. TP was at its highest during Autumn 2019 ( $96.30 \pm 82.96$   $\mu\text{g L}^{-1}$ ) and winter 2019 ( $93.33 \pm 66.4$   $\mu\text{g L}^{-1}$ ) and 2020 ( $87.13 \pm 59.54$   $\mu\text{g L}^{-1}$ ). During winter 2019, TSI indicator chl a ( $6.17 \pm 4.24$   $\mu\text{g L}^{-1}$ ) had a mesotrophic status, with all other seasons and years experiencing eutrophic status (8.28 – 19.13  $\mu\text{g L}^{-1}$ ) with some variation ( $\pm 3.49$  –  $\pm 6.27$   $\mu\text{g L}^{-1}$ ) (Figure 3.5C).

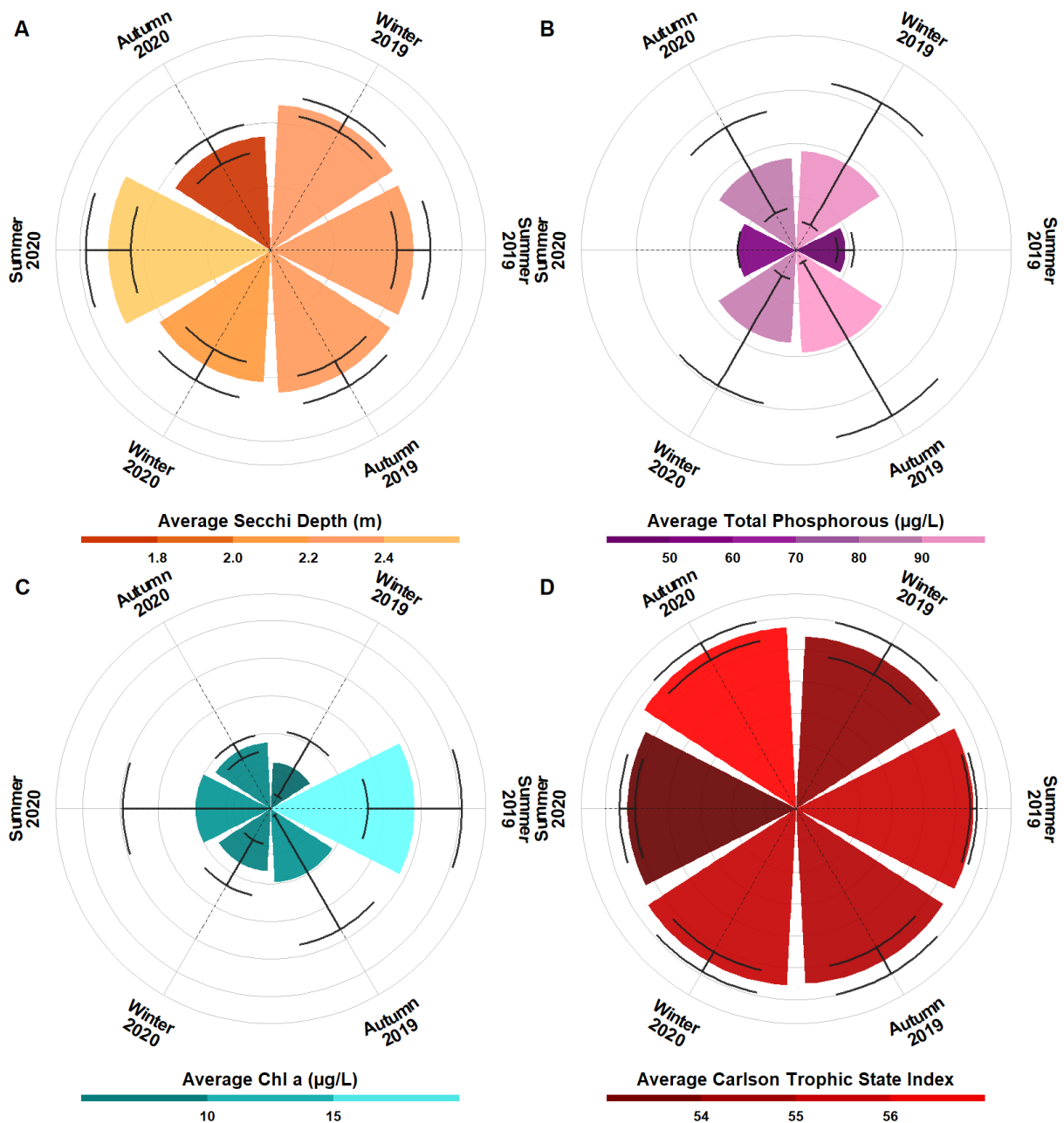


Figure 3.5: Four circular bar plots representing means and standard deviation error bars for the Alaw reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.

### 3.3.3.2 – Alwen

The CTSI for the Alwen reservoir was generally mesotrophic throughout the study period ( $47.69 \pm 0.00$  –  $49.15 \pm 1.91$ ) but transitioned to a eutrophic status during summer 2019 ( $51.04 \pm 0.00$ ) (Figure 3.6D). TSI indicator SD reflected a mesotrophic status throughout all seasons and years from the study period (3.04 – 3.58 m) with little to no variation (autumn 2019 –  $\pm 0.28$  m) (Figure 3.6A). Similarly, TP also reflected a mesotrophic classification of the reservoir throughout all seasons and years (43 – 55  $\mu\text{g L}^{-1}$ ) with no variation around the mean (Figure 3.6B). The average TP was at its lowest during summer 2019 (43  $\mu\text{g L}^{-1}$ ). From all TSI indicators, Chl *a* gave the most substantial fluctuations influencing the trophic status to transition from mesotrophic to eutrophic as witnessed during summer 2019 ( $7.8 \pm 0.00$   $\mu\text{g L}^{-1}$ ) (Figure 3.6C). All other average TP concentrations from seasons and years except summer 2019 were deemed to be either oligotrophic (2.0  $\mu\text{g L}^{-1}$ ) for winter 2019, winter 2020 and spring 2020 or mesotrophic (3.8  $\mu\text{g L}^{-1}$ ) in autumn 2019 with no standard deviation.

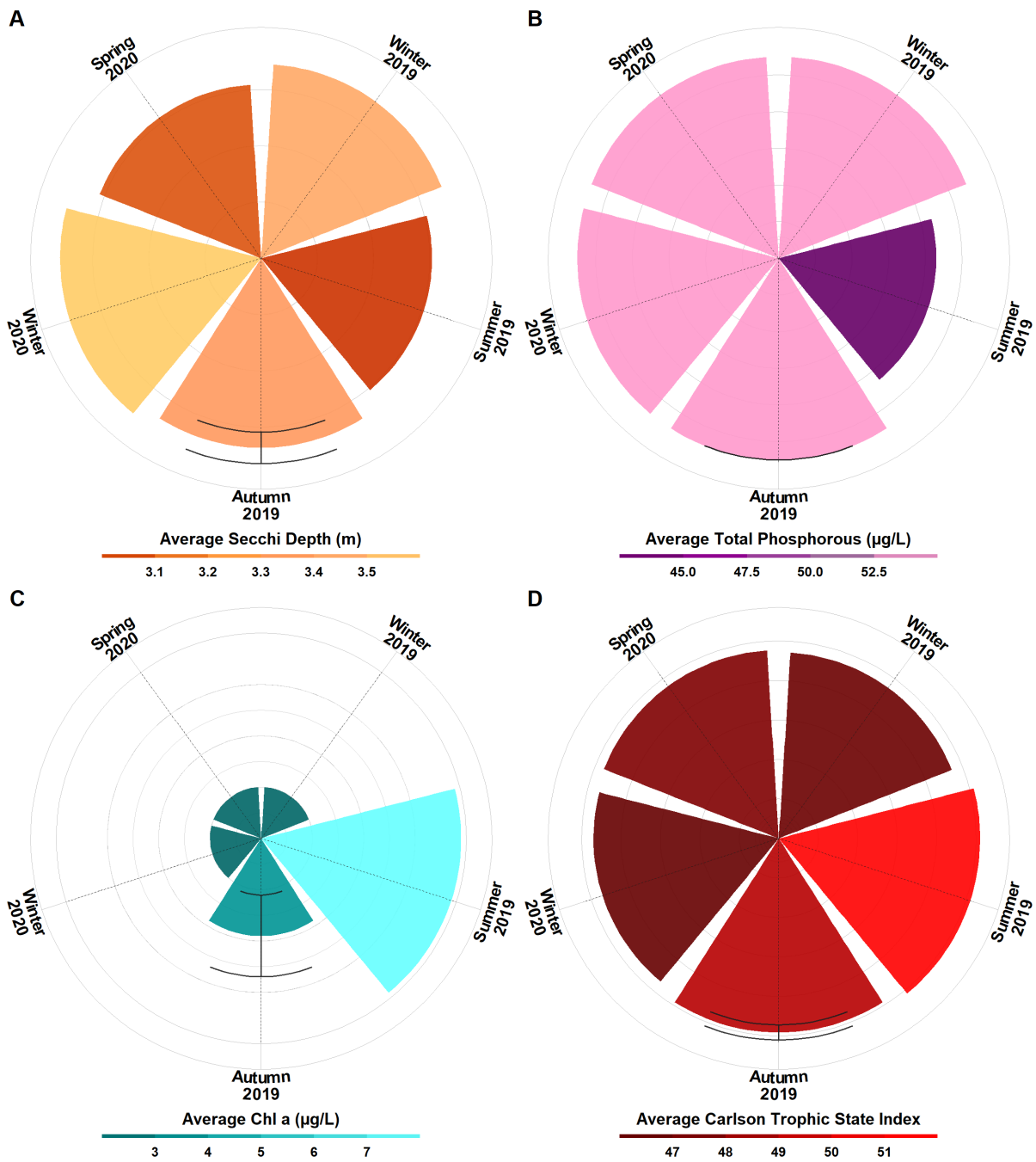


Figure 3.6: Four circular bar plots representing means and standard deviation error bars for the Alwen reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.

### 3.3.3.3 – Cefni

CTSI values for the Cefni reservoir reflected a eutrophic status for each season and year recorded throughout the study period ( $50.67 \pm 1.07 - 64.36 \pm 7.62$ ) (Figure 3.7D). The highest CTSI values were confined to summer from both studied years (summer 2019 –  $57.88 \pm 2.17$  and summer 2020 –  $57.88 \pm 2.17$ ). Although CTSI indicated an overall eutrophic status, the TSI values for the SD indicator revealed a general mesotrophic classification ( $2.05 \pm 0.43 - 2.79 \pm 0.29$  m) (Figure 3.7A). For the TP indicator, average concentrations were typically classified as eutrophic for all seasons and years ( $43.33 \pm 0.58 - 73.83 \pm 3.37 \mu\text{g L}^{-1}$ ) apart from summer 2020 that had a hypereutrophic status during summer 2020 ( $147.67 \mu\text{g L}^{-1}$ ) with a sizeable standard deviation ( $\pm 104.4 \mu\text{g L}^{-1}$ ) revealing possible spikes in TP during summer 2020 (Figure 3.7B). Chl *a* concentrations showed the greatest variations in revealing trophic status (Figure 3.7C). During winter 2019, autumn 2019, winter 2020, spring 2020 and autumn 2020 Chl *a* concentrations revealed a mesotrophic classification ( $2.85 \pm 1.80 - 5.00 \pm 3.04 \mu\text{g L}^{-1}$ ). The highest Chl *a* concentrations were confined to the summer seasons, as shown in summer 2019 displaying a eutrophic classification ( $46.33 \pm 21.39 \mu\text{g L}^{-1}$ ) and summer 2020 ( $116.34 \pm 147.77 \mu\text{g L}^{-1}$ ) revealing a hypereutrophic classification with substantial variations in phytoplankton abundance.

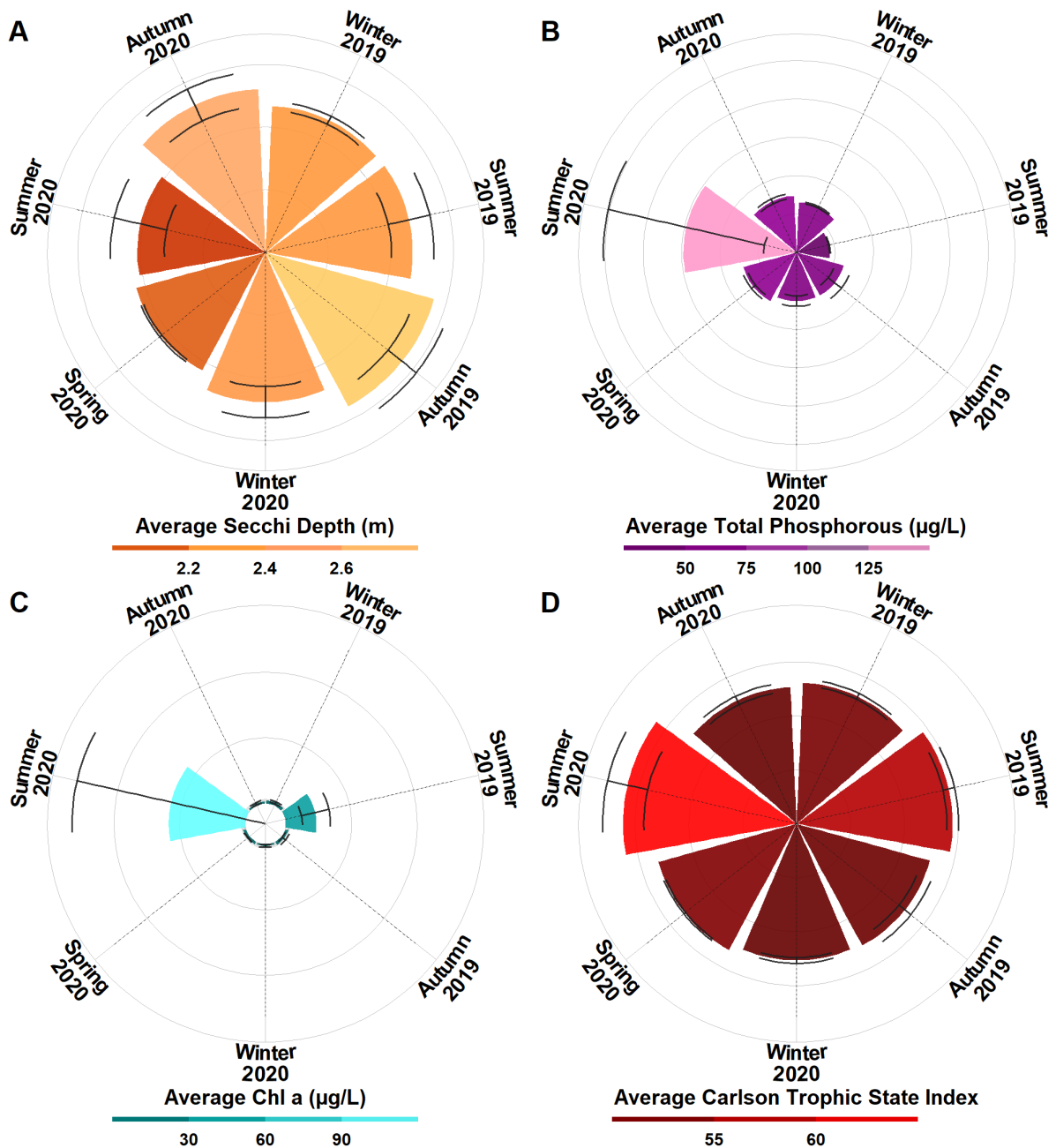


Figure 3.7: Four circular bar plots representing means and standard deviation error bars for the Cefni reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.



### 3.3.3.4 – Dolwen

Dolwen reservoir was mainly eutrophic according to the CTSI values ( $54.49 \pm 6.89 - 58.39 \pm 2.13$ ) except for winter 2019 ( $49.95 \pm 0.00$ ), experienced mesotrophic classification (Figure 3.8D). The TSI indicator SD exhibited mesotrophic classification throughout all studied seasons and years ( $2.19 \pm 0.07 - 2.56 \pm 0.00$  m) (Figure 3.8A). Average TP concentrations indicated a eutrophic classification for all studied seasons and years ( $43 \pm 0.00 - 86 \pm 8.49 \mu\text{g L}^{-1}$ ) (Figure 3.8B), with the highest average TP recorded in winter 2020. The Chl a TSI indicator noted fluctuations in phytoplankton biomass that revealed a mesotrophic status during both years for winter (2019 –  $2.9 \pm 0.00 \mu\text{g L}^{-1}$  and 2020 –  $7.25 \pm 2.19 \mu\text{g L}^{-1}$ ) and a eutrophic status for summer 2019 ( $24.00 \pm 11.31 \mu\text{g L}^{-1}$ ), autumn 2019 ( $24.53 \pm 35.93 \mu\text{g L}^{-1}$ ) and autumn 2020 ( $40.00 \pm 0.00 \mu\text{g L}^{-1}$ ) (Figure 3.8C).

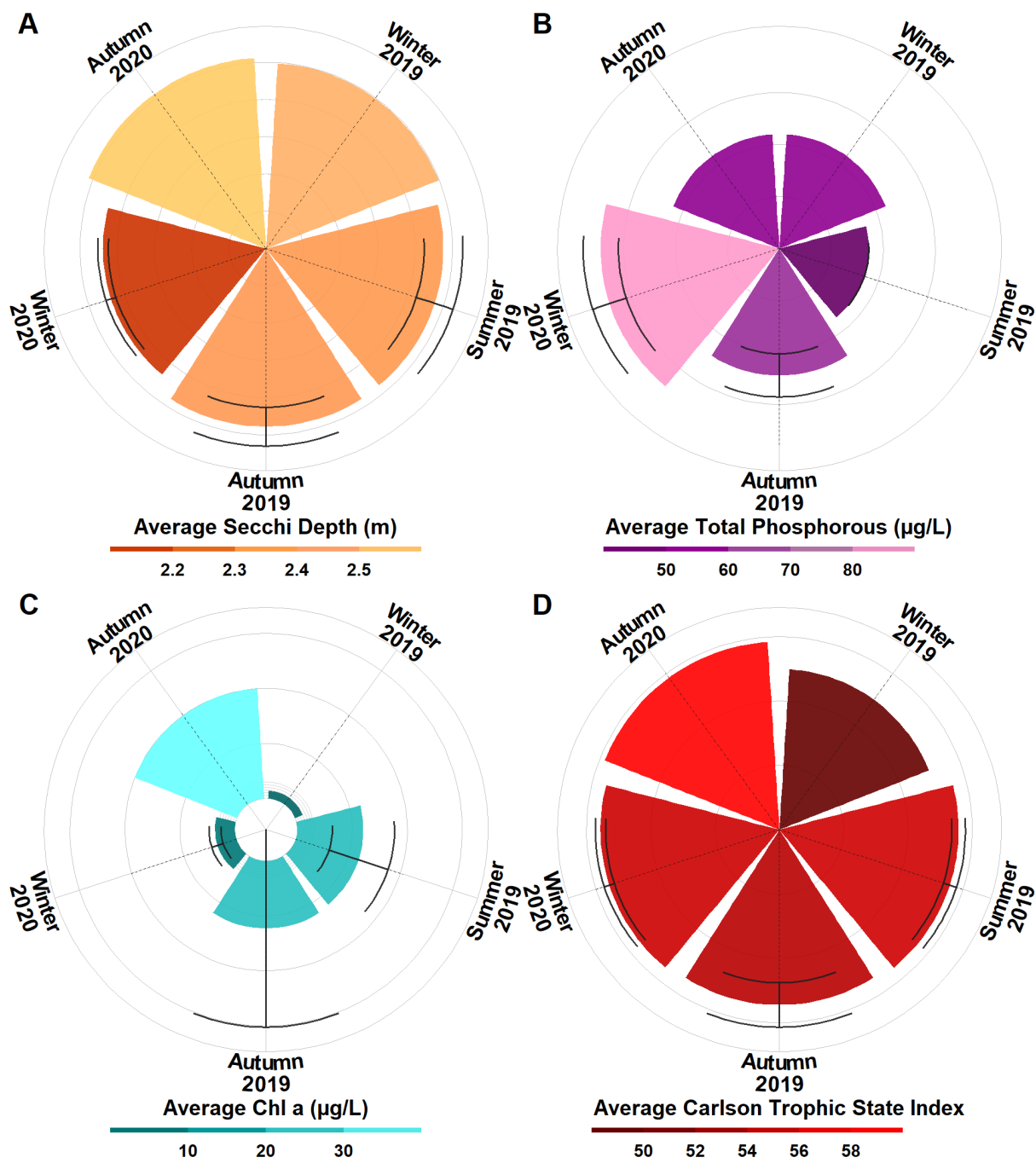


Figure 3.8: Four circular bar plots representing means and standard deviation error bars for the Dolwen reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.

### 3.3.3.5 – Llandegfedd

The CTSI values for the Llandegfedd reservoir demonstrated a transition from mesotrophic status during winter ( $49.49 \pm 1.91$ ) and autumn 2019 ( $49.86 \pm 2.22$ ) to eutrophic during all other seasons and years ( $51.67 \pm 2.19 - 54.18 \pm 7.24$ ) (Figure 3.9D). Recorded average SD measurements for all seasons and years showed mesotrophic classification ( $2.34 \pm 0.78 - 3.00 \pm 0.24$  m) (Figure 3.9A). The TSI indicator TP showed the highest average concentration during winter 2020 ( $412.5 \pm 397.9 \mu\text{g L}^{-1}$ ), revealing a hypereutrophic status during this period but with a substantial standard deviation (Figure 3.9B). All other TP averages for the season and year of the study revealed eutrophic status ( $45.38 \pm 6.72 - 59.00 \pm 6.93 \mu\text{g L}^{-1}$ ). Chl *a* measurements were mesotrophic during winter 2019 ( $2.37 \pm 0.64 \mu\text{g L}^{-1}$ ), autumn 2019 ( $3.60 \pm 1.97 \mu\text{g L}^{-1}$ ) and spring 2020 ( $5.67 \pm 2.71 \mu\text{g L}^{-1}$ ) and gained a eutrophic status during summer 2019 ( $13.23 \pm 8.88 \mu\text{g L}^{-1}$ ). During winter 2020, phytoplankton biomass was at its lowest, as indicated by the average Chl *a* measurement ( $2.00 \pm 0.00 \mu\text{g L}^{-1}$ ) gaining oligotrophic status (Figure 3.9C).

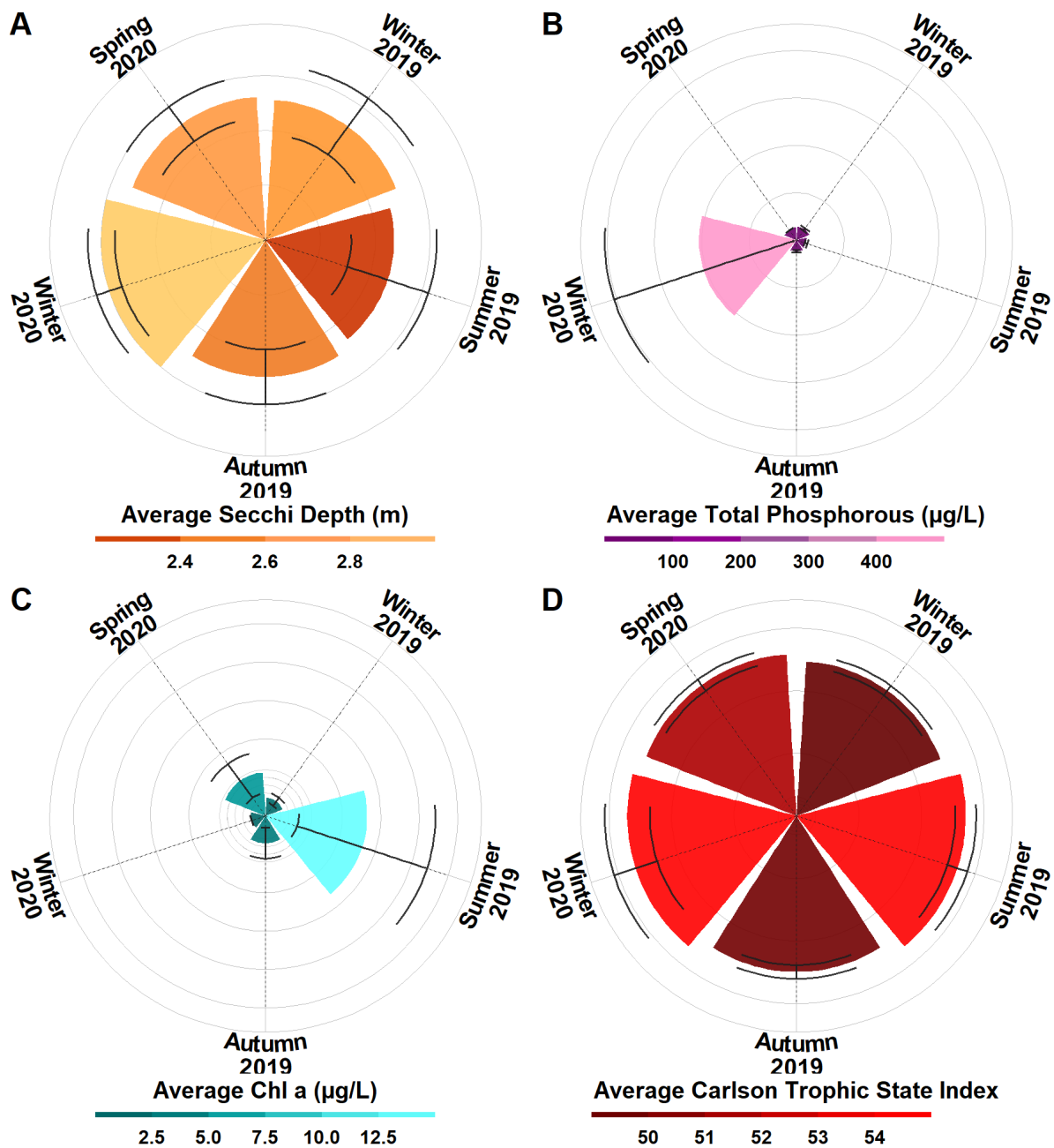


Figure 3.9: Four circular bar plots representing means and standard deviation error bars for the Llandegfedd reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.

### 3.3.3.6 – Llwyn On

The trophic status of Llwyn On reservoir was generally mesotrophic ( $47.93 \pm 0.11 - 49.20 \pm 1.62$ ), apart from the summer months from both studied years that were eutrophic ( $50.30 \pm 1.12 - 50.82 \pm 1.37$ ) (Figure 3.10D). Average SD measurements from all seasons and years revealed mesotrophic status classification ( $2.56 \pm 0.00 - 3.31 \pm 0.46$  m) (Figure 3.10A). All average TP measurements were classified as eutrophic ( $43.00 \pm 0.00 - 62.00 \pm 9.9 \mu\text{g L}^{-1}$ ) (Figure 3.10B). The greatest variation in TSI indicators came from the averaged Chl *a* concentrations (Figure 3.10C). During spring 2020 ( $2.00 \pm 0.00 \mu\text{g L}^{-1}$ ), winter 2020 ( $2.00 \pm 0.00 \mu\text{g L}^{-1}$ ) and winter 2019 ( $2.10 \pm 0.14 \mu\text{g L}^{-1}$ ), Chl *a* concentrations were oligotrophic. During autumn 2019 ( $3.31 \pm \mu\text{g L}^{-1}$ ), summer 2020 ( $5.16 \pm 1.97 \mu\text{g L}^{-1}$ ) and summer 2019 ( $6.37 \pm 1.55 \mu\text{g L}^{-1}$ ), average Chl *a* concentrations were considered to be mesotrophic.

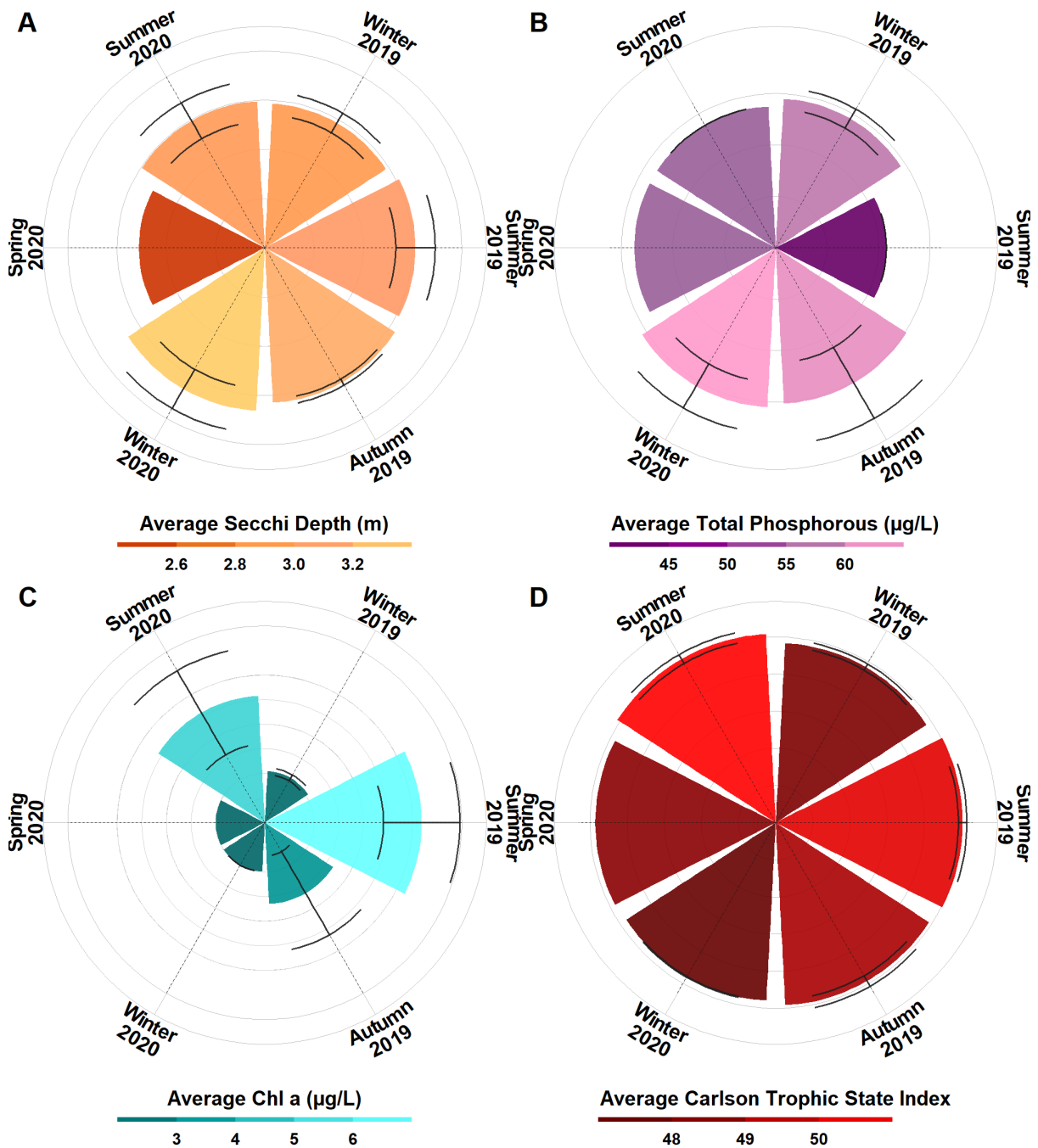


Figure 3.10: Four circular bar plots representing means and standard deviation error bars for the Llwyn On reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.

### 3.3.3.7 – Pentwyn

Pentwyn reservoir transitioned from a mesotrophic status governed by CTSI values during both winter years ( $49.38 \pm 0.01 - 49.41 \pm 0.04$ ) and spring 2020 ( $49.67 \pm 0.02$ ) to eutrophic status during both autumn years ( $50.50 \pm 2.25 - 50.69 \pm 1.62$ ) and both summer years ( $52.30 \pm 1.59 - 52.86 \pm 0.85$ ) (Figure 3.11D). For the TSI indicator TP, all concentrations stayed at a constant value of  $55.00 \pm 0.00 \mu\text{g L}^{-1}$  except for the average TP concentration taken for summer 2019 ( $43 \pm 0.00 \mu\text{g L}^{-1}$ ) which was lower (Figure 3.11B); all values revealed a eutrophic status. The overall CTSI for seasons and years was influenced by the average Chl *a* concentrations; both winter years ( $2.00 \pm 0.00 - 2.00 \pm 0.00 \mu\text{g L}^{-1}$ ) and spring 2020 ( $2.00 \pm 0.00 \mu\text{g L}^{-1}$ ) were classified as mesotrophic and both autumn years ( $4.35 \pm 2.78 - 4.71 \pm 3.71 \mu\text{g L}^{-1}$ ) and both summer years ( $6.38 \pm 2.50 - 10.33 \pm 4.38 \mu\text{g L}^{-1}$ ) were classified as eutrophic (Figure 3.11C).

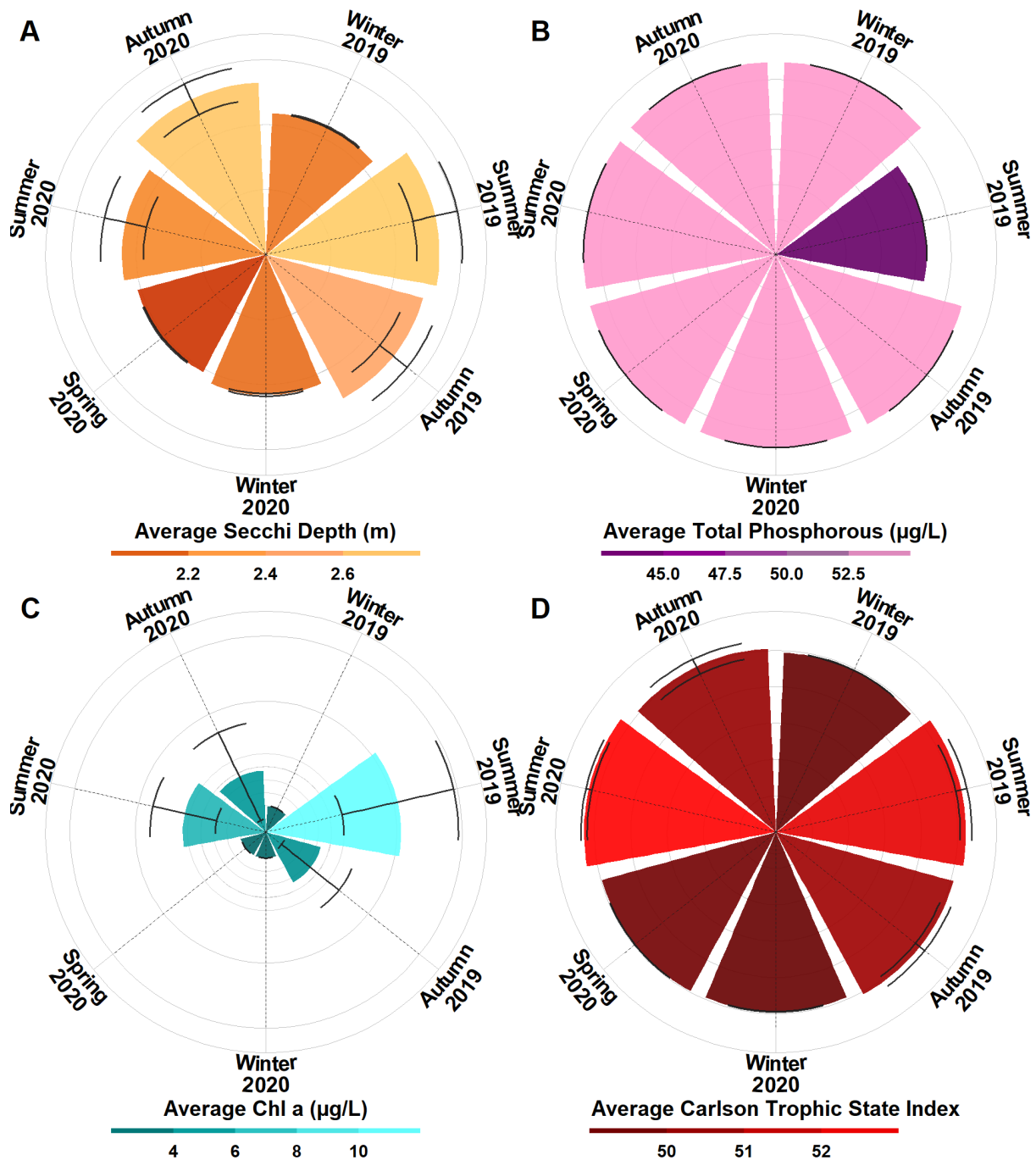


Figure 3.11: Four circular bar plots representing means and standard deviation error bars for the Pentwyn reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.



### 3.3.3.8 – Plas Uchaf

In the Plas Uchaf reservoir, the trophic status given by the CTSI values for each season and year of this study was deemed eutrophic ( $50.37 \pm 0.48 - 56.95 \pm 14.06$ ) (Figure 3.12D). SD measurements were all classified as mesotrophic ( $2.22 \pm 0.16 - 2.81 \pm 0.07$  m) (Figure 3.12A). Measurements for the TP determinant of the overall CTSI value were mainly classified as eutrophic ( $46.63 \pm 6.61 - 81.57 \pm 28.59 \mu\text{g L}^{-1}$ ), except for autumn 2019 that had a hypereutrophic average concentration ( $239.50 \mu\text{g L}^{-1}$ ) with a large standard deviation ( $583.44 \mu\text{g L}^{-1}$ ) (Figure 3.12B). Average Chl *a* concentrations for winter 2019 were classified as mesotrophic ( $3.13 \pm 0.56 \mu\text{g L}^{-1}$ ) (Figure 3.12C). Chl *a* concentrations during winter 2020 ( $8.70 \pm 2.03 \mu\text{g L}^{-1}$ ), autumn 2020 ( $10.50 \pm 2.12 \mu\text{g L}^{-1}$ ) and summer 2019 ( $13.99 \pm 10.04 \mu\text{g L}^{-1}$ ) were all classified as eutrophic. Phytoplankton biomass indicated by the average Chl *a* concentrations was at its highest during autumn 2019 ( $529.94 \mu\text{g L}^{-1}$ ) and was considered hypereutrophic, although the standard deviation was substantial ( $1640.97 \mu\text{g L}^{-1}$ ).

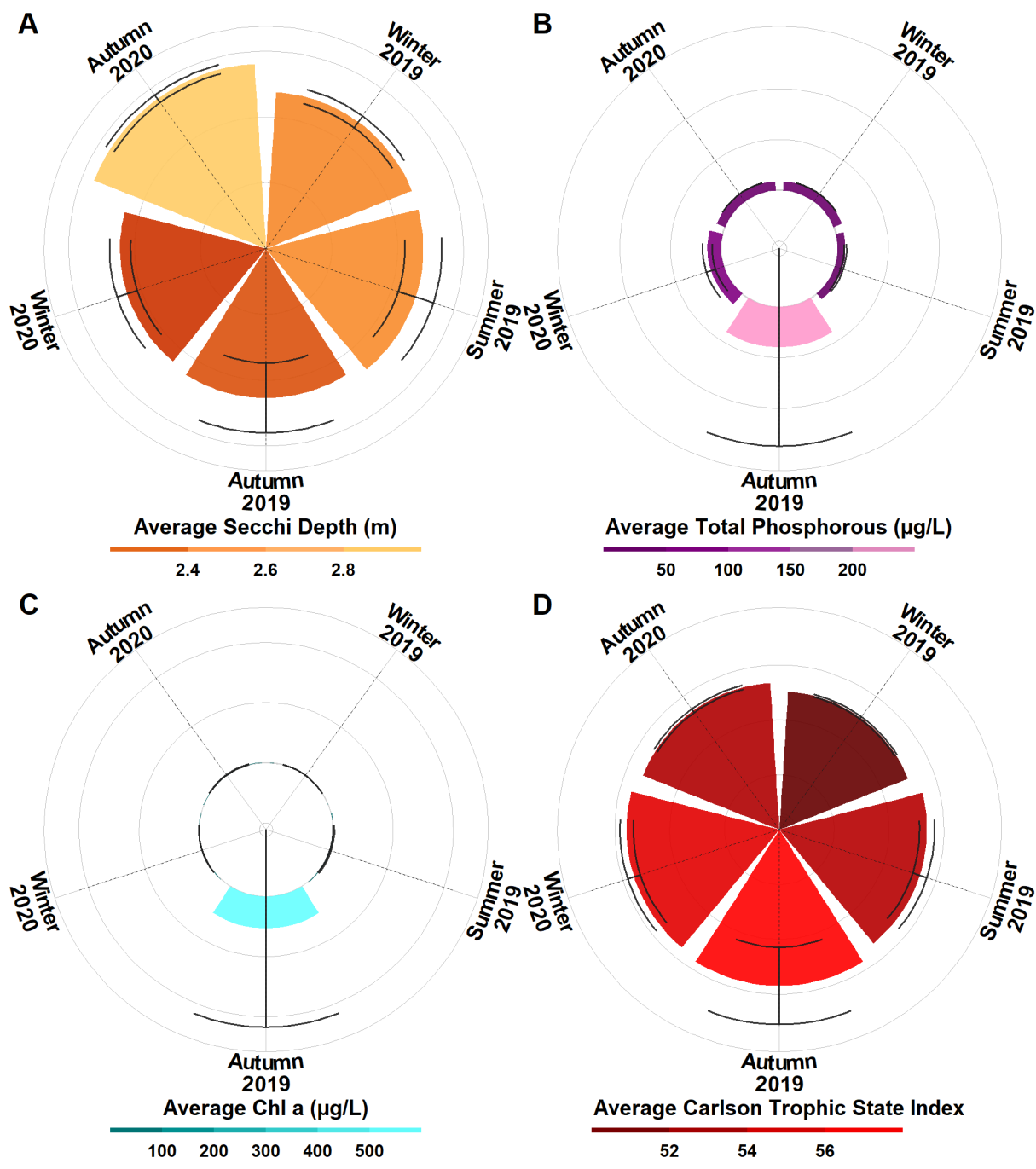


Figure 3.12: Four circular bar plots representing means and standard deviation error bars for the Plas Uchaf reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.

### 3.3.3.9 – Pontsticill

The CTSI for the Pontsticill reservoir was mesotrophic during winter–spring 2020 ( $48.24 \pm 0.55 - 48.73 \pm 0.19$ ) and eutrophic for the remaining seasons and years of this study ( $50.05 \pm 2.64 - 52.12 \pm 2.29$ ) (Figure 3.13D). Average TP concentrations for all seasons and years were eutrophic ( $43.00 \pm 0.00 - 60.38 \pm 15.20 \mu\text{g L}^{-1}$ ) (Figure 3.13B). Average Chl *a* concentrations for winter and spring 2020 were the same and were classified as oligotrophic ( $2.00 \pm 0.00 \mu\text{g L}^{-1}$ ) (Figure 3.13C). For winter 2019 ( $3.30 \pm 1.84 \mu\text{g L}^{-1}$ ), autumn 2020 ( $5.88 \pm 2.55 \mu\text{g L}^{-1}$ ) and summer 2020 ( $6.43 \pm 1.11 \mu\text{g L}^{-1}$ ), average Chl *a* concentrations were considered to be mesotrophic. Summer 2019 ( $7.35 \pm 3.75 \mu\text{g L}^{-1}$ ) and autumn 2019 ( $10.48 \pm 14.82 \mu\text{g L}^{-1}$ ) were considered to be eutrophic in relation to Chl *a* concentrations.

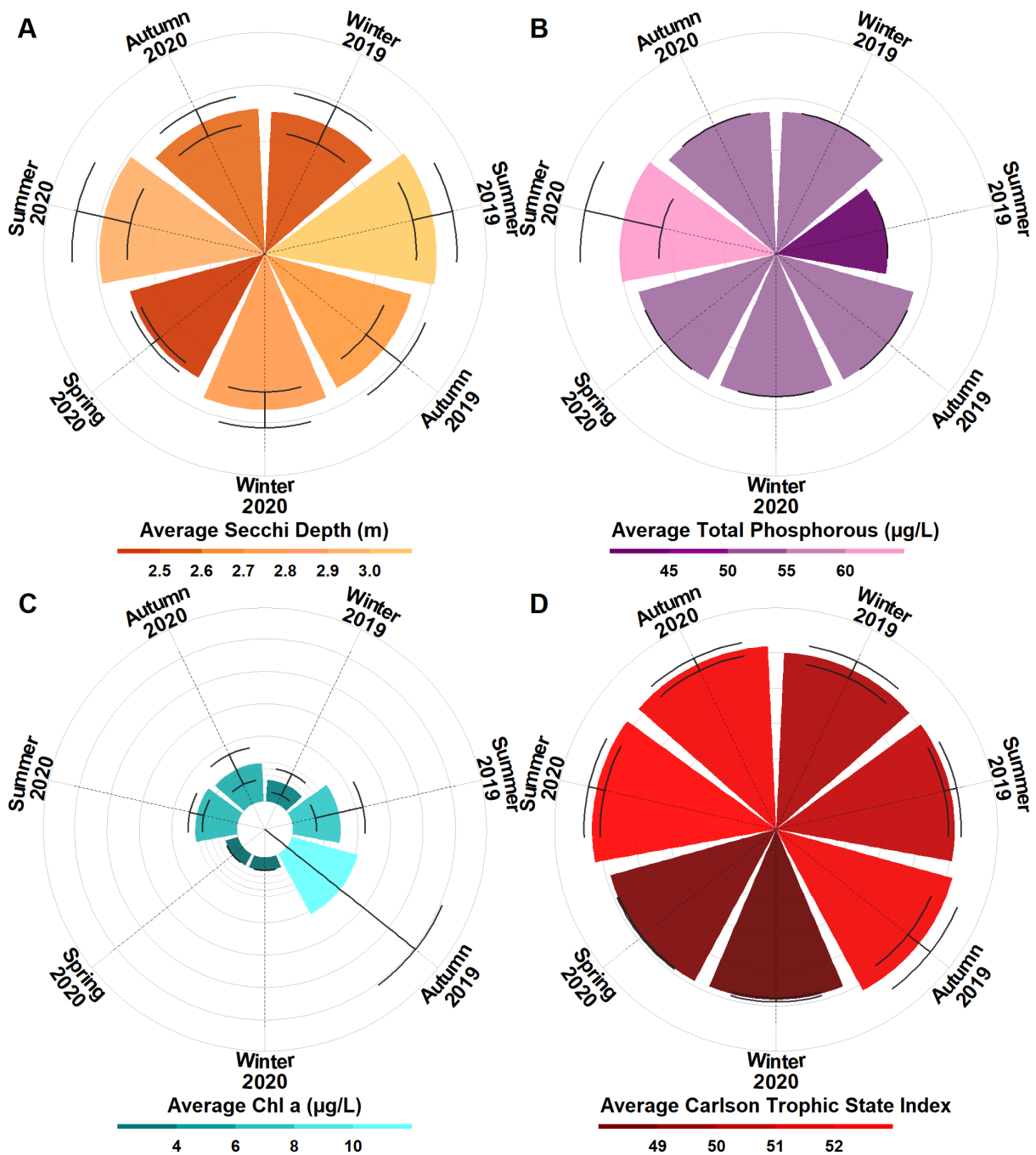


Figure 3.13: Four circular bar plots representing means and standard deviation error bars for the Pontsticill reservoir for A) average Secchi depth measurement (m) per season and year, B) average total phosphorous ( $\mu\text{g L}^{-1}$ ) per season and year, C) average Chl a ( $\mu\text{g L}^{-1}$ ) per season and year and D) average CTSI measurements per season and year.

### 3.3.4 – Significant associations between physical and chemical water parameters and T&O levels using PCA

#### 3.3.4.1 – Alaw – geosmin and 2-MIB concentrations

In Alaw's reservoir locations, samples showed slight clustering according to geosmin levels (Figure 3.14 – top). The slight clustering in geosmin levels followed a gradient from high, medium, and low concentrations. High and medium geosmin concentrations tended to cluster more tightly than low geosmin data points. High geosmin concentrations were influenced by pH and sulphate (SO<sub>4</sub>) loadings. In contrast, nitrate (NO<sub>3</sub>) ammonium (NH<sub>4</sub>) dissolved reactive silicate (DSil) and turbidity loaded in the direction of low geosmin concentrations.

Alaw's tributary locations showed a small cluster of low geosmin concentrations (Figure 3.14 – middle), with medium geosmin concentrations showing a higher dispersion within the PCA with a slight overlap between the two levels. All variables are loaded in the direction of medium geosmin concentrations. For 2-MIB concentrations, low concentrations formed a cluster, although some data points fell outside of the ellipse associated with low concentrations (Figure 3.14 – bottom). Total phosphorous (TP) loading was associated in the direction of one datapoint representing a high 2-MIB concentration.

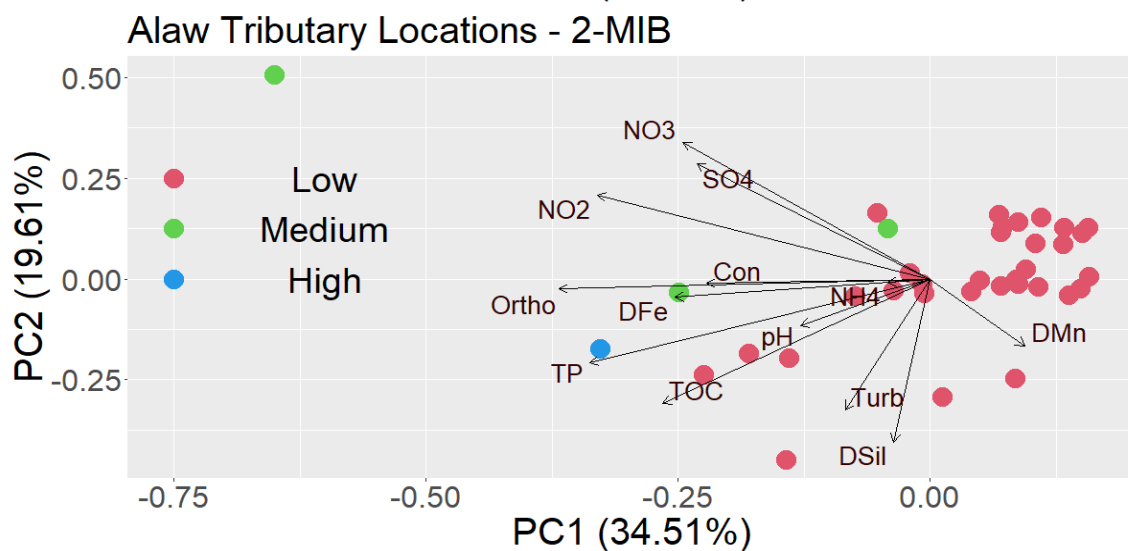
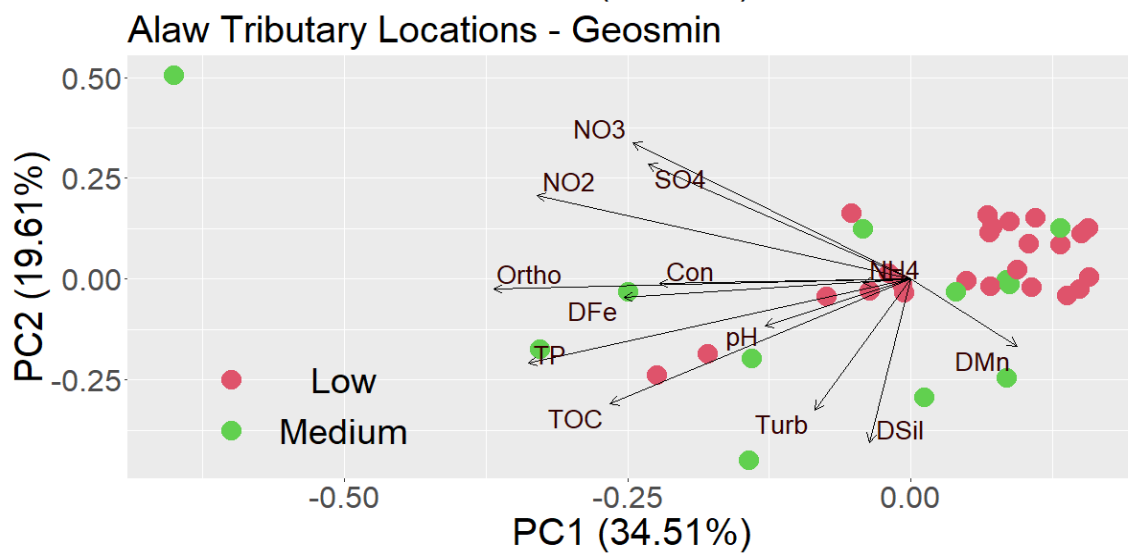
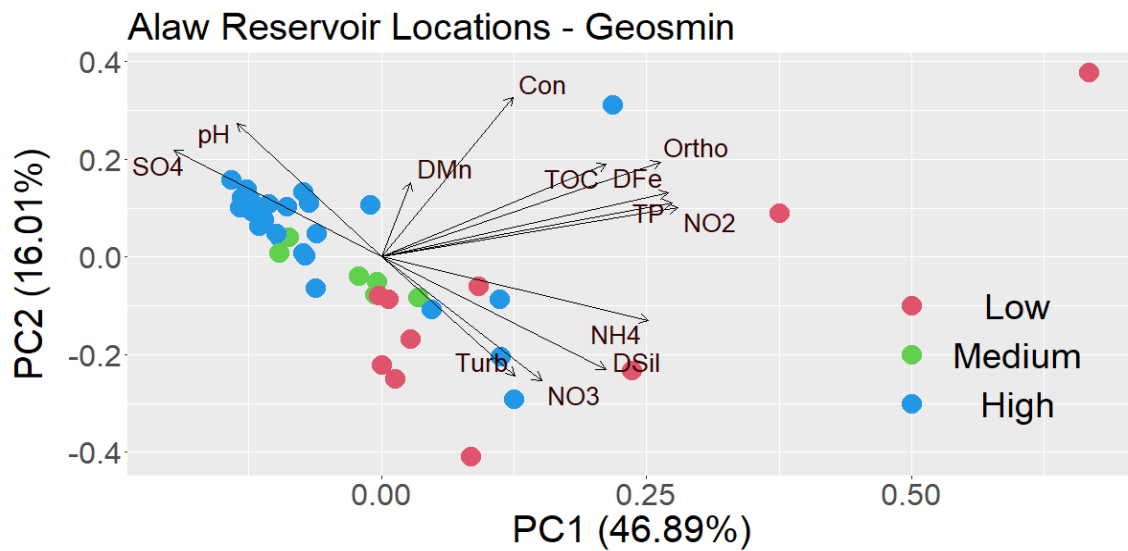


Figure 3.14: PCA biplot of Alaw reservoir and tributary locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned T&O level.

### 3.3.4.2 – Alwen – 2-MIB concentrations

For 2-MIB concentrations in Alwen's tributary locations, low concentrations of 2-MIB formed a cluster (Figure 3.15). Only one data point was at a medium 2-MIB level classification, which was not distinctly separated from low 2-MIB concentrations. However, TP loading is in the direction of this one high 2-MIB concentration datapoint.

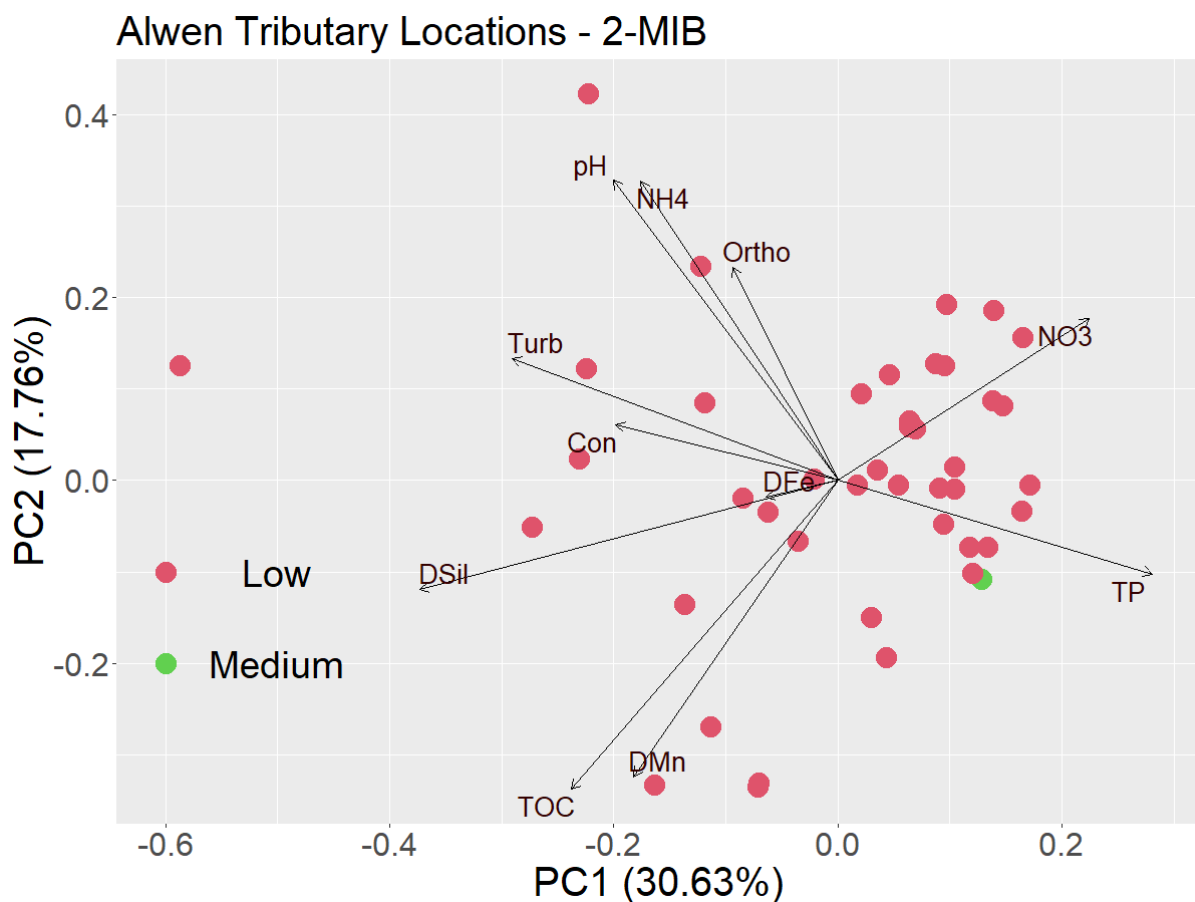


Figure 3.15: PCA biplot of Alwen reservoir's tributary locations using component 1 (PC1) and component 2 (PC2) with individual points coloured according to the assigned 2-MIB level.

### 3.3.4.3 – Cefni – geosmin concentrations

For geosmin concentrations in Cefni's reservoir locations, low and medium concentrations showed no distinct separation in clustering (Figure 3.16).

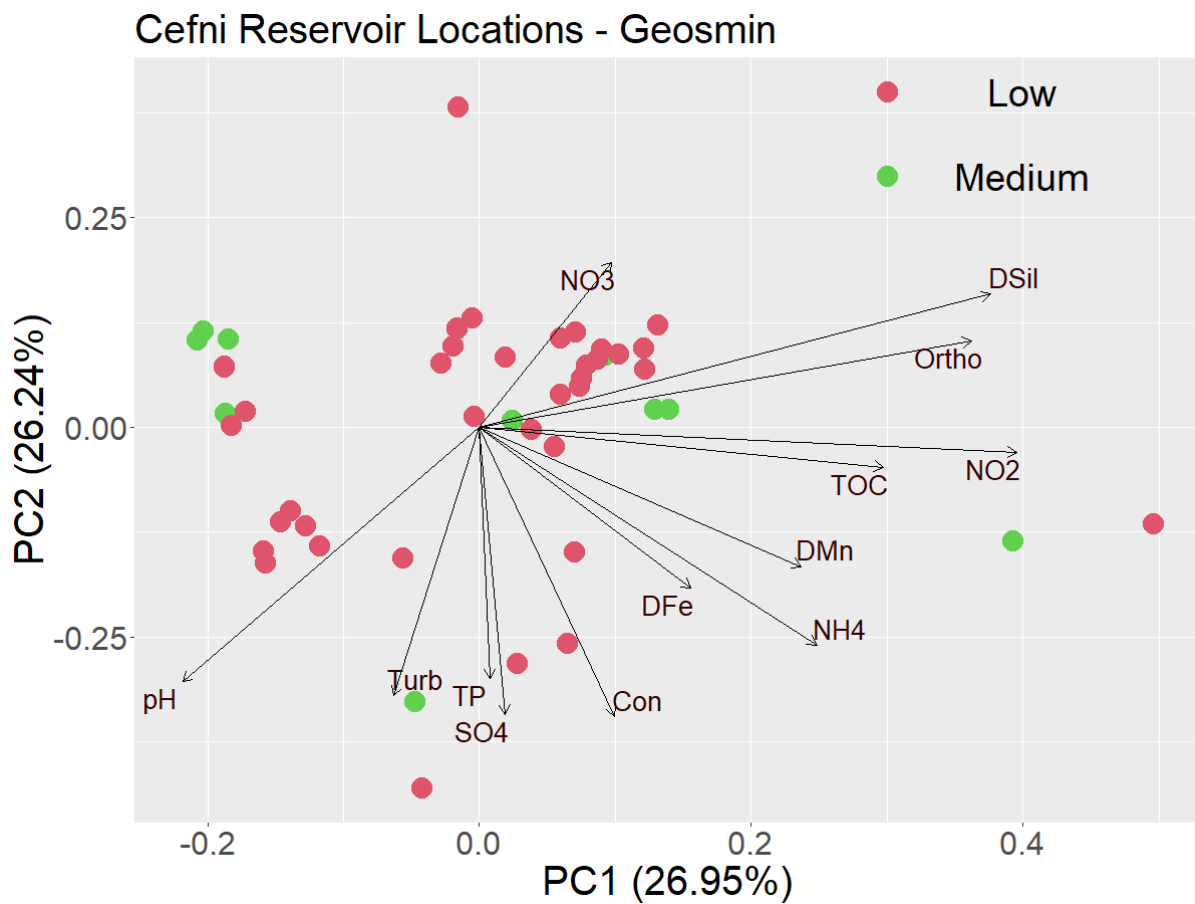


Figure 3.16: PCA biplot of Cefni reservoir's reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level.



### 3.3.4.4 – Dolwen – geosmin concentrations

In Dolwen's reservoir locations, low geosmin concentrations can be seen to cluster (Figure 3.17). Medium geosmin concentrations showed no overlap with low geosmin concentrations, although they were highly dispersed within the PCA.

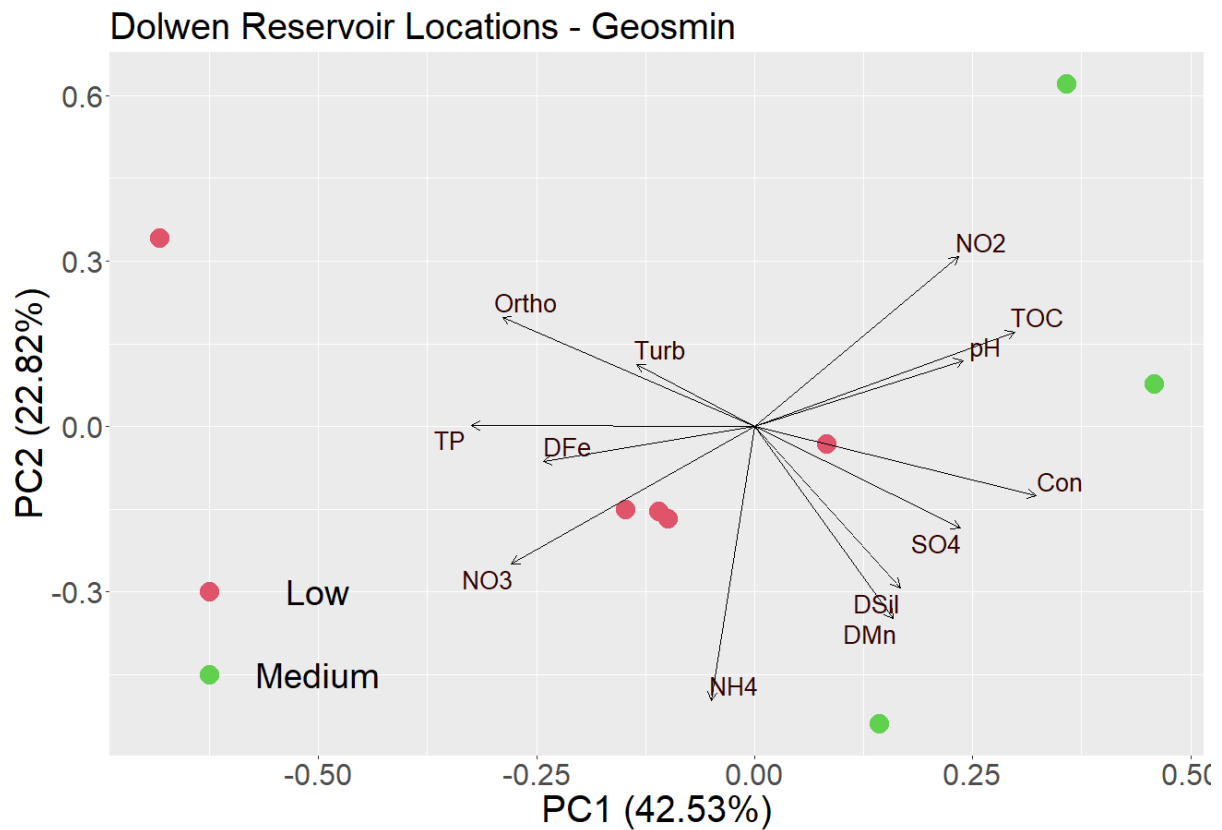


Figure 3.17: PCA biplot of Dolwen reservoir's reservoir locations using component 1 (PC1) and component 2 (PC2) with individual points coloured according to the assigned geosmin level.

### 3.3.4.5 – Llandegfedd – 2-MIB concentrations

2-MIB concentrations in Llandegfedd’s tributary locations showed a distinct clustering of low concentrations apart from one datapoint (Figure 3.18). The only medium 2-MIB concentration datapoint was separated from the low concentrations, with Orthophosphate (Ortho), dissolved iron (DFe) and turbidity (Turb) loading in the same direction.

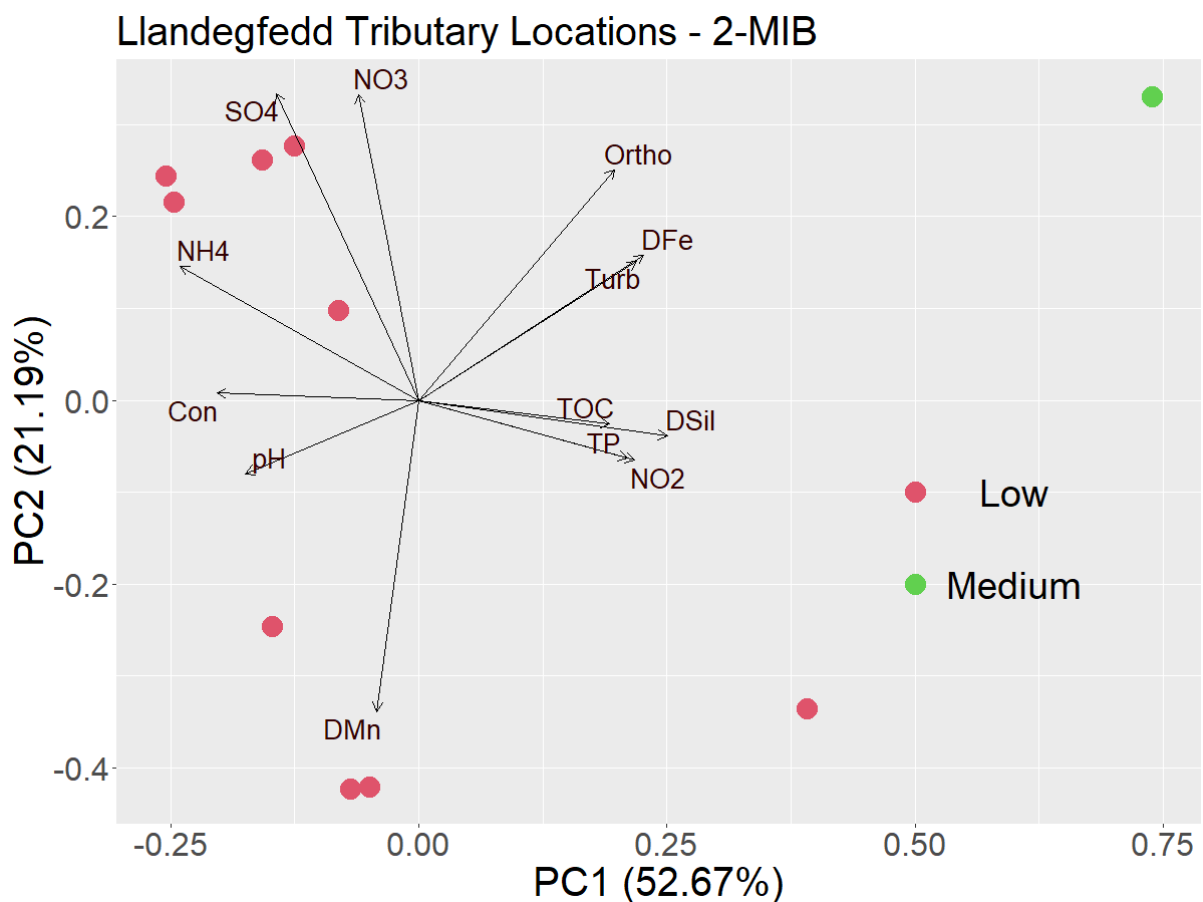


Figure 3.18: PCA biplot of Llandegfedd reservoir’s tributary locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned 2-MIB level.

### 3.3.4.6 – Llwyn On – geosmin concentrations

Low geosmin concentrations in Llwyn On’s reservoir locations showed the highest dispersity within the PCA, overlapping medium and high geosmin concentration data points (Figure 3.19). Both medium and high geosmin datapoints formed tight clusters with dissolved manganese (DMn) influencing high geosmin concentrations.

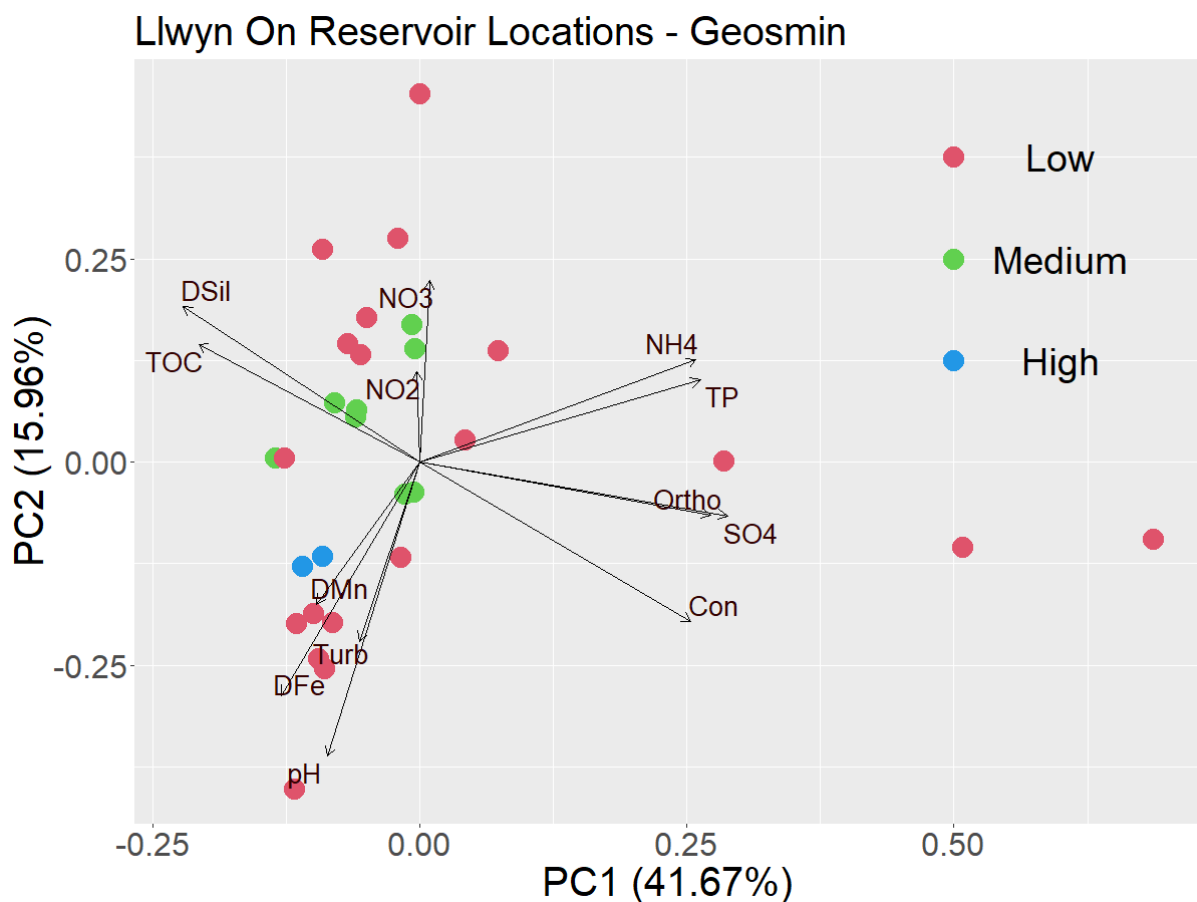


Figure 3.19: PCA biplot of Llwyn On reservoir’s reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level.

### 3.3.4.7 – Pentwyn – 2-MIB concentrations

2-MIB concentrations in Pentwyn’s reservoir locations showed significant overlap in assigned 2-MIB concentration levels (Figure 3.20). Low concentrations should have the largest dispersity within the PCA, with ellipses becoming tighter for medium and high concentrations. Conductivity (Con) and pH are loaded in the direction of a distinct grouping of high 2-MIB concentration data points.

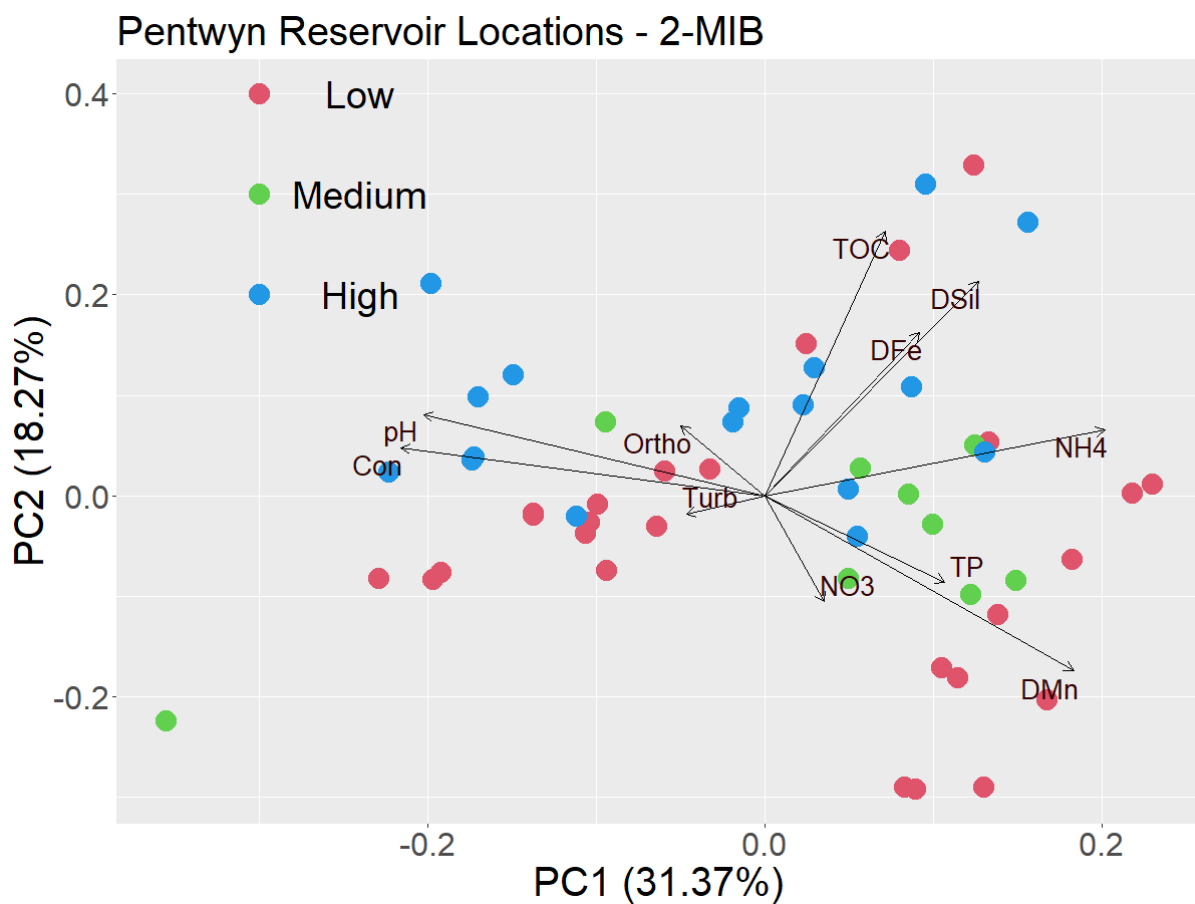


Figure 3.20: PCA biplot of Pentwyn reservoir’s reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned 2-MIB level.

### 3.3.4.8 – Plas Uchaf – geosmin concentrations

Geosmin concentrations in Plas Uchaf's reservoir locations showed no distinct clustering of low and medium concentrations (Figure 3.21).

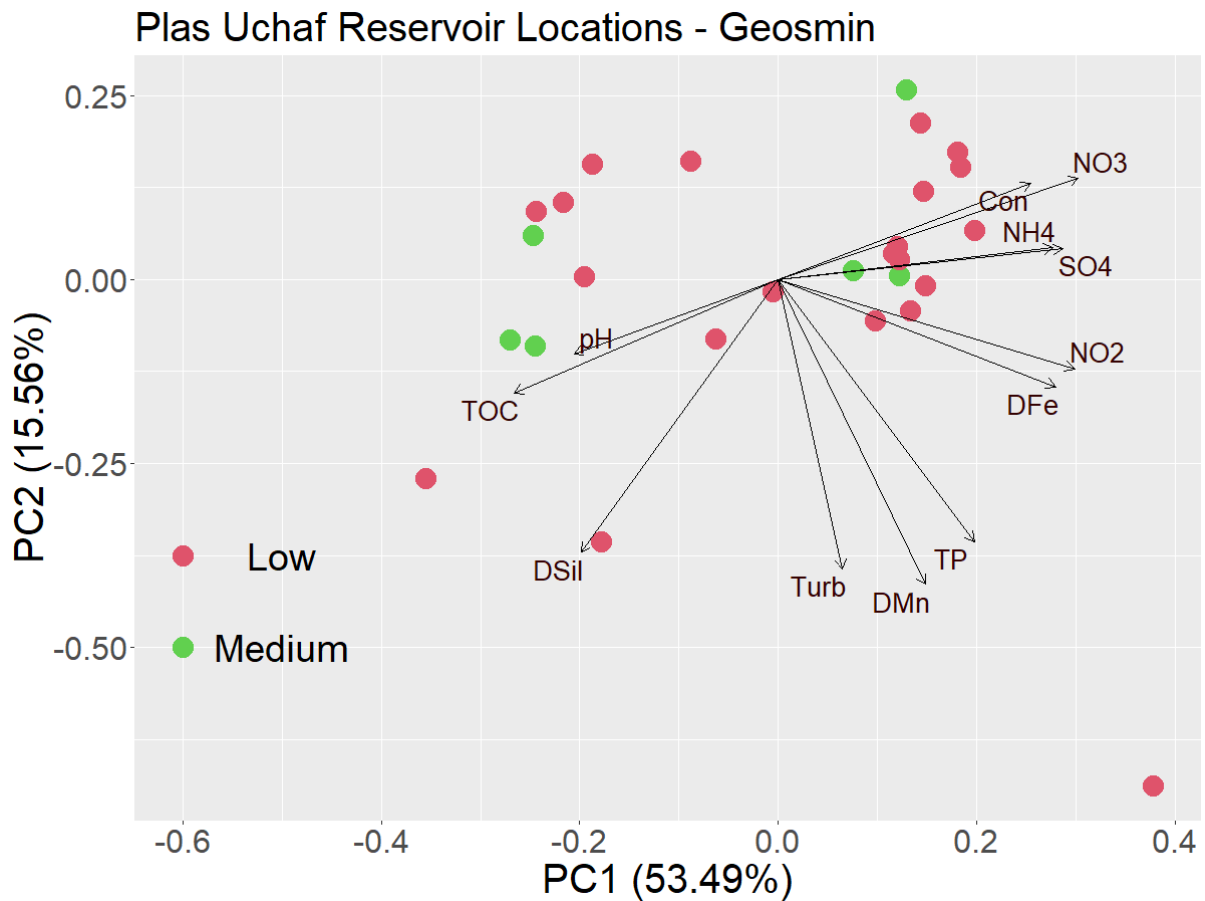


Figure 3.21: PCA biplot of Plas Uchaf reservoir's reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level.

#### 3.3.4.9 – Pontsticill – geosmin and 2-MIB concentrations

For geosmin (Figure 3.22 – top) and 2-MIB concentrations (Figure 3.22 – bottom) in Pontsticill’s reservoir locations, there were no distinct groupings for either T&O compound based on the chemical and physical properties of the water.

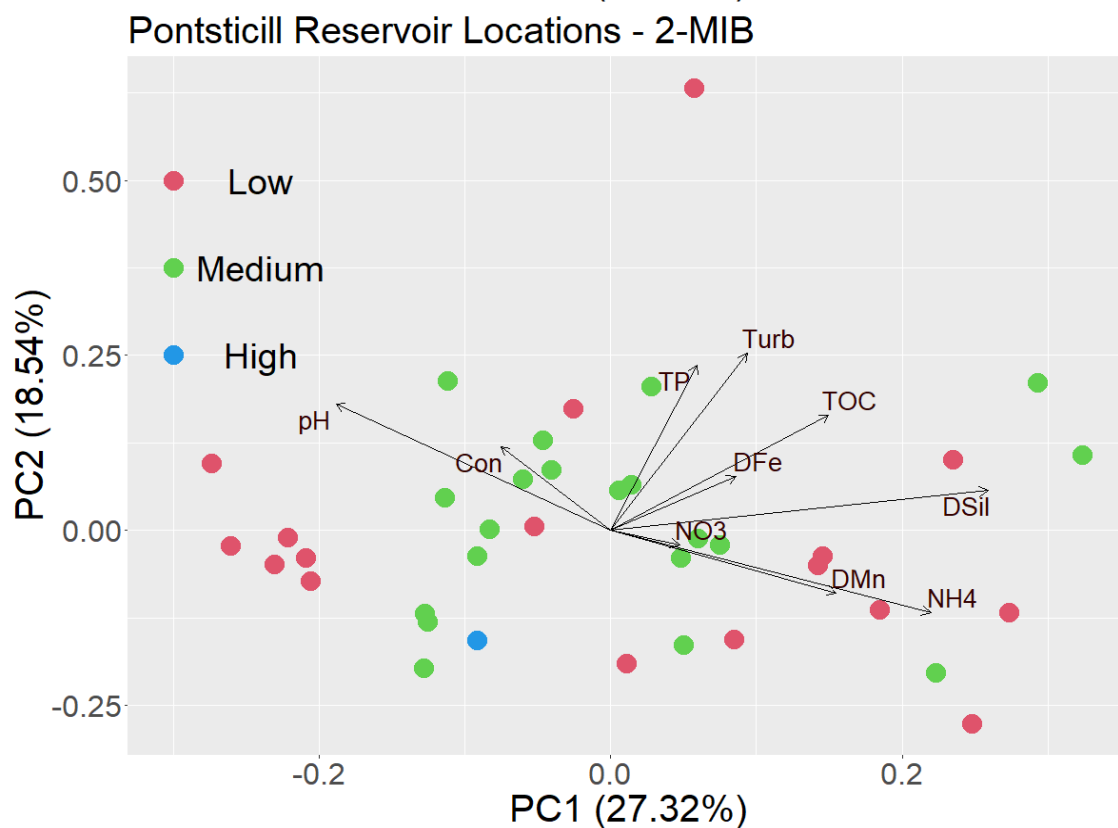
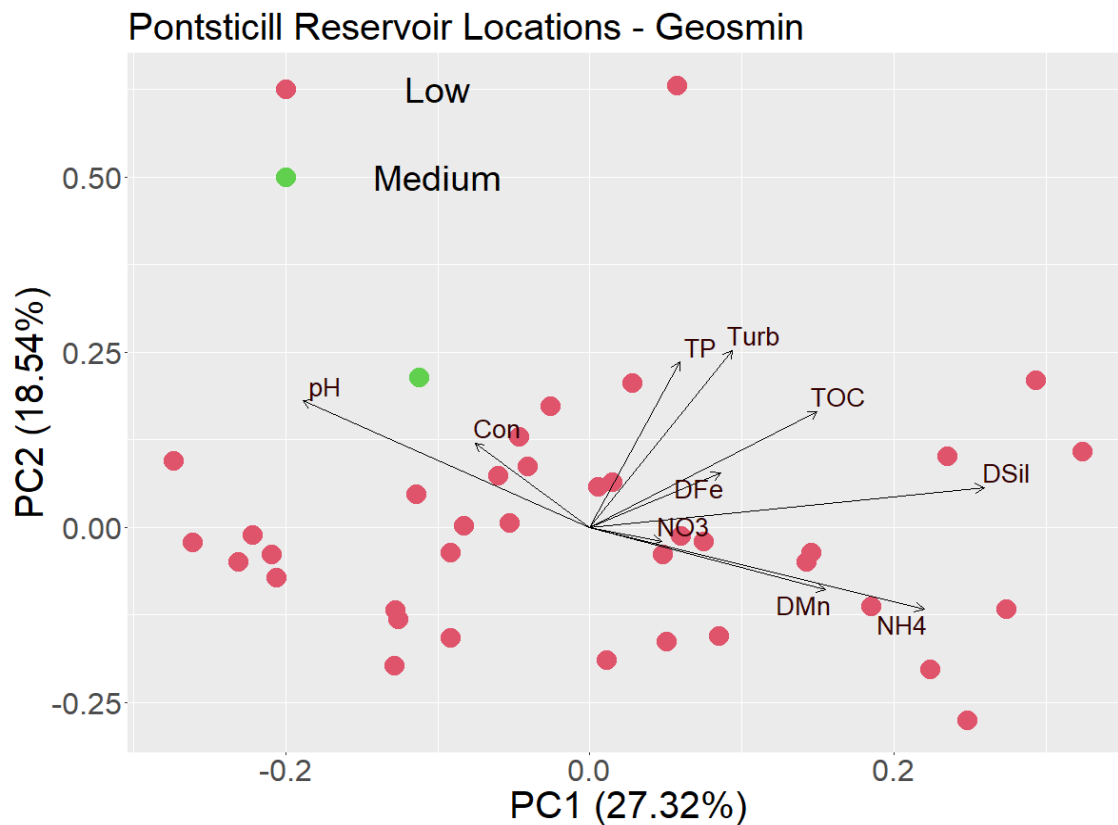


Figure 3.22: PCA biplot of Pontsticill reservoir's reservoir locations using components 1 (PC1) and 2 (PC2) with individual points coloured according to the assigned geosmin level.

### 3.3.4 – Monthly changes in nutrients associated with monthly changes in T&O concentrations

#### 3.3.4.1 – Alaw – geosmin

For the Alaw reservoir, the greatest increase in the monthly changes in geosmin concentration was witnessed in the Res-1 location during month six (+396.2 ng L<sup>-1</sup>), which occurred a month after an increase in NH<sub>4</sub><sup>+</sup> during month five (+0.97 mg L<sup>-1</sup>) (Figure 3.23 – A). During this increase in geosmin concentrations, NO<sub>3</sub><sup>-</sup> slightly increased (+0.16 mg L<sup>-1</sup>), with further increases during month seven (+ 0.90 mg L<sup>-1</sup>) that coincided with a considerable decrease in geosmin concentrations (- 397.50 ng L<sup>-1</sup>) (Figure 3.23 – B). Around the time of heightened NH<sub>4</sub><sup>+</sup> during month five, TP and Ortho also increased simultaneously (+ 0.25 and +0.16 mg L<sup>-1</sup>, respectively) a month before the spike in geosmin concentration (month six) (Figure 3.24 – A and B). Other high monthly changes in geosmin concentrations were confined to Res-2 (+296.20 ng L<sup>-1</sup> – month 13), Res-4 (+261.6 ng L<sup>-1</sup> – month 13) and Res-5 (+215.00 and +146.00 ng L<sup>-1</sup> – months three and 14). Although data from previous months were missing and monthly change was calculated by the last available sampled month (Table 3.3); so, interpretation needs to be treated with caution. During the second month, the Res-1 location had elevated monthly change in geosmin concentration (+128.00 ng L<sup>-1</sup>), although monthly changes in NH<sub>4</sub><sup>+</sup>, NO<sub>3</sub><sup>-</sup> and Ortho concentrations stayed constant. Monthly changes in TP slightly decreased during this period (-0.018 mg L<sup>-1</sup>).



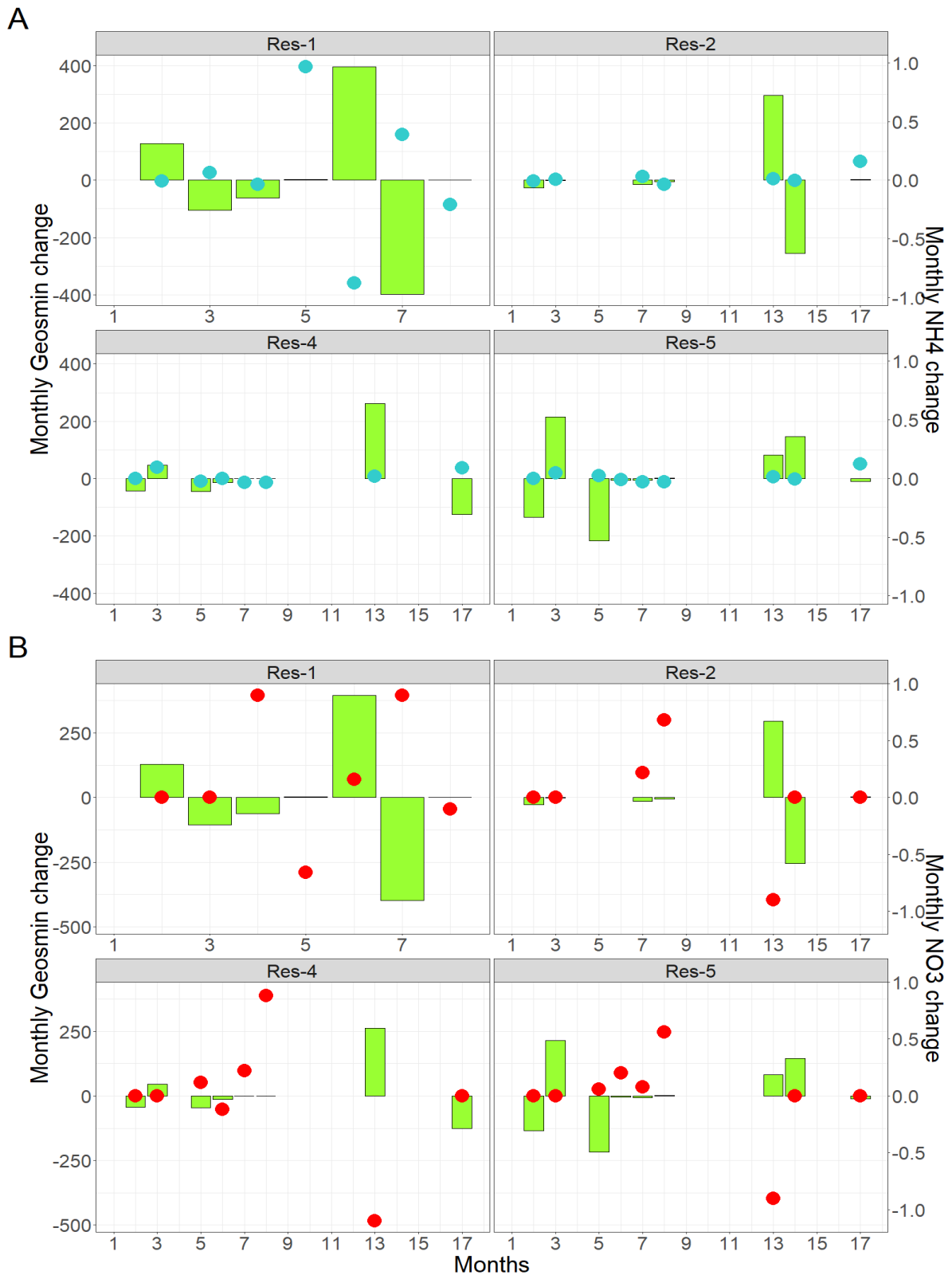


Figure 3.23: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in NH<sub>4</sub><sup>+</sup> concentrations (mg L<sup>-1</sup>) (blue dots) and B – monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (mg L<sup>-1</sup>) (red dots) given from each consecutive sampled month of the study period in Alaw reservoir locations.

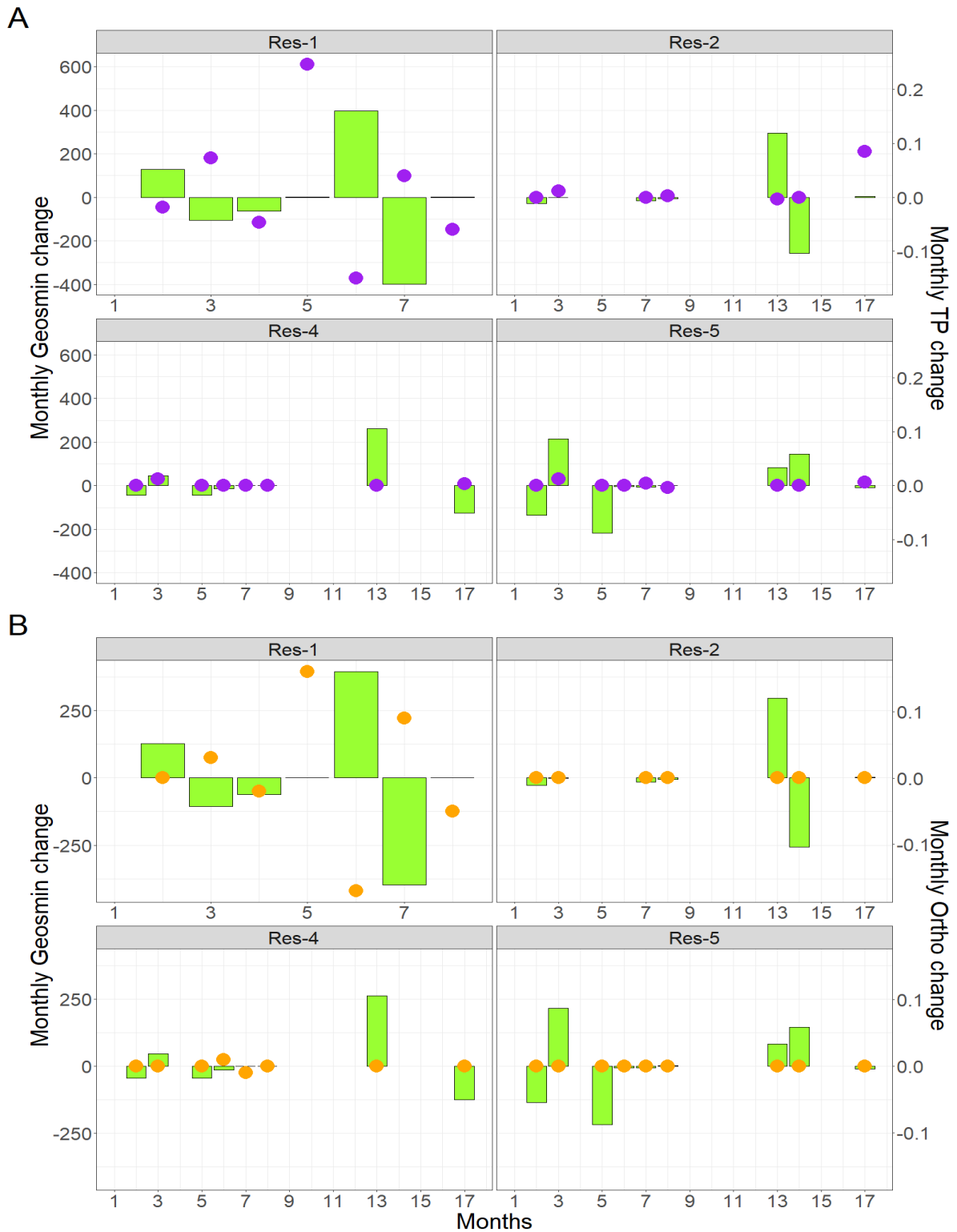


Figure 3.24: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) and B – monthly changes in orthophosphate concentrations ( $\text{mg L}^{-1}$ ) (orange dots) given from each consecutive sampled month of the study period in Alaw reservoir locations.

### 3.3.4.2 – Cefni – geosmin

Monthly changes in geosmin concentrations for the Cefni reservoir (-8.20 ng L<sup>-1</sup> to + 11.40 ng L<sup>-1</sup>) were not as extreme as monthly changes observed in the Alaw reservoir (-397.50 ng L<sup>-1</sup> to +396.20 ng L<sup>-1</sup>), this is graphically illustrated in Appendix 3.1 – 3.2 where the y-axis has been scaled to reflect Alaw's monthly geosmin changes. The most considerable positive changes observed in monthly geosmin concentrations for each reservoir location were during month three (Res-3: 9.10 ng L<sup>-1</sup>, Res-4: 9.30 ng L<sup>-1</sup> and Res-5: +11.40 ng L<sup>-1</sup>) (Figures 3.25 – 3.26). These increases in monthly geosmin concentrations for all reservoir locations occurred a month after slight increases in the monthly changes in NH<sub>4</sub><sup>+</sup> concentrations during month 2 (Res-3: +0.02 mg L<sup>-1</sup>, Res-4: +0.02 mg L<sup>-1</sup> and Res-5: +0.05 mg L<sup>-1</sup>) (Figure 3.25 – A). During these increases in monthly geosmin concentrations, decreases in the monthly NH<sub>4</sub><sup>+</sup> concentrations were noted (Res-3: - 0.03 mg L<sup>-1</sup>, Res-4: - 0.03 mg L<sup>-1</sup> and Res-5: -0.09 mg L<sup>-1</sup>). During month two, the monthly changes in NO<sub>3</sub><sup>-</sup> concentrations remained constant in Res-3 and Res-4 locations and decreased in Res-5 (-0.44 mg L<sup>-1</sup>) (Figure 3.25 – B). Monthly changes in TP during month two before the increases in monthly geosmin concentrations were all positive (Res-3: 0.047 mg L<sup>-1</sup>, Res-4: 0.005 mg L<sup>-1</sup> and Res-5: 0.001 mg L<sup>-1</sup>) (Figure 3.26 – A). During the time of heightened monthly changes in geosmin concentrations (month three), reservoir locations Res-4 and Res-5 experienced a slight increase in monthly changes in TP concentrations (0.006 mg L<sup>-1</sup> and 0.11 mg L<sup>-1</sup>, respectively), whereas Res-3 experienced a decrease in monthly changes in TP concentrations (-0.035 mg L<sup>-1</sup>). During months two and three, no monthly changes were observed for Ortho as it remained at a constant concentration (Figure 3.26 – B).

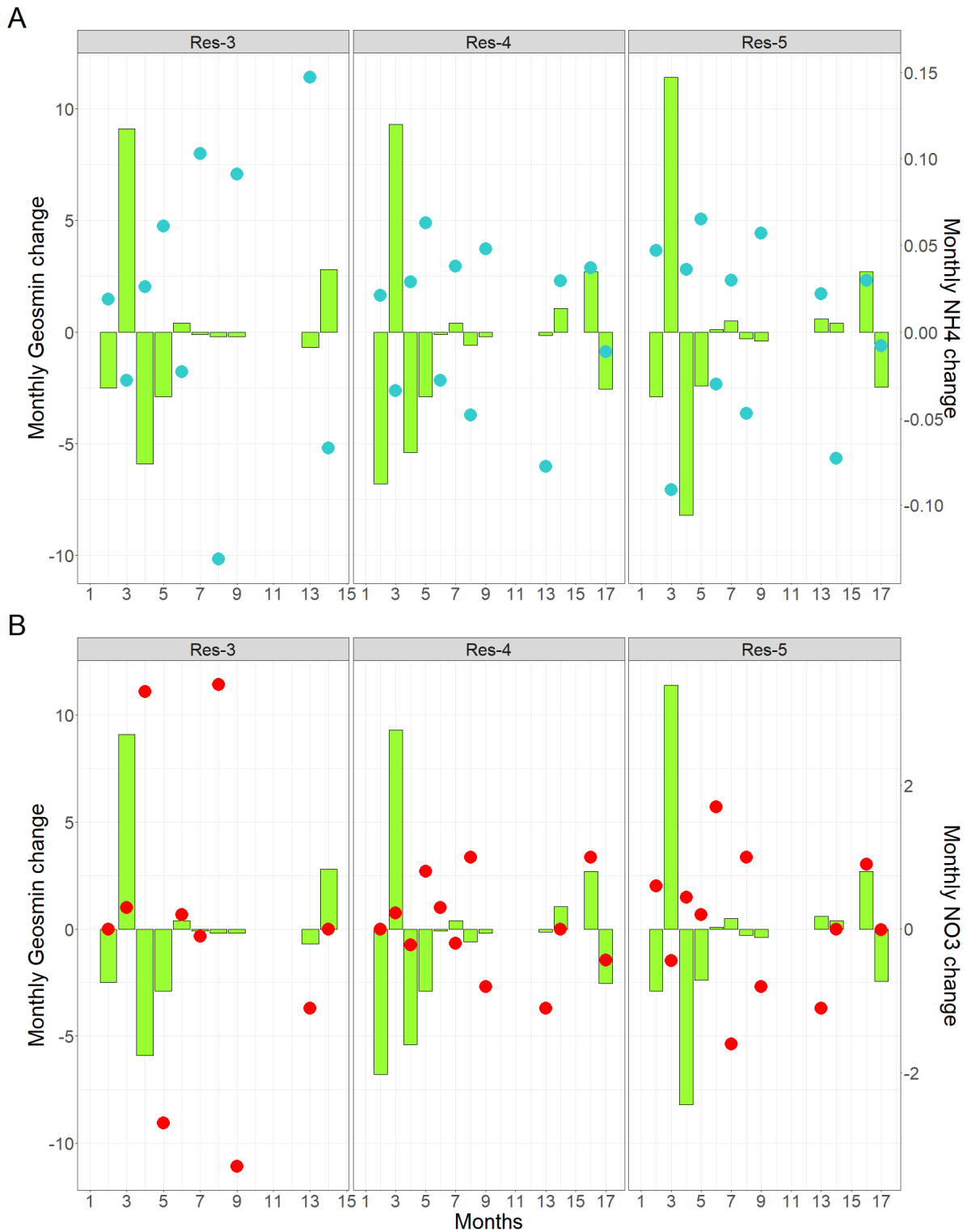


Figure 3.25: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Cefni reservoir locations.

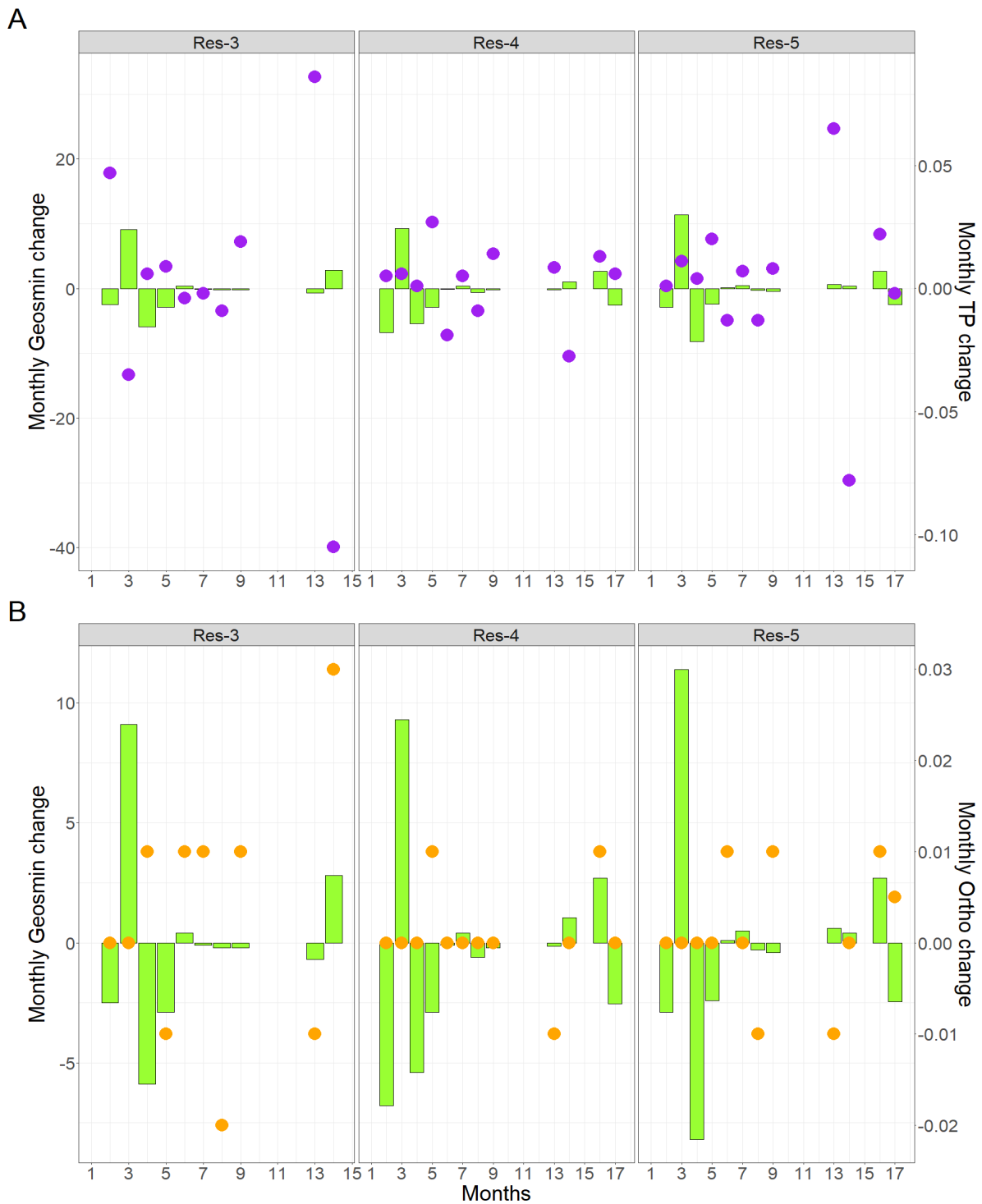


Figure 3.26: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) and B – monthly changes in orthophosphate concentrations ( $\text{mg L}^{-1}$ ) (orange dots) given from each consecutive sampled month of the study period in Cefni reservoir locations.

### 3.3.4.3 – Dolwen – geosmin

Monthly changes in geosmin concentrations for the Dolwen reservoir (-23.10 ng L<sup>-1</sup> to +14.00 ng L<sup>-1</sup>) were not as extreme as monthly changes observed in the Alaw reservoir (-397.50 ng L<sup>-1</sup> to +396.20 ng L<sup>-1</sup>), this is graphically illustrated in Appendix 3.3 – 3.4 where the y-axis has been scaled to reflect Alaw's monthly geosmin changes. Heightened monthly changes in geosmin concentrations in the Res-1 location in Dolwen reservoir were apparent during months two and three (9.50 and 14.00 ng L<sup>-1</sup>, respectively) (Figures 3.27 – 3.28). During month two, there was a slight reduction in NH<sub>4</sub><sup>+</sup> concentration (-0.097 mg L<sup>-1</sup>), followed by an increase during month three (0.034 mg L<sup>-1</sup>), coinciding with a further increase in the monthly geosmin change in concentration (Figure 3.27 – A). During month two, the monthly change in NO<sub>3</sub><sup>-</sup> concentration increased slightly (+0.90 mg L<sup>-1</sup>). During month three NO<sub>3</sub><sup>-</sup> concentration decreased (-1.52 mg L<sup>-1</sup>), this reduction coincided with the largest positive change in geosmin concentration (Figure 3.27 – B). The most positive monthly change in NO<sub>3</sub><sup>-</sup> concentration (+2.62 mg L<sup>-1</sup>) during month four simultaneously occurred with the largest negative monthly change in geosmin concentration (-23.10 ng L<sup>-1</sup>). The monthly change in TP concentration was constant during month two but increased during month three (+0.03 mg L<sup>-1</sup>) when the monthly change in geosmin concentration was at its highest (14.00 ng L<sup>-1</sup>) (Figure 3.28 – A). There was no monthly change in orthophosphate concentration around the period of increases in the monthly change of geosmin concentrations (Figure 3.28 – B).

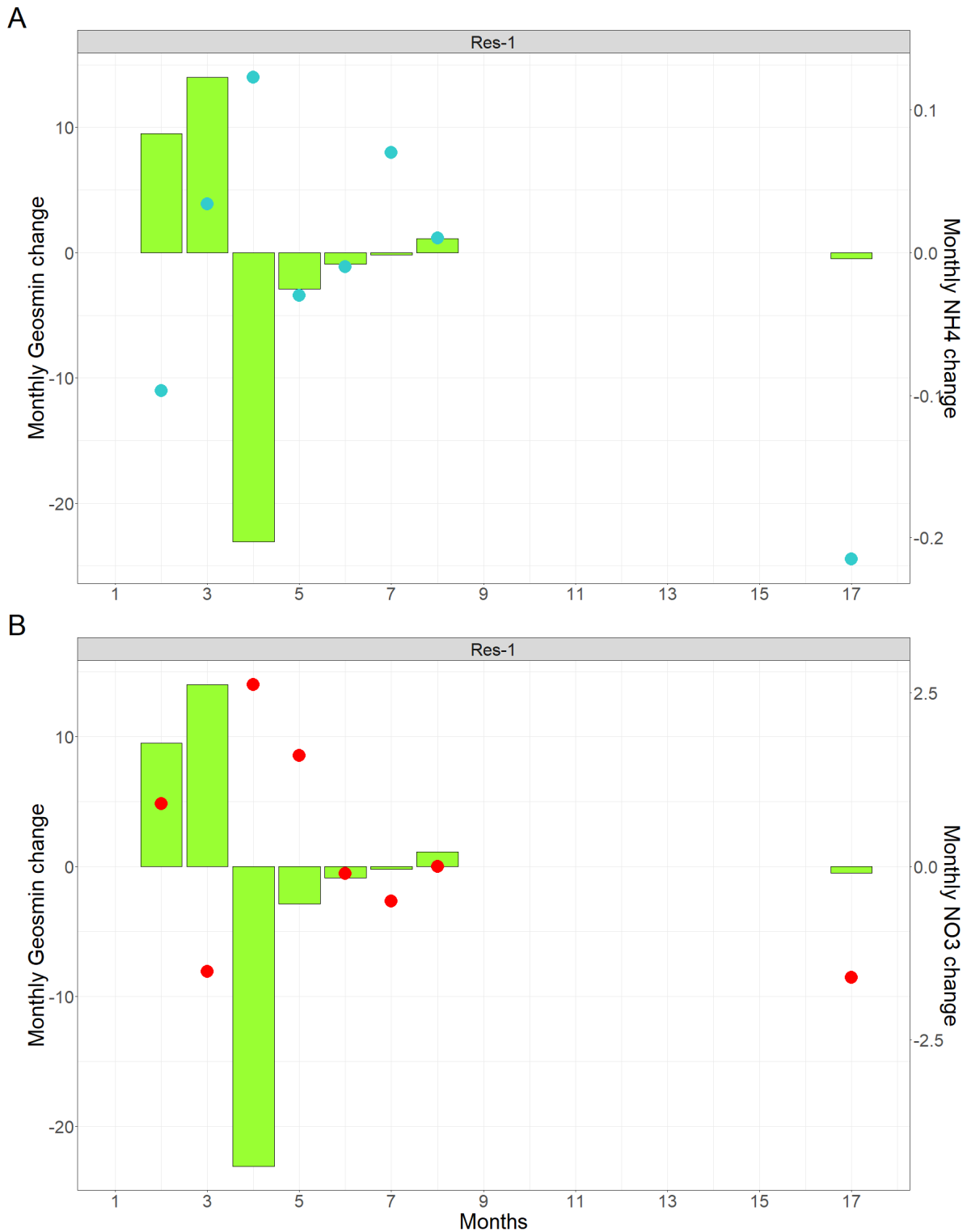


Figure 3.27: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in NH<sub>4</sub><sup>+</sup> concentrations (mg L<sup>-1</sup>) (blue dots) and B – monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (mg L<sup>-1</sup>) (red dots) given from each consecutive sampled month of the study period in Dolwen reservoir locations.

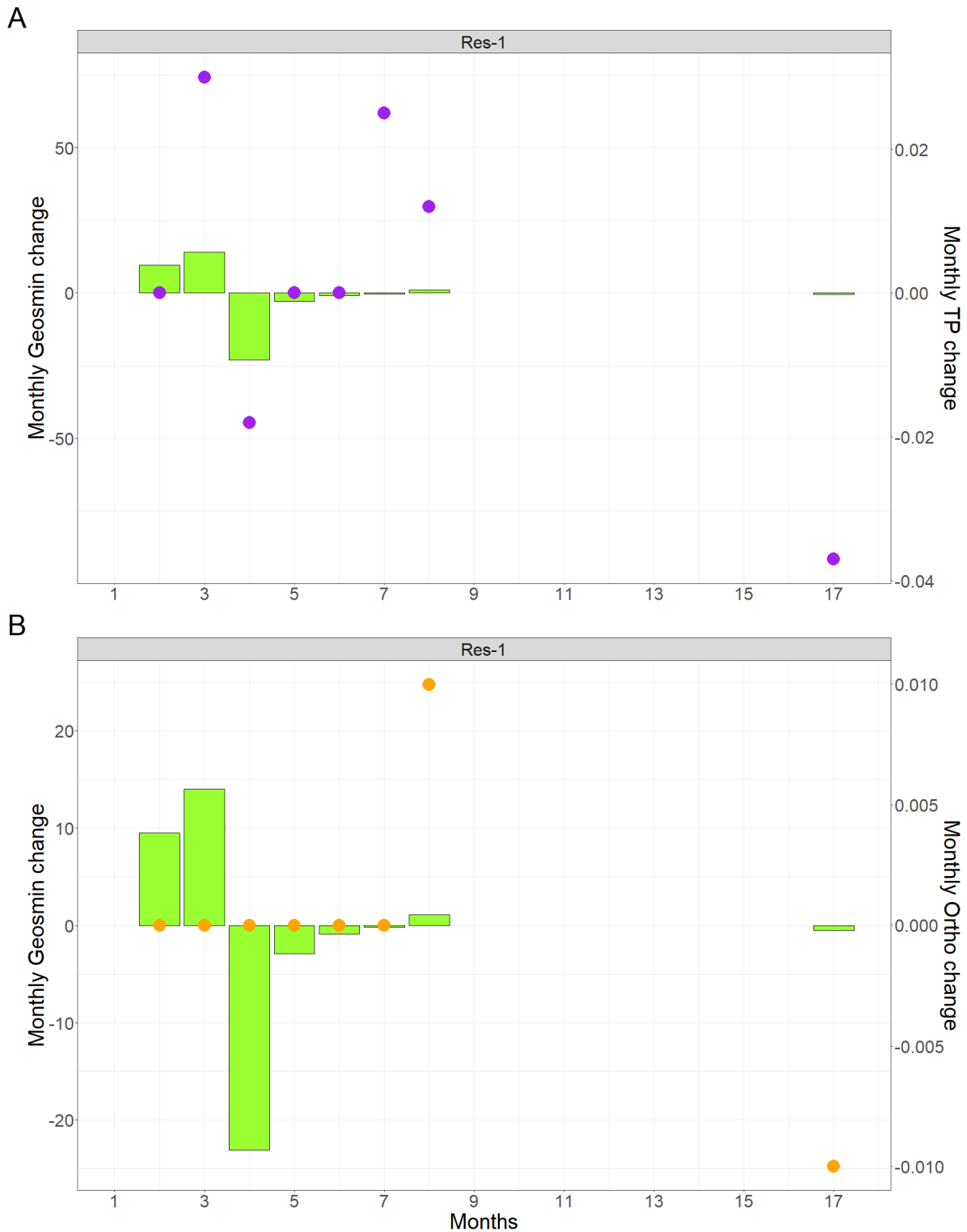


Figure 3.28: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) and B – monthly changes in orthophosphate concentrations (mg L<sup>-1</sup>) (orange dots) given from each consecutive sampled month of the study period in Dolwen reservoir locations.



#### 3.3.4.4 – Llwyn On – geosmin

Monthly changes in geosmin concentrations for Llwyn On reservoir (-13.65 ng L<sup>-1</sup> to +17.15 ng L<sup>-1</sup>) were not as extreme as monthly changes observed in Alaw reservoir (-397.50 ng L<sup>-1</sup> to +396.20 ng L<sup>-1</sup>), this is graphically illustrated in Appendix 3.5 – 3.6 where the y-axis has been scaled to reflect Alaw's monthly geosmin changes. Monthly increases in geosmin concentrations in both locations were confined to month 13, which had previous missing data points (Res-1: 14.35 ng L<sup>-1</sup> and Res-2: 17.15 ng L<sup>-1</sup>) (Figures 3.29 – 3.30). Res-1 location month 13 was calculated using data from sampled months seven and 13. For Res-2 location month 13 was calculated using data from sampled months nine and 13 (Table 3.3). During month 13, for both locations, there was an increase in NH<sub>4</sub><sup>+</sup> (Res-1: +0.0045 mg L<sup>-1</sup> and Res-2: +0.0050 mg L<sup>-1</sup>) (Figure 3.29 – A). During month 13, there was also a decrease in the monthly change in NO<sub>3</sub><sup>-</sup> concentration at the Res-1 location (-0.360 mg L<sup>-1</sup>), at the Res-2 location there was no change in NO<sub>3</sub><sup>-</sup> concentration (Figure 3.29 – B). No TP and Ortho concentrations changes were witnessed during month 13, staying at constant values (Figure 3.30 – A and B, respectively).

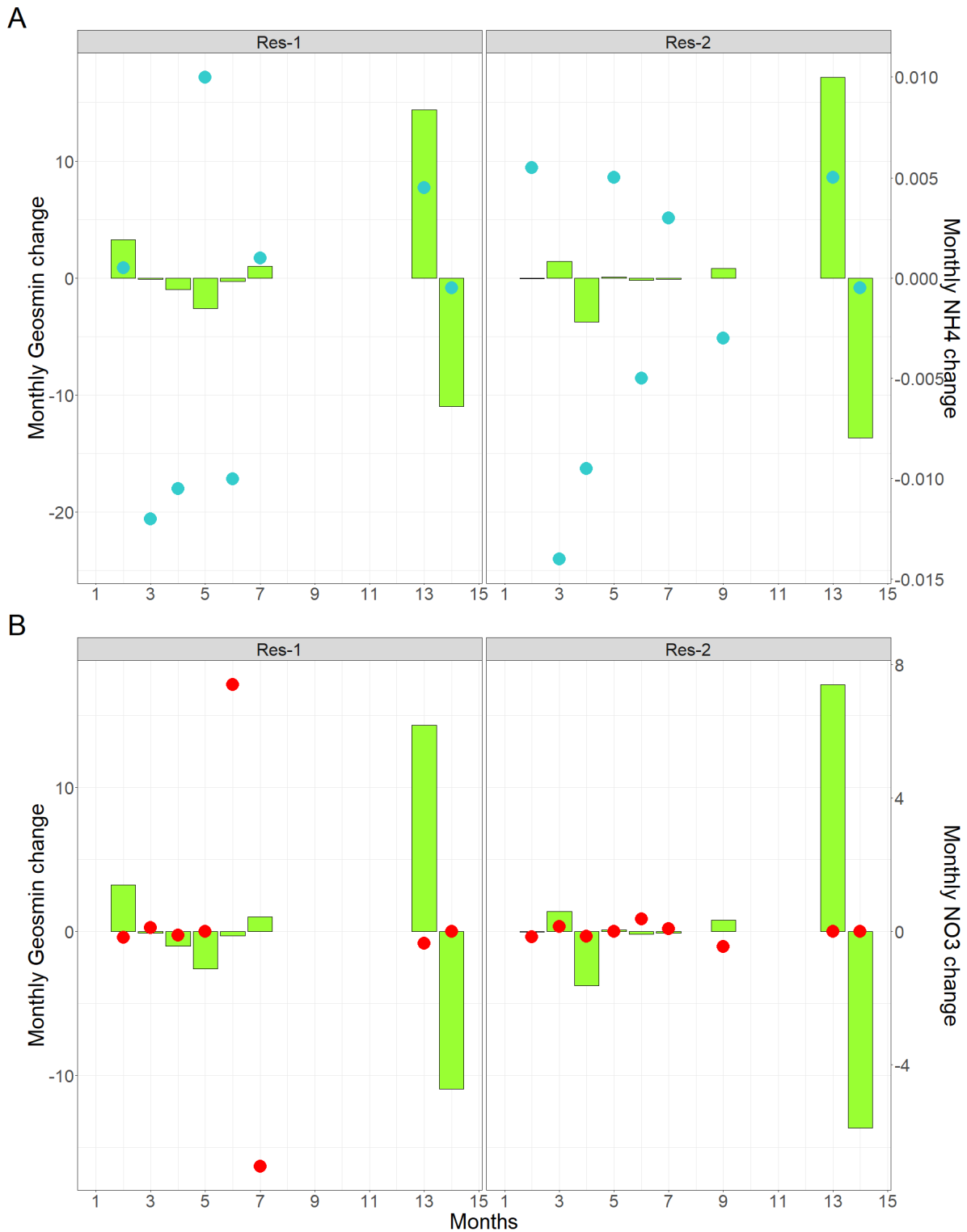


Figure 3.29: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Llwyn On reservoir locations.

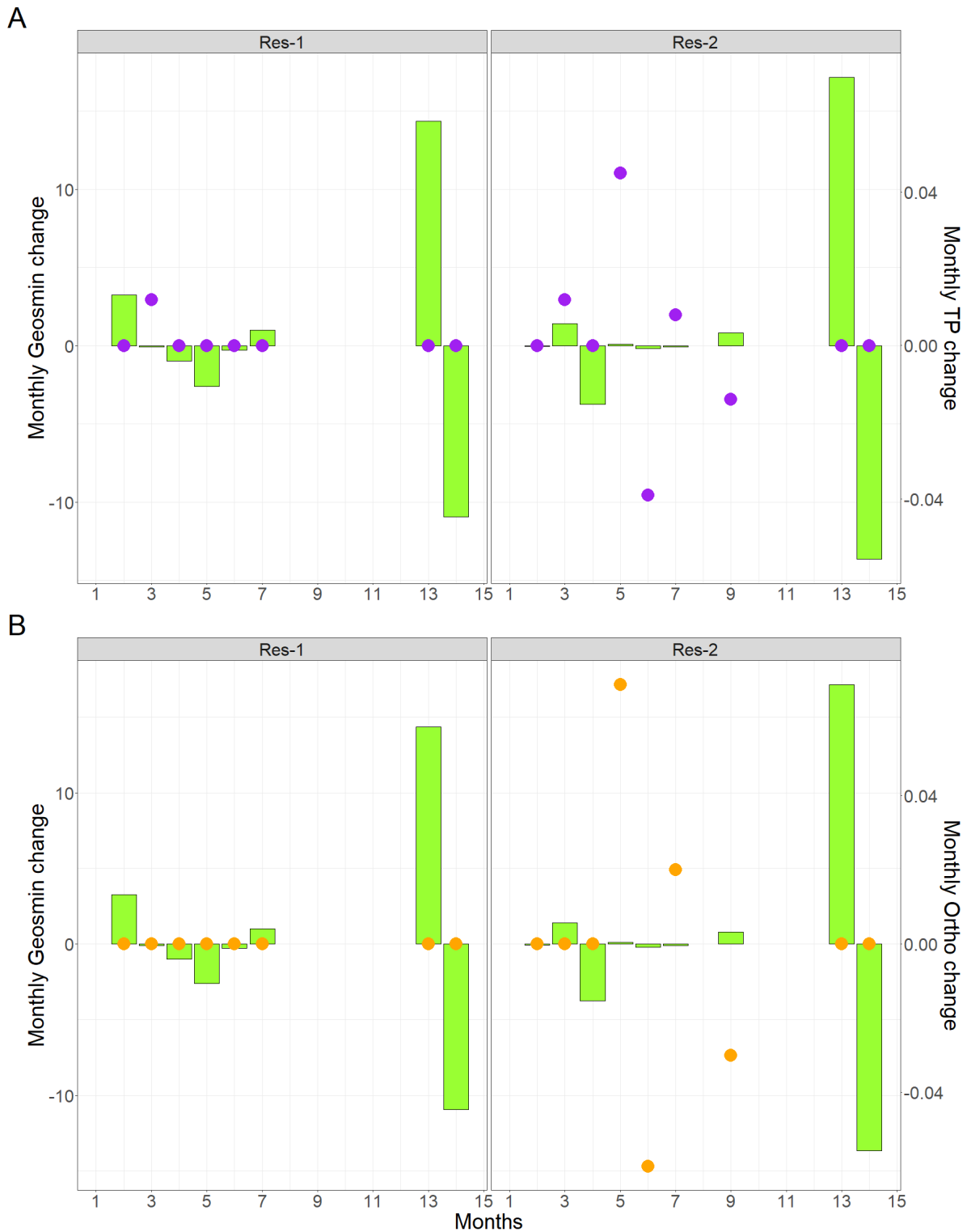


Figure 3.30: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) and B – monthly changes in orthophosphate concentrations ( $\text{mg L}^{-1}$ ) (orange dots) given from each consecutive sampled month of the study period in Llwyn On reservoir locations.

### 3.3.4.5 – Plas Uchaf – geosmin

Monthly changes in geosmin concentrations for Plas Uchaf (-6.30 ng L<sup>-1</sup> to +3.80 ng L<sup>-1</sup>) reservoir were not as extreme as monthly changes observed in the Alaw reservoir (-397.50 ng L<sup>-1</sup> to +396.20 ng L<sup>-1</sup>), this is graphically illustrated in Appendix 3.7 – 3.8 where the y-axis has been scaled to reflect Alaw's monthly geosmin changes. Heightened monthly changes in geosmin concentrations were witnessed during month two for Res-1 (+3.80 ng L<sup>-1</sup>) and Res-3 (+3.50 ng L<sup>-1</sup>) and month five for Res-2 (+3.4 ng L<sup>-1</sup>) (Figures 3.31 – 3.32). The most considerable reduction in monthly geosmin change occurred at Res-2 during month four (-6.30 ng L<sup>-1</sup>). During month two in Res-1 and Res-3 locations, monthly changes in geosmin concentrations increased during reduced monthly changes in NH<sub>4</sub><sup>+</sup> (Res-1: -0.032 mg L<sup>-1</sup> and Res-3: -0.059 mg L<sup>-1</sup>) (Figure 3.31 – A). During month two at Res-1 and Res-3 locations, there were increases in the monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (Res-1: +0.61 and Res-3: +0.50 mg L<sup>-1</sup>) (Figure 3.31 – B). During month two at the Res-1 location, there was no monthly change in TP concentration, but there was a slight increase at the Res-3 location (+0.005 mg L<sup>-1</sup>) (Figure 3.32). At the Res-2 location during month four, there was an increase in the monthly change in NH<sub>4</sub><sup>+</sup> concentration (+0.088 mg L<sup>-1</sup>) before the increase in the monthly change in geosmin concentration during month five (+3.40 ng L<sup>-1</sup>). During the increase in the monthly change in geosmin concentration during month five at the Res-2 location, the monthly change in NH<sub>4</sub><sup>+</sup> concentration subsequently decreased (-0.016 mg L<sup>-1</sup>). Monthly changes in TP concentration did not fluctuate during months four or five at the Res-2 location. Orthophosphate data was removed from this analysis due to no apparent monthly concentration changes throughout the study period.

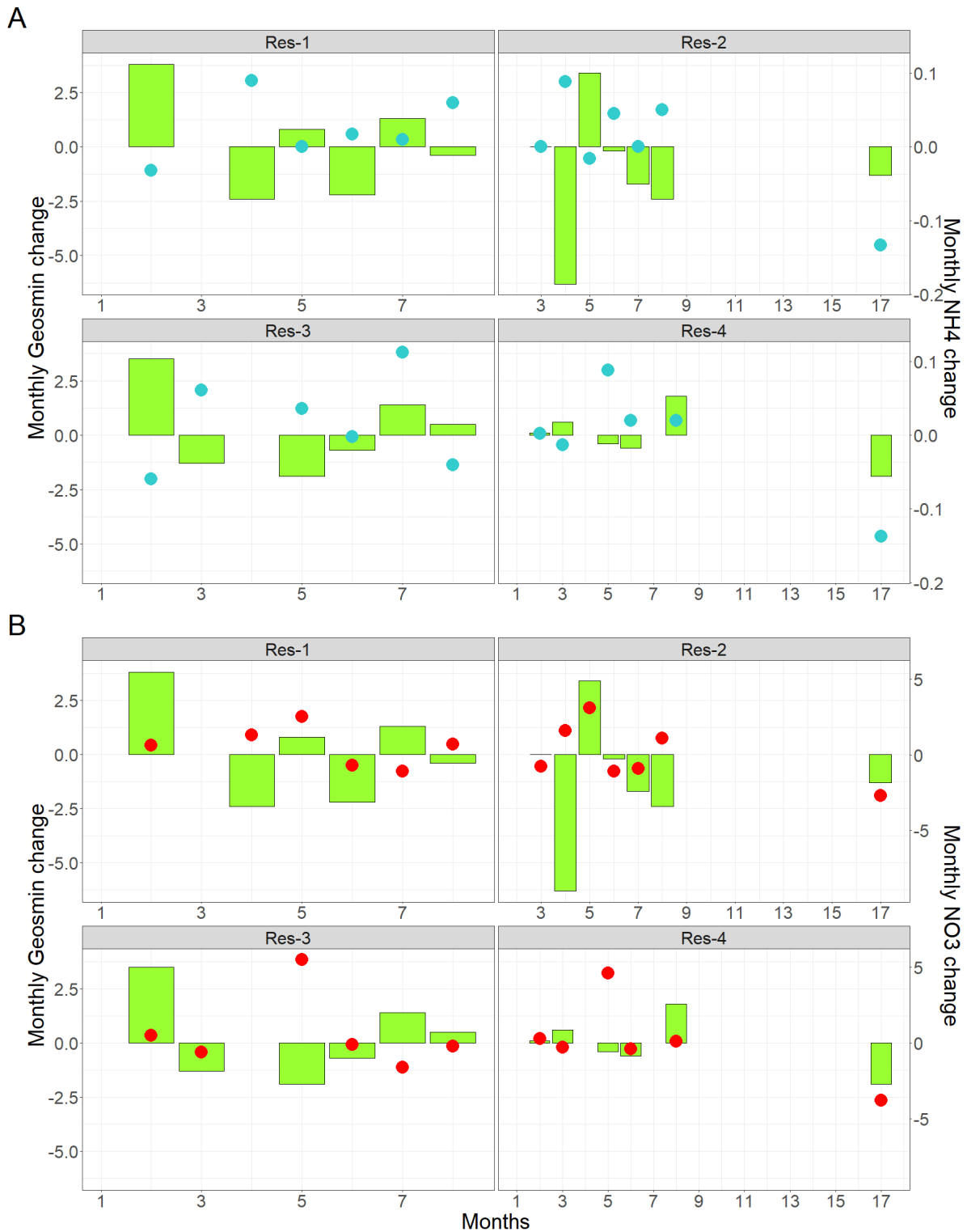


Figure 3. 31: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Plas Uchaf reservoir locations.

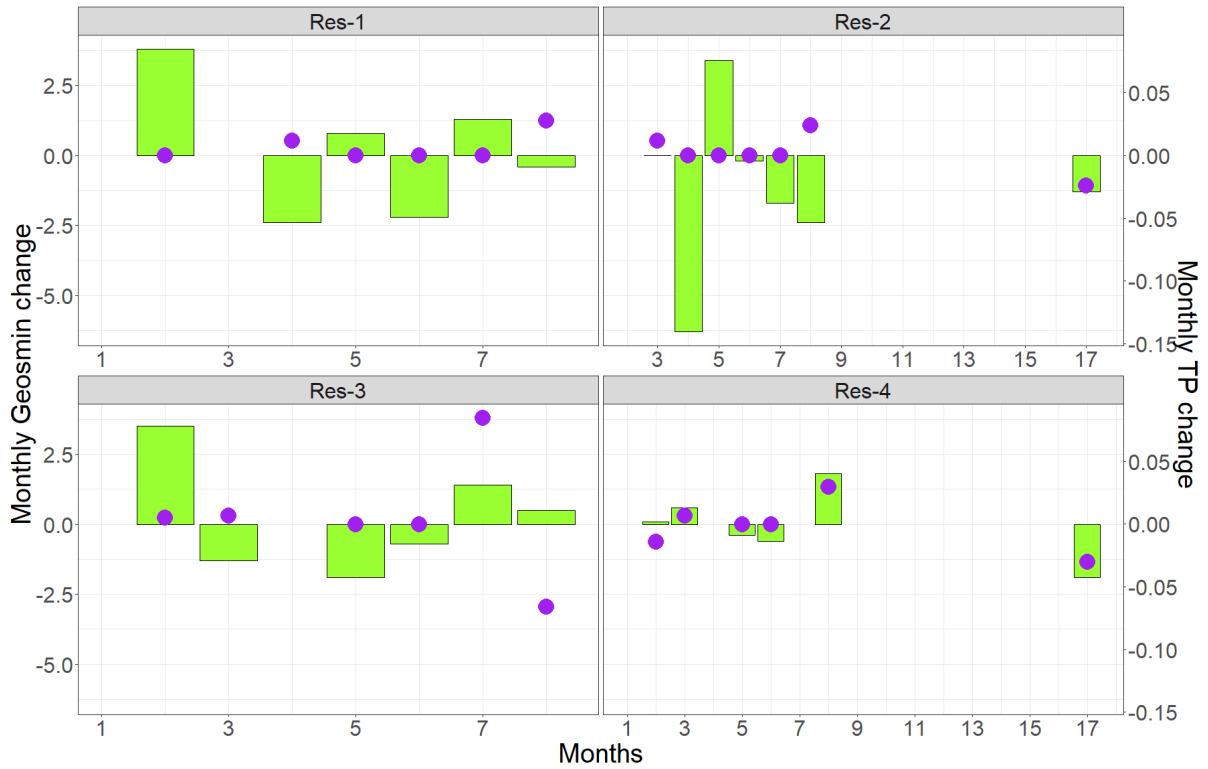


Figure 3.32: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) given from each consecutive sampled month of the study period in Plas Uchaf reservoir locations.

### 3.3.4.6 – Pontsticill – geosmin

Monthly changes in geosmin concentrations for Pontsticill (-6.35 ng L<sup>-1</sup> to +7.07 ng L<sup>-1</sup>) reservoir were not as extreme as monthly changes observed in the Alaw reservoir (-397.50 ng L<sup>-1</sup> to +396.20 ng L<sup>-1</sup>), this is graphically illustrated in Appendix 3.9 – 3.10 where the y-axis has been scaled to reflect Alaw's monthly geosmin changes. There was only one incident of a heightened monthly change in geosmin concentration at the Res-1 location during month three (+7.07 ng L<sup>-1</sup>), with a subsequent reduction during month four (-6.35 ng L<sup>-1</sup>) (Figures 3.33 – 3.34). Before the increase in monthly change in geosmin concentration at Res-1 during month two, there was a slight increase in the monthly change in NH<sub>4</sub><sup>+</sup> concentration (+0.001 mg L<sup>-1</sup>) (Figure 3.33 – A). During the heightened monthly change in geosmin concentration at Res-1 during month three, NH<sub>4</sub><sup>+</sup> decreased (-0.003 mg L<sup>-1</sup>). Monthly changes in NO<sub>3</sub><sup>-</sup> concentration during month three at the Res-1 location increased (+0.100 mg L<sup>-1</sup>) and consequently decreased during month four when the monthly geosmin concentration decreased (-6.35 ng L<sup>-1</sup>) (Figure 3.33 – B). Monthly changes in NH<sub>4</sub><sup>+</sup> increased during month four (+0.084 mg L<sup>-1</sup>), and monthly changes in NO<sub>3</sub><sup>-</sup> concentration remained at a constant concentration, yet monthly changes in geosmin concentrations continued to decrease from month five (-0.100 ng L<sup>-1</sup>) to month six (-0.500 ng L<sup>-1</sup>). During the heightened monthly change in geosmin concentration, there was an increase in the monthly change in TP concentration (+0.012 mg L<sup>-1</sup>) (Figure 3.34). No monthly changes in TP concentrations occurred from months four to six. Orthophosphate data was removed from this analysis due to no apparent monthly concentration changes throughout the study period.

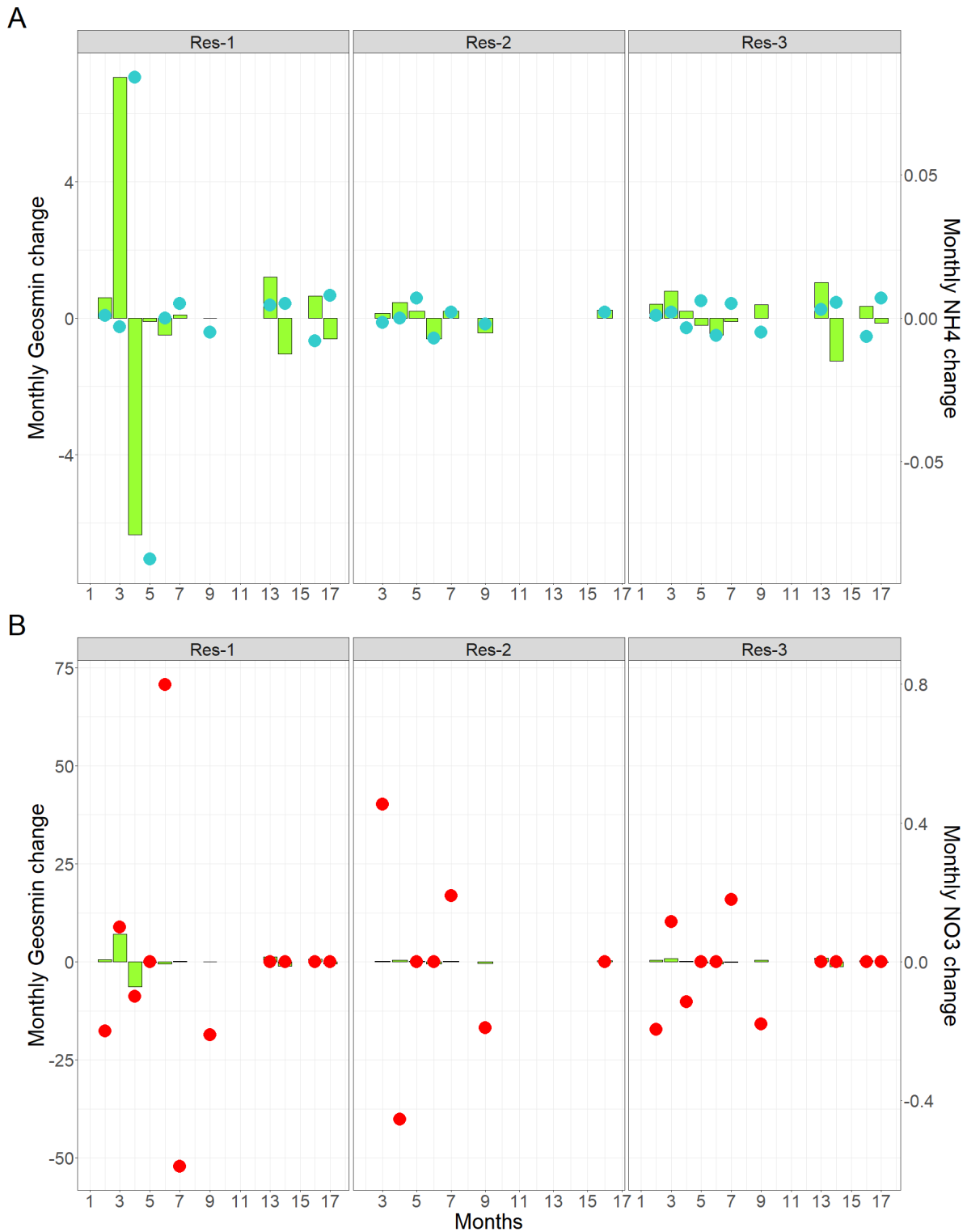


Figure 3. 33: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations.



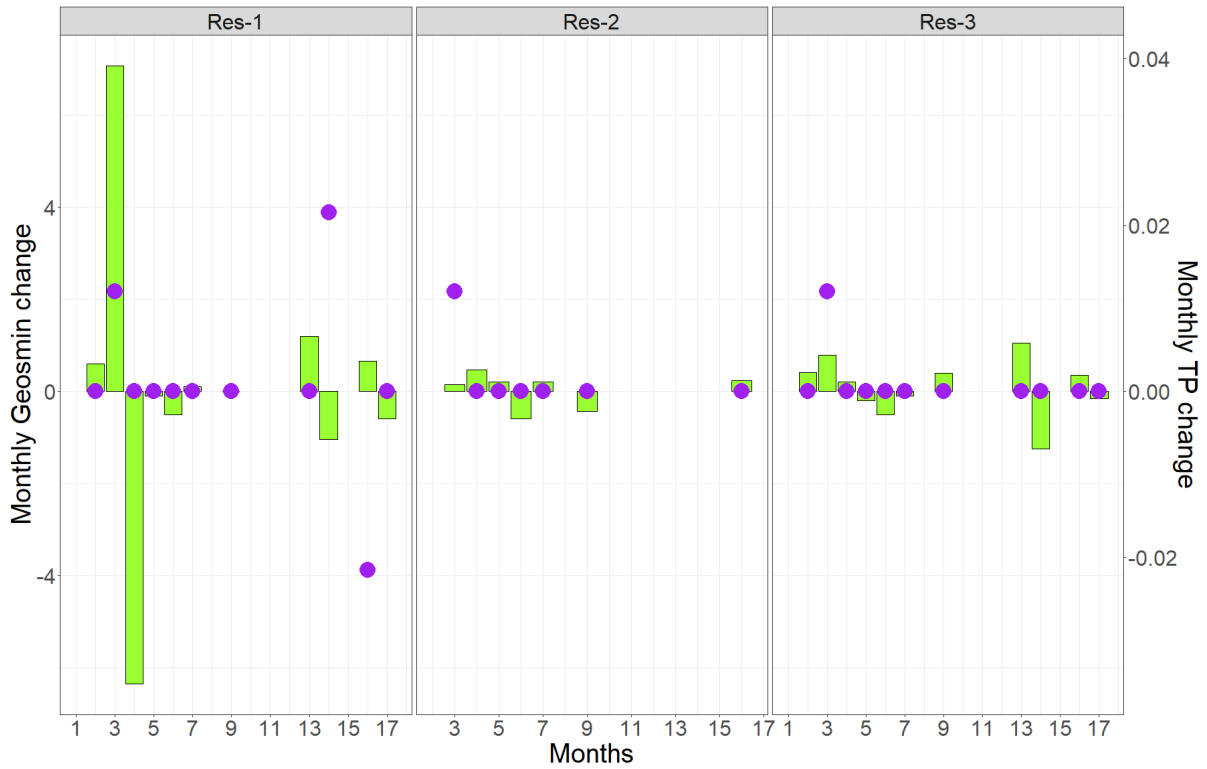


Figure 3.34: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations.

### 3.3.4.7 – Pentwyn – 2-MIB

For the Pentwyn reservoir, the greatest increases in the monthly changes in 2-MIB concentrations were witnessed in the Res-1 location during month two (+12.50 ng L<sup>-1</sup>) and month 16 (+10.32 ng L<sup>-1</sup>) and at the Res-2 location during month 16 (+10.08 ng L<sup>-1</sup>) (Figures 3.35 – 3.36). The monthly increase in 2-MIB concentration witnessed at Res-1 during month 2 coincided with a decrease in NH<sub>4</sub><sup>+</sup> (-0.15 mg L<sup>-1</sup>) (Figure 3.35 – A) whilst the monthly change in NO<sub>3</sub><sup>-</sup> concentration remained the same (Figure 3.35 – B). During this increase in monthly 2-MIB concentration, the monthly change in TP remained constant, increasing during month three (+0.012 mg L<sup>-1</sup>) (Figure 3.36). This slight increase in the monthly change in TP concentration coincided with a decrease in the monthly change in 2-MIB concentration (-5.50 ng L<sup>-1</sup>). For month 16 in the Res-1 location, the monthly change calculations utilised data present for month 14 (Table 3.3). During month 14 in the Res-1 location, there was a slight increase in the monthly change in NH<sub>4</sub><sup>+</sup> concentration (+0.002 mg L<sup>-1</sup>), which subsequently decreased during month 16 (-0.0015 mg L<sup>-1</sup>) when there was an increase in the monthly change for 2-MIB concentration (+10.32 ng L<sup>-1</sup>); no observed monthly NO<sub>3</sub><sup>-</sup> or TP concentration changes existed during this period. Similarly, for the Res-2 location, there was an increase in the monthly change in NH<sub>4</sub><sup>+</sup> concentration (+0.003 mg L<sup>-1</sup>) during month 15 before an increase in the monthly change in 2-MIB concentration (+10.08 mg L<sup>-1</sup>) during month 16. No change was noted for the monthly changes in NO<sub>3</sub><sup>-</sup> concentrations or TP concentrations around this increase in the monthly change in 2-MIB concentration. Orthophosphate data was removed from this analysis due to no apparent monthly concentration changes throughout the study period.

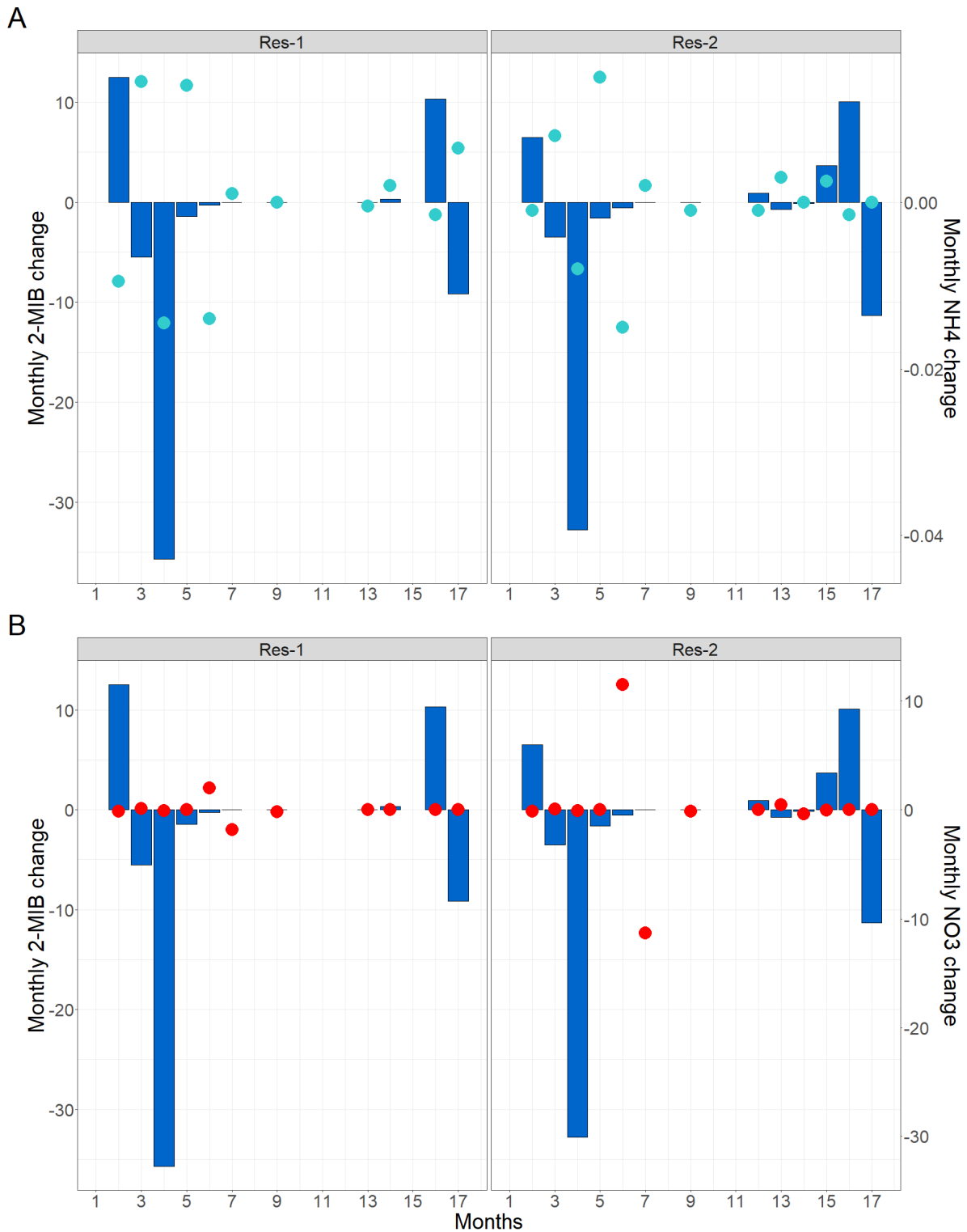


Figure 3.35: Monthly changes in 2-MIB concentrations ( $\text{ng L}^{-1}$ ) (dark blue bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (light blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Pentwyn reservoir locations.

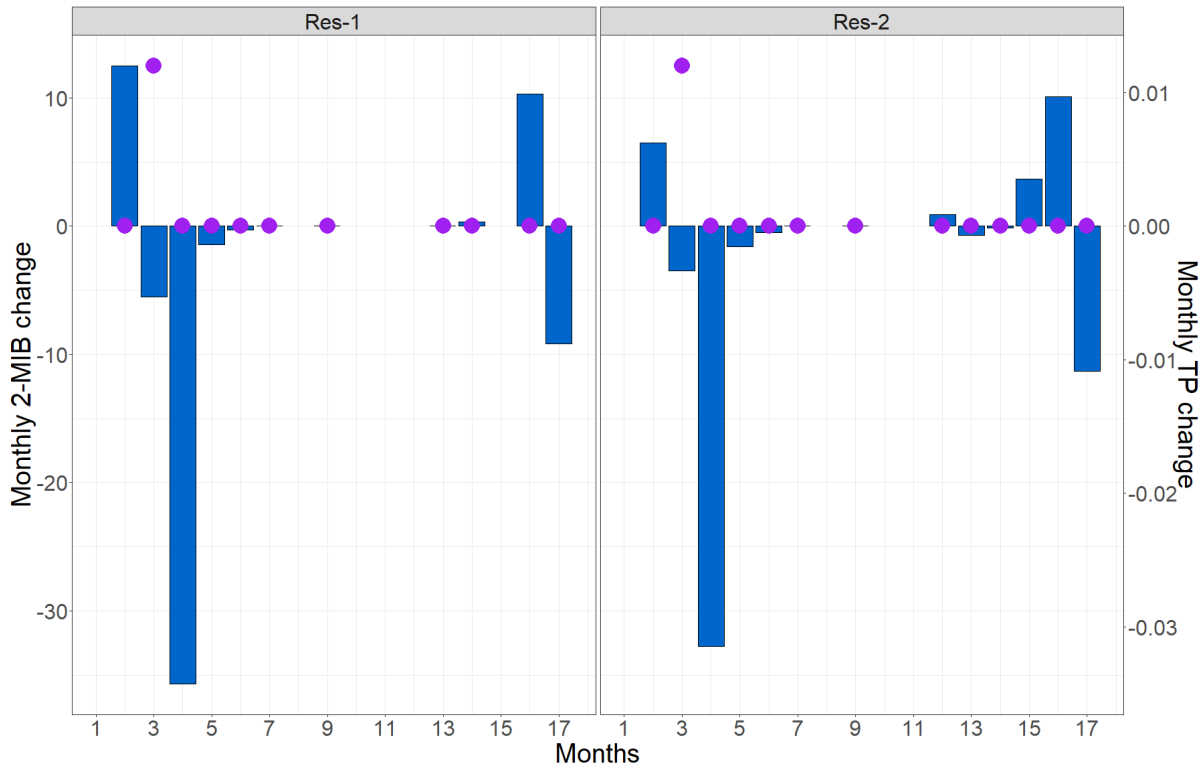


Figure 3.36: Monthly changes in 2-MIB concentrations (ng L<sup>-1</sup>) (dark blue bars) and monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) given from each consecutive sampled month of the study period in Pentwyn reservoir locations.

#### 3.3.4.8 – Pontsticill – 2-MIB

For the Pontsticill reservoir, the greatest increases in the monthly changes in 2-MIB concentrations were witnessed during month 16 at reservoir locations: Res-3 (+3.88 ng L<sup>-1</sup>) and Res-1 (+3.07 ng L<sup>-1</sup>) (Figures 3.37 – 3.38). For both incidences of increased monthly changes in 2-MIB concentrations, data was not available for sampled month 15 and monthly change for month 16 was calculated using sampled month 14 data (Table 3.3). At both locations during month 14, there were increases in the monthly change in NH<sub>4</sub><sup>+</sup> concentrations (Res-1: +0.005 mg L<sup>-1</sup> and Res-3: 0.006 mg L<sup>-1</sup>) that decreased (Res-1: -0.008 mg L<sup>-1</sup> and Res-3: -0.007 mg L<sup>-1</sup>) during month 16 (Figure 3.37 – A). During months 14 to 16, no monthly changes in NO<sub>3</sub><sup>-</sup> concentrations were recorded (Figure 3.37 – B); this was also true for TP concentrations for Res-3 (Figure 3.38). For the Res-1 location during month 14, a substantial positive monthly change in TP concentration (+0.022 mg L<sup>-1</sup>)

subsequently decreased ( $-0.022 \text{ mg L}^{-1}$ ) when the 2-MIB monthly change increased during month 16. Orthophosphate data was removed from this analysis due to no apparent monthly concentration changes throughout the study period.

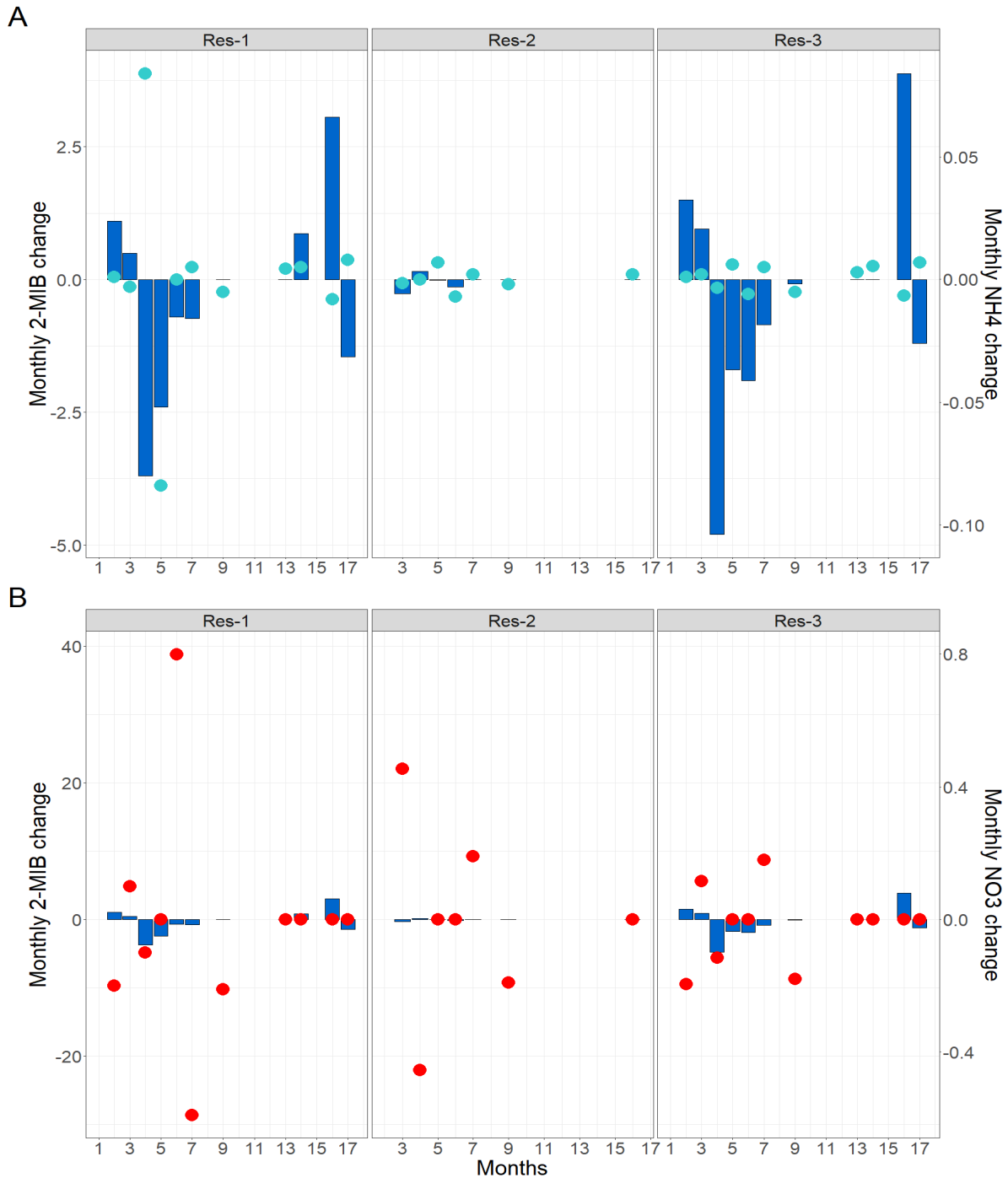


Figure 3.37: Monthly changes in 2-MIB concentrations ( $\text{ng L}^{-1}$ ) (dark blue bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (light blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations.

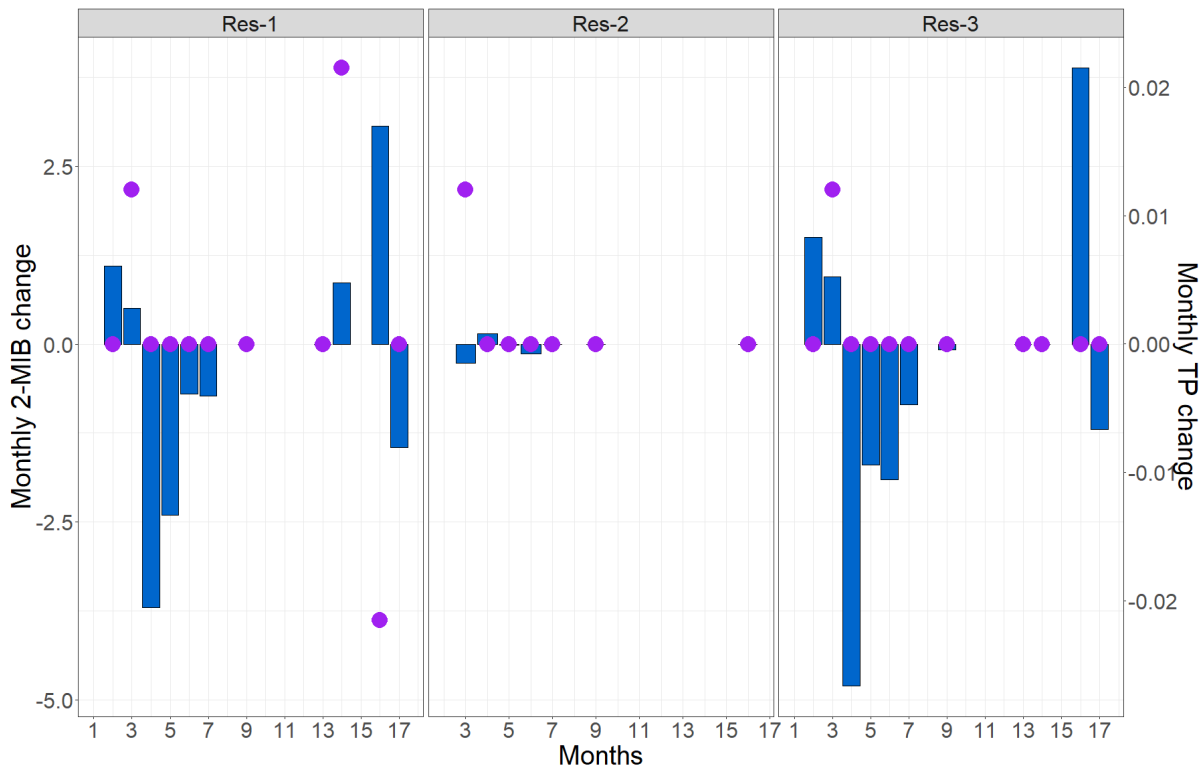


Figure 3.38: Monthly changes in 2-MIB concentrations (ng L<sup>-1</sup>) (dark blue bars) and monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations.

### 3.3.5 – Correlation matrices of monthly changes in nutrients associated with monthly changes in T&O concentrations

Monthly changes in geosmin concentrations were lagged by one consecutive sampled month to distinguish any time-lapsed relationships with the monthly nutrient changes for all reservoirs experiencing a geosmin event (Figure 3.39). Here, monthly changes in NH<sub>4</sub><sup>+</sup> concentrations were found to be significantly positively associated with all reservoir's lagged monthly changes in geosmin concentrations ( $r = 0.908$ ,  $p < 0.001$ ). Although this relationship was most likely driven by Alaw's association with lagged monthly changes in geosmin concentrations and monthly changes in NH<sub>4</sub><sup>+</sup> ( $r = 0.965$ ,  $p < 0.005$ ). Lagged monthly changes in geosmin concentrations and monthly changes in NO<sub>3</sub><sup>-</sup> concentrations did not reveal any significant relationships ( $r = -0.046$ ,  $p = 0.77$ ). TP was also significantly positively associated with all reservoir's lagged geosmin concentrations ( $r = 0.858$ ,  $p < 0.001$ ); however, this relationship was

most likely driven by Alaw's correlation ( $r = 0.931$ ,  $p < 0.01$ ). Monthly changes in Ortho were significant to all reservoirs' lagged geosmin concentration ( $r = 0.885$ ,  $p < 0.001$ ), although this was likely influenced by Alaw's correlation ( $r = 0.923$ ,  $p < 0.01$ ). Lagged monthly changes in geosmin concentrations were significantly negatively impacted by increases in the monthly changes in  $\text{SO}_4^{2-}$  for Alaw ( $r = -0.838$ ,  $p < 0.05$ ), although  $\text{SO}_4^{2-}$  was positively influential on lagged geosmin in Plas Uchaf reservoir ( $r = 0.971$ ,  $p < 0.01$ ). Overall lagged monthly changes in geosmin concentrations had no significant relationship with monthly changes in DFe, although monthly changes in DFe significantly affected Plas Uchaf reservoir ( $r = -0.891$ ,  $p < 0.05$ ) geosmin concentrations.

Using the corresponding monthly changes in geosmin concentrations without a month lag revealed a significant negative relationship with all reservoir's monthly changes in  $\text{NH}_4^+$  ( $r = -0.611$ ,  $p < 0.001$ ) (Figure 3.40). Monthly changes in  $\text{NH}_4^+$  concentrations had the most significant effects in Cefni ( $r = -0.486$ ,  $p < 0.05$ ) and Plas Uchaf ( $r = -0.845$ ,  $p < 0.05$ ) on geosmin concentrations. Monthly changes in TP concentrations were negatively significant against monthly changes in geosmin concentrations for all reservoirs ( $r = -0.427$ ,  $p < 0.005$ ), having the main effects in the Pontsticill reservoir ( $r = -0.797$ ,  $p < 0.1$ ). Overall monthly changes in geosmin concentrations were not significantly associated with monthly changes in  $\text{NO}_3^-$  concentrations, although they were significantly negatively associated in the Dolwen reservoir ( $r = -0.784$ ,  $p < 0.05$ ). Monthly changes in DFe were significantly associated with monthly changes in geosmin concentrations, negatively in the Cefni reservoir ( $r = -0.409$ ,  $p < 0.05$ ) and positively in Dolwen ( $r = 0.781$ ,  $p < 0.05$ ). For the monthly changes in DSil, there was an overall significantly negative relationship with monthly changes in geosmin concentrations ( $r = -0.279$ ,  $p < 0.05$ ), with the greatest effects in the Cefni reservoir ( $r = -0.692$ ,  $p < 0.001$ ).



Figure 3.39: Correlogram of monthly changes in nutrients and monthly changes in geosmin concentrations with a one-month lag for all reservoirs. Colour-coded by the individual reservoir. Pearson correlation is displayed on the right-hand side of the figure, and variable distribution is available on the diagonal. Correlation significance codes: ‘.’  $p < 0.1$ , ‘\*’  $p < 0.05$ , ‘\*\*’  $p < 0.01$ , ‘\*\*\*’  $p < 0.001$ .





Figure 3.40: Correlogram of monthly changes in nutrients and monthly changes in geosmin concentrations for all reservoirs. Colour-coded by the individual reservoir. Pearson correlation is displayed on the right-hand side of the figure, and variable distribution is available on the diagonal. Correlation significance codes: ‘.’  $p < 0.1$ ,  $p < 0.05$  ‘\*’, ‘\*\*’  $p < 0.01$ , ‘\*\*\*’  $p < 0.001$ .

## 3.4 – Discussion

Out of nine reservoirs, only one (Alaw) had severe geosmin “events”; five other reservoirs (Cefni, Dolwen, Llwyn On, Plas Uchaf and Pontsticill) experienced milder geosmin concentrations that exceeded the “event” level threshold ( $>10 \text{ ng L}^{-1}$ ). For 2-MIB, two reservoirs exceeded the “event” level threshold ( $>5 \text{ ng L}^{-1}$ ), although Pentwyn had higher concentrations than Pontsticill ( $58 \text{ ng L}^{-1}$  and  $7.90 \text{ ng L}^{-1}$ , respectively). Here, the trophic status of each reservoir was assessed by season and year, and the seasonality of geosmin and 2-MIB concentrations was evaluated by the studied year. Seasonality trends in geosmin and 2-MIB differed between reservoirs and in some cases, between seasonal years. There was also evidence of one geosmin event during winter 2019 in Alaw reservoir, indicating that T&O events cannot only be considered during warmer months. Using Chl *a* TSI values as a proxy for biomass suggested T&O production is productivity driven rather than biomass related, highlighted by low Chl *a* TSI values during a winter geosmin event in Alaw. PCA analysis of raw chemistry and physical water parameters confirmed the complexity of the data and potential non-linear relationships between T&O production and some nutrients, namely  $\text{NH}_4^+$ . Monthly change data highlighted a time lag in the monthly increases in  $\text{NH}_4^+$ , TP, and orthophosphate in relation to a monthly reduction in  $\text{NO}_3^-$  in the case of a reservoir experiencing severe geosmin “events” (Alaw). The significance of increases in  $\text{NH}_4^+$ , TP, and Ortho and decreases in  $\text{SO}_4^{2-}$  in association with monthly lagged geosmin concentrations was later confirmed with correlation analysis for the Alaw reservoir. Correlation analysis also revealed the importance of direct negative monthly influences of DSil and  $\text{NO}_3^-$  on geosmin monthly change in other reservoirs experiencing geosmin concentrations above the event level threshold, along with significant monthly influences of  $\text{Fe}^{2+}$ .

### 3.4.1 – Seasonality of T&O compounds in relation to CTSI and TSI indicators

Geosmin and 2-MIB production has typically only been considered problematic during warmer seasons. For example, Westerhoff et al. (2005) found that 2-MIB concentrations increased from spring to late summer in three different

water supply reservoirs, consistent with previous findings (Bruchet, 1999; Yagi, 2005). Although, it has been suggested that T&O events are not necessarily confined to summer months (Dzialowski et al., 2009; Jüttner and Watson, 2007; Wang et al., 2005; Watson et al., 2001). From this study, T&O events occurring in colder months can be reaffirmed by geosmin concentrations during winter 2019 in Alaw reservoir (400 ng L<sup>-1</sup>). CTSI values for Alaw were categorised as eutrophic throughout this study, although the TP TSI for winter 2019 had the second highest value. Increased TP concentrations are in accordance with previous findings that suggest that increasing TP and hence lowering the total nitrogen to total phosphorous (TN:TP) ratio will promote cyanobacterial productivity and dominance in freshwater ecosystems (Espinosa et al., 2021b; Harris et al., 2016; Olsen et al., 2016), and hence geosmin events. In contrast, the Chl a TSI revealed winter 2019 in the Alaw reservoir to have the lowest values giving it a mesotrophic trophic status. A low Chl a TSI suggests that geosmin production was productivity driven instead of biomass driven. The production of geosmin during periods of high productivity may explain why past studies have failed to see a correlation between T&O concentrations and cyanobacterial biomass (Graham et al., 2010). In addition, Xuwei et al. (2019) found that *Oscillatoria* were responsible for T&O production in a 'non-blooming' reservoir in China, which is indicative that high biomass is not always associated with geosmin and 2-MIB concentrations. Another study showed maximal secondary metabolite productivity (amount of secondary metabolite per unit biomass) happened when biomass was lowest (Kim et al., 2018). Slavin (2020) reported that peaks in cyanobacterial biomass did not always correlate with peaks in geosmin or 2-MIB concentrations over time. These findings confirm that T&O production is most likely productivity rather than biomass driven.

Dzialowski et al. (2009) concluded that trophic state alone was a poor predictor of geosmin concentrations because the highest geosmin concentrations were observed in a reservoir experiencing the lowest nutrient and Chl a concentrations. Overall Trophic status using CTSI values was a poor indicator of T&O outcome by season and year; this was made evident by Alwen and Llandegfedd reservoirs gaining eutrophic status in some seasons and years, but not yielding T&O concentrations above the event level threshold. Although, all reservoirs

that experienced geosmin and 2-MIB events occurred when the water column either remained eutrophic or transitioned from a mesotrophic to eutrophic status.

### 3.4.2 – Indirect influences of changes in nutrients on the change in T&O concentrations

PCA from Alaw's reservoir locations revealed sulphate and pH to influence high geosmin concentrations. Higher pH values are likely caused by increased cyanobacterial productivity and is not a direct factor in geosmin production. High photosynthetic activity decreases dissolved CO<sub>2</sub> in the water column, leading to increased pH in the water (Visser et al., 2016), and *Cyanobacteria* are known to stay productive in more alkaline environments (da Silva Brito et al., 2018). Sulphate reduction to sulphide can limit Fe<sup>2+</sup> diffusion rates from anoxic sediments because of insoluble iron sulphide formation (Carignan and Tessier, 1988); availability of Fe<sup>2+</sup> has been proposed to regulate the ability of *Cyanobacteria* to compete with its eukaryotic competitors. A high rate of iron sulphide formation limits the formation of insoluble Fe phosphate (due to sulphide's higher solubility product), in turn permitting higher internal P loading, thus favouring *Cyanobacteria* by altering the TN:TP ratio (Molot et al., 2014).

The non-direct influences of nutrients governed by PCA analysis highlight a non-direct relationship between nutrients and geosmin production, especially in the case of the Alaw reservoir. Similarly, the non-direct influences of nutrients demonstrated in this study follow reports by Slavin (2020) that used Self-Organising maps (SOMs) to compare chemical and physical water parameters with geosmin and 2-MIB concentrations. SOMs can be considered as a non-linear generalisation of PCA, which recognises lags within time series data. This non-direct relationship could be due to the frequency of sampling (monthly); one study demonstrated that a significant change in T&O concentration might decline by 12% within a week (Pochiraju et al., 2021). Thus, monthly data may also miss spikes in nutrients, especially when associated with slurry-spreading activities.

### 3.4.3 – Monthly changes in N and P fractions associated with monthly changes in T&O concentrations

Restrictions of monthly data and the indirect influences nutrients had on T&O concentrations led to looking at monthly changes in T&O concentrations and the monthly changes in nutrients. A distinct relationship existed in the Alaw reservoir, which demonstrated that a month before a significant increase in monthly change in geosmin, there were increases in  $\text{NH}_4^+$ , TP, and orthophosphate with a simultaneous decrease in the monthly change in  $\text{NO}_3^-$ . Spikes in  $\text{NH}_4^+$  are in accordance with previous findings from Perkins et al. (2019) that identified  $\text{NH}_4^+$  as a key trigger in the production of T&O compounds.  $\text{NH}_4^+$  is the most reduced form of nitrogen and requires less cellular energy to assimilate (Flores and Herrero, 2005; Harris et al., 2016). As a result, rates of assimilation of  $\text{NH}_4^+$  are faster than oxidised fractions of nitrogen ( $\text{NO}_3^-$  and  $\text{NO}_2^-$ ) because assimilation of these oxidised nitrogen fractions requires nitrate and/ or nitrite reductase (Collos and Berges, 2003). For example, Saadoun et al. (2001) found maximal geosmin synthesis in *Dolichospermum* (previously named *Anabaena*) to be correlated with high  $\text{NH}_4^+$  concentrations ( $r^2 = 0.89$ ) and low  $\text{NO}_3^-$  concentrations. *Cyanobacteria* assimilate  $\text{NH}_4^+$  more efficiently than  $\text{NO}_3^-$  (Hempel et al., 2018; Su et al., 2019). Although, there is an extensive range in the ability between taxa to uptake and assimilate  $\text{NO}_3^-$  and  $\text{NH}_4^+$  (Glibert et al., 2016), which subsequently affects the productivity of *Cyanobacteria* and hence T&O production.  $\text{NO}_3^-$  assimilation requires a 2-step reduction to  $\text{NH}_4^+$ , which requires cellular energy. High  $\text{NH}_4^+$  relative to  $\text{NO}_3^-$  most likely stimulates greater productivity in *Cyanobacteria* along the MEP pathway (Seto et al., 1996), increasing T&O compound production. Increased  $\text{NH}_4^+$  concentrations stimulating cyanobacterial productivity could also explain why some studies have failed to find a correlation between geosmin/ 2-MIB concentrations and cyanobacterial biomass (Graham et al., 2010; Watson et al., 2008).

Direct correlations between the monthly changes in  $\text{NH}_4^+$ , TP and Ortho reflected an inverse relationship with monthly changes in geosmin concentrations in all reservoirs. Direct monthly changes in  $\text{NH}_4^+$  had a significant inverse relationship, especially in Cefni and Plas Uchaf reservoirs. Due to monthly data, this inverse relationship could represent  $\text{NH}_4^+$  and TP being used up by the *Cyanobacteria* for

energetic costs required in T&O compound production. Direct monthly TP changes were considered significantly negative for all reservoirs experiencing geosmin concentrations above the event level threshold, particularly for the Pontsticill reservoir. A previous study hypothesised that phosphorous is required for cyanobacterial growth and needs to be considered in management, although phosphorous availability may not be the critical trigger for T&O production in *Cyanobacteria* (Slavin, 2020). TP not being a critical trigger for T&O production is in accordance with findings from Howard (2020) that reported low TN:TP ratios favoured cyanobacterial dominance and growth, while high  $\text{NH}_3^+:\text{NO}_3^-$  ratios were associated with increased T&O production. However, direct monthly reductions in  $\text{NO}_3^-$  were only seen to negatively impact monthly changes in geosmin concentrations in the Dolwen reservoir; this confirms the direct inhibition of geosmin production in the Dolwen reservoir that is consistent with previous findings (Domingues et al., 2011; Saadoun et al., 2001), yet conflicting results from other studies indicate that geosmin production can occur in the presence of high  $\text{NO}_3^-$  (Oh et al., 2017; Schrader et al., 2013).

#### 3.4.4 – Time lagged monthly changes in geosmin concentrations in a reservoir experiencing severe geosmin events with monthly changes in N and P fractions

Trends in the Alaw reservoir suggested a month time lapse in the monthly changes in geosmin concentrations in relation to monthly changes in nitrogen and phosphorous fractions. Cross-correlation results from a previous study revealed that T&O concentrations reached their maximum three weeks after increases in Ortho and TP concentrations, and increases in the  $\text{NH}_4^+:\text{NO}_3^-$  ratio occurred one week before T&O events (Slavin, 2020). Espinosa et al. (2021a) predicted that after nutrient enrichment of the water column, especially with phosphorous, geosmin would be released into the water column between seven to 15 days later. The residence time of the water column plays a significant role in the productivity and hence the T&O outcome of *Cyanobacteria* (Clercic and Druschel, 2019), suggesting that *Cyanobacteria* would undergo a time lag depending on the reservoirs residence time before the production of T&O compounds. Introducing a time lag in the monthly changes in geosmin concentrations in the Alaw reservoir revealed a significant

positive relationship with monthly changes in  $\text{NH}_4^+$ , TP, and Ortho with the most profound effects on low and high geosmin concentrations. Results from this study support the findings that  $\text{NH}_4^+$  is a crucial trigger responsible for T&O compound production (Perkins et al., 2019) and highlight the need to investigate the importance of  $\text{NH}_4^+$  in relation to  $\text{NO}_3^-$  further. Higher concentrations of T&O compounds have been associated with lower  $\text{NH}_4^+:\text{NO}_3^-$  ratios (Harris et al., 2016), because if  $\text{NH}_4^+$  is present at elevated concentrations T&O production would be inhibited if  $\text{NO}_3^-$  was present in higher concentrations. In addition, increases in the monthly changes in TP and Ortho would affect the TN:TP ratio. A TN:TP threshold of 16:1 was defined by Redfield (1958), but Smith (1983) suggested that *Cyanobacteria* dominance was likely at ratios <29. The TN:TP ratio has been used to monitor and predict T&O events (Keene, 2002). Although, in eutrophic waters where nutrients are abundant and phytoplankton are more light-limited, the TN:TP ratio is less predictive of *Cyanobacterial* productivity and hence T&O outcome (Huisman et al., 2004). A previous study demonstrated that orthophosphate uptake in bloom-forming *Cyanobacteria* is instant (commencing within 15-25 minutes) and implies that the elevated uptake rates promote higher productivity rates (Aubriot and Bonilla, 2012). In this study, fluctuations in orthophosphate concentrations may have been missed due to monthly sampling.

### 3.4.5 – Monthly changes in sulphate, dissolved iron and dissolved reactive silicate and implications on monthly changes in T&O concentrations

*Cyanobacteria* have been reported to cause T&O issues in oligotrophic waters having very low phosphorous inputs (Jahnichen et al., 2011). Something other than P kinetics is implicated in the production of T&O compounds. It is known that *Cyanobacteria* have a higher Fe requirement than their eukaryotic counterparts, with nitrogen fixation imposing a high Fe demand (Molot et al., 2014). Previous literature suggests that the availability of  $\text{Fe}^{2+}$  gives *Cyanobacteria* a competitive advantage over other algae taxa, and even ferric iron ( $\text{Fe}^{3+}$ ) can be transported and used through siderophores that some *Cyanobacteria* acquire (Wilhelm and Trick, 1994). In addition, low dissolved oxygen near the sediment-water interface can promote the reduction of iron ( $\text{Fe}^{3+}$  to  $\text{Fe}^{2+}$ ) and release iron-bound phosphate into the labile pool

(Bostrom et al., 1988). Internal phosphorus loading from the sediments can alter the TN:TP ratio, favouring *Cyanobacterial* productivity and hence T&O production. Lagged monthly changes in geosmin concentrations revealed an inversely related relationship between  $\text{Fe}^{2+}$  for the Plas Uchaf reservoir. Similarly, a direct negative relationship existed between monthly changes in geosmin concentrations with monthly changes in  $\text{Fe}^{2+}$  in the Cefni reservoir. Indicating that  $\text{Fe}^{2+}$  was being utilised whilst geosmin and 2-MIB concentrations were elevated. In contrast, an increase in monthly changes in  $\text{Fe}^{2+}$  was positively related to increases in the monthly changes in geosmin concentrations but only in the Dolwen reservoir. This could be indicative of an increase in the reduction of  $\text{Fe}^{3+}$  liberating free labile phosphate for *Cyanobacterial* secondary metabolite production. Previous findings have demonstrated that the uptake of phosphate in bloom-forming *Cyanobacteria* is rapid (commencing within 15-25 minutes) and suggests that the elevated rates of uptake promote higher productivity rates (Aubriot and Bonilla, 2012). Thus, the importance of monthly changes in phosphate (orthophosphate) would not be picked up in this study.

In addition to  $\text{Fe}^{2+}$ , sulphate reduction to sulphide can limit  $\text{Fe}^{2+}$  diffusion rates from anoxic sediments because of the insoluble iron sulphide formation (Carignan and Tessier, 1988). Hence, lakes with a high sulphide formation rate preventing  $\text{Fe}^{2+}$  diffusion should not experience heightened cyanobacterial productivity. A higher rate of iron sulphide formation limits the formation of insoluble Fe phosphate (due to sulphides' higher solubility product), in turn permitting higher internal loading (Molot et al., 2014). For the Alaw reservoir, a significant negative relationship existed between the lagged monthly changes in geosmin concentration and the monthly change in sulphate. Although, this relationship was deemed positive in the case of the Plas Uchaf reservoir. No studies to date have evaluated monthly changes in sulphate and  $\text{Fe}^{2+}$  concerning redox potential indicators for T&O production.

As silicate consumption is related to diatom growth and formation (Wasmund et al., 2013), monthly changes in dissolved reactive silicate (DSil) were analysed to assess community composition change from *Cyanobacteria* to diatoms in relation to T&O production. Interestingly, monthly changes in DSil were significantly negatively associated with monthly changes in geosmin for all reservoirs, especially in the Cefni reservoir. This indicates that geosmin production occurred when DSil was used in



diatom formation. Olsen et al. (2016) found a stronger correlation between 2-MIB concentrations and diatom biovolume and suggested that there may be a link between T&O production and diatoms. Later studies confirmed that the direct or indirect influences of diatoms on 2-MIB and geosmin production by *Cyanobacteria* should not be ignored (Olsen et al., 2017). In addition, other studies have found a connection between *Synedra* spp. and 2-MIB production (Izaguirre and Taylor, 1998; Schrader et al., 2011; Sugiura et al., 2004, 1998). Although *Synedra* has not been shown to produce 2-MIB directly, it has been identified as a substrate for the growth of *Streptomyces* spp., a 2-MIB-producing actinomycete (Sugiura et al., 2004). It is known that *Cyanobacterial* production of the cyanotoxin, cylindrospermopsin, causes the secretion of alkaline phosphatase (APase) by other phytoplankton, thus increasing the amount of inorganic phosphate available to the *Cyanobacteria*, allowing it to outcompete other taxa in environments with limited inorganic phosphate (Bar-Yosef et al., 2010). Interactions between *Cyanobacteria* and diatoms may be beneficial for *Cyanobacterial* growth and productivity and hence attributing to T&O events. This study warrants investigating the community dynamics of bacterial and algal communities to try to better understand relationships that may exist.

### 3.4.6 – Limitations of the study

*Cyanobacteria* can be planktonic in the water column or grow as benthic populations in the sediments. T&O events produced by planktonic *Cyanobacteria* are well documented compared to those originating from benthic *Cyanobacteria* (Taylor et al., 2006). However, benthic *Cyanobacteria* are becoming recognised as significant sources of unexplained odours (Chen et al., 2010). Burlingame et al. (1986) reported elevated levels of geosmin in Philadelphia during 1985 and concluded that the source of one event was a localised bed of benthic *Cyanobacteria*. Both geosmin and 2-MIB have been linked to T&O events originating from benthic *Cyanobacteria* (Izaguirre and Taylor, 2007; Sugiura et al., 1998; Watson and Ridal, 2004). Chen et al. (2010) predicted that odorant emission rates based on column surface area for geosmin were 18–190 ng h<sup>-1</sup> cm<sup>-2</sup> and 4.2–4.4 ng h<sup>-1</sup> cm<sup>-2</sup> for 2-MIB. Chen et al. (2010) identified a substantial emission source of T&O compounds that could be volatilising from benthic origin. In this study, only open water samples were taken, and no bottom sediment samples were taken, which may

have missed potential relationships in the sedimental regions giving rise to T&O production. It is known that increases in cyanobacterial growth and hence productivity can occur within hours of increased nutrient availability (Örnólfssdóttir et al., 2004). Therefore, fluctuations and spikes in nutrients may be missed due to the frequency of water sample measurements taken in this study (monthly). Similarly, Paerl et al. (2022) demonstrated that a large change in geosmin concentration can occur on a week to week basis during spring – summer, and Pochiraju et al. (2021) reported that geosmin concentrations may decline by 12% within a week. With the rapid uptake of nutrients and the quick volatilization of the T&O compounds, monthly monitoring must be biweekly to limit the fluctuations in both nutrients and T&O compounds.

This study would benefit from adding molecular data to better understand the relationships between diatoms and Cyanobacteria and how community composition reacts to nutrient ratios, particularly TN:TP. In two reservoirs where efforts were made to reduce agricultural sources of N and P pollution, it was followed by ameliorating the water quality (Jeppesen et al., 2005; Shatwell and Köhler, 2019). Long-term (37 years) surveys highlighted changes in the cyanobacterial community composition towards diazotrophic (heterocyst-forming *Cyanobacteria*) species as a consequence of N decrease, with P remaining constant in one reservoir (Schindler et al., 2008). In the other reservoir, reduction in both N and P decreased the proportion of N<sub>2</sub> fixers in the phytoplankton community (Shatwell and Köhler, 2019). Knowing the community composition of *Cyanobacteria* present will help understand what triggers are causing T&O events.

### 3.5 – Conclusions

In this study, it can be confirmed that the seasonality of T&O production cannot necessarily be confined to warmer months. During winter 2019, when a geosmin event occurred in a reservoir experiencing severe events (Alaw), the TP TSI indicator had the second highest score favouring cyanobacterial productivity. During this period, the Chl *a* indicator had the lowest value compared to all studied seasons and years; this indicates that T&O production is productivity rather than biomass related. Overall, CTSI was a poor indicator of T&O outcome by season and

year, as shown by seasons gaining eutrophic status in two reservoirs (Alwen and Llandegfedd) but not yielding T&O compounds above the event level threshold. Monthly lagged changes in geosmin concentrations in the Alaw reservoir revealed significant positive influences from monthly changes in  $\text{NH}_4^+$ , TP, and orthophosphate, whilst monthly changes in  $\text{NO}_3^-$  had non-significant negative effects.  $\text{NH}_4^+$  was identified as a critical trigger in monthly lagged geosmin production in a reservoir experiencing severe geosmin events (Alaw).  $\text{NH}_4^+$  should be considered in relation to reductions in  $\text{NO}_3^-$  due to the inhibitory effects of oxidised nitrogen on T&O production; further research should consider the  $\text{NH}_4^+:\text{NO}_3^-$  ratio. In conjunction with the  $\text{NH}_4^+:\text{NO}_3^-$  ratio, the TN:TP ratio should also be considered in future work, as monthly changes in TP were positively related to monthly time-lagged changes in geosmin concentrations. Monthly changes in dissolved iron concentrations were negatively associated with monthly changes in the Cefni reservoir and lagged geosmin in the Plas Uchaf reservoir. Although this relationship was unclear as direct monthly changes in dissolved iron positively influenced monthly changes in geosmin concentrations in the Dolwen reservoir. In addition, monthly changes in sulphate had a negative influence on lagged monthly changes in geosmin in the Alaw reservoir yet had positive influences in the Plas Uchaf reservoir. The role of ferrous ( $\text{Fe}^{2+}$ ) and ferric iron ( $\text{Fe}^{3+}$ ) and sulphate reduction potential with T&O production needs further research. A direct inverse relationship existed between monthly changes in dissolved reactive silicate and monthly changes in geosmin, especially for the Cefni reservoir. The negative association between dissolved reactive silicate and geosmin indicates a link between diatom formation and T&O production. Further work needs to be done in evaluating the bacterial and algal community composition and how this is structured with changing  $\text{NH}_4^+:\text{NO}_3^-$  and TN:TP ratios. Community analysis could also aid the entanglement of the potential relationship between *Cyanobacteria* and diatoms.

# Chapter 4: Negatively co-occurring taxa with *Cyanobacteria* inducing T&O events: modelling T&O using indicative taxa



## 4.1 – Introduction

Historically, cyanobacterial blooms were considered the leading cause of the production of nuisance T&O compounds (Adams et al., 2021; Bruder et al., 2014; Hayes and Burch, 1989; Newton et al., 2015). However, periods of high productivity have been associated with high geosmin and 2-MIB production in *Cyanobacteria* (Zimmerman et al., 1995). The production of geosmin and 2-MIB during periods of high productivity could explain how some studies fail to correlate geosmin and 2-MIB concentrations with cyanobacterial biomass (Graham et al., 2010; Watson et al., 2008). For example, Kim et al. (2018) found that the amount of secondary metabolite per unit biomass was maximal when biomass was lowest, at 15°C. Similarly, there have been rises in the frequency and magnitude of winter T&O events (Dzialowski et al., 2009), indicating that T&O events cannot be limited to summer months when cyanobacterial biomass is at its highest. Along with *Cyanobacteria*, the bacterioplankton community composition of freshwater environments is typically dominated by *Actinobacteria* (Keshri et al., 2018; Tanaka et al., 2017), *Bacteroidetes* (Schmidt et al., 2016), or *Proteobacteria* (Olapade, 2018; Salmaso et al., 2018) bacterioplankton. Extensive uses of molecular tools in exploring microbial diversity have enabled the characterisation of novel organisms capable of producing T&O compounds (Churro et al., 2020; Otten et al., 2016). However, no studies to date have reported interactions between the bacterioplankton and phytoplankton communities and their relation to the production of T&O events. To better predict the onset of a T&O event, a better understanding of the phytoplankton and bacterioplankton community ecology is required to learn what drives cyanobacterial T&O production.

Numerous heterotrophic bacteria are proven to be directly associated with *Cyanobacteria*, and studies have shown that the associated bacteria can have an important impact on cyanobacterial growth (Cai et al., 2014). Cyanobacterial-heterotrophic bacterial associations are commonly observed inside cyanobacterial colonies/aggregates and within extracellular polymers outside the *Cyanobacteria* cell walls – collectively, these microhabitats constitute the cyanobacterial phycosphere. Within the phycosphere, bacteria and algae can live freely, attached to the cyanobacterial surface or extracellular products (Jasti et al., 2005). *Cyanobacteria*

excrete an abundance of extracellular organic matter that provides energy for associated taxa (Worm and Søndergaard, 1998). In turn, associated partners in the phycosphere may provide CO<sub>2</sub>, nitrogen, phosphorous, sulphur and trace elements to the *Cyanobacteria* (Havens, 2008). Reports have shown that *Cyanobacteria* form relationships within the phycosphere; one common obligate mutualism exists between heterocyst-forming *Cyanobacteria* and diatoms whereby the *Cyanobacteria* provide fixed nitrogen to the diatom in exchange for amino acids and organic carbon (Foster et al., 2011; Hilton et al., 2013; Thompson et al., 2012). The hepatotoxin cylindrospermopsin produced by certain *Cyanobacteria*, e.g. *Aphanizomenon ovalisporum*, promote phosphorous supply by inducing APase secretion by other eukaryotic phytoplankton that the *Cyanobacteria* use for growth under phosphorous deficiency (Bar-Yosef et al., 2010). Bacterial phylotypes belonging to *Proteobacteria*, *Bacteroidetes* and *Actinobacteria* in the bacterial community are often associated with *Microcystis* spp. (Cai et al., 2014; Parveen et al., 2013; Shi et al., 2012, 2009). The associated bacterial flora is said to acquire energy from organic substrates excreted by *Microcystis* and play a vital role in stimulating cyanobacterial growth and colony formation of *Microcystis* (Shen et al., 2011; Xie et al., 2016). It has been hypothesised that phycosphere-associated taxa may detoxify cyanobacterial extracellular metal (Fe) chelates that may (under anoxic conditions) be autotoxic to the host (Paerl and Pinckney, 1996). In addition, associated bacteria may help lower oxygen tension near cyanobacterial cells and filaments where oxygen-sensitive biochemical processes (photosynthesis, nitrogen fixation) occur (Paerl and Kellar, 1978). Recently, studies have identified taxa in the cyanobacterial phycosphere that participated in forming phosphorous cycling co-pathways as their functional links to *Cyanobacteria* (Shi et al., 2022). The presence of these functional taxa in the cyanobacterial phycosphere may accelerate nutrient cycling and facilitate cyanobacterial growth and productivity, hence T&O production. To date, no studies have evaluated the cyanobacterial phycosphere concerning T&O events.

Identifying the cyanobacterial phycosphere can identify implicated organisms involved in T&O production. Relationships between *Cyanobacteria* and other taxa can be determined to see if and how they influence T&O production — in recent years, applying correlation-based network analysis to explore microbial communities' co-occurrence and co-exclusion patterns has been successful (Barberán et al., 2012;

Eiler et al., 2012; Ju and Zhang, 2015). This approach reveals potential interspecies interactions and examines co-occurrence patterns of residents based on relative abundances of the marker gene (16S rRNA and *rbcL*) amplicon community profiling data (Faust and Raes, 2012). These co-occurrence patterns can directly or indirectly reflect growth inhibition or facilitation outcomes of one or more negative or positive interactions, respectively. Although on a larger scale, nitrogen and phosphorous fractions entering waterways play a paramount role in cyanobacterial productivity (Jarosiewicz et al., 2012), nutrient recycling on a micro-scale within the cyanobacterial phycosphere may help explain fluctuations in T&O concentrations. Previous studies show that high molecular weight organic compounds in a cyanobacterial phycosphere are degraded by bacterial communities from large-size aggregates resulting in small-size aggregates (Cai et al., 2014). This study concluded that bacteria within the phycosphere could efficiently provide nutrients and trace elements to the *Cyanobacteria* by recycling the organic matter provided by the *Cyanobacteria*. Louati et al. (2015) found that the proportion of *Nitrosomonadales* increased during a *Microcystis* bloom; this group encompass bacteria involved in nitrification. This specific functional N-recycling group had already been highlighted in several studies but only associated with nitrogen-fixing *Cyanobacteria*, marine blooms in the Baltic Sea, where no nitrification activity could be detected (Hietanen et al., 2002; Tuomainen et al., 2003). There is a clear need for further investigation into the nutrient recycling processes that occur within a cyanobacterial phycosphere and how taxa associated with the phycosphere can influence cyanobacterial productivity and hence, T&O production.

Here, we employed Next Generation Sequencing using the 16S rRNA and *rbcL* genes to examine community profiles of both bacterial and algal communities from nine reservoirs around Wales, U.K.. Bacterial and algal community compositions in each studied reservoir were used as an informative response to nutrient influxes and environmental variables occurring in the water bodies. Firstly, exploring temporal (seasonal) and spatial differences in bacterial phyla over the study period was performed to reveal the window for cyanobacterial growth. Co-occurrence analysis was performed to assess significantly co-occurring bacterial and eukaryotic taxa that could be associated in the cyanobacterial phycosphere (the “cyanosphere” hereafter), potentially attributing to T&O production. Significantly co-



occurring taxa with *Cyanobacteria* taxa were further analysed through ordination plots evaluating if the significantly co-occurring 16S rRNA and *rbcL* taxa could attribute to high geosmin and 2-MIB concentrations. To concisely pinpoint indicative genera associated with assigned T&O concentration levels, an ensemble learning method was applied to target indicative taxa associated with high geosmin and 2-MIB concentrations. This strategy is the first to date to include *rbcL* data with 16S rRNA community data and to apply machine learning to identify indicative genera for T&O events whilst considering possible biogeochemical recycling occurring within the “cyanosphere”.

## 4.2 – Materials and Methods

### 4.2.1 – Sample collection

Samples were collected from sample site locations within reservoirs as illustrated in Chapter 2 – Materials and Methods Sections 2.1.1 – 2.1.7, as described in Chapter 2 – Materials and Methods section 2.1. The collection of samples was disrupted during the sampling period, as depicted in Chapter 2 – Materials and Methods Section 2.1.8.

### 4.2.2 – eDNA extraction and community analysis using the 16S rRNA and *rbcL* genes

Water was filtered, and eDNA was extracted according to Chapter 2 – Materials and Methods, Section 2.4. Extracted eDNA then underwent an initial PCR (Chapter 2 – Materials and Methods, Section 2.4.1), post-PCR clean-up (Chapter 2 – Materials and Methods, Section 2.4.2), a secondary PCR (Chapter 2 – Materials and Methods, Section 2.4.3) and normalisation (Chapter 2 – Materials and Methods, Section 2.4.4). Post normalisation, samples were pooled together to form a library that was used for Next Generation Sequencing (Chapter 2 – Materials and Methods, Section 2.4.5). Raw sequences were then put through a bioinformatics pipeline according to Chapter 2 – Material and Methods, Section 2.4.6.

### 4.2.3 – Water chemistry and physical analysis

The water samples' chemistry and physical properties were determined through analysis described in Chapter 2 – Materials and Methods, Section 2.3. The determination of geosmin and 2-MIB concentrations are detailed in Section 2.31 and nutrient analysis in Section 2.3.2.

### 4.2.4 – Categorising T&O concentration levels

The categorisation of T&O concentrations into low, medium, and high levels is explained in Chapter 2 – Materials and Methods, Section 2.2.

### 4.2.5 – Spatial and temporal changes in 16S rRNA phyla community composition

Normalisation was required to convert raw read counts into informative measurements to obtain the proportion of each ASV. The most common normalisation procedure consists in dividing by the total number of reads to obtain the proportion of each ASV (Zemb et al., 2020). This method creates a link between the ASVs (as the sum is constant) and converts each abundance to a ratio providing relative abundance, thereby introducing ambiguity to interpret an increase in the relative abundance of an ASV as an enrichment of this ASV. This method was applied to the raw count data and multiplied by 100 to get proportional abundance percentages for all ASVs for 16S rRNA.

Exploratory bar graphs were created for each reservoir through 'phyloseq' and 'ggplot2' (Wickham, 2016) to evaluate cyanobacterial seasonal succession between years, facet wrapped by site locations to determine any spatial pattern similarities or differences. Taxa within each reservoir's phyloseq 16S rRNA object were agglomerated at the phylum level, and the top 20 occurring phyla were obtained by calculating the total phyla present and taking the 20 highest phyla recorded. The top 20 phyla in each reservoir were plotted using 'ggplot2' in a bar graph against season and year and facet wrapped by reservoir location to assess dominant phyla seasonally and spatially.

#### 4.2.6 – Co-occurrence of 16S rRNA and *rbcL* genera and network visualisation

To assess the co-occurrence between 16S rRNA and *rbcL* genera, firstly, 16S rRNA and *rbcL* datasets were made proportional separately using the total number of reads and calculating the proportional percentage of genera present, as discussed in Section 4.2.5. The two datasets were subsequently merged using the sample's unique identifier and transformed into binomial data, with presence (1) being recorded as a genus proportion  $\geq 1\%$  and absence (0)  $< 1\%$ . The binomial transformation was applied to these data to allow 16S rRNA sequence data to be compared to *rbcL* sequence data to account for any community biases between datasets. The dataset was then subsetted by the individual reservoir, and genera that had a total of 0 in each reservoir's dataset were removed. The total genera present at a  $\geq 1\%$  proportion per reservoir is detailed below in Table 4.1.

Co-occurrence analysis was performed using the R package 'cooccur' (Veech, 2013). Using the probabilistic co-occurrence model, the `cooccur()` function calculated  $p$ -values associated with pairwise co-occurrences. This algorithm calculates the observed and expected frequencies of co-occurrence between each pair of present taxa in each studied reservoir over the study period. Negative relationships were established by the  $p$ -value  $< 0.01$  associated with the probability that two genera would co-occur at a frequency less than the noted number of co-occurrence sites if the two species were distributed randomly of one another. Positive values were determined by the  $p$ -value  $< 0.01$  associated with the probability of co-occurrence at a frequency greater than the observed frequency (Griffith et al., 2016).

Co-occurring 16S rRNA and *rbcL* genera were filtered to reveal only significant positive and negative relationships using  $p$ -values listed for each reservoir in Table 4.2. Networks were subsequently established for each reservoir using the R package 'visNetwork' (Almende et al., 2019). The nodes represented the genera present, and edges were constructed using the positive and negative associated  $p$ -values established from the 'cooccur' package. Nodes for cyanobacterial genera were highlighted in green, and T&O degrading genera were highlighted in purple for

a clear distinction. Edges illustrated negative relationships with a black dashed line and positive relationships with a solid black line. Note that in some reservoirs, the total taxa present in the network reflected sample size, and to achieve the network, the  $p$ -value has been adjusted to allow significant taxa to become identifiable.

Table 4.1: Total 16S rRNA and *rbcL* taxa present at  $\geq 1\%$  proportion.

| Reservoir   | Total Taxa Present |
|-------------|--------------------|
| Alaw        | 200                |
| Alwen       | 105                |
| Cefni       | 185                |
| Dolwen      | 119                |
| Llandegfedd | 165                |
| Llwyn On    | 168                |
| Pentwyn     | 149                |
| Plas Uchaf  | 154                |
| Pontsticill | 133                |

Table 4.2:  $p$ -values used for filtering significant positive and negative co-occurring 16S rRNA and *rbcL* genera and the number of samples used for analysis.

| Reservoir   | $p$ -value used | n  |
|-------------|-----------------|----|
| Alaw        | 0.01            | 33 |
| Alwen       | 0.05            | 7  |
| Cefni       | 0.01            | 32 |
| Dolwen      | 0.1             | 6  |
| Llandegfedd | 0.01            | 23 |
| Llwyn On    | 0.01            | 19 |
| Pentwyn     | 0.01            | 30 |
| Plas Uchaf  | 0.01            | 22 |
| Pontsticill | 0.01            | 17 |

#### 4.2.7 – Non-metric Multi-dimensional Scaling of 16S rRNA and *rbcL* communities

Non-metric Multi-Dimensional Scaling (NMDS) was applied to 16S rRNA and *rbcL* sequencing data for individual reservoir's reservoir locations that experienced T&O concentrations above the defined "event" level classification (see Section 4.2.4). Both gene datasets were normalised to the total number of reads for 16S rRNA and *rbcL* as detailed in Section 4.2.5. Ordination was achieved through the 'Vegan' 2.6-2 (Oksanen, 2013) R package using the metaMDS function with a bray-curtis distance and plotted using 'ggplot2' (Wickham, 2016). Each point was shaped by site location and coloured by geosmin or 2-MIB concentration categories (see Section 4.2.4).

The envfit function was later used on each reservoir's NMDS output to identify variables driving the pattern. Vectors from the envfit function were added to each reservoir's NMDS plots. For 16S rRNA the first NMDS vector plot (A) represents significant co-occurring cyanobacterial genera identified in Section 4.2.6. The second NMDS vector plot (B) illustrates significant co-occurring geosmin and 2-MIB degrading genera identified in Section 4.2.6. The last NMDS vector plot (C) illustrates environmental variables of interest (MIB = 2-MIB concentrations ng L<sup>-1</sup>, GEO = geosmin concentrations ng L<sup>-1</sup>, SO4 = Sulphate mg L<sup>-1</sup>, TIN.TP = total inorganic nitrogen to total phosphorous ratio, Ortho = orthophosphate ng L<sup>-1</sup>, NH4.NO3 = ammonium to nitrate ratio, Dfe = dissolved iron mg L<sup>-1</sup>, DMn = dissolved manganese mg L<sup>-1</sup>, Sil = Dissolved reactive silicate mg L<sup>-1</sup>, TOC = total organic carbon mg L<sup>-1</sup>, pH, Con = conductivity µS cm<sup>-1</sup> at 20°C and Turb = turbidity NTU).

For *rbcL* significant genera were identified by using the *p*-values obtained through NMDS analysis for the first NMDS vector plot (A). The second plot contained significantly co-occurring algae with *Cyanobacteria* and geosmin and 2-MIB degrading taxa (B) (identified through using Section 4.2.6) were used for the second plot. Environmental variables of interest (as listed above for 16S rRNA) were used for the final *rbcL* NMDS plot (C).

The NMDS plots for 16S rRNA and *rbcL* for each reservoir were compiled together using the R package 'ggpubr' (<https://github.com/kassambara/ggpubr>).

#### 4.2.8 – Random Forest to identify indicator genera from significant co-occurring organisms in reservoirs experiencing T&O “events”

Random forests (RFs) were first devised by Breiman (2001) and were employed in this study to identify significant co-occurring genera (as identified through Section 4.2.6) from reservoirs experiencing T&O events that were indicative of high T&O concentrations (geosmin:  $>20 \text{ ng L}^{-1}$  and 2-MIB:  $>10 \text{ ng L}^{-1}$ ). The R package ‘randomForest’ was utilised (<http://www.r-project.org>).

First, settings were optimised for each reservoir’s significantly co-occurring genera. The final number of trees was selected by running an RF with a large number of trees (all three reservoirs – 100000) and plotting the outcome, choosing a point on the x-axis where the error had reached its asymptote. The number of variables ( $m$ ) used was optimised for the building trees by running a RF Cross-Validation for the feature selection function. This function showed the cross-validated prediction performance of models with a sequentially reduced number of predictors (ranked by variable importance) via a nested cross-validation procedure. After refining the number of trees (Alaw – 5000, Llwyn On – 20000, and Pentwyn – 20000) and variables (Alaw – 28, Llwyn On – 44, and Pentwyn – 62) in the RF, an ordination plot was performed using the ‘Vegan’ package (Oksanen, 2013) to reveal trends in T&O levels highlighted through the RF analysis.

To identify indicative organisms relating to T&O levels, a subset of the top 20 most important taxa ( $m$ ) was taken by sorting taxa in the rank of descending order of Mean Decrease in Accuracy (MDA). This dataset was then taken, and a final RF was performed, optimising the number of trees (all three reservoirs – 20000) and assessing the effectiveness of using the 20 variables as described above. A final ordination plot was then constructed.

## 4.3 – Results

### 4.3.1 – 16S rRNA seasonal succession of phyla abundance between years

The seasonal succession of 16S rRNA phyla was assessed spatially per year in each reservoir. Here, we present data from the Alaw reservoir as an example; other reservoirs showed similar trends with the highest cyanobacterial abundance confined to the summer and autumn seasons (Appendix 4.1 – 4.8).

16S rRNA phyla in the Alaw reservoir were mainly dominated by *Proteobacteria*, *Cyanobacteria*, *Bacteroidota*, *Verrucomicrobiota*, and *Actinobacteriota* throughout the study period. Overall abundance was highest in sites during autumn 2019 except for a low abundance recorded in the Res-2 location (Figure 4.1). The season with the second highest abundance occurred during summer and was consistent during both years across all sites, except for summer 2020 in the Res-5 location which had a similar abundance to autumn 2019. The lowest phyla abundance across sites in Alaw occurred in winter and was consistently low during both years. *Cyanobacteria* abundance was highest during both summer seasons and during autumn 2019 at the Res-5 location. *Proteobacteria* abundance followed a similar pattern to *Bacteroidota* abundance; both were generally highest during autumn 2019 except for the Res-2 location, and lowest during winter for both years.

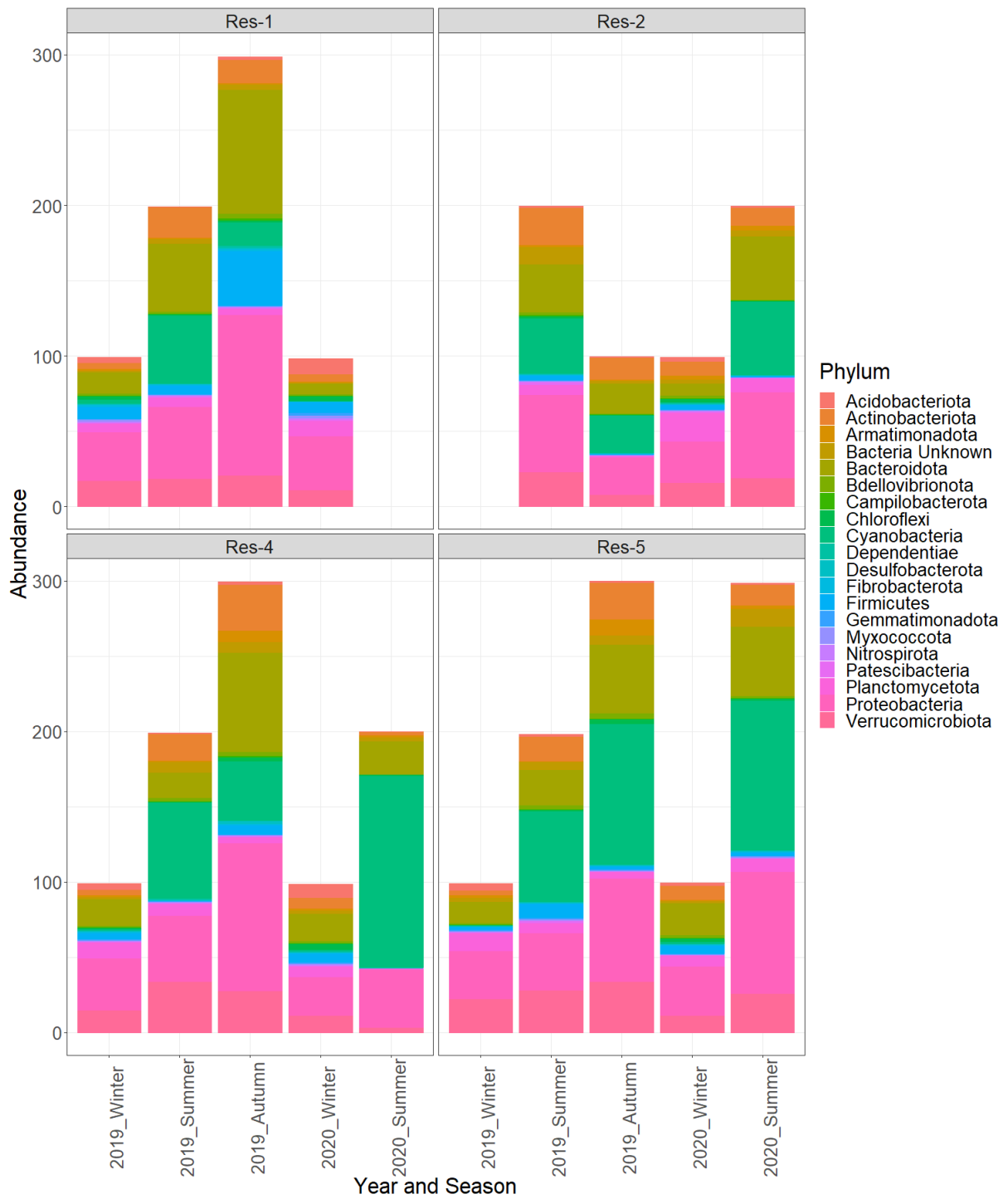


Figure 4.1: Alaw top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.



### 4.3.2 – Binomial co-occurrence analysis of 16S rRNA and *rbcL* communities

Binomial 16S rRNA and *rbcL* co-occurrence networks were established for all reservoirs and are graphically illustrated online through RPubS by RStudio via links provided in Table 4.3. Additional figures displaying the cyanospheres captured from the online networks from each reservoir are included below for ease of reading (Figures 4.2 – 4.28).

Table 4.3: Online binomial 16S rRNA and *rbcL* networks for each reservoir and the *p*-value used to determine negative and positive relationships between genera.

| Reservoir   | <i>p</i> -value used | Chrome file   |
|-------------|----------------------|---|
| Alaw        | 0.01                 | <a href="https://rpubs.com/ASHooper/978472">https://rpubs.com/ASHooper/978472</a> |
| Alwen       | 0.05                 | <a href="https://rpubs.com/ASHooper/978480">https://rpubs.com/ASHooper/978480</a> |
| Cefni       | 0.01                 | <a href="https://rpubs.com/ASHooper/978493">https://rpubs.com/ASHooper/978493</a> |
| Dolwen      | 0.1                  | <a href="https://rpubs.com/ASHooper/978497">https://rpubs.com/ASHooper/978497</a> |
| Llandegfedd | 0.01                 | <a href="https://rpubs.com/ASHooper/978520">https://rpubs.com/ASHooper/978520</a> |
| Llwyn On    | 0.01                 | <a href="https://rpubs.com/ASHooper/978527">https://rpubs.com/ASHooper/978527</a> |
| Pentwyn     | 0.01                 | <a href="https://rpubs.com/ASHooper/978532">https://rpubs.com/ASHooper/978532</a> |
| Plas Uchaf  | 0.01                 | <a href="https://rpubs.com/ASHooper/978536">https://rpubs.com/ASHooper/978536</a> |
| Pontsticill | 0.01                 | <a href="https://rpubs.com/ASHooper/978544">https://rpubs.com/ASHooper/978544</a> |

#### 4.3.2.1 – Alaw

In Alaw reservoir, 199 pairwise comparisons yielded statistically significant co-occurrence ( $<0.01$ ); 29 were negatively associated, and 170 co-occurring taxa were positively associated. The network can be accessed online via the link contained within Table 4.3. In general, there were six networks tightly compacted into the overall network; five of these networks were comprised of between two – four nodes and did not contain any cyanobacterial taxa. The largest network contained the majority of nodes (186) and all four significantly co-occurring cyanobacterial taxa, showing overlaps in each cyanosphere. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.4.

Figure 4.2 illustrates the connectivity within co-occurring taxa in the *Aphanizomenon* NIES81 cyanosphere. Within this cyanosphere, 10 significant co-occurring taxa negatively influenced *Aphanizomenon* NIES81, and only one genus, *Fluviicola*, had a positive association. Figure 4.3 depicts the connectivity within the *Cyanobium* PCC-6307 cyanosphere. Four taxa can be seen to be positively associated with *Cyanobium* PCC-6307, one of which is another *Cyanobacteria*, *Snowella* OTU37S04. For *Snowella* OTU37S04, 10 co-occurring taxa can be found within this cyanosphere (Figure 4.4), five positively co-occurring and five negatively co-occurring. The *Nostocaceae* family had three positively co-occurring nodes, as displayed in Figure 4.5.

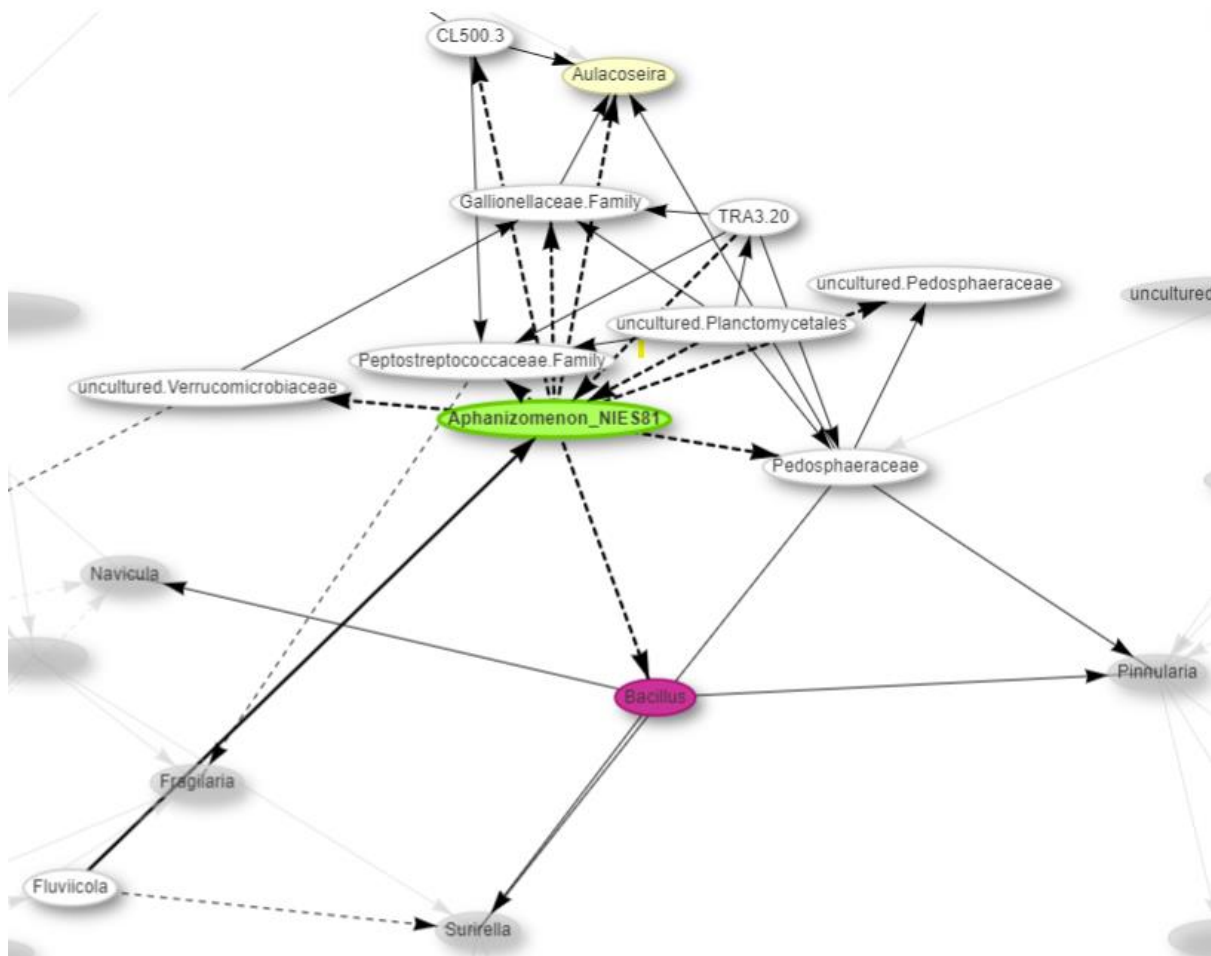


Figure 4.2: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Aphanizomenon* NIES81 node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978472>.

Table 4.4: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Alaw reservoir.

| <i>Cyanobacteria</i>           | Co-occurring taxon   | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|--------------------------------|--|-------------------|--------------------|---|
| <i>Aphanizomenon</i><br>NIES81 | Uncultured <i>Verrucomicrobiaceae</i>                      | 16S               | -ve                | Associated with rhizospheres and endospheres found in plant roots and soil. Implicated in root carbon metabolism. Has also been connected to <i>Lyngbya</i> (non-heterocyst-forming) blooms. (Newitt, 2020; Rajaneesh et al., 2020)   |
|                                | <i>Pedosphaeraceae</i> / uncultured <i>Pedosphaeraceae</i> | 16S               | -ve                | Previously found to be a keystone species in rhizospheres found in soils. They play an important role in sediment nutrient circulation with high metal resistance potentials. (Yuan et al., 2022)   |
|                                | <i>CL500-3</i>   | 16S               | -ve                | Constitutes abundant bacterioplankton groups characteristic for oxygenated hypolimnion waters. (Okazaki et al., 2017)   |
|                                | <i>Bacillus</i>  | 16S               | -ve                | Shown to have an antagonistic relationship with geosmin-producing strains of <i>Streptomyces</i> , inhibiting growth. Also, a well-known geosmin and 2-MIB degrader. (Ma et al., 2015; Zhi et al., 2016)  |
|                                | <i>Gallionellaceae</i> Family                              | 16S               | -ve                | Chemolithoautotrophic and neutrophilic ferrous iron-oxidising bacteria secrete extracellular fibres that attract iron hydroxides and many trace metals. (Hallbeck and Pederson, 2014)   |
|                                | <i>Peptostreptococcaceae</i> Family                        | 16S               | -ve                | Indicator of nutrient enrichment in urban lakes. Also, a common commensal bacterial taxon is thought to play a role in maintaining gut homeostasis in fish. (Numberger et al., 2022)  |
|                                | <i>Aulacoseira</i>   | <i>rbcL</i>       | -ve                | Centric diatoms belonging to more enriched waters with sensitivities to pH rise and silicon depletion. Grouped into Group B codon from Reynolds classification. Have been associated with rises in cyanobacterial cell numbers. Colonial growth habitats grow in linear colonies joined together by linking spines. (Dae-Kyun et al., 2001; Reynolds et al., 2002; Stoermer and Julius, 2003) |
|                                | Uncultured <i>Planctomycetales</i>                         | 16S               | -ve                | Detected in media samples collected from biologically active filters. Capable of anammox (anaerobic denitrification) converting nitrate and nitrite to N <sub>2</sub> . (Maldonado et al., 2012; T. Zhang et al., 2017)   |

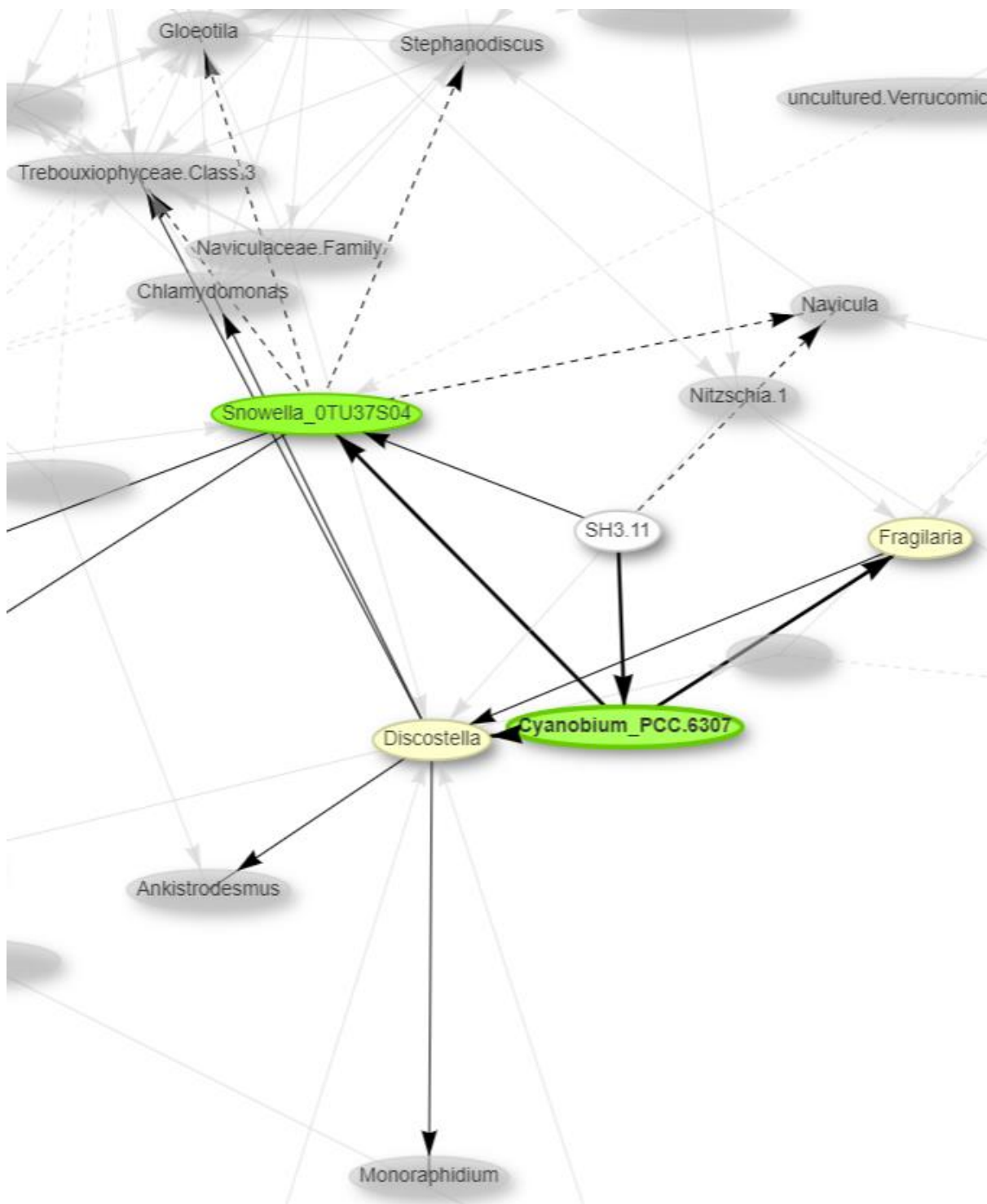


Figure 4.3: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978472>.

Table 4.4 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Alaw reservoir.

| <i>Cyanobacteria</i>        | Co-occurring taxon       | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|-----------------------------|--------------------------|-------------------|--------------------|---|
| <i>Aphanizomenon NIES81</i> | <i>TRA3-20</i>           | 16S               | -ve                | Known to be a heterotrophic iron-oxidising bacteria. Also described as a biomarker positively associated with the higher potential of microbial denitrification, C fixation and P accumulation in gullies. (Xiao et al., 2022)  |
|                             | <i>Fluviicola</i>        | 16S               | +ve                | Chemoorganotrophic Bacteroidetes-related bacterium, usually found in shallow lakes associated with high levels of dissolved organic C derived from algal/cyanobacterial blooms. Has been reported to have positive correlations with <i>Synechococcus</i> . (Farkas et al., 2020; Guedes et al., 2018)  |
| <i>Cyanobium PCC-6307</i>   | <i>Snowella</i> OTU37S04 | 16S               | +ve                | Not a known producer of geosmin but can produce $\beta$ -ionone. Cells are arranged radially at the ends of the mucilaginous stalks, with an average cell diameter of 3.2 $\mu$ m. (Loar, 2009; Rajaniemi-Wacklin et al., 2006)   |
|                             | <i>Fragilaria</i>        | <i>rbcL</i>       | +ve                | This genus can be planktonic or benthic, existing as colonial and non-colonial species. Colonial species tend to form ribbon-like colonies. Has been associated with increases in the cell number of <i>Cyanobacteria</i> and biofilms with <i>Cyanobacteria</i> . (Dae-Kyun et al., 2001; Espinosa et al., 2020; Heudre et al., 2019)  |
|                             | <i>Discostella</i>       | <i>rbcL</i>       | +ve                | Increases in this small cyclotelloid taxon are indicative of increased intensity and duration of thermal stratification during summer and are associated with increased temperatures. Is characteristic of oligotrophic conditions and can take up nutrients efficiently, often showing a negative relationship to nutrients. (Chen et al., 2021)   |
|                             | <i>SH3-11</i>            | 16S               | +ve                | Has been reported to be significantly and positively correlated with KO9819 (manganese/ iron transport system), K17225 (SOXC; sulphane dehydrogenase subunit), K17222 (SOXA; sulphur-oxidising protein SOxA), K04758 (feoA; ferrous iron transport protein A), and K04759 (feoB; ferrous iron transport protein B), and K01011 (Thiosulfate sulphurtransferase). (Nagarajan et al., 2022) |

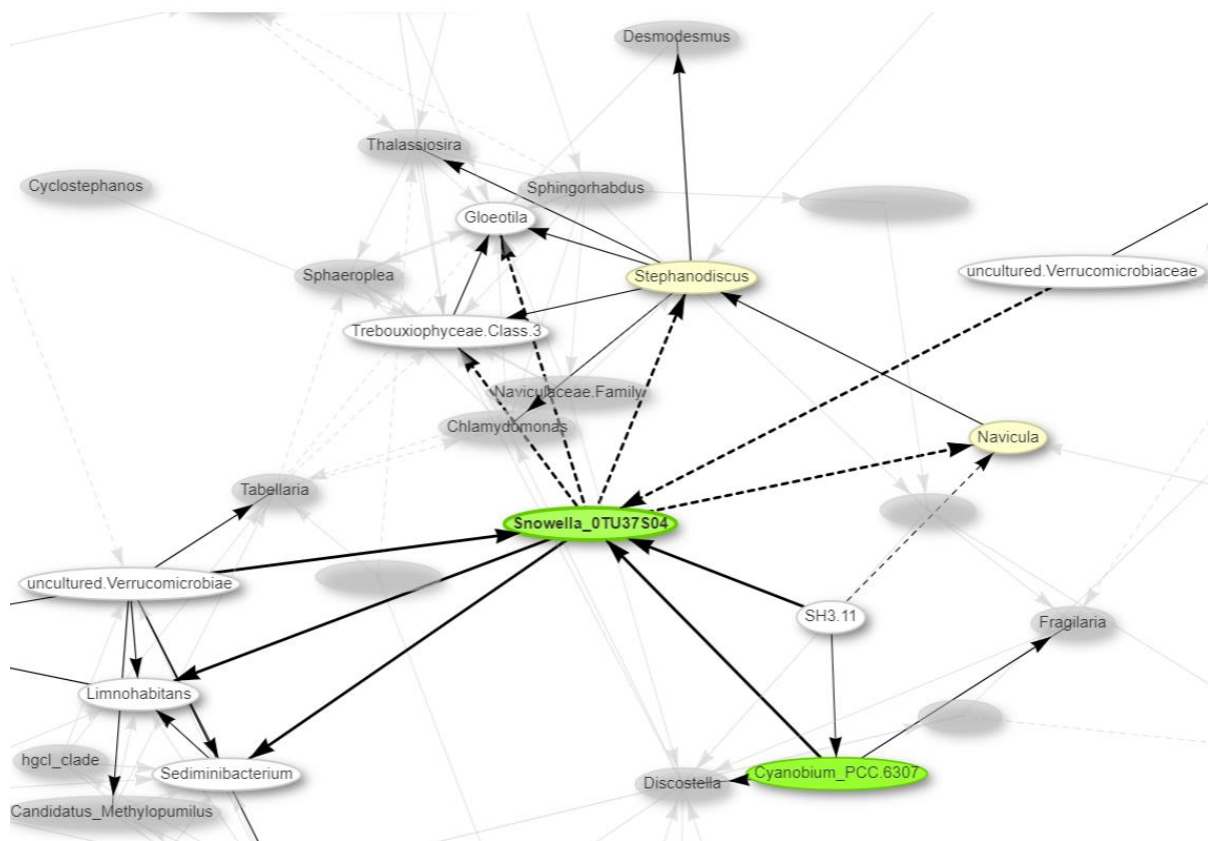


Figure 4.4: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Snowella* OTU37S04 node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978472>.

Table 4.4 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Alaw reservoir.

| <i>Cyanobacteria</i>        | Co-occurring taxon               | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|-----------------------------|----------------------------------|-------------------|--------------------|--|
| <i>Snowella</i><br>OTU37S04 | <i>Navicula</i>                  | <i>rbcL</i>       | -ve                | Raphed, pennate diatom that may exist singly or in colonial ribbons often found in nutrient-rich waters, usually bottom-dwelling. Highly motile and can adhere to surfaces and drift in response to fluctuations in nutrients. Continual secretion of adhesive mucilage. Has been associated in biofilms with <i>Cyanobacteria</i> . (Chen et al., 2019; Sabater et al., 2003) |
|                             | <i>Stephanodiscus</i>            | <i>rbcL</i>       | -ve                | Typically, colonial and placed in Reynolds codon D classification; mostly found in shallow, nutrient-enriched, well-ventilated waters, liable to be turbid. Small-celled ( $\leq 103 \mu\text{m}^3$ in volume) and fast-growing. (Reynolds et al., 2002)   |
|                             | <i>Trebouxiophyceae</i><br>Class | <i>rbcL</i>       | -ve                | Some members of this class can produce mycoporine-like amino acids, used as photoprotective substances. (Karsten et al., 2005)   |
|                             | <i>Gloeotila</i>                 | <i>rbcL</i>       | -ve                | Has synonymously been recognised as <i>Geminella</i> . Found in deep, well-mixed epilimnia. Tolerant to light deficiency and sensitive to nutrient deficiency. Assigned to codon T according to Reynolds classification. (Fernandes et al., 2021; Reynolds et al., 2002)   |
|                             | <i>Sediminibacterium</i>         | 16S               | +ve                | Previous genome sequencing of a cyanosphere exposed <i>Sediminibacterium</i> to be a resident. Hypothesised to offer protection against colonisation of opportunistic bacteria by producing bacteriocin and toxoflavin, whilst providing <i>Cyanobacteria</i> with inorganic nutrients for cyanobacterial growth. (Sethuraman et al., 2022; Yang et al., 2022)                 |
|                             | <i>Limnohabitans</i>             | 16S               | +ve                | Found to thrive along with cyanobacterial blooms in spring or summer, using phytoplankton-derived organic material as the substrate for growth. Culture-independent studies have revealed this genus to be the second most abundant (11%) taxa in the cyanosphere. (Luo et al., 2022; Seok Jea Youn et al., 2020)  |



Table 4.4 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Alaw reservoir.

| <i>Cyanobacteria</i>      | Co-occurring taxon  | 16S / <i>rbcl</i> | Co-occur influence | Importance of co-occurring taxon  |
|---------------------------|---|-------------------|--------------------|---|
| <i>Snowella OTU37S04</i>  | Uncultured <i>Verrucomicrobiaceae</i> / <i>Verrucomicrobiae</i> | 16S               | -ve / +ve          | Associated with rhizospheres and endospheres found in plant roots and soil. Implicated in root carbon metabolism. Has also been connected to <i>Lyngbya</i> (non-heterocyst-forming) blooms. (Newitt, 2020; Rajaneesh et al., 2020)   |
|                           | <i>SH3-11</i>   | 16S               | +ve                | Has been reported to be significantly and positively correlated with KO9819 (manganese/ iron transport system), K17225 (SOXC; sulphane dehydrogenase subunit), K17222 (SOXA; sulphur-oxidising protein SOxA), K04758 (feoA; ferrous iron transport protein A), and K04759 (feoB; ferrous iron transport protein B), and K01011 (Thiosulfate sulfurtransferase). (Nagarajan et al., 2022)  |
|                           | <i>Cyanobium PCC-6307</i>                                       | 16S               | +ve                | Type of picocyanobacteria; species vary from oval to cylindrical cells, with little or no evident mucilage exterior. Not a known producer of geosmin, although can produce 2-MIB and microcystins. (Clerc et al., 2022; Hojun et al., 2021; Śliwińska-Wilczewska et al., 2018)  |
| <i>Nostocaceae</i> Family | <i>Sphingorhabdus</i>   | 16S               | +ve                | Suspected geosmin degrader from the family <i>Sphingomonadaceae</i> , which includes other known geosmin degraders. Characterised by harbouring glycosphingolipids. Many sphingomonads degrade polycyclic aromatic compounds and xenobiotics. Many geosmin-degrading genera, like <i>Sphingopyxis</i> , have been reclassified as <i>Spingorhabdus</i> and some members of this genus have been reclassified as the geosmin-degrading genus <i>Novosphingobium</i> . (Jogler et al., 2013; Sharma et al., 2021) |
|                           | <i>Thalassiosira</i>  | <i>rbcl</i>       | +ve                | Large centric diatom capable of producing polyunsaturated aldehydes (PUAs), that has antiproliferative activity in zooplankton. The liberation of embryotoxic PUAs by damaged diatom cells could be considered an activated defence of these diatoms on the population level. (Fink, 2007)  |
|                           | <i>Chlamydomonadaceae</i> Family                                | <i>rbcl</i>       | +ve                | Includes the genus <i>Chlamydomonas</i> which is known to produce polyolglycosides, sterols, phenols, and fatty acids along with other compounds. Thought to use these compounds in chemotaxonomic relationships with other lineages. Additions of <i>Aphanizomenon</i> or cylindrospermopsin to <i>Chlamydomonas</i> induce genes typically upregulated under P limitation and give rise to extracellular APase activity that <i>Cyanobacteria</i> can use. (Bar-Yosef et al., 2010; Khanh Tran et al., 2019)  |

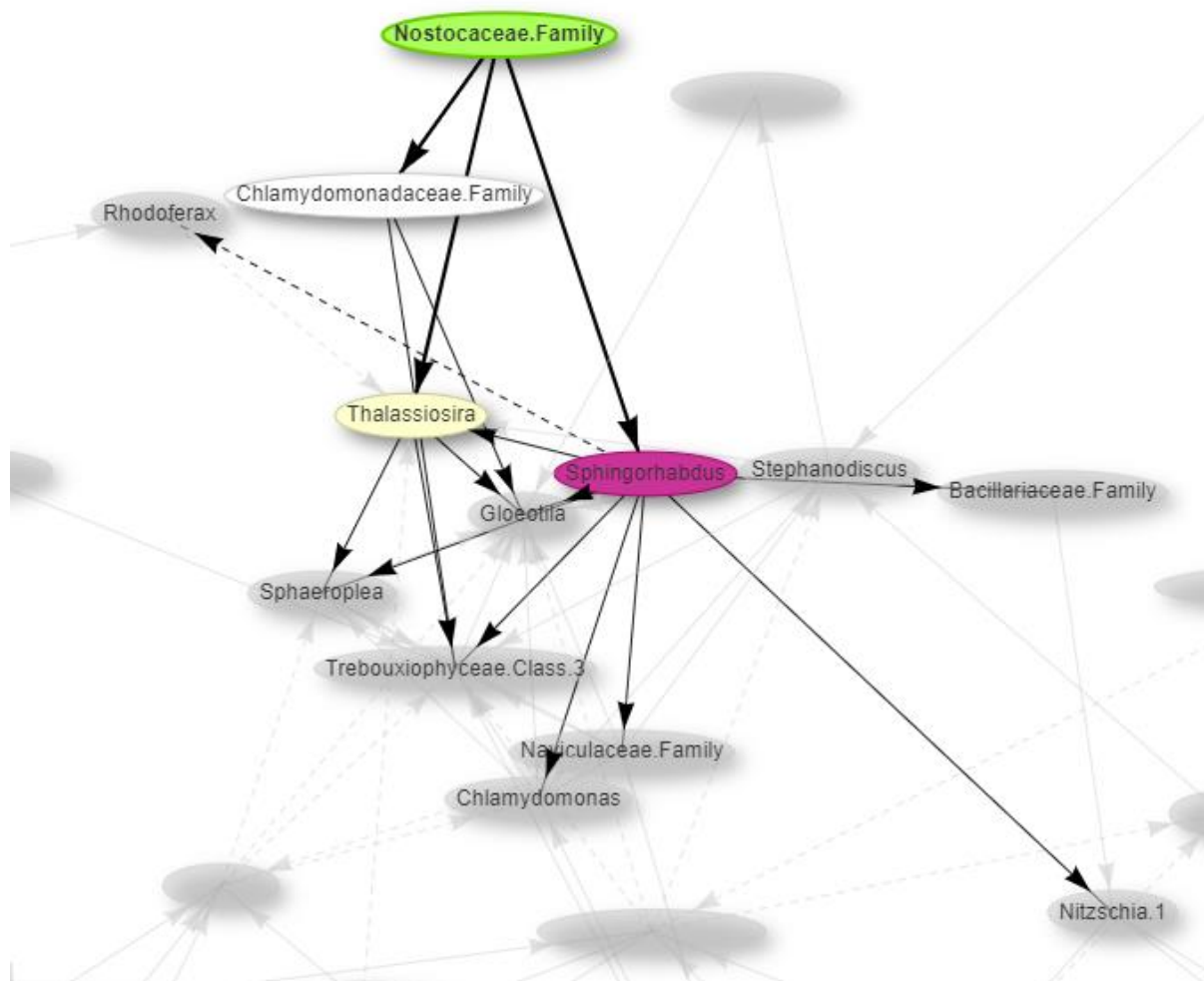


Figure 4.5: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Nostocaceae* family node during the entirety of this study in the Alaw reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978472>.

#### 4.3.2.2 – Alwen

In the Alwen reservoir, only 36 pairwise comparisons yielded statistically significant co-occurrence ( $<0.05$ ); 13 of which were negatively associated, and 23 co-occurring taxa were positively associated. The network can be accessed online via the link in Table 4.3. 9 disconnected network communities were contained within the overall network, and no overlapping cyanospheres were present. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.5.

Figure 4.6 represents the *Aphanizomenon* NIES81 cyanosphere; three contained taxa have negative co-occurrences with *Aphanizomenon* NIES81, and uncultured *Pedosphaeraceae* had a positive co-occurring relationship. For *Dolichospermum* NIES41, only one taxon was found to be significantly positively co-occurring, *Sediminibacterium* (Figure 4.7). In the *Nostocaceae* family cyanosphere, there were only two significantly negatively co-occurring taxa (Figure 4.8).

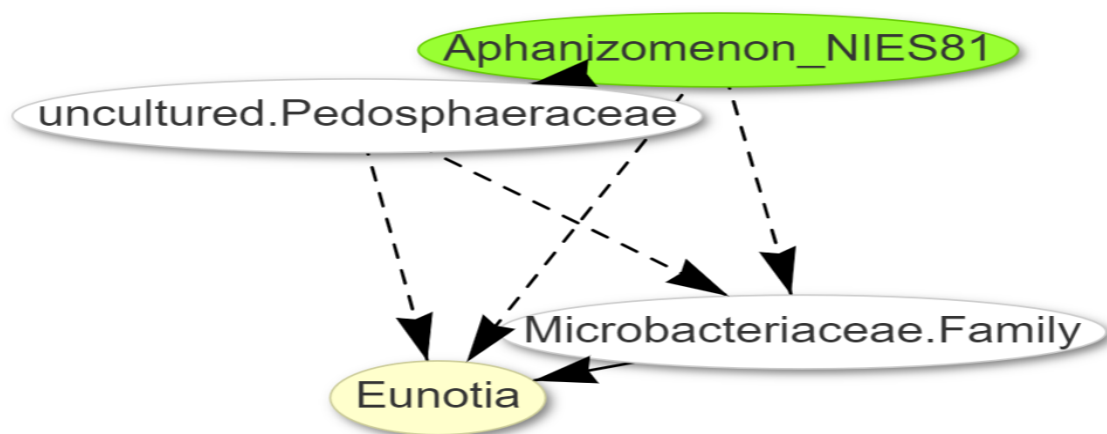


Figure 4.6: Subset of the significant ( $< 0.05$ ) co-occurring taxa associated with the *Aphanizomenon* NIES81 node during the entirety of this study in the Alwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link:

<https://rpubs.com/ASHooper/978480>.

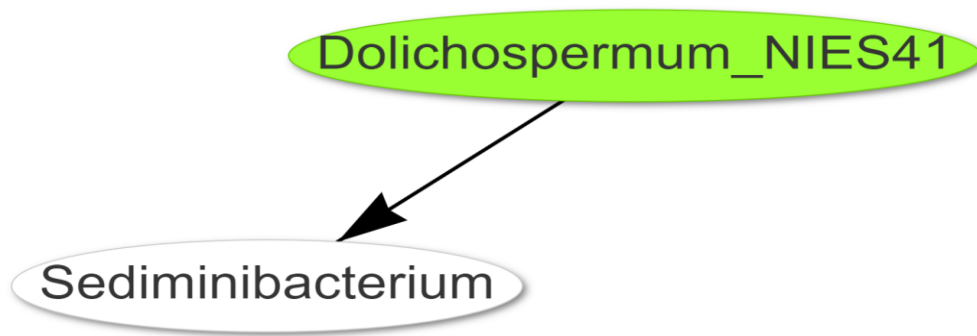


Figure 4.7: Subset of the significant ( $< 0.05$ ) co-occurring taxa associated with the *Dolichospermum* NIES41 node during the entirety of this study in the Alwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978480>.

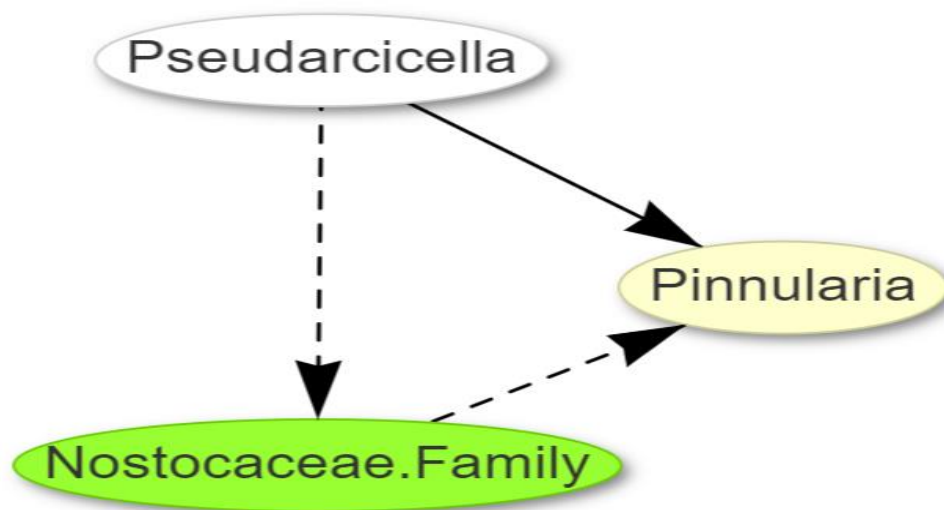


Figure 4.8: Subset of the significant ( $< 0.05$ ) co-occurring taxa associated with the *Nostocaceae* family node during the entirety of this study in the Alwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978480>.

Table 4.5: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Alwen reservoir.

| <i>Cyanobacteria</i>            | Co-occurring taxon                | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|---------------------------------|-----------------------------------|-------------------|--------------------|--|
| <i>Aphanizomenon</i><br>NIES81  | <i>Microbacteriaceae</i> Family   | 16S               | -ve                | Family in the order <i>Actinomycetales</i> that contain known geosmin producers. Previous studies have identified this family to correlate with levels of <i>geoA</i> gene copies. The family embraces mesophilic/ psychrophilic bacteria, which are obligately aerobic to facultatively anaerobic. (Mie B. Lukassen et al., 2019)                             |
|                                 | <i>Eunotia</i>                    | <i>rbcL</i>       | -ve                | Abundant in the epiphyton of oligotrophic waters with low conductivity levels and in highly transparent waters. Most species are acidobiontic or acidophilic organisms. (Luo et al., 2019)   |
|                                 | uncultured <i>Pedosphaeraceae</i> | 16S               | +ve                | Previously found to be a keystone species in rhizospheres found in soils. They play an important role in sediment nutrient circulation with high metal resistance potentials. (Yuan et al., 2022)  |
| <i>Dolichospermum</i><br>NIES41 | <i>Sediminibacterium</i>          | 16S               | +ve                | Previous genome sequencing of a cyanosphere exposed <i>Sediminibacterium</i> to be a resident. Hypothesised to offer protection against colonisation of opportunistic bacteria by producing bacteriocin and toxoflavin, whilst providing <i>Cyanobacteria</i> with inorganic nutrients for cyanobacterial growth. (Sethuraman et al., 2022; Yang et al., 2022) |
| <i>Nostocaceae</i><br>Family    | <i>Pinnularia</i>                 | <i>rbcL</i>       | -ve                | A raphid benthic pennate diatom, known to produce EPS which covers the siliceous cell wall and forms adhesive trails with atomic force microscopy. Cells contain cavities and channels in the complex secondary plastid which harbours symbiotic bacteria. (Chiovitti et al., 2003; Souffreau et al., 2011)  |
|                                 | <i>Pseudarcicella</i>             | 16S               | -ve                | Known bacterial indicator of good water quality. (Guo et al., 2021)  |

#### 4.3.2.3 – Cefni

In the Cefni reservoir, 375 pairwise comparisons yielded statistically significant co-occurrence (<0.01); 64 were negatively associated, and 311 co-occurring taxa were positively associated. The network can be accessed online via the link contained within Table 4.3. There was one extensive network with one independent network containing two nodes (*Aeromonas* and *Lindavia*). Eight significantly co-occurring *Cyanobacteria* were present in the extensive network and overlapped in all cyanospheres. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen illustrated in Table 4.6.

Figure 4.9 represents the *Aphanizomenon* MDT14a cyanosphere, which encompasses 10 implicated taxa, nine of which are positively co-occurring and one that was negatively co-occurring uncultured *Rhizobiales Incertae Sedis*. *Pseudanabaena* PCC-7429 can be seen connected to three nodes, all of which were positively influential (Figure 4.10). Similarly, *Aphanizomenon* NIES81 can be witnessed to be positively joined to four taxa (Figure 4.11), and *Cyanobium* PCC-6307 had positive associations with five taxa (Figure 4.12). In contrast, *Gloeotrichia* PYH6 had seven associated taxa within the cyanosphere, three of which had a negative co-occurrence and four positive co-occurrence (Figure 4.13). *Microcystis* PCC-7914 had six positively co-occurring taxa within its cyanosphere, and one negatively co-occurring taxa (*NS11-12 marine group*) (Figure 4.14). The *Nostocaceae* family had the largest cyanosphere, comprising over 21 taxa, 12 of which were negatively associated and nine positively associated (Figure 4.15). *Geitlerinema* LD9 had a cyanosphere containing seven taxa, all of which were positively associated, apart from *Flavobacterium*, which was negatively associated with *Geitlerinema* LD9 (Figure 4.16).

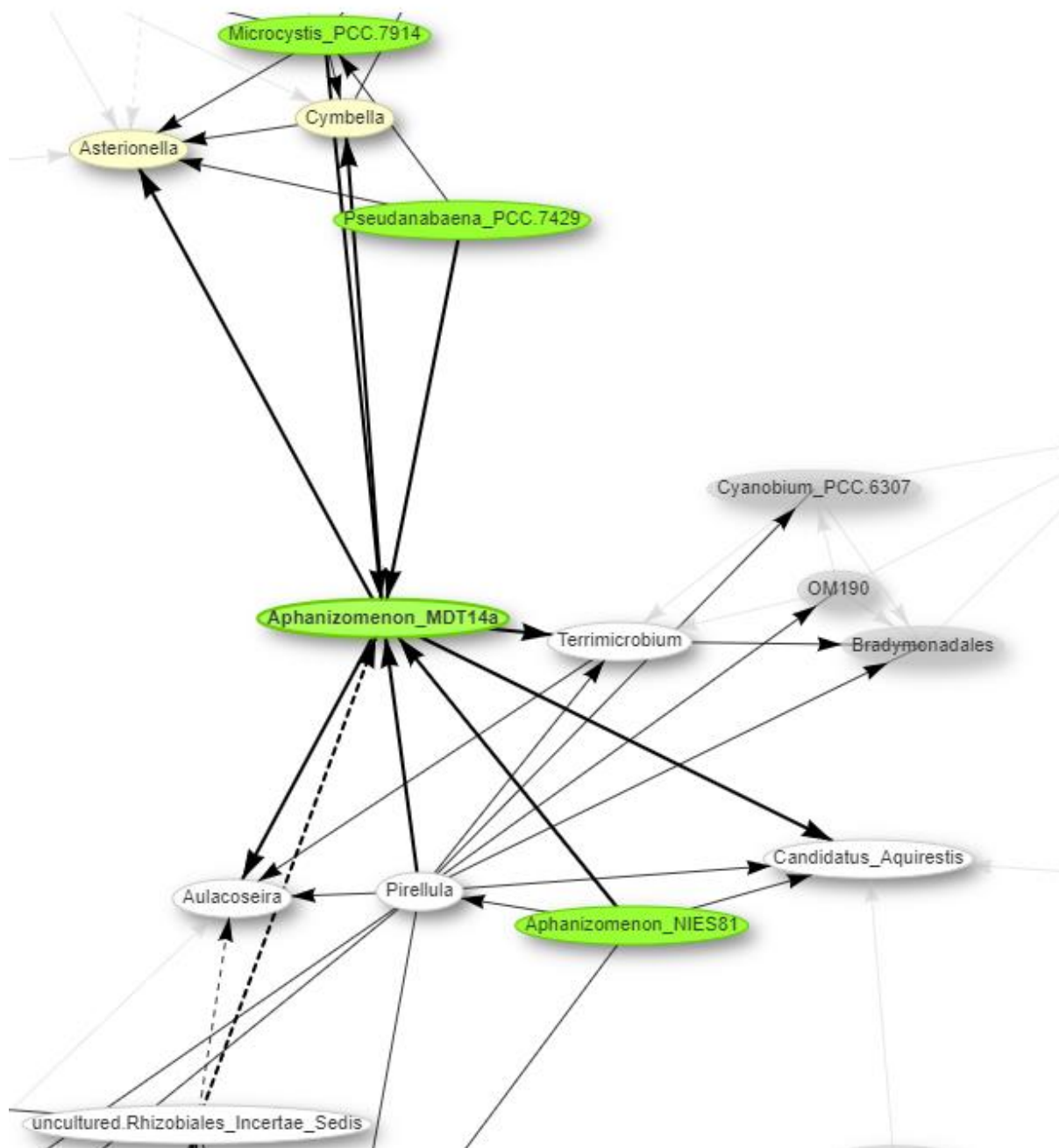


Figure 4.9: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Aphanizomenon* MDT14a node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.

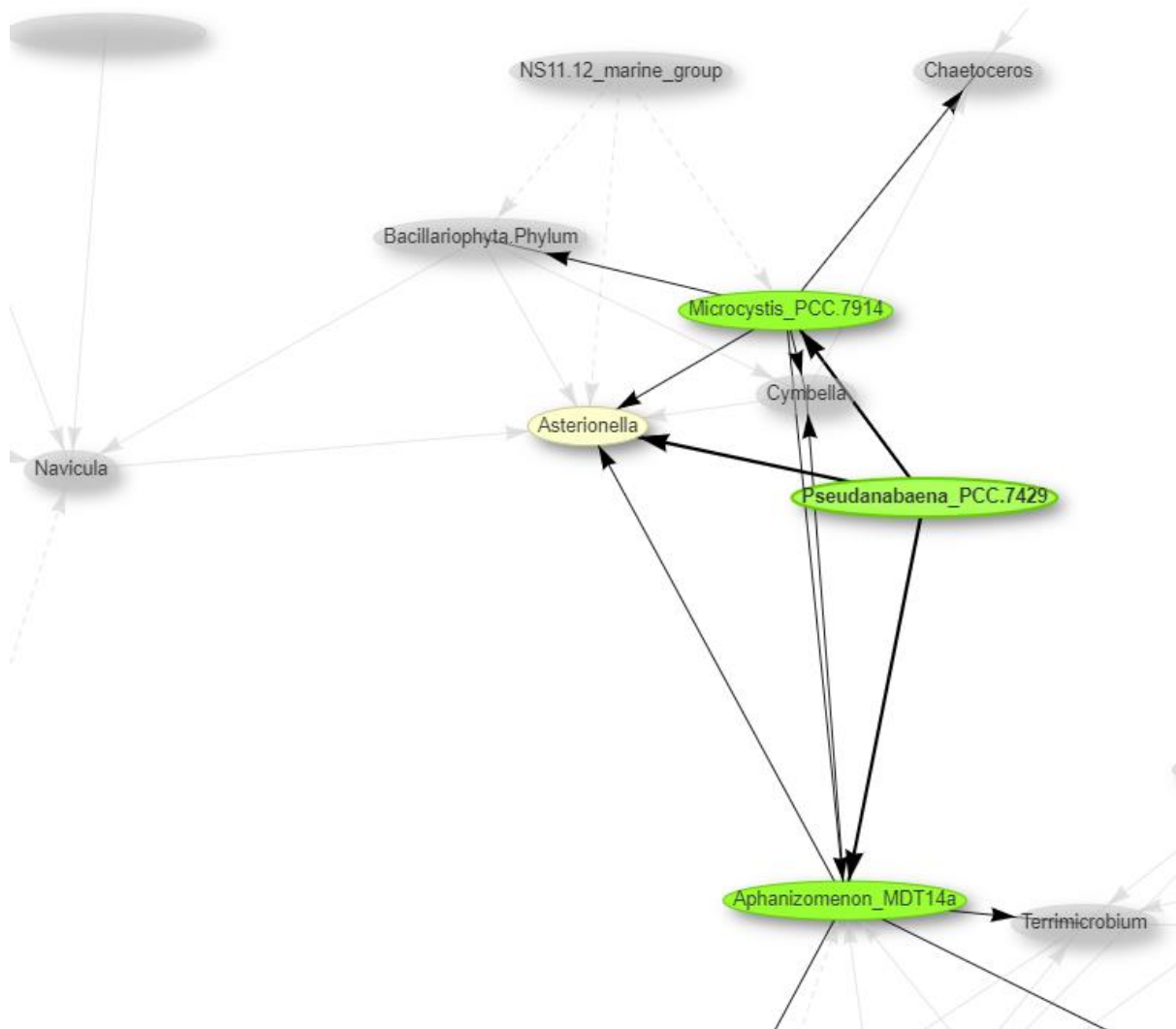


Figure 4.10: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Pseudanabaena* PCC-7429 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.



Table 4.6: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i>           | Co-occurring taxon           | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|--------------------------------|------------------------------|-------------------|--------------------|---|
| <i>Aphanizomenon</i><br>MDT14a | <i>Terrimicrobium</i>        | 16S               | +ve                | Non-motile, anaerobic rods. Ferments various sugars, but not fatty acids, alcohols (i.e., geosmin) or amino acids. Previous research has identified this genus with <i>Nostoc</i> epibacteria. (Qiu et al., 2014; Satjarak et al., 2021)  |
|                                | <i>Candidatus Aquirestis</i> | 16S               | +ve                | Filamentous chemoorganotrophic bacterium. Primarily inhabit the pelagic zone of nonacidic stagnant inland waters and show a strong spring peak, with minor peaks in summer and early autumn. Presence is often related to the high level of dissolved organic carbon concentration derived from algal/ cyanobacterial blooms. (Farkas et al., 2020; Nguyen et al., 2018)                                  |
|                                | <i>Cymbella</i>              | <i>rbcL</i>       | +ve                | Can be epiphytic on surfaces. Some species of this genus can be potentially harmful. Produces mucilage, said to be more dependent on physiochemical conditions rather than its cell density. Mucilage production is strongly related to P and temperature levels. (Zamorano et al., 2019)   |
|                                | <i>Aulacoseira</i>           | <i>rbcL</i>       | +ve                | Colonial growth habitat, found in mixed, eutrophic small medium lakes. Part of codon B and C classification set by Reynolds (one species belonging to codon P, found in eutrophic epilimnia). High tolerances to light and C deficiencies. This can be indicative of very phosphorous-rich systems. Sensitive to stratification and sedimentation due to the large size of cells. (Reynolds et al., 2002) |
|                                | <i>Asterionella</i>          | <i>rbcL</i>       | +ve                | Belongs to codon C in the Reynolds classification. Characteristic member of nutrient-rich temperate lakes. Rapid increases of this species during March - April leads to the extraction of major anions (nitrate, phosphate, and silicate), resulting in nutrient depletion and altering nutrient ratios. (Krivtsov et al., 2000; Reynolds et al., 2002)  |

Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i>          | Co-occurring taxon                           | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|-------------------------------|--|-------------------|--------------------|--|
| <i>Aphanizomenon MDT14a</i>   | Uncultured <i>Rhizobiales Incertae Sedis</i> | 16S               | -ve                | family of nitrogen-fixing bacteria. Members of the <i>Rhizobiales</i> have been reported to degrade cyanobacterial-derived particulate organic matter. (Huang et al., 2022; Millar et al., 2022)   |
|                               | <i>Pseudanabaena</i> PCC-7429                | 16S               | +ve                | Benthic dwelling picocyanobacteria, not a known producer of geosmin but can produce 2-MIB. Does not produce siderophores but encodes part of the siderophore uptake system. (Enzingmüller-bleyl et al., 2021)  |
|                               | <i>Aphanizomenon</i> NIES81                  | 16S               | +ve                | Filamentous capable of nitrogen fixation found in eutrophic waters, assigned to codon H in Reynolds classification. Known geosmin and microcystin producer. (Reynolds et al., 2002)  |
|                               | <i>Pirellula</i>                             | 16S               | +ve                | Species of this genus are found usually attached to filamentous algae and <i>Cyanobacteria</i> by a holdfast located at the distal end of the fascicle or at the nonreproductive pole of the cell. All species are said to be tachyletic. (Clum et al., 2009)  |
|                               | <i>Microcystis</i> PCC-7914                  | 16S               | +ve                | Non-heterocyst-forming <i>Cyanobacteria</i> found in eutrophic shallow waters. Assigned to codon M in Reynolds classification, tolerant to high insolation and sensitive to flushing and low light. Upregulates P scavenging genes and can outcompete other phytoplankton in P-limited environments. (Berry et al., 2017; Reynolds et al., 2002)         |
| <i>Pseudanabaena</i> PCC-7429 | <i>Microcystis</i> PCC-7914                  | 16S               | +ve                | Non-heterocyst-forming <i>Cyanobacteria</i> found in eutrophic shallow waters. Assigned to codon M in Reynolds classification, tolerant to high insolation and sensitive to flushing and low light. Upregulates P scavenging genes and can outcompete other phytoplankton in P-limited environments. (Berry et al., 2017; Reynolds et al., 2002)         |
|                               | <i>Aphanizomenon</i> MDT14a                  | 16S               | +ve                | Filamentous capable of nitrogen fixation found in eutrophic waters, assigned to codon H in Reynolds classification. Known geosmin and microcystin producer. (Reynolds et al., 2002)  |
|                               | <i>Asterionella</i>                          | <i>rbcL</i>       | +ve                | Belongs to codon C in the Reynolds classification. Characteristic member of nutrient-rich temperate lakes. Rapid increases of this species during March - April leads to the extraction of major anions (nitrate, phosphate, and silicate), resulting in nutrient depletion and altering nutrient ratios. (Krivtsov et al., 2000; Reynolds et al., 2002) |

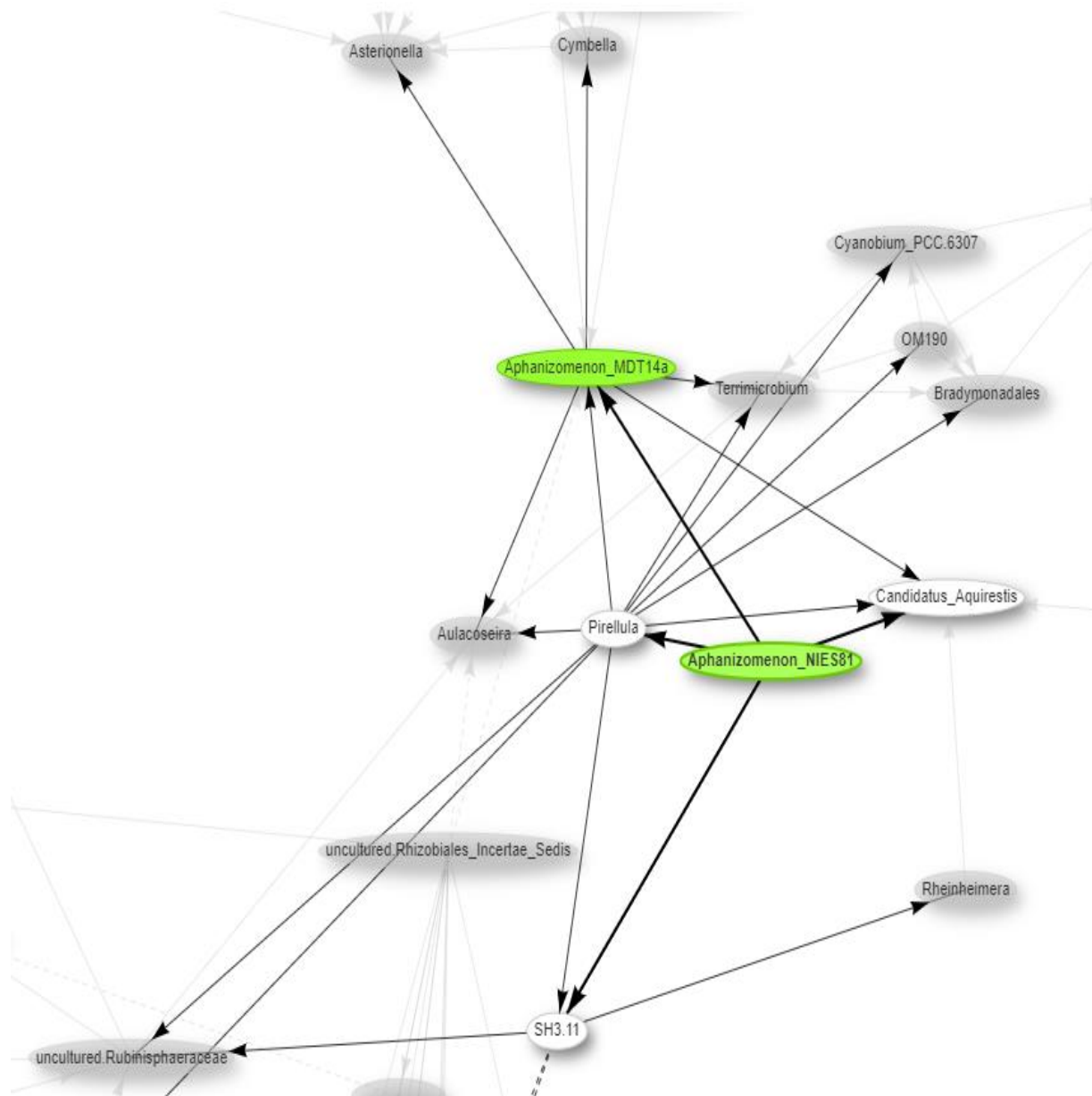


Figure 4.11: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Aphanizomenon* NIES81 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.

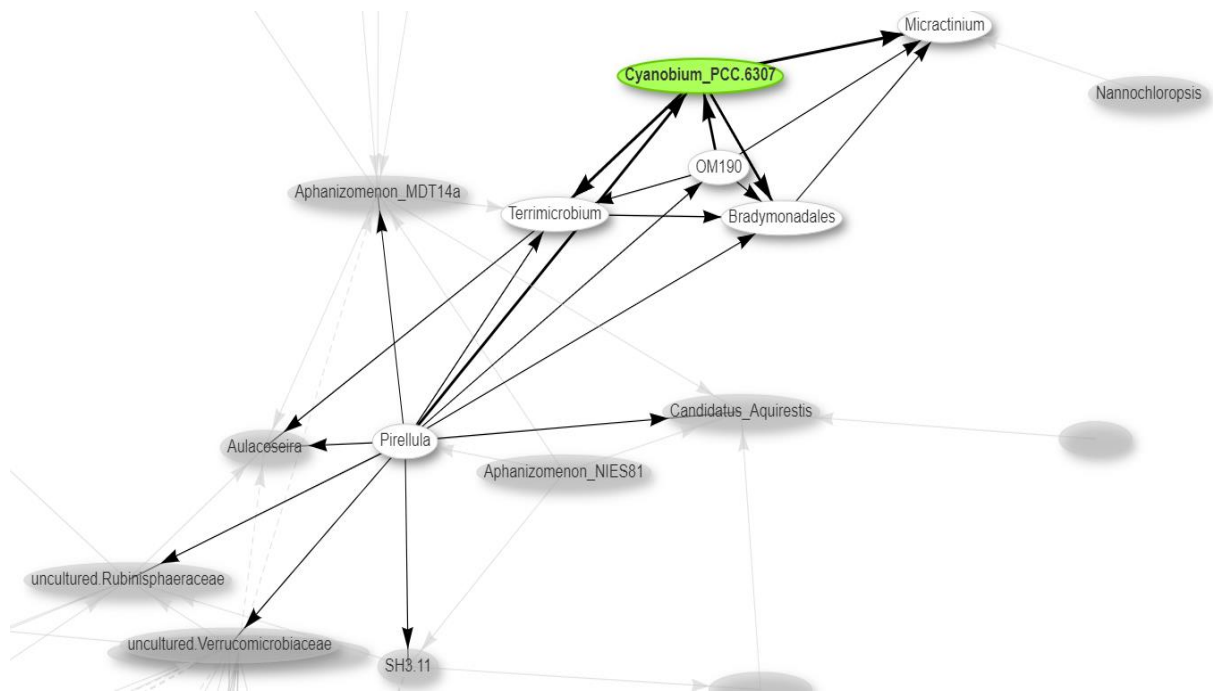


Figure 4.12: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.

Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i>           | Co-occurring taxon                     | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|--------------------------------|--|-------------------|--------------------|--|
| <i>Aphanizomenon</i><br>NIES81 | <i>Pirellula</i>                       | 16S               | +ve                | Species of this genus are found usually attached to filamentous algae and <i>Cyanobacteria</i> by a holdfast located at the distal end of the fascicle or at the nonreproductive pole of the cell. All species are said to be tachylectic. (Clum et al., 2009)   |
|                                | <i>SH3-11</i>                          | 16S               | +ve                | Has been reported to be significantly and positively correlated with KO9819 (manganese/ iron transport system), K17225 (SOXC; sulphane dehydrogenase subunit), K17222 (SOXA; sulphur-oxidising protein SOxA), K04758 (feoA; ferrous iron transport protein A), and K04759 (feoB; ferrous iron transport protein B), and K01011 (Thiosulfate sulfurtransferase). (Nagarajan et al., 2022) |
|                                | <i>Aphanizomenon</i><br>MDT14a         | 16S               | +ve                | Filamentous capable of nitrogen fixation found in eutrophic waters, assigned to codon H in Reynolds classification. Known geosmin and microcystin producer. (Reynolds et al., 2002)  |
|                                | <i>Candidatus</i><br><i>Aquirestis</i> | 16S               | +ve                | Chemoorganotrophic bacteria found in stagnant shallow freshwater habitats, often related to high levels of dissolved organic C concentration derived from algal and cyanobacterial blooms. (Farkas et al., 2020)   |
| <i>Cyanobium</i><br>PCC-6307   | <i>Terrimicrobium</i>                  | 16S               | +ve                | Non-motile, anaerobic rods. Ferments various sugars, but not fatty acids, alcohols (i.e., geosmin) or amino acids. Previous research has identified this genus with <i>Nostoc</i> epibacteria. (Qiu et al., 2014; Satjarak et al., 2021)   |
|                                | <i>Bradymonadales</i>                  | 16S               | +ve                | Possesses a transitional survival mode, alternating between "obligate" and so-called facultative predators (facultatively prey-dependent). They store nutrients as polymers in cells during predation and can accumulate polyphosphate in the phosphate-rich zone, using it as an energy source. (Mu et al., 2020)   |
|                                | <i>Micractinium</i>                    | <i>rbcL</i>       | +ve                | Chlorella-like genus, which has spherical-to-ovoid cells and possesses a parietal and cup-shaped chloroplasts with a pyrenoid. Species tend to be mesophilic and tolerant to low temperatures with optimal growth at 20°C. Contains high concentrations of polyunsaturated fatty acids. (Hong et al., 2015)  |
|                                | <i>Pirellula</i>                       | 16S               | +ve                | Species of this genus are found usually attached to filamentous algae and <i>Cyanobacteria</i> by a holdfast located at the distal end of the fascicle or at the nonreproductive pole of the cell. All species are said to be tachylectic. (Clum et al., 2009)   |
|                                | <i>OM190</i>                           | 16S               | +ve                | Members of this clade are usually associated with microalgae. Produces secondary metabolites, including antimicrobial compounds. It has been shown that diatoms produce fucose-containing sulphated polysaccharides which this genus can supposedly use as an energy source. (Pushpakumara et al., 2022)   |

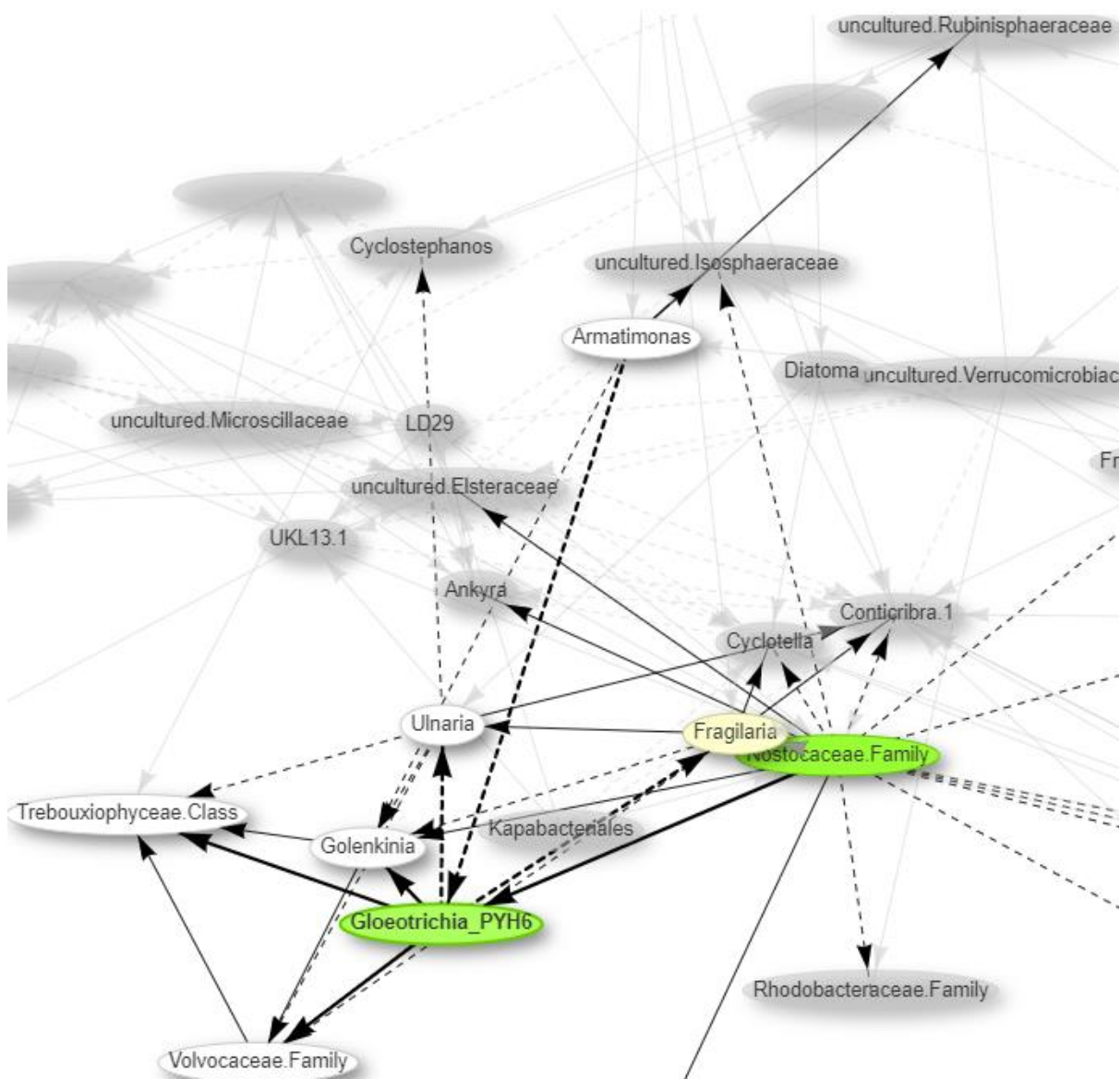


Figure 4.13: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Gloeotrichia* PYH6 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.

Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i>     | Co-occurring taxon            | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|--------------------------|-------------------------------|-------------------|--------------------|--|
| <i>Gloeotrichia</i> PYH6 | <i>Fragilaria</i>             | <i>rbcL</i>       | -ve                | The genus can be planktonic or benthic, existing as colonial and non-colonial species. Colonial species tend to form ribbon-like colonies. Has been associated with increases in the cell number of <i>Cyanobacteria</i> and in biofilms with <i>Cyanobacteria</i> . (Dae-Kyun et al., 2001; Espinosa et al., 2020; Heudre et al., 2019) |
|                          | <i>Ulnaria</i>                | <i>rbcL</i>       | -ve                | Present in stable meso to eutrophic water columns, with circumneutral to slightly alkaline conditions. Previously included in the large genus <i>Synedra</i> . Species are mesoaprobic. Normally associated with increased productivity. (Thacker and Karthick, 2022)  |
|                          | <i>Golenkinia</i>             | <i>rbcL</i>       | +ve                | Green algae are known for their production of lipids, sugars, and carotenoids. Industrial interests in the genus for their phytoremediation of chemical oxygen demand, N and P. (Sisman-Aydin and Simsek, 2022)  |
|                          | <i>Volvocaceae</i> Family     | <i>rbcL</i>       | +ve                | Group of multicellular colonial algae frequently encountered in eutrophic waters. Previously found to be associated with <i>Nostoc</i> as epibionts. (Satjarak et al., 2021)   |
|                          | <i>Trebouxiophyceae</i> Class | <i>rbcL</i>       | +ve                | Some members of this class can produce mycoporine-like amino acids, used as photoprotective substances. (Karsten et al., 2005)   |
|                          | <i>Armatimonas</i>            | 16S               | -ve                | Aerobic, non-motile, ovoid to rod-shaped often found in mesophilic waters. Previously shown to be positively correlated with TP. (Xu et al., 2018)   |
|                          | <i>Nostocaceae</i> Family     | 16S               | +ve                | Genera of this family are heterocyst forming. Includes many genera that possess <i>geoA</i> , including <i>Anabaena</i> and <i>Nostoc</i> species. (Yamada et al., 2015)   |

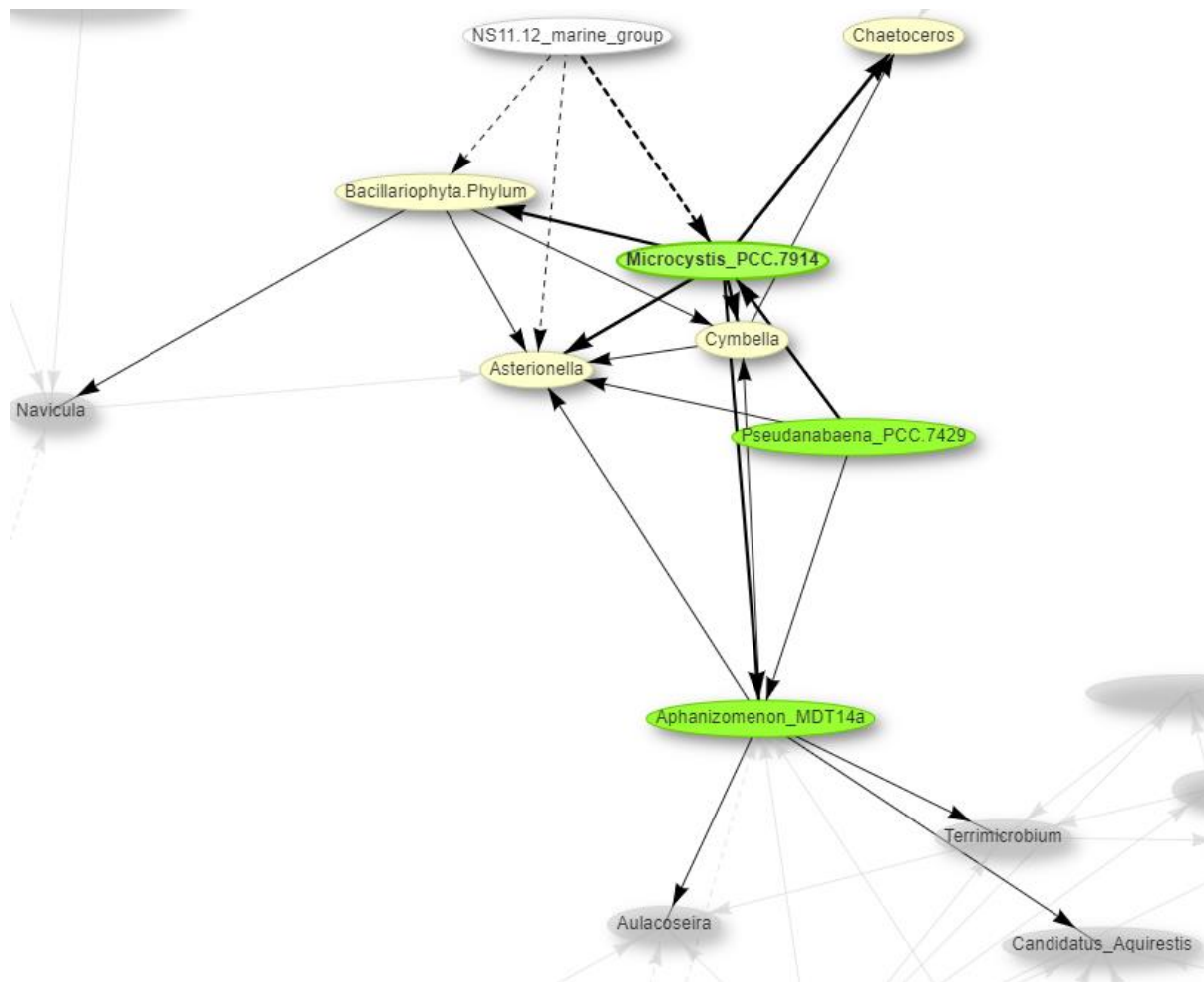


Figure 4.14: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Microcystis* PCC-7914 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.



Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i>        | Co-occurring taxon            | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|-----------------------------|-------------------------------|-------------------|--------------------|--|
| <i>Microcystis</i> PCC-7914 | <i>Aphanizomenon</i> MDT14a   | 16S               | +ve                | Filamentous capable of nitrogen fixation found in eutrophic waters, assigned to codon H in Reynolds classification. Known geosmin and microcystin producer. (Reynolds et al., 2002)  |
|                             | <i>Bacillariophyta</i> Phylum | <i>rbcL</i>       | +ve                | Unicellular or colonial coccoid algae diatoms, that can be benthic or planktonic.  |
|                             | <i>Cymbella</i>               | <i>rbcL</i>       | +ve                | Can be epiphytic on surfaces. Some species of this genus can be potentially harmful. Produces mucilage, said to be more dependent on physiochemical conditions rather than its cell density. Mucilage production is strongly related to P and temperature levels. (Zamorano et al., 2019)  |
|                             | <i>Asterionella</i>           | <i>rbcL</i>       | +ve                | Belongs to codon C in the Reynolds classification. Characteristic member of nutrient-rich temperate lakes. Rapid increases of this species during March - April leads to the extraction of major anions (nitrate, phosphate, and silicate), resulting in nutrient depletion and altering nutrient ratios. (Krivtsov et al., 2000; Reynolds et al., 2002) |
|                             | <i>Chaetoceros</i>            | <i>rbcL</i>       | +ve                | Benthic chain-forming diatom usually forms epiphytically diatom-cyanobiont symbioses with <i>Calothrix</i> . Has a high iron uptake and storage capacity. (Iwade et al., 2006; Tuo et al., 2017)   |
|                             | NS11-12 marine group          | 16S               | -ve                | Planktonic chemoorganotrophic bacteria indicative of high levels of dissolved organic carbon derived from algal/ cyanobacterial blooms. Mainly identified in marine environments but also freshwater, unfortunately without clear ecological implications. (Farkas et al., 2020)   |
|                             | <i>Pseudanabaena</i> PCC-7429 | 16S               | +ve                | Benthic dwelling picocyanobacteria, not a known producer of geosmin but can produce 2-MIB. Does not produce siderophores but encodes part of the siderophore uptake system. (Enzingmüller-bleyl et al., 2021)  |

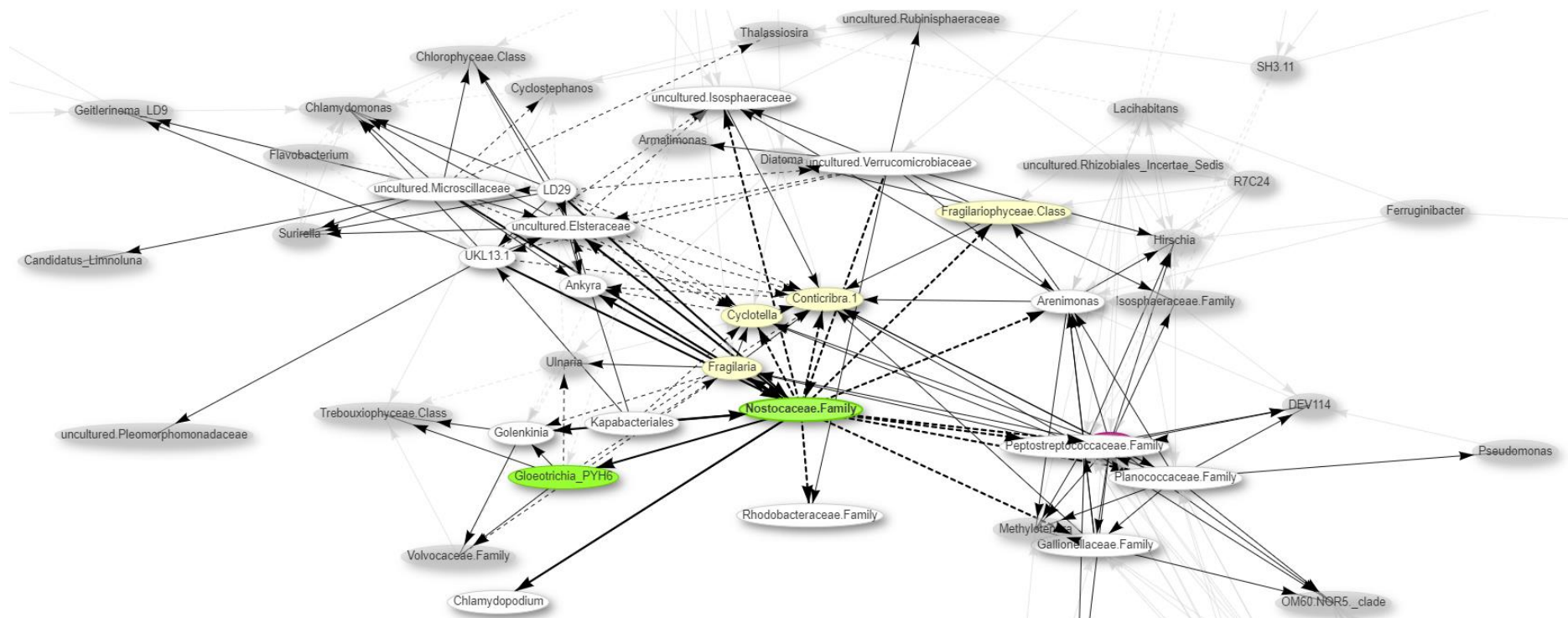


Figure 4.15: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Nostocaceae* family node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each genus. Link: <https://rpubs.com/ASHooper/978493>.

Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i>         | Co-occurring taxon                     | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|------------------------------|--|-------------------|--------------------|---|
| <i>Nostocaceae</i><br>Family | <i>Rhodobacteraceae</i><br>Family      | 16S               | -ve                | Mainly consists of aerobic photo- and chemoheterotrophs, purple non-sulphur bacteria that perform photosynthesis, and can also be anaerobic. Play roles in sulphur, nitrogen, and carbon cycles, decomposing various compounds and generating secondary metabolites. Produces EPS that retains large amounts of water, some species synthesise ectoine (protein protectant), and other species can synthesise polyhydroxyalkanoates (PHAs) (compounds accumulated by many microorganisms). (Azpiazu-muniozguren et al., 2022) |
|                              | <i>Bacillus</i>                        | 16S               | -ve                | Shown to have an antagonistic relationship with geosmin-producing strains of <i>Streptomyces</i> , inhibiting growth. Also, a well-known geosmin and 2-MIB degrader. (Ma et al., 2015; Zhi et al., 2016)  |
|                              | <i>Planococcaceae</i><br>Family        | 16S               | -ve                | This family is a taxonomically heterogeneous assemblage of 13 genera, many of which are polyphyletic. There is a lack of known characteristics exclusive to all members of this family. (Gupta and Patel, 2020)   |
|                              | <i>Gallionellaceae</i> Family          | 16S               | -ve                | Chemolithoautotrophic and neutrophilic ferrous iron-oxidising bacteria secrete extracellular fibres that attract iron hydroxides and many trace metals. (Hallbeck and Pederson, 2014)   |
|                              | <i>Peptostreptococcaceae</i><br>Family | 16S               | -ve                | Indicator of nutrient enrichment in urban lakes. Also, a common commensal bacterial taxon is thought to play a role in maintaining gut homeostasis in fish. (Numberger et al., 2022)  |
|                              | <i>Arenimonas</i>                      | 16S               | -ve                | Found to be a core node in previous microbial networks and was responsible for maintaining interactions between autotrophic and heterotrophic nitrifying bacteria. (Wu et al., 2022)  |
|                              | Uncultured<br><i>Isosphaeraceae</i>    | 16S               | -ve                | This family accommodates stalk-free planctomycetes with spherical cells, which can assemble in short chains, long filaments, or aggregates. All members do not produce motile swarmer cells as typical for other planctomycetes. Genes encode metabolic pathways common for chemoorganotrophic bacteria (glycolysis, citrate cycle, pentose-phosphate pathway, and oxidative phosphorylation). (Ivanova et al., 2017)   |
|                              | <i>Fragilariophyceae</i><br>Class      | <i>rbcL</i>       | -ve                | Pennate diatoms without a raphe, Important components of both planktonic and periphytic communities in freshwater environments. (Dunck et al., 2012)  |

Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i> | Co-occurring taxon                    | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|----------------------|---------------------------------------|-------------------|--------------------|--|
| Nostocaceae Family   | <i>Cyclotella</i>                     | <i>rbcL</i>       | -ve                | A common constituent of spring blooms in temperate lakes which can produce sulphur-based T&Os. Contains high intracellular unsaturated and polyunsaturated fatty acids associated with fish odours in water. Placed in codon A according to Reynolds classification; found in clear, often well-mixed reservoirs. Tolerant to nutrient deficiency and sensitive to pH rises. (Kehoe et al., 2015; Reynolds et al., 2002) |
|                      | <i>Conticribra</i>                    | <i>rbcL</i>       | -ve                | Centric diatoms with high growth rates, are very robust and contain vast accumulated lipids. High tolerance to nutrient limitations. Has been shown to grow best with urea as the major nitrogen source. Essential elements for the cultivation of this genera are calcium phosphate, sodium metasilicate, and biotin. (Couto et al., 2021)  |
|                      | Uncultured <i>Elsteraceae</i>         | 16S               | +ve                | -  |
|                      | <i>Gloeotrichia</i> PYH6              | 16S               | +ve                | Benthic, nitrogen-fixing cyanobacterium that is well known for blooming in eutrophic lakes. Capable of microcystin production. (Carey et al., 2007)  |
|                      | <i>Chlamydomodium</i>                 | <i>rbcL</i>       | +ve                | Tends to grow fixed or adhered to surfaces. Has potential biostimulant and antimicrobial activities. Phenotypic plasticity shows this genus acclimatizes to different nutrient conditions by modulating its biomass. Has been shown to have effective N removal rates. (Touloupakis et al., 2020; Xiong et al., 2022)  |
|                      | <i>Golenkinia</i>                     | <i>rbcL</i>       | +ve                | Green algae are known for their production of lipids, sugars, and carotenoids. Industrial interests in the genus for their phytoremediation of chemical oxygen demand, N and P. (Sisman-Aydin and Simsek, 2022)  |
|                      | <i>Ankyra</i>                         | <i>rbcL</i>       | +ve                | Species of this genera dominate in the clear water stages of stagnant waters. Can endure high UV radiation. (Saini et al., 2022)   |
|                      | Uncultured <i>Verrucomicrobiaceae</i> | 16S               | -ve                | Associated with rhizospheres and endospheres found in plant roots and soil. Implicated in root carbon metabolism. Has also been connected to Lyngbya (non-heterocyst-forming) blooms. (Newitt, 2020; Rajaneesh et al., 2020)   |

Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i>         | Co-occurring taxon                | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|------------------------------|-----------------------------------|-------------------|--------------------|--|
| <i>Nostocaceae</i><br>Family | <i>LD29</i>                       | 16S               | +ve                | Metagenomic analysis has shown their genetic potential to utilise carbon sources and sulphated polysaccharides which are commonly found in phytoplankton. Has been found in association with algal and cyanobacterial cells, suggesting these genera may play an important role in the utilisation of phytoplankton-derived organic matter. (Mohapatra et al., 2020)   |
|                              | <i>Kapabacteriales</i>            | 16S               | +ve                | chemolithotrophic heterotrophic bacteria that has been found co-occurring with <i>Dolichospermum</i> . Genome analysis has revealed that this genus contains a cluster for sulphur assimilation and metabolism, this genus also contains transporters for cobalt, copper, ferrous iron, phosphate, phosphonate, and ammonia. (Al-Saud et al., 2020)  |
|                              | Uncultured <i>Microscillaceae</i> | 16S               | +ve                | This family can degrade carbohydrates and their derivatives to provide energy for cell growth. Has also been shown to be positively correlated with <i>Microcystis</i> , implying a tight interaction. (Chun et al., 2020; Lan et al., 2020)   |
|                              | <i>UKL13-1</i>                    | 16S               | +ve                | Contains more lipid metabolism genes compared with most bacteria. Genomes contain anoxygenic photosynthesis and reaction centre genes, as well as genes for bacteriochlorophyll and carotenoid synthesis. Thus, they're heterotrophs that use phototrophy to drive ATP and NAD(P)H production. Associated with bloom-forming <i>Cyanobacteria</i> , said to benefit through ammonium and organic nutrients released by the cyanobacterium. (Driscoll et al., 2017) |
|                              | <i>Fragilaria</i>                 | <i>rbcL</i>       | -ve                | The genus can be planktonic or benthic, existing as colonial and non-colonial species. Colonial species tend to form ribbon-like colonies. Has been associated with increases in the cell number of <i>Cyanobacteria</i> and biofilms with <i>Cyanobacteria</i> . (Dae-Kyun et al., 2001; Espinosa et al., 2020; Heudre et al., 2019)  |

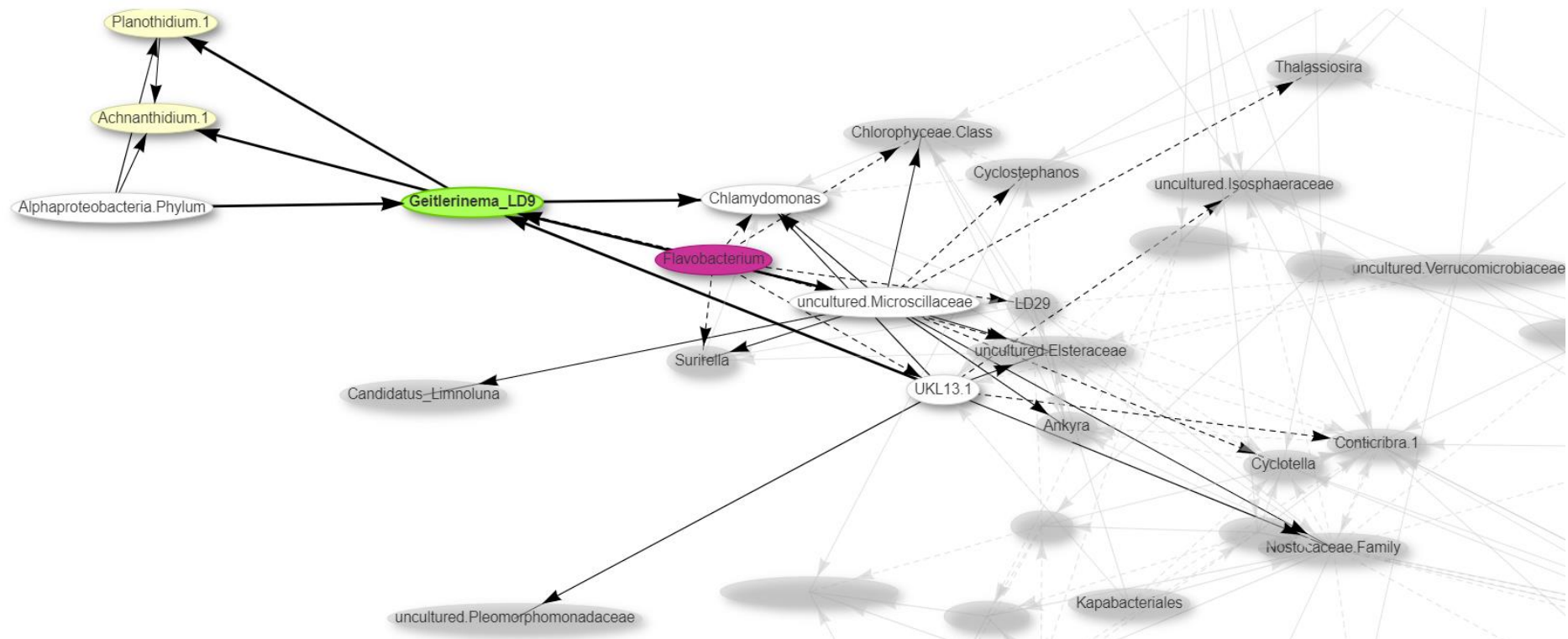


Figure 4.16: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Geitlerinema* LD9 node during the entirety of this study in the Cefni reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978493>.

Table 4.6 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Cefni reservoir.

| <i>Cyanobacteria</i> | Co-occurring taxon                | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|----------------------|-----------------------------------|-------------------|--------------------|--|
| Geitlerinema LD9     | <i>Planothidium</i>               | <i>rbcL</i>       | +ve                | Cells of this genus are usually solitary, living on aquatic plants, algae/ <i>Cyanobacteria</i> , or inorganic substratum. (Stancheva, 2019)   |
|                      | <i>Achnantheidium</i>             | <i>rbcL</i>       | +ve                | Prostrate diatoms that often reported from the littoral zone of oligotrophic lakes, living in the epipelon, epipsammon, or epilithon, preferring alkaline waters. Normally inhabit depressions and cervices on grains where they obtain protection. (Kingston, 2003)   |
|                      | <i>Chlamydomonas</i>              | <i>rbcL</i>       | +ve                | Known to produce polyolglycosides, sterols, phenols, and fatty acids along with other compounds. Thought to use these compounds in chemotaxonomic relationships with other lineages. Additions of <i>Aphanizomenon</i> or cylindrospermopsin to <i>Chlamydomonas</i> induce genes typically upregulated under P limitation and give rise to extracellular APase activity that <i>Cyanobacteria</i> can use. (Bar-Yosef et al., 2010; Khanh Tran et al., 2019)      |
|                      | <i>Flavobacterium</i>             | 16S               | -ve                | Strictly aerobic, usually found inhabiting cyanobacterial aggregates during blooms. Known 2-MIB degrading genus said to maintain cyanobacterial blooms dominance by catalysing the turnover of complex organic matters released by the cyanobacteria to recycle previously loaded nutrient sources. (Cai et al., 2018)   |
|                      | Uncultured <i>Microscillaceae</i> | 16S               | +ve                | This family can degrade carbohydrates and their derivatives to provide energy for cell growth. Has also been shown to be positively correlated with <i>Microcystis</i> , implying a tight interaction. (Chun et al., 2020; Lan et al., 2020)   |
|                      | UKL 13-1                          | 16S               | +ve                | Contains more lipid metabolism genes compared with most bacteria. Genomes contain anoxygenic photosynthesis and reaction centre genes, as well as genes for bacteriochlorophyll and carotenoid synthesis. Thus, they're heterotrophs that use phototrophy to drive ATP and NAD(P)H production. Associated with bloom-forming <i>Cyanobacteria</i> , said to benefit through ammonium and organic nutrients released by the cyanobacterium. (Driscoll et al., 2017) |
|                      | <i>Alphaproteobacteria</i> Phylum | 16S               | +ve                | Contains many genera that are known geosmin degraders, with the potential to be used in the bioremediation of geosmin and 2-MIB in aquaculture and water treatment facilities, often in bacterial consortiums in specific niches. (Churro et al., 2020)  |

#### 4.3.2.4 – Dolwen

In the Dolwen reservoir, 33 pairwise comparisons yielded statistically significant co-occurrence (<0.01); seven of which were negatively associated, and 26 co-occurring taxa were positively associated. The network can be accessed online via the link contained within Table 4.3. Typically, there were seven disconnected networks in the overall network; the two largest disconnected networks were comprised of five taxa. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.4.

One of the disconnected networks contained two cyanobacterial taxa, *Nostocaceae* family and *Aphanizomenon* MDT14a; three of the associated nodes had positive co-occurrences, and one node belonging to Bacteria unknown had a negative co-occurrence (Figure 4.17). *Dolichospermum* NIES41 was contained in a much smaller cyanosphere, consisting of three nodes, and two genera had negative co-occurrences (Figure 4.18).

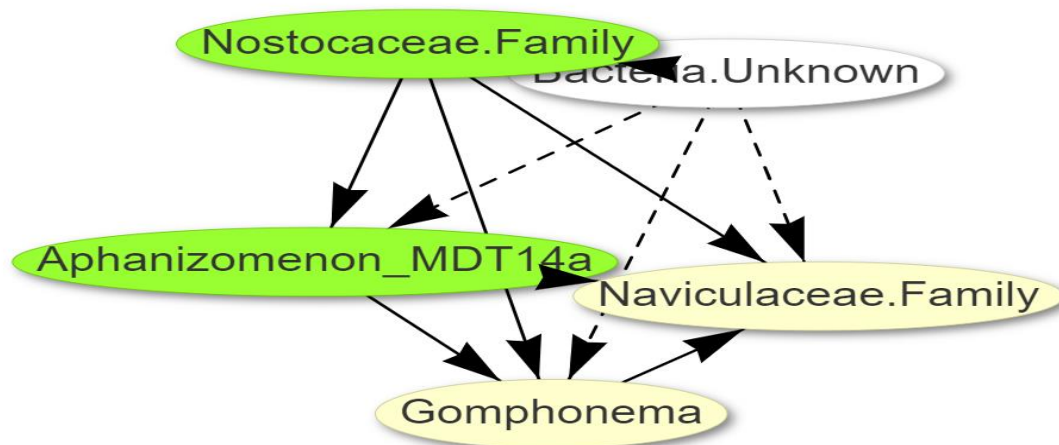


Figure 4.17: Subset of the significant (< 0.1) co-occurring taxa associated with the *Aphanizomenon* MDT14a and *Nostocaceae* family nodes during the entirety of this study in the Dolwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978497>.



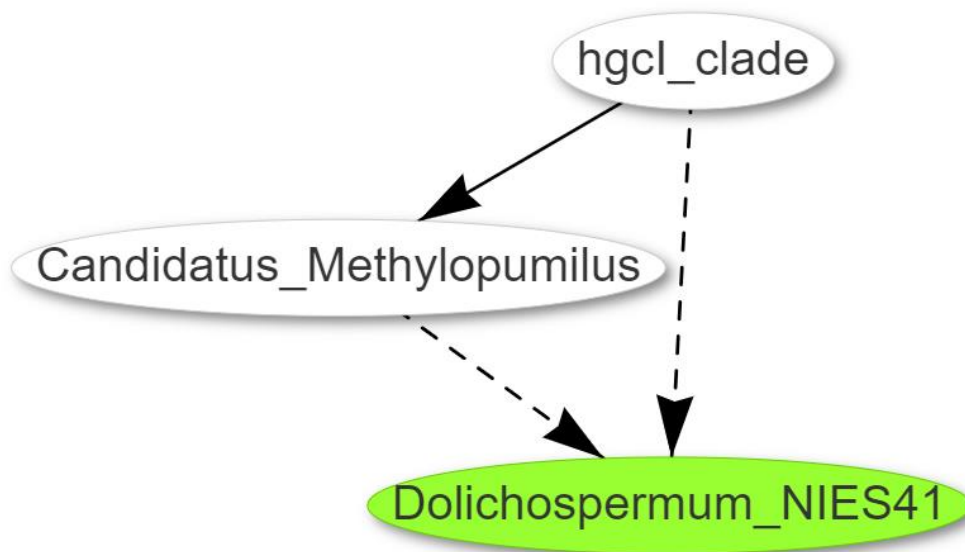


Figure 4.18: Subset of the significant ( $< 0.1$ ) co-occurring taxa associated with the *Dolichospermum* NIES41 node during the entirety of this study in the Dolwen reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978497>.

Table 4.7: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Dolwen reservoir.

| <i>Cyanobacteria</i>         | Co-occurring taxon                | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|------------------------------|-----------------------------------|-------------------|--------------------|---|
| <i>Aphanizomenon</i> MDT14a  | <i>Gomphonema</i>                 | <i>rbcL</i>       | +ve                | Produces brown algal pheromones (hormosirene and two dictyopterens), however, functions of these volatile hydrocarbons in freshwater diatoms are unknown. Shown to be less favoured by ammonia abundance but have been shown to be closely correlated to the waterflow and nitrate abundance. (Veraart et al., 2008)  |
|                              | <i>Naviculaceae</i> Family        | <i>rbcL</i>       | +ve                | An Ochrophyta family common in moderate to good quality streams. General term for many motile diatom species. Members of this family have been found in a symbiotic association with Cyanobacteria (e.g., <i>Braarudosphaera</i> and nitrogen-fixing cyanobacterium, <i>Candidatus A. thalassa</i> UCYN-A). (Moore et al., 2019)  |
|                              | <i>Nostocaceae</i> Family         | 16S               | +ve                | Genera of this family are heterocyst forming. Includes many genera that possess <i>geoA</i> , including <i>Anabaena</i> and <i>Nostoc</i> species. (Yamada et al., 2015)  |
| <i>Nostocaceae</i> Family    | <i>Aphanizomenon</i> MDT14a       | 16S               | +ve                | Filamentous capable of nitrogen fixation found in eutrophic waters, assigned to codon H in Reynolds classification. Known geosmin and microcystin producer. (Reynolds et al., 2002)   |
|                              | <i>Gomphonema</i>                 | <i>rbcL</i>       | +ve                | Produces brown algal pheromones (hormosirene and two dictyopterens), however, functions of these volatile hydrocarbons in freshwater diatoms are unknown. Shown to be less favoured by ammonia abundance but have been shown to be closely correlated to the waterflow and nitrate abundance. (Veraart et al., 2008)  |
|                              | <i>Naviculaceae</i> Family        | <i>rbcL</i>       | +ve                | An Ochrophyta family common in moderate to good quality streams. General term for many motile diatom species. Members of this family have been found in a symbiotic association with Cyanobacteria (e.g., <i>Braarudosphaera</i> and nitrogen-fixing cyanobacterium, <i>Candidatus A. thalassa</i> UCYN-A). (Moore et al., 2019)  |
| <i>Dolichospermum</i> NIES41 | <i>hgcl</i> clade                 | 16S               | -ve                | Strong genetic ability to take carbohydrate and N-rich organic compounds. Also, the potential to utilize sunlight via actinorhodopsin which might promote anaplerotic carbon fixation, indicative of both heterotrophic and autotrophic lifestyles. Often found in epilimnetic waters and favours high water transparency and low DOC concentrations. This clade has been related to the decomposition of <i>Cyanobacteria</i> . (Luo et al., 2022) |
|                              | <i>Candidatus Methylopusillus</i> | 16S               | -ve                | Abundant in cold hypolimnion waters, growing psychrophilically. Has been shown to be inversely proportional to cyanobacterial growth. Capable of nitrification and denitrification and metabolising methane and sulphur. (Wei et al., 2022)   |

#### 4.3.2.5 – Llandegfedd

In the Llandegfedd reservoir, 121 pairwise comparisons yielded statistically significant co-occurrence (<0.01); 37 were determined to be negatively associated, and 84 co-occurring taxa were deemed positively associated. The network can be accessed online via the link contained within Table 4.3. One extensive network contained most nodes (109), and then five tight networks surrounded the more extensive network containing a maximum of three nodes. The two Cyanobacteria genera and T&O degrading bacteria were constituents of the extensive network. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.4.

Both cyanospheres interlinked, forming a positive co-occurrence from *Cyanobium* PCC-6307 to *Aphanizomenon* MDT14a. In the *Aphanizomenon* MDT14a cyanosphere, there were six associated co-occurring taxa, five of the associations were positive, and one belonging to the T&O-degrading bacteria *Pseudomonas* was negative (Figure 4.19). Similarly, in the *Cyanobium* PCC-6307 cyanosphere there were five co-occurring taxa, four of which were positively influential and one negatively influential belonging to *Schlesneria* (Figure 4.20).

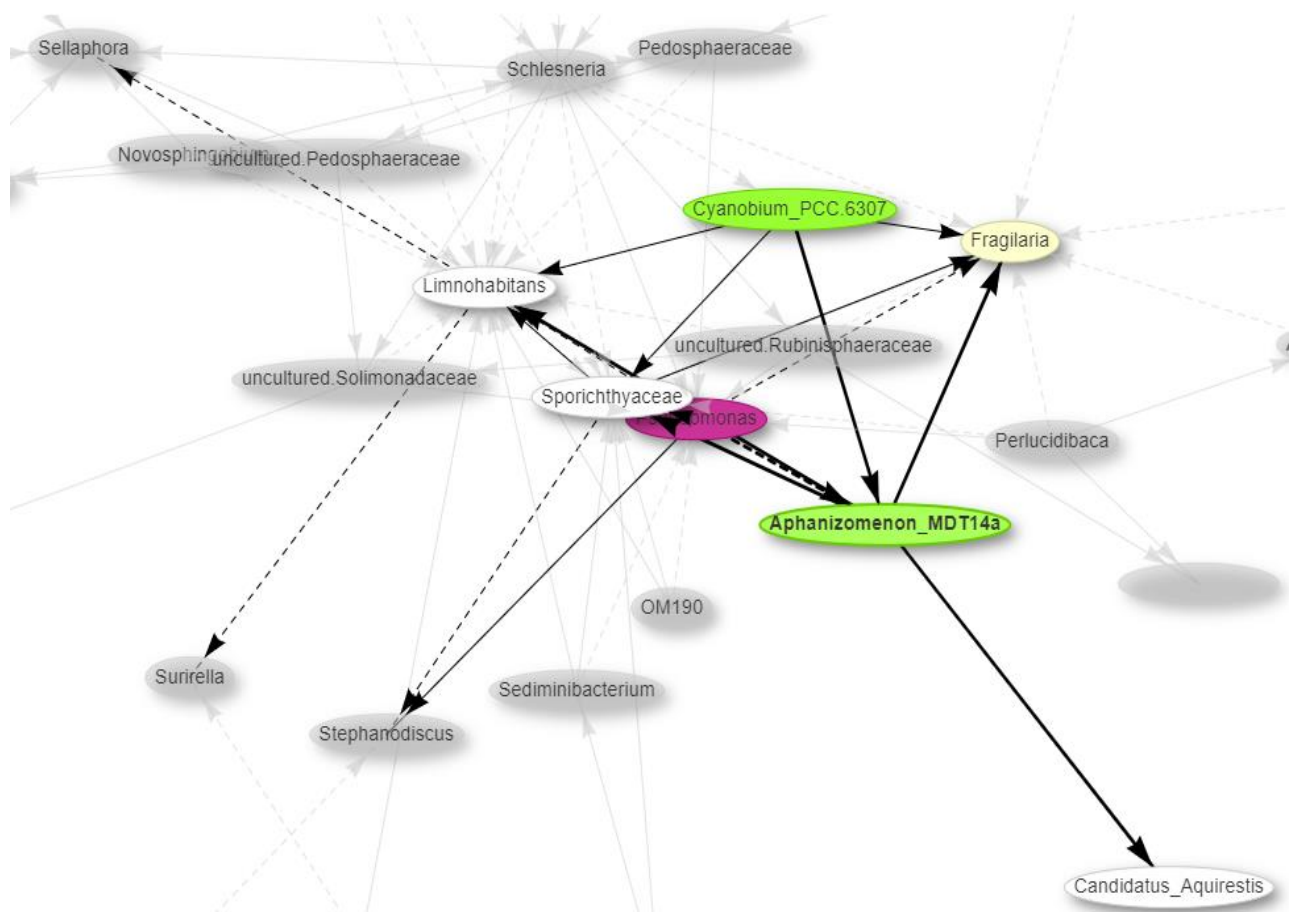


Figure 4.19: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Aphanizomenon* MDT14a node during the entirety of this study in the Llandegfedd reservoir. Cyanobacterial nodes are illustrated as green, T&O degrading bacterial nodes are indicated as purple and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978520>.

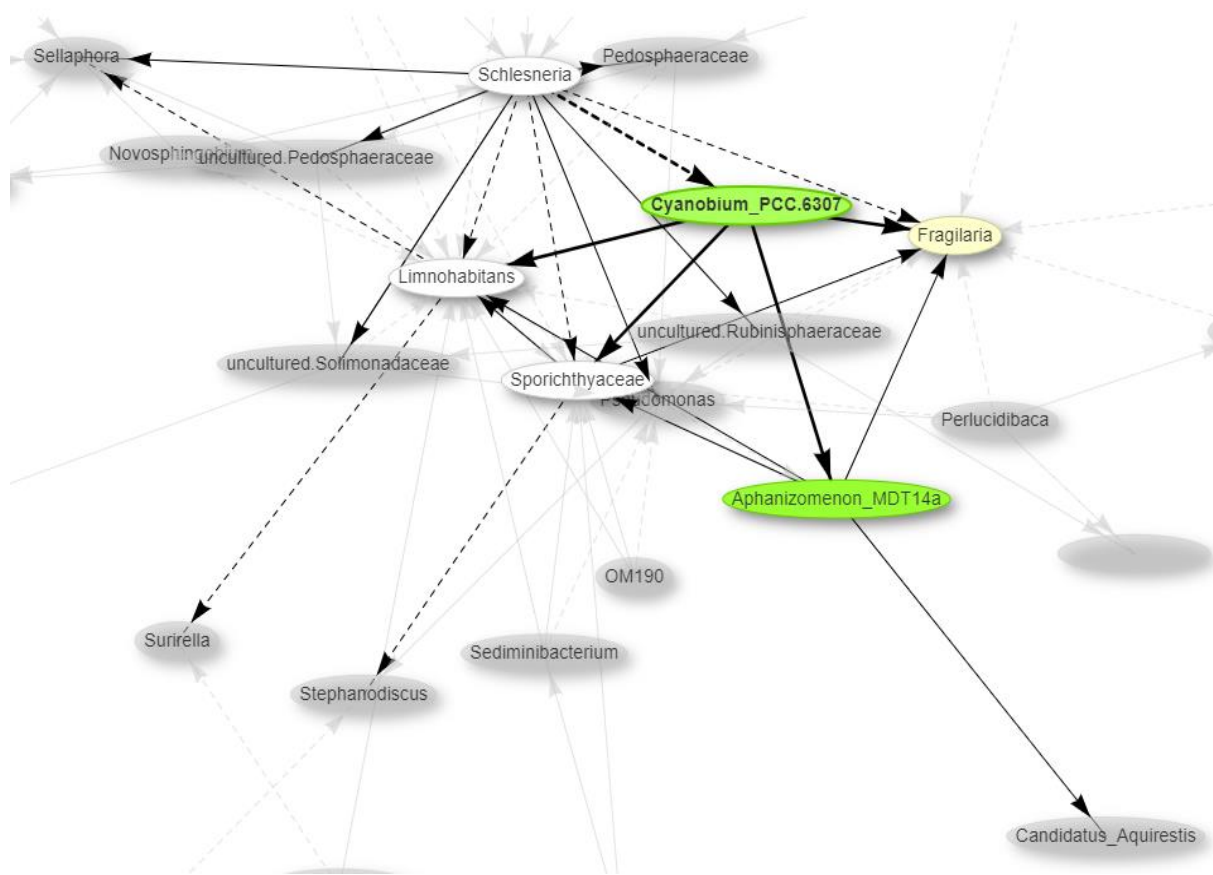


Figure 4.20: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Llandegfedd reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978520>.

Table 4.8: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Llandegfedd reservoir.

| <i>Cyanobacteria</i>           | Co-occurring taxon            | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|--------------------------------|-------------------------------|-------------------|--------------------|---|
| <i>Aphanizomenon</i><br>MDT14a | <i>Candidatus Aquirestis</i>  | 16S               | +ve                | Chemoorganotrophic bacteria found in stagnant shallow freshwater habitats, often related to high levels of dissolved organic C concentration derived from algal and cyanobacterial blooms. (Farkas et al., 2020)  |
|                                | <i>Sporichthyaceae</i> Family | 16S               | +ve                | This family only contains the genus <i>Sporichthya</i> , which is comprised of two species. Members are motile facultative anaerobes and have a cell wall containing large amounts of LL-diaminopimelic acid. Produce aerial hyphae made up of rod-shaped spores. No substrate mycelium. The main habitat of this family appears to be soil. (Tamura, 2014)   |
|                                | <i>Limnohabitans</i>          | 16S               | +ve                | Found to thrive along with cyanobacterial blooms in spring or summer, using phytoplankton-derived organic material as they substrate for growth. Culture-independent studies have revealed this genus to be the second most abundant (11%) taxa in the cyanosphere. (Luo et al., 2022; Seok Jea Youn et al., 2020)  |
|                                | <i>Fragilaria</i>             | <i>rbcL</i>       | +ve                | The genus can be planktonic or benthic, existing as colonial and non-colonial species. Colonial species tend to form ribbon-like colonies. Has been associated with increases in the cell number of <i>Cyanobacteria</i> and in biofilms with <i>Cyanobacteria</i> . (Dae-Kyun et al., 2001; Espinosa et al., 2020; Heudre et al., 2019)  |
|                                | <i>Pseudomonas</i>            | 16S               | -ve                | Known geosmin and 2-MIB degrading genus. Converting geosmin to several oxidation compounds, like, ketogeosmins. Contains 2-methyl-2-bornene, an enzyme capable of 2-MIB degradation. Three strains of <i>Pseudomonas</i> have been identified to possess the <i>geoA</i> , although it is unknown if they are effective geosmin producers. (Churro et al., 2020; Eaton and Sandusky, 2010; Shao and Du, 2020) |
|                                | <i>Cyanobium</i> PCC-6307     | 16S               | +ve                | Type of picocyanobacteria; species vary from oval to cylindrical cells, with little or no evident mucilage exterior. Not a known producer of geosmin, although can produce 2-MIB and microcystins. (Clerc et al., 2022; Hojun et al., 2021; Śliwińska-Wilczewska et al., 2018)  |

Table 4.8 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Llandegfedd reservoir.

| <i>Cyanobacteria</i>      | Co-occurring taxon            | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|---------------------------|-------------------------------|-------------------|--------------------|---|
| <i>Cyanobium</i> PCC-6307 | <i>Aphanizomenon</i> MDT14a   | 16S               | +ve                | Filamentous capable of nitrogen fixation found in eutrophic waters, assigned to codon H in Reynolds classification. Known geosmin and microcystin producer. (Reynolds et al., 2002)   |
|                           | <i>Sporichthyaceae</i> Family | 16S               | +ve                | This family only contains the genus <i>Sporichthya</i> , which is comprised of two species. Members are motile facultative anaerobes and have a cell wall containing large amounts of LL-diaminopimelic acid. Produce aerial hyphae made up of rod-shaped spores. No substrate mycelium. The main habitat of this family appears to be soil. (Tamura, 2014) |
|                           | <i>Limnohabitans</i>          | 16S               | +ve                | Found to thrive along with cyanobacterial blooms in spring or summer, using phytoplankton-derived organic material as they substrate for growth. Culture-independent studies have revealed this genus to be the second most abundant (11%) taxa in the cyanosphere. (Luo et al., 2022; Seok Jea Youn et al., 2020)  |
|                           | <i>Fragilaria</i>             | <i>rbcL</i>       | +ve                | The genus can be planktonic or benthic, existing as colonial and non-colonial species. Colonial species tend to form ribbon-like colonies. Has been associated with increases in the cell number of <i>Cyanobacteria</i> and biofilms with <i>Cyanobacteria</i> . (Dae-Kyun et al., 2001; Espinosa et al., 2020; Heudre et al., 2019)                       |
|                           | <i>Schlesneria</i>            | 16S               | -ve                | Moderately acidophilic <i>Planctomycetes</i> growing at pH values between 4 and 7, with an optimum at pH 5 - 6. Chemo-organotrophic facultative aerobes, positive for dissimilatory nitrate reduction. (Kulichevskaya et al., 2007)   |

#### 4.3.2.6 – Llwyn On

In the Llwyn On reservoir 31 pairwise comparisons yielded statistically significant co-occurrence ( $<0.01$ ); six of which were determined to be negatively associated, and 25 co-occurring taxa were deemed positively associated. The network can be accessed online via the link contained within Table 4.3. In general, there were 15 disconnected networks in the overall network. The largest network contained the two significantly co-occurring *Cyanobacteria* genera, *Aphanizomenon* NIES81 and *Cyanobium* PCC-6307. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.4.

The *Aphanizomenon* NIES81 cyanosphere included two positively co-occurring taxa (Figure 4.21) including *Cyanobium* PCC-6307; the *Cyanobium* PCC-6307 cyanosphere involved four co-occurring taxa, two of which were positive and two which were negatively influential (Figure 4.22).

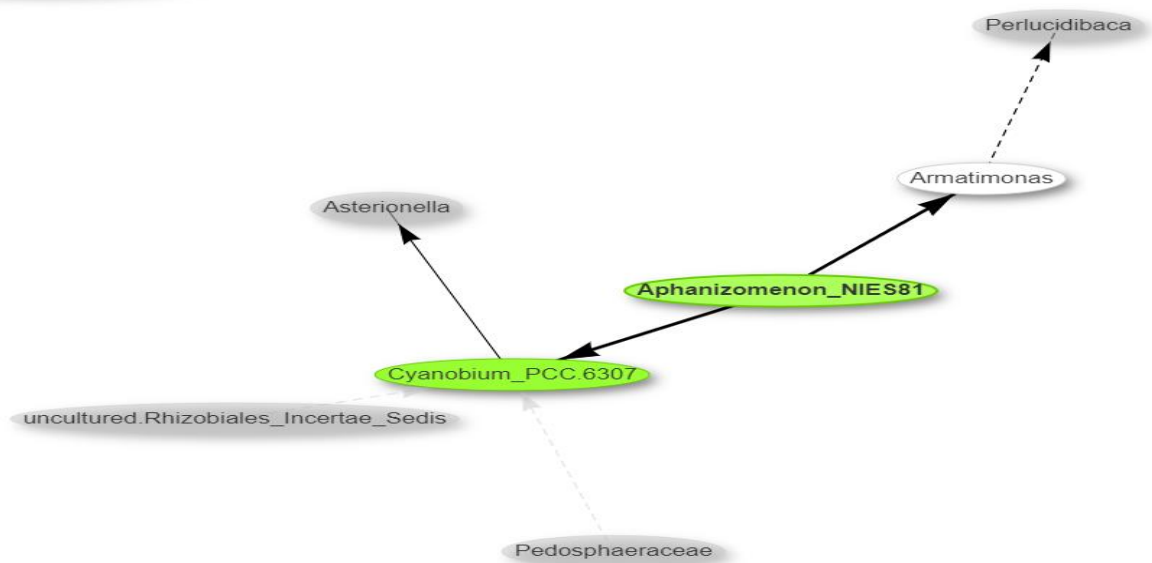


Figure 4.21: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Aphanizomenon* NIES81 node during the entirety of this study in the Llwyn On reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978527>.



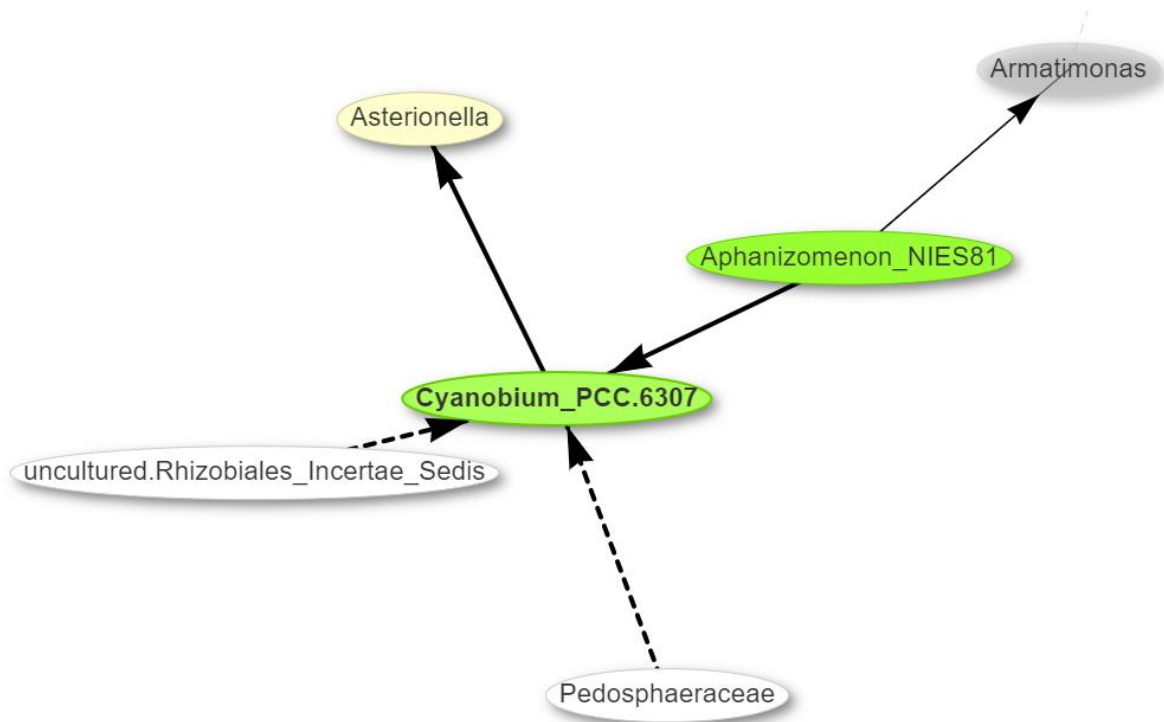


Figure 4.22: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Llwyn On reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978527>.

Table 4.9: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within Llwyn On reservoir.

| <i>Cyanobacteria</i>        | Co-occurring taxon                           | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|-----------------------------|--|-------------------|--------------------|--|
| <i>Aphanizomenon</i> NIES81 | <i>Armatimonas</i>                           | 16S               | +ve                | Aerobic, non-motile, ovoid to rod-shaped often found in mesophilic waters. Previously shown to be positively correlated with TP. (Xu et al., 2018)   |
|                             | <i>Cyanobium</i> PCC-6307                    | 16S               | +ve                | Type of picocyanobacteria; species vary from oval to cylindrical cells, with little or no evident mucilage exterior. Not a known producer of geosmin, although can produce 2-MIB and microcystins. (Clerc et al., 2022; Hojun et al., 2021; Śliwińska-Wilczewska et al., 2018)   |
| <i>Cyanobium</i> PCC-6307   | <i>Asterionella</i>                          | <i>rbcL</i>       | +ve                | Belongs to codon C in the Reynolds classification. Characteristic member of nutrient-rich temperate lakes. Rapid increases of this species during March - April leads to the extraction of major anions (nitrate, phosphate, and silicate), resulting in nutrient depletion and altering nutrient ratios. (Krivtsov et al., 2000; Reynolds et al., 2002) |
|                             | Uncultured <i>Rhizobiales Incertae Sedis</i> | 16S               | -ve                | family of nitrogen-fixing bacteria. Members of the <i>Rhizobiales</i> have been reported to degrade cyanobacterial-derived particulate organic matter. (Huang et al., 2022; Millar et al., 2022)   |
|                             | <i>Pedosphaeraceae</i>                       | 16S               | -ve                | Previously found to be a keystone species in rhizospheres found in soils. They play an important role in sediment nutrient circulation with high metal resistance potentials. (Yuan et al., 2022)  |
|                             | <i>Aphanizomenon</i> NIES81                  | 16S               | +ve                | Filamentous capable of nitrogen fixation found in eutrophic waters, assigned to codon H in Reynolds classification. Known geosmin and microcystin producer. (Reynolds et al., 2002)  |

#### 4.3.2.7 – Pentwyn

In the Pentwyn reservoir, 135 pairwise comparisons yielded statistically significant co-occurrence ( $<0.01$ ); 48 were determined to be negatively associated, and 87 co-occurring taxa were deemed positively associated. The network can be accessed online via the link contained within Table 4.3. In general, there were five networks tightly compacted into the overall network; four were comprised of two nodes, one of which contained *Dolichospermum* NIES41 and the diatom *Melosira* (Figure 4.24). The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.4.

*Cyanobium* PCC-6307 had more taxa involved within its cyanosphere (Figure 4.23) than *Dolichospermum* NIES41 (Figure 4.24). Eight taxa were directly included within the *Cyanobium* PCC-6307 cyanosphere, five of which were negatively associated and three were positively co-occurring.

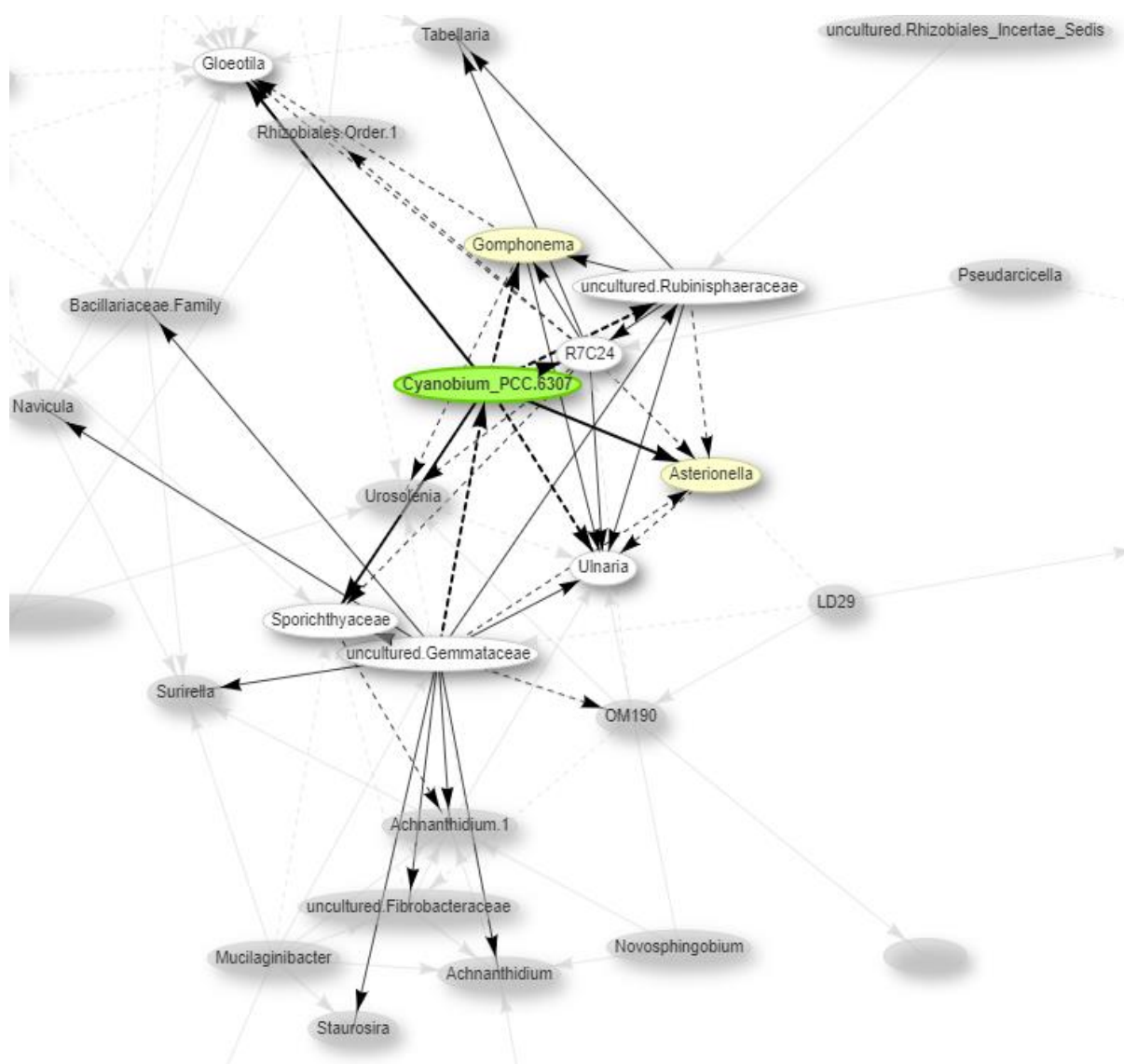


Figure 4.23: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Pentwyn reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978532>.

Table 4.10: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Pentwyn reservoir.

| <i>Cyanobacteria</i>         | Co-occurring genus                  | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring genus  |
|------------------------------|-------------------------------------|-------------------|--------------------|---|
| <i>Cyanobium</i><br>PCC-6307 | Uncultured <i>Rubinisphaeraceae</i> | 16S               | -ve                | A heterotrophic bacterium that has been reported in methanotrophic accumulating poly(3-hydroxybutyrate) systems. Not known to perform anammox. (Cattaneo et al., 2022; Khairunisa et al., 2022)   |
|                              | <i>R7C24</i>                        | 16S               | -ve                | -   |
|                              | <i>Gomphonema</i>                   | <i>rbcL</i>       | -ve                | Produces brown algal pheromones (hormosirene and two dictyopterens), however, functions of these volatile hydrocarbons in freshwater diatoms are unknown. Shown to be less favoured by ammonia abundance but have been shown to be closely correlated to the water flow and nitrate abundance. (Veraart et al., 2008)                                       |
|                              | <i>Ulnaria</i>                      | <i>rbcL</i>       | -ve                | Present in stable meso to eutrophic water columns, with circumneutral to slightly alkaline conditions. Previously included in the large genus <i>Synedra</i> . Species are mesoaprobic. Normally associated with increased productivity. (Thacker and Karthick, 2022)   |
|                              | <i>Sporichthyaceae</i> Family       | 16S               | +ve                | This family only contains the genus <i>Sporichthya</i> , which is comprised of two species. Members are motile facultative anaerobes and have a cell wall containing large amounts of LL-diaminopimelic acid. Produce aerial hyphae made up of rod-shaped spores. No substrate mycelium. The main habitat of this family appears to be soil. (Tamura, 2014) |
|                              | <i>Asterionella</i>                 | <i>rbcL</i>       | +ve                | Belongs to codon C in the Reynolds classification. Characteristic member of nutrient-rich temperate lakes. Rapid increases of this species during March – April leads to the extraction of major anions (nitrate, phosphate, and silicate), resulting in nutrient depletion and altering nutrient ratios. (Krivtsov et al., 2000; Reynolds et al., 2002)    |
|                              | <i>Gloeotila</i>                    | <i>rbcL</i>       | +ve                | Has synonymously been recognised as <i>Geminella</i> . Found in deep, well-mixed epilimnia. Tolerant to light deficiency and sensitive to nutrient deficiency. Assigned to codon T according to Reynolds classification. (Fernandes et al., 2021; Reynolds et al., 2002)  |

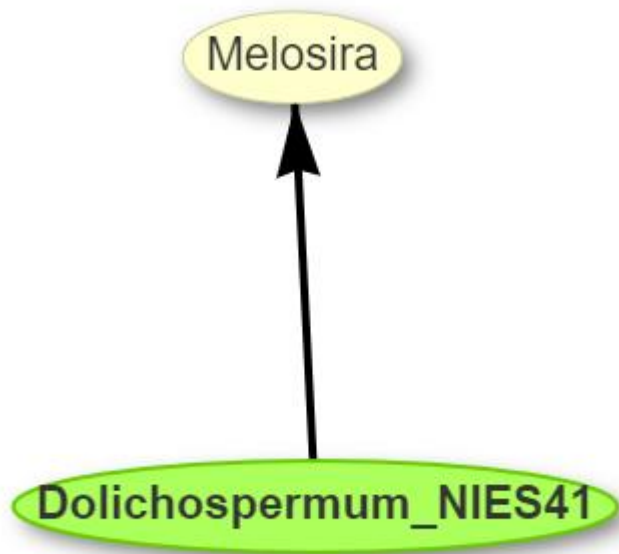


Figure 4.24: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Dolichospermum* NIES41 node during the entirety of this study in the Pentwyn reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978532>.

Table 4.9 cont'd: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Pentwyn reservoir.

| <i>Cyanobacteria</i>         | Co-occurring taxon            | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|------------------------------|-------------------------------|-------------------|--------------------|--|
| <i>Cyanobium</i> PCC-6307    | Uncultured <i>Gemmataceae</i> | 16S               | -ve                | Strictly aerobic, chemo-organotrophic <i>Planctomycetes</i> which occur singly or in pairs or are assembled in large rosette-like clusters and dendriform-like structures. All members of this family have large genome sizes, and hold repertoires of carbohydrate-active enzymes, including many unclassified putative glycoside hydrolases; high glycolytic potential/ utilises polysaccharides such as xylan, laminarin, lichenin and starch. Some members can degrade chitin and utilise it as a source of nitrogen. (Kulichevskaya et al., 2020) |
| <i>Dolichospermum</i> NIES41 | <i>Melosira</i>               | <i>rbcL</i>       | +ve                | Generally benthic, occurring in naturally eutrophic and polluted streams and lakes, commonly entrained into the plankton. Chain-forming diatom genus. Reportedly contains high intracellular unsaturated and polyunsaturated fatty acids associated with fish/ oily/ cucumber odours in waters. (Van de Vijver and Crawford, 2020)   |

#### 4.3.2.8 – Plas Uchaf

In the Plas Uchaf reservoir, 86 pairwise comparisons yielded statistically significant co-occurrence ( $<0.01$ ), 14 were determined to be negatively associated, and 72 co-occurring taxa were deemed positively associated. The network can be accessed online via the link contained within Table 4.3. In general, there were seven networks tightly compacted into the overall network; six were comprised of between two – four nodes and did not contain any cyanobacterial taxa. The largest network contained the majority of nodes (71), and both significantly co-occurring cyanobacterial taxa. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.4.

In the *Nostocaceae* family cyanosphere, four associated taxa had a negative influence (Figure 4.25). For *Microcystis* PCC-7914, only two genera were included in the cyanosphere, both of which positively influenced *Microcystis* PCC-7914 (Figure 4.26).



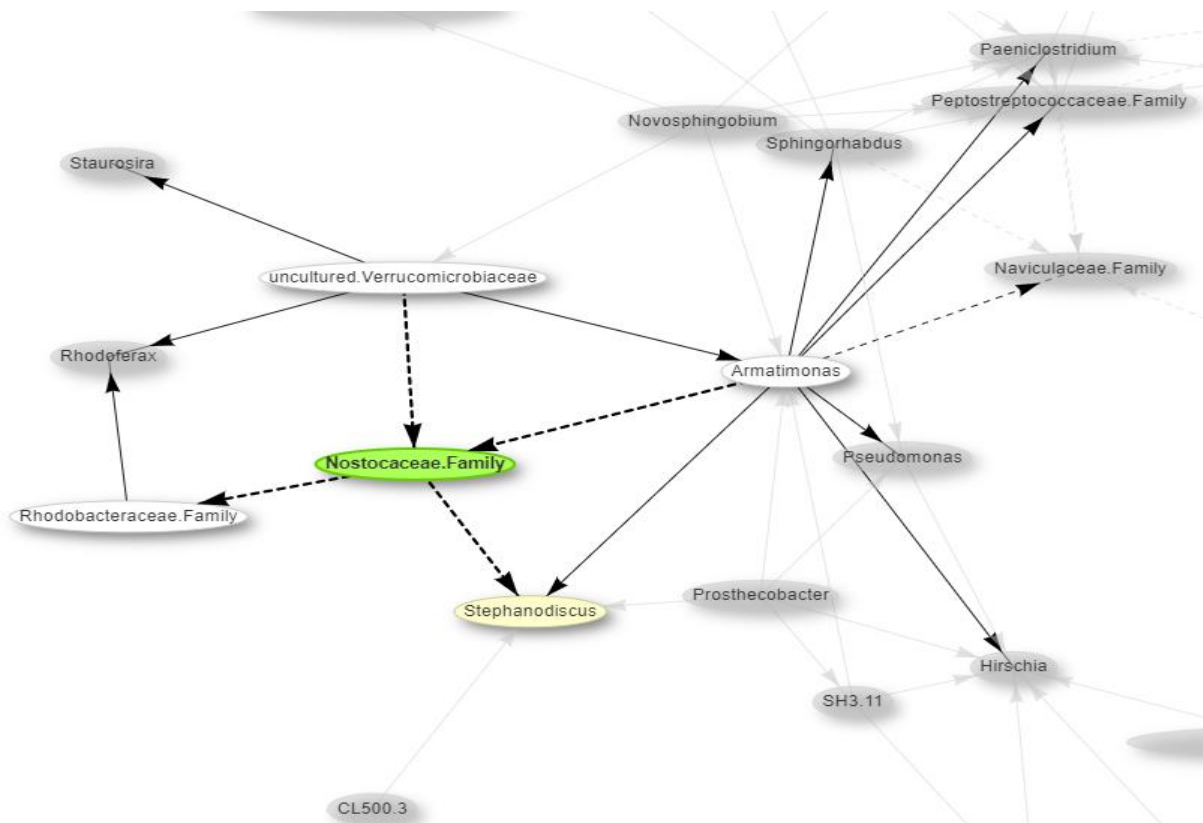


Figure 4.25: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Nostocaceae* family node during the entirety of this study in the Plas Uchaf reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each genus. Link:

<https://rpubs.com/ASHooper/978536>.

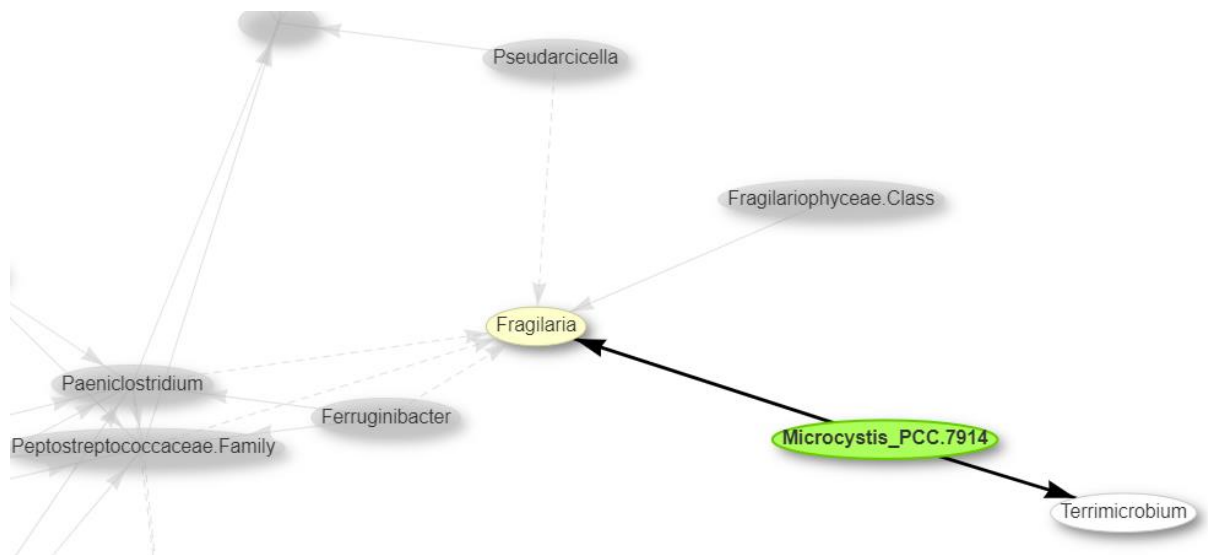


Figure 4.26: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Microcystis* PCC-7914 node during the entirety of this study in the Plas Uchaf reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978536>.

Table 4.11: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Plas Uchaf reservoir.

| <i>Cyanobacteria</i>         | Co-occurring taxon                       | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon   |
|------------------------------|--|-------------------|--------------------|--|
| <i>Nostocaceae</i><br>Family | <i>Rhodobacteraceae</i> Family           | 16S               | -ve                | Mainly consists of aerobic photo- and chemoheterotrophs, purple non-sulphur bacteria that perform photosynthesis, and can also be anaerobic. Play roles in sulphur, nitrogen, and carbon cycles, decomposing various compounds and generating secondary metabolites. Produces EPS that retains large amounts of water, some species synthesise ectoine (protein protectant), and other species can synthesise polyhydroxyalkanoates (PHAs) (compounds accumulated by many microorganisms). (Azpiazu-muniozgueren et al., 2022) |
|                              | <i>Stephanodiscus</i>                    | <i>rbcL</i>       | -ve                | Typically, colonial and placed in Reynolds codon D classification; mostly found in shallow, nutrient-enriched, well-ventilated waters, liable to be turbid. Small-celled ( $\leq 103 \mu\text{m}^3$ in volume) and fast-growing. (Reynolds et al., 2002)   |
|                              | Uncultured<br><i>Verrucomicrobiaceae</i> | 16S               | -ve                | Associated with rhizospheres and endospheres found in plant roots and soil. Implicated in root carbon metabolism. Has also been connected to <i>Lyngbya</i> (non-heterocyst-forming) blooms. (Newitt, 2020; Rajaneesh et al., 2020)  |
|                              | <i>Armatimonas</i>                       | 16S               | -ve                | Aerobic, non-motile, ovoid to rod-shaped often found in mesophilic waters. Previously shown to be positively correlated with TP. (Xu et al., 2018)   |
| <i>Microcystis</i> PCC-7914  | <i>Terrimicrobium</i>                    | 16S               | +ve                | Non-motile, anaerobic rods. Ferments various sugars, but not fatty acids, alcohols (i.e., geosmin) or amino acids. Previous research has identified this genus with <i>Nostoc</i> epibacteria. (Qiu et al., 2014; Satjarak et al., 2021)   |
|                              | <i>Fragilaria</i>                        | <i>rbcL</i>       | +ve                | The genus can be planktonic or benthic, existing as colonial and non-colonial species. Colonial species tend to form ribbon-like colonies. Has been associated with increases in the cell number of <i>Cyanobacteria</i> and biofilms with <i>Cyanobacteria</i> . (Dae-Kyun et al., 2001; Espinosa et al., 2020; Heudre et al., 2019)  |

#### 4.3.2.9 – Pontsticill

In the Pontsticill reservoir, 63 pairwise comparisons yielded statistically significant co-occurrence ( $<0.01$ ), 16 of which were determined to be negatively associated, and 47 co-occurring taxa were deemed positively associated. The network can be accessed online via the link contained within Table 4.3. In general, nine networks were tightly compacted into the overall network; eight were comprised of between two – five. The largest network contained the majority of nodes (41), and one significantly co-occurring cyanobacterial taxa. There was no evidence of cyanospheres overlapping. The potential importance of the significantly co-occurring taxa directly connected to cyanobacterial nodes can be seen as illustrated in Table 4.4.

The *Cyanobium* PCC-6307 cyanosphere contained four co-occurring taxa, three of which were positively influential, and one that was negatively influential, namely, uncultured *Rubinisphaeraceae* (Figure 4.27). The other two cyanospheres were much smaller and only had two occupants. The *Dolichospermum* NIES41 cyanosphere had positive associations with *DEV114* and *Cyclostephanos* (Figure 4.28). The *Snowella* 0TU37S04 cyanosphere had one positive association with *Tabellaria* and one negative association with a member from the *Sporichthyaceae* family (Figure 4.29).

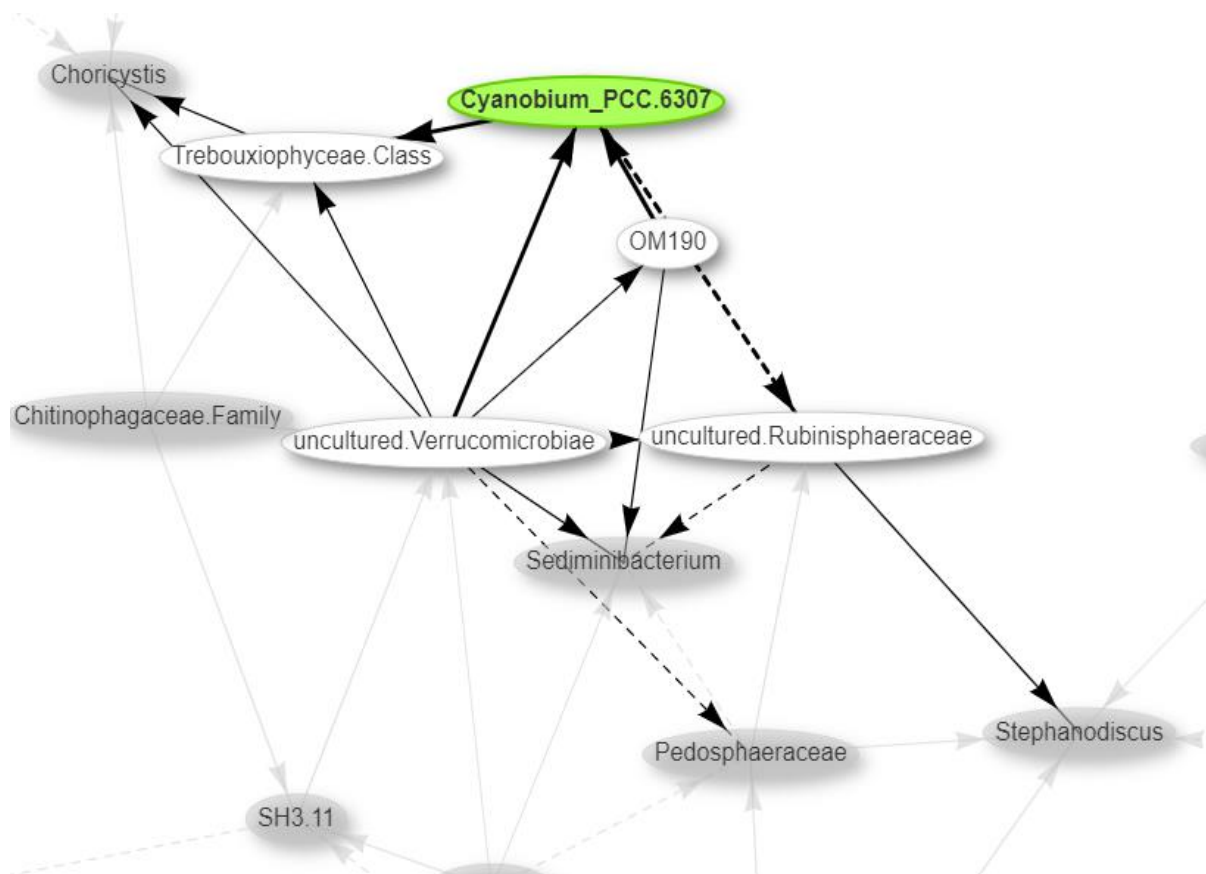


Figure 4.27: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Cyanobium* PCC-6307 node during the entirety of this study in the Pontsticill reservoir. Cyanobacterial nodes are illustrated in green. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978544>.

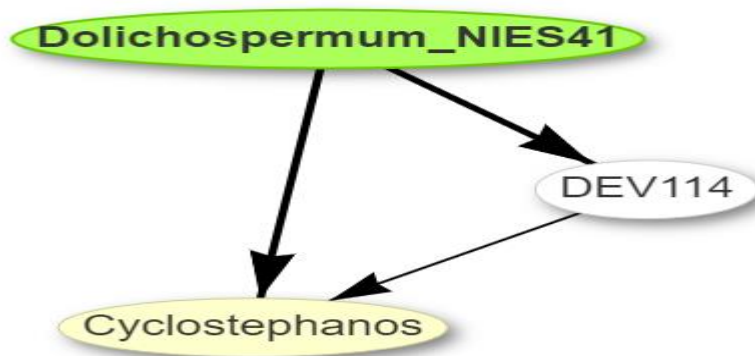


Figure 4.28: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Dolichospermum* NIES41 node during the entirety of this study in the Pontsticill reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978544>.

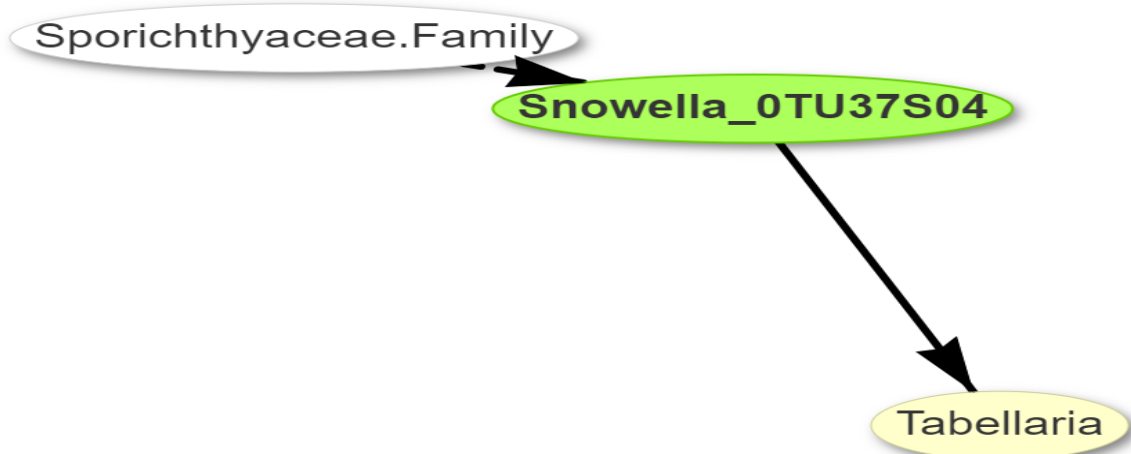


Figure 4.29: Subset of the significant ( $< 0.01$ ) co-occurring taxa associated with the *Snowella* out37S04 node during the entirety of this study in the Pontsticill reservoir. Cyanobacterial nodes are illustrated as green and diatom nodes are coloured yellow. The edges connecting the nodes show negative (indicated with dashed black lines) and positive (straight black lines) associations between each taxon. Link: <https://rpubs.com/ASHooper/978544>.

Table 4.12: Representation of taxa nodes associated with cyanobacterial nodes and their influence and importance in the cyanosphere within the Pontsticill reservoir.

| <i>Cyanobacteria</i>            | Co-occurring taxon                  | 16S / <i>rbcL</i> | Co-occur influence | Importance of co-occurring taxon  |
|---------------------------------|-------------------------------------|-------------------|--------------------|---|
| <i>Cyanobium</i><br>PCC-6307    | Uncultured <i>Rubinisphaeraceae</i> | 16S               | -ve                | A heterotrophic bacterium that has been reported in methanotrophic accumulating poly(3-hydroxybutyrate) systems. Not known to perform anammox. (Cattaneo et al., 2022; Khairunisa et al., 2022)   |
|                                 | <i>Trebouxiophyceae</i> Class       | <i>rbcL</i>       | +ve                | Some members of this class can produce mycoporine-like amino acids, used as photoprotective substances. (Karsten et al., 2005)  |
|                                 | Uncultured <i>Verrucomicrobiae</i>  | 16S               | +ve                | Associated with rhizospheres and endospheres found in plant roots and soil. Implicated in root carbon metabolism. Has also been connected to <i>Lyngbya</i> (non-heterocyst-forming) blooms. (Newitt, 2020; Rajaneesh et al., 2020)   |
|                                 | <i>OM190</i>                        | 16S               | +ve                | Members of this clade are usually associated with microalgae. Produces secondary metabolites, including antimicrobial compounds. It has been shown that diatoms produce fucose-containing sulphated polysaccharides which this genus can supposedly use as an energy source. (Pushpakumara et al., 2022)  |
| <i>Dolichospermum</i><br>NIES41 | <i>DEV114</i>                       | 16S               | +ve                | Genus belonging to the <i>Pedosphaeraceae</i> family. Has been found to be negatively correlated with the concentration of Ammonia. (Choi et al., 2022)   |
|                                 | <i>Cyclostephanos</i>               | <i>rbcL</i>       | +ve                | Large single-celled diatoms indicative of mesotrophic to eutrophic conditions, normally sensitive to nutrient enrichment. A superior genus in phosphate uptake and thus most resistant to nutrient depletion. Also, contain adaptation for floating in the water column and reducing sinking losses. (Simiyu and Kurmayer, 2022)                            |
| <i>Snowella</i><br>OTU37S04     | <i>Tabellaria</i>                   | <i>rbcL</i>       | +ve                | Assigned to codon N from Reynolds classification. Often found in mesotrophic epilimnia, tolerant to nutrient deficiency and sensitive to stratification, and pH rise. (Reynolds et al., 2002)   |
|                                 | <i>Sporichthyaceae</i> Family       | 16S               | -ve                | This family only contains the genus <i>Sporichthya</i> , which is comprised of two species. Members are motile facultative anaerobes and have a cell wall containing large amounts of LL-diaminopimelic acid. Produce aerial hyphae made up of rod-shaped spores. No substrate mycelium. The main habitat of this family appears to be soil. (Tamura, 2014) |

### 4.3.3 – NMDS analysis for 16S rRNA genera

#### 4.3.3.1 – Alaw

16S rRNA genera composition in the Alaw reservoir revealed a distinct clustering of samples according to high geosmin levels (Figure 4.30). Samples experiencing low and medium geosmin levels did not cluster as distinctively as high geosmin levels and expressed overlaps. Significant co-occurring *Cyanobacteria* (as identified in Section 4.3.3) showed all *Cyanobacteria* vectors pointing toward high geosmin level data points (Figure 4.30A). *Aphanizomenon* NIES81 was the most significantly associated genus in relation to high geosmin levels. Significant co-occurring T&O degrading bacteria (as identified in Section 4.3.3) displayed *Sphingorhabdus* and *Bacillus* to be most closely related to high geosmin levels, with *Sphingorhabdus* being the most significant (Figure 47B). *Flavobacterium* and *Rhodoferrax* vectors were directed more towards low and medium geosmin levels. Environmental vectors influencing the ordination of samples revealed dissolved manganese (DMn), pH, sulphate (SO<sub>4</sub>), and both T&O compounds to be associated with high geosmin levels (Figure 47C). Orthophosphate (Ortho), NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup>, TIN:TP, dissolved reactive silicate (Sil) and dissolved iron (DFe) were most influential on low and medium geosmin levels.



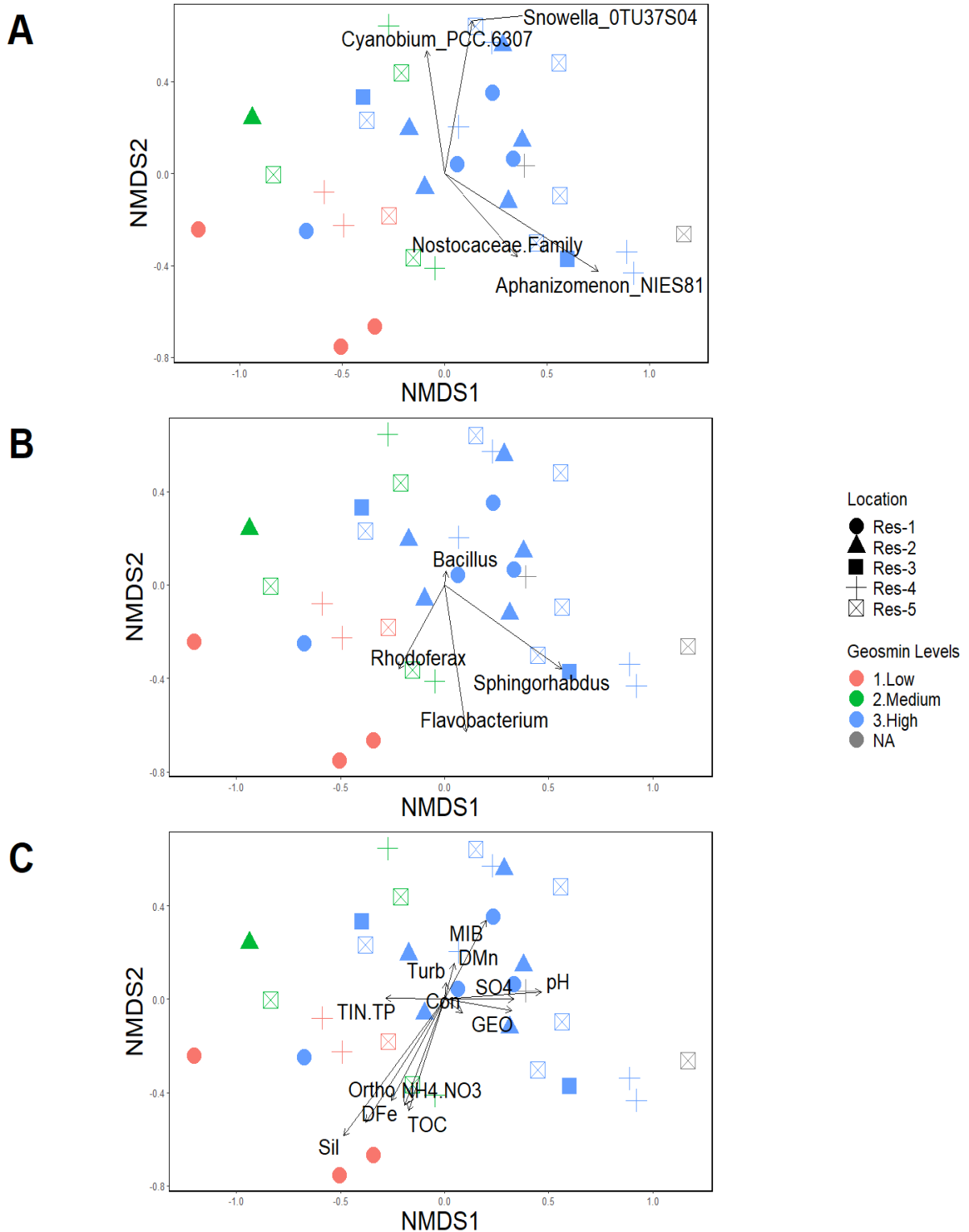


Figure 4.30: Three 16S rRNA NMDS ordination of samples from Alaw (n = 34) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring *Cyanobacteria* from co-occurrence analysis B) significant co-occurring T&O degrading bacteria from co-occurrence analysis C) environmental variables of interest.

### 4.3.3.2 – Cefni

For 16S rRNA genera composition in the Cefni reservoir, samples did not expose any clustering in relation to geosmin levels (Figure 4.31).

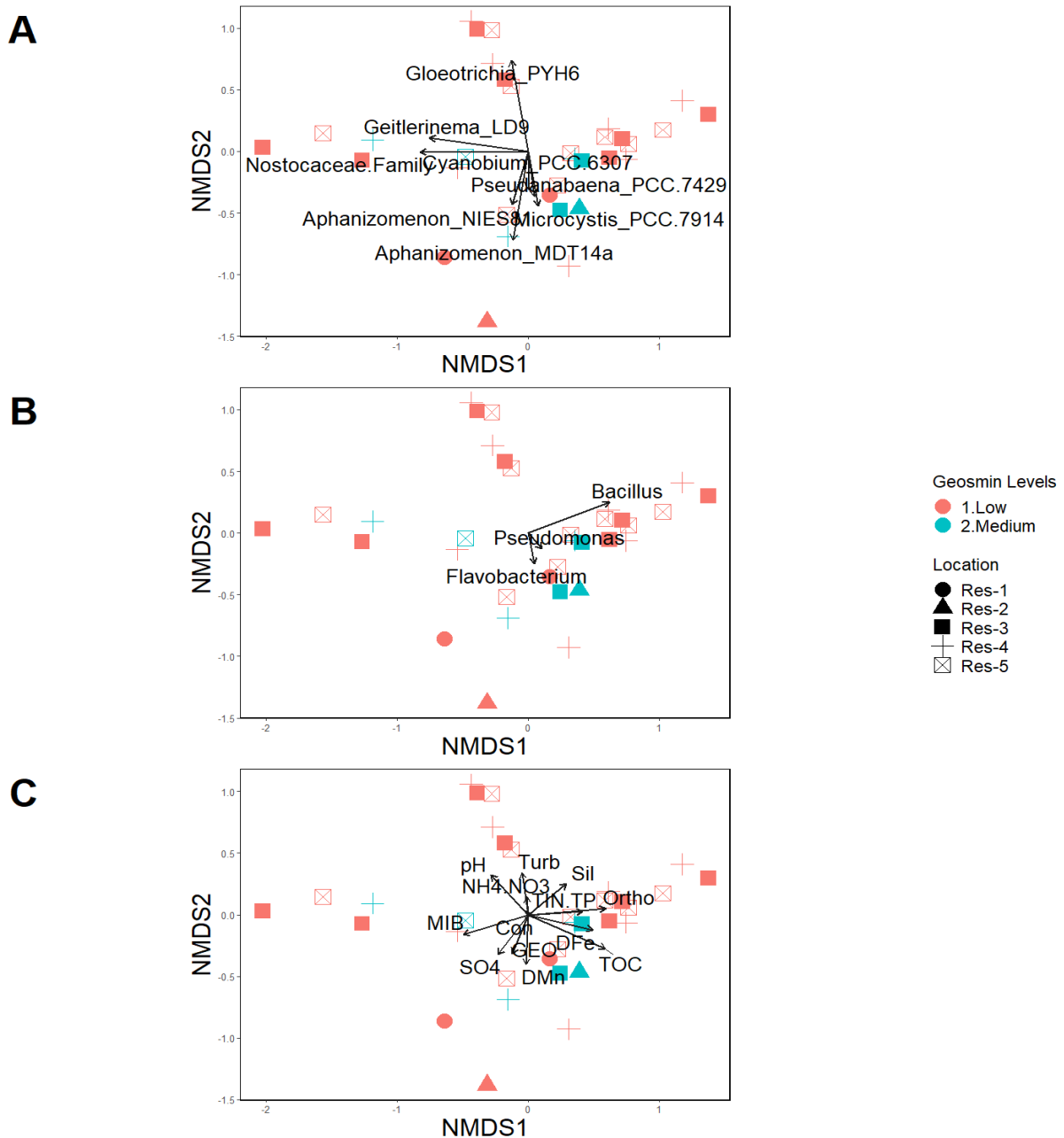
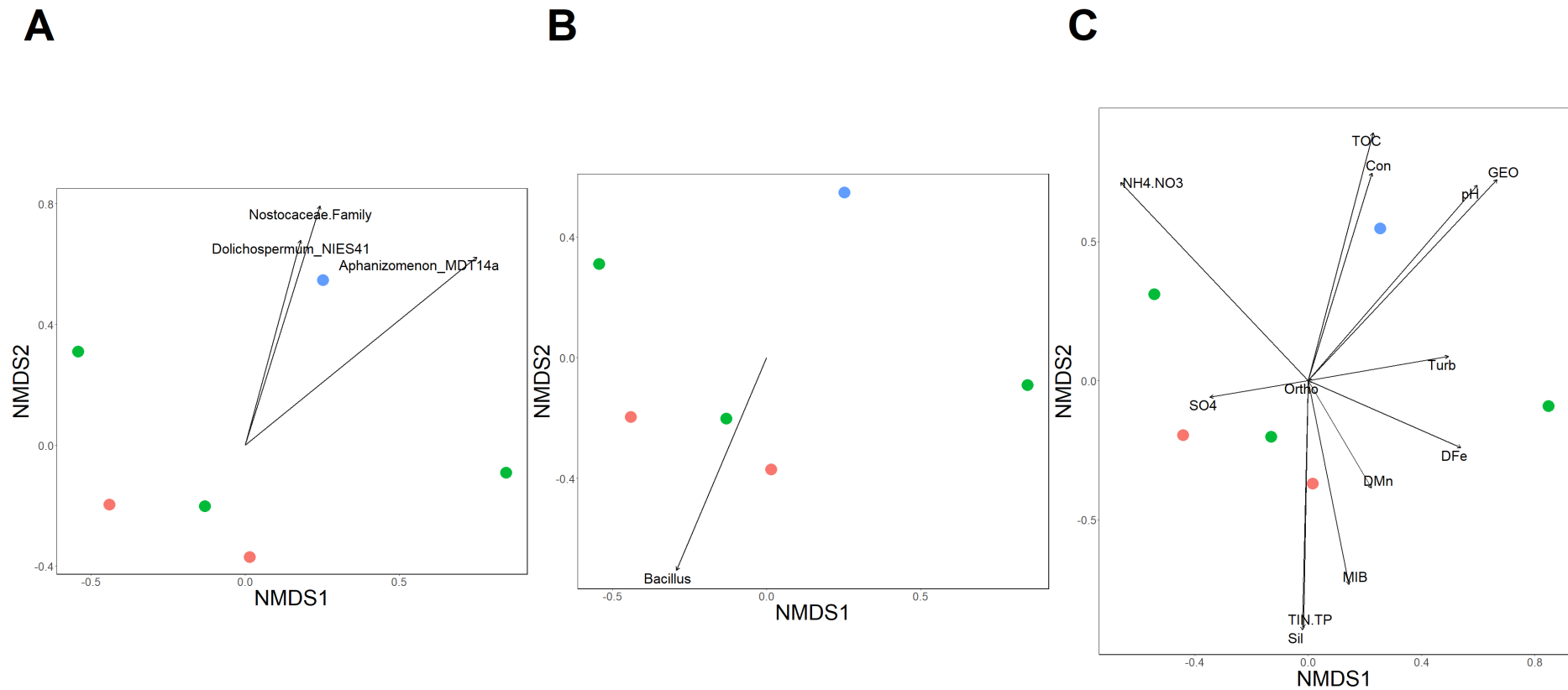


Figure 4.31: Three 16S rRNA NMDS ordination of samples from Cefni (n = 33) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring *Cyanobacteria* from co-occurrence analysis B) significant co-occurring T&O degrading bacteria from co-occurrence analysis C) environmental variables of interest.

#### 4.3.3.3 – Dolwen

16S rRNA genera composition for the Dolwen reservoir revealed a gradient of clustering, from low to high geosmin levels (Figure 4.32). Although a small sample size (n=6) must be considered when interpreting results. For significant co-occurring *Cyanobacteria*, as identified in Section 4.3.3, all three *Cyanobacteria* vectors were significantly directed towards high geosmin levels (Figure 4.32A). The significant co-occurring geosmin degrading genus, *Bacillus*, was significantly related to samples experiencing low and medium geosmin levels (Figure 4.32B). Significant environmental variables influential of high geosmin levels were pH, geosmin (GEO), conductivity (Con) and total organic carbon (TOC) (Figure 4.32C).  $\text{NH}_4^+:\text{NO}_3^-$  was significantly associated with medium geosmin levels, and TIN:TP and dissolved reactive silicate were significantly associated with low geosmin levels.



Location • Res-1 Geosmin Levels • 1.Low • 2.Medium • 3.High

Figure 4.32: Three 16S rRNA NMDS ordination of samples from Dolwen (n = 6) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring *Cyanobacteria* from co-occurrence analysis B) significant co-occurring T&O degrading bacteria from co-occurrence analysis C) environmental variables of interest.

#### 4.3.3.4 – Llwyn On

16S rRNA genera composition in Llwyn On reservoir showed apparent clustering of samples according to geosmin levels (Figure 4.33). Significant co-occurring *Cyanobacteria* (identified from Section 4.3.3) vectors pointed towards samples experiencing high geosmin levels (Figure 4.33A). Due to the absence of any significant co-occurring T&O degrading bacteria from Section 4.3.3, significant 16S rRNA genera as determined by NMDS analysis ( $p \leq 0.001$ ) were plotted (Figure 4.33B). Two T&O degrading bacteria were identified, *Flavobacterium* and *Sphingorhabdus*; *Sphingorhabdus* was closely related to high geosmin levels and closely related to *Aphanizomenon* NIES81. Environmental variable vectors revealed only geosmin (GEO) to be closely related to the samples subjected to high geosmin levels (Figure 4.33C).

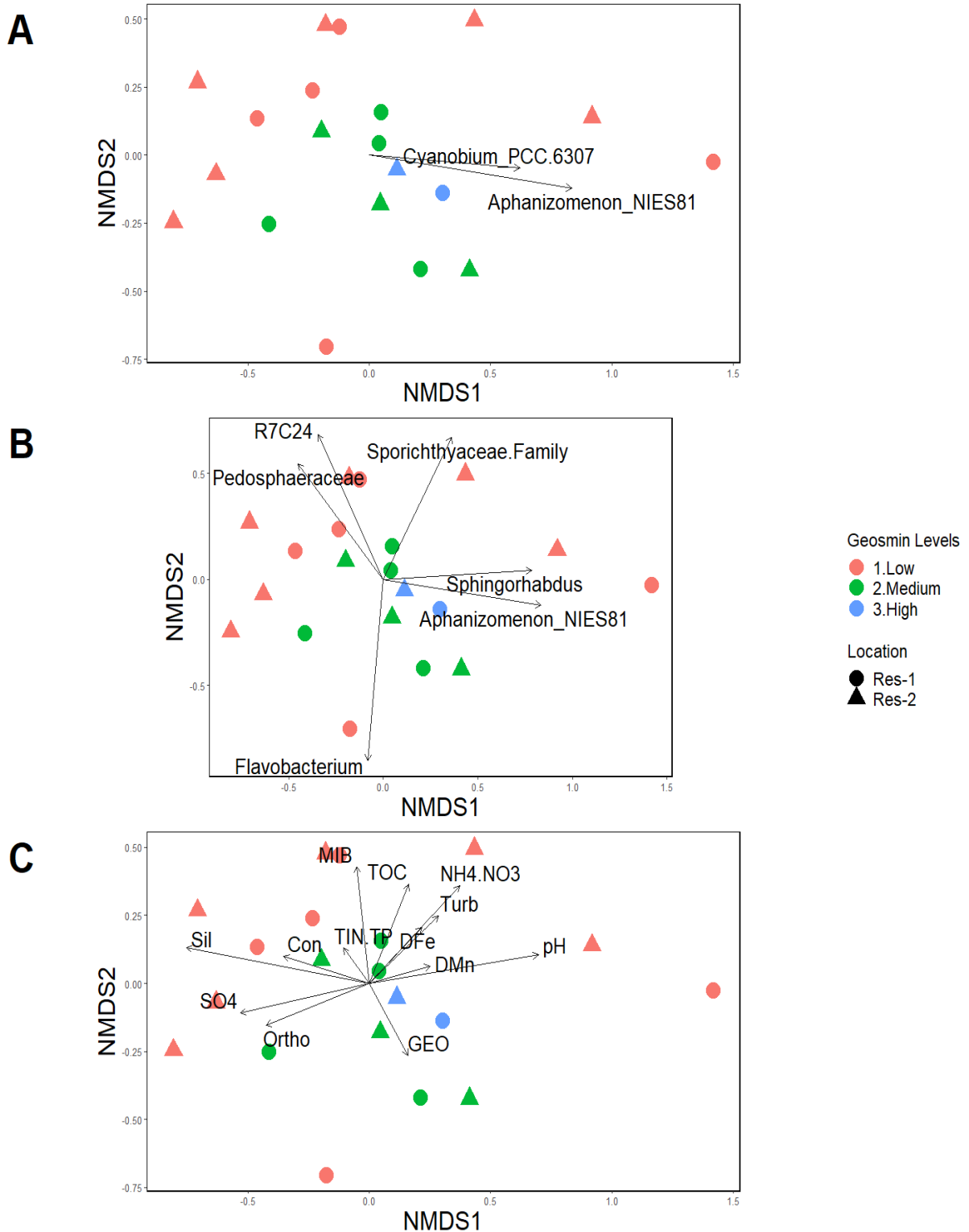


Figure 4.33: Three 16S rRNA NMDS ordination of samples from Llwyn On ( $n = 20$ ) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring *Cyanobacteria* from co-occurrence analysis B) significant 16S rRNA genera identified by NMDS analysis ( $p \leq 0.001$ ) C) environmental variables of interest.

#### 4.3.3.5 – Pentwyn

16S rRNA genera composition in Pentwyn followed distinct clustering according to assigned 2-MIB levels (Figure 4.34), apart from one spurious medium 2-MIB level present in the low 2-MIB cluster. Significant co-occurring *Cyanobacteria*, identified in Section 4.3.3, exposed *Cyanobium* PCC-6307 to be significantly associated with high 2-MIB levels (Figure 4.34A). For significant co-occurring T&O degrading bacteria, *Novosphingobium* was the vector considered to be significant and towards low 2-MIB levels (Figure 4.34B). The only environmental variable considered to be significant towards high 2-MIB levels was 2-MIB concentrations (MIB) (Figure 4.34C). pH, conductivity (Con), dissolved iron (DFe) and dissolved reactive silicate (Sil) were all significantly related to low 2-MIB levels.

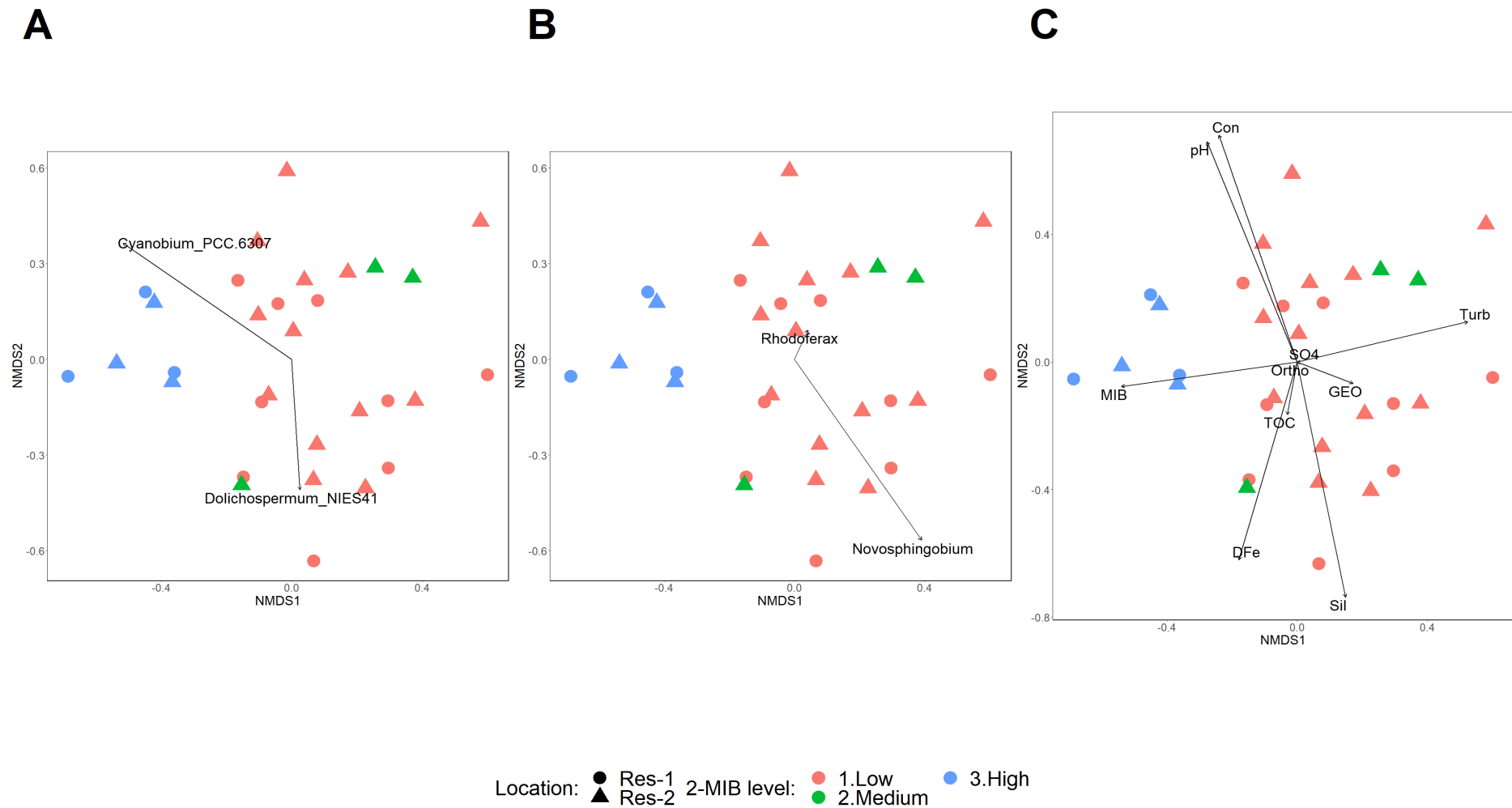


Figure 4.34: Three 16S rRNA NMDS ordination of samples from Pentwyn (n = 31) reservoir coloured by 2-MIB level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring *Cyanobacteria* from co-occurrence analysis B) significant co-occurring T&O degrading bacteria from co-occurrence analysis C) environmental variables of interest.



#### 4.3.3.6 – Plas Uchaf

16S rRNA genera composition in the Plas Uchaf reservoir did not reveal any distinctive clustering of samples according to geosmin levels (Figure 4.35). Although, overlaid vectors of significant co-occurring *Cyanobacteria*, as determined in Section 4.3.3, revealed *Microcystis* PCC-7914 to be closely related to most samples that experienced medium geosmin levels (Figure 4.35A).

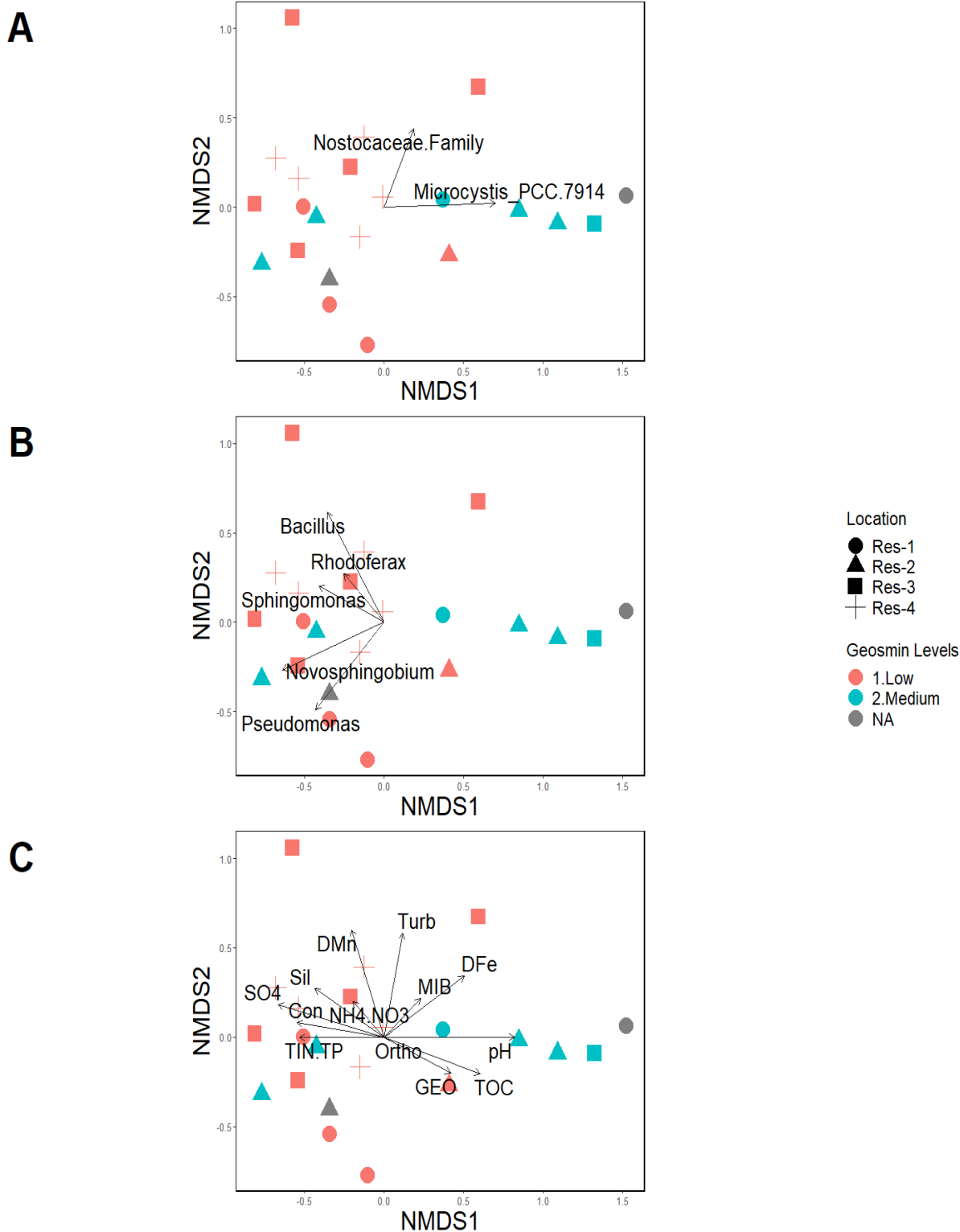


Figure 4.35: Three 16S rRNA NMDS ordination of samples from Plas Uchaf (n = 22) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant co-occurring *Cyanobacteria* from co-occurrence analysis B) significant co-occurring T&O degrading bacteria from co-occurrence analysis C) environmental variables of interest.

#### 4.3.3.7 – Pontsticill

For the 16S rRNA genera composition in the Pontsticill reservoir, no distinct separation from the low and the one medium geosmin event occurred (Figure 4.36 – left). For 2-MIB concentrations, a clear separation existed between low and medium levels (Figure 4.36 – right), apart from one spurious low-level datapoint. Significant co-occurring *Cyanobacteria* genera identified in Section 4.3.3 revealed *Cyanobium* PCC-6307 and *Snowe/out*OTU37S04 vectors to be significantly directed towards the cluster of medium 2-MIB level data points (Figure 4.36B). The only significant co-occurring T&O degrading bacteria, *Pseudomonas*, was significantly associated with low geosmin and 2-MIB levels (Figure 4.36C and D, respectively). Environmental variables associated with medium 2-MIB levels were geosmin (GEO), 2-MIB (MIB), TIN:TP and  $\text{NH}_4^+:\text{NO}_3^-$  (Figure 4.36F).

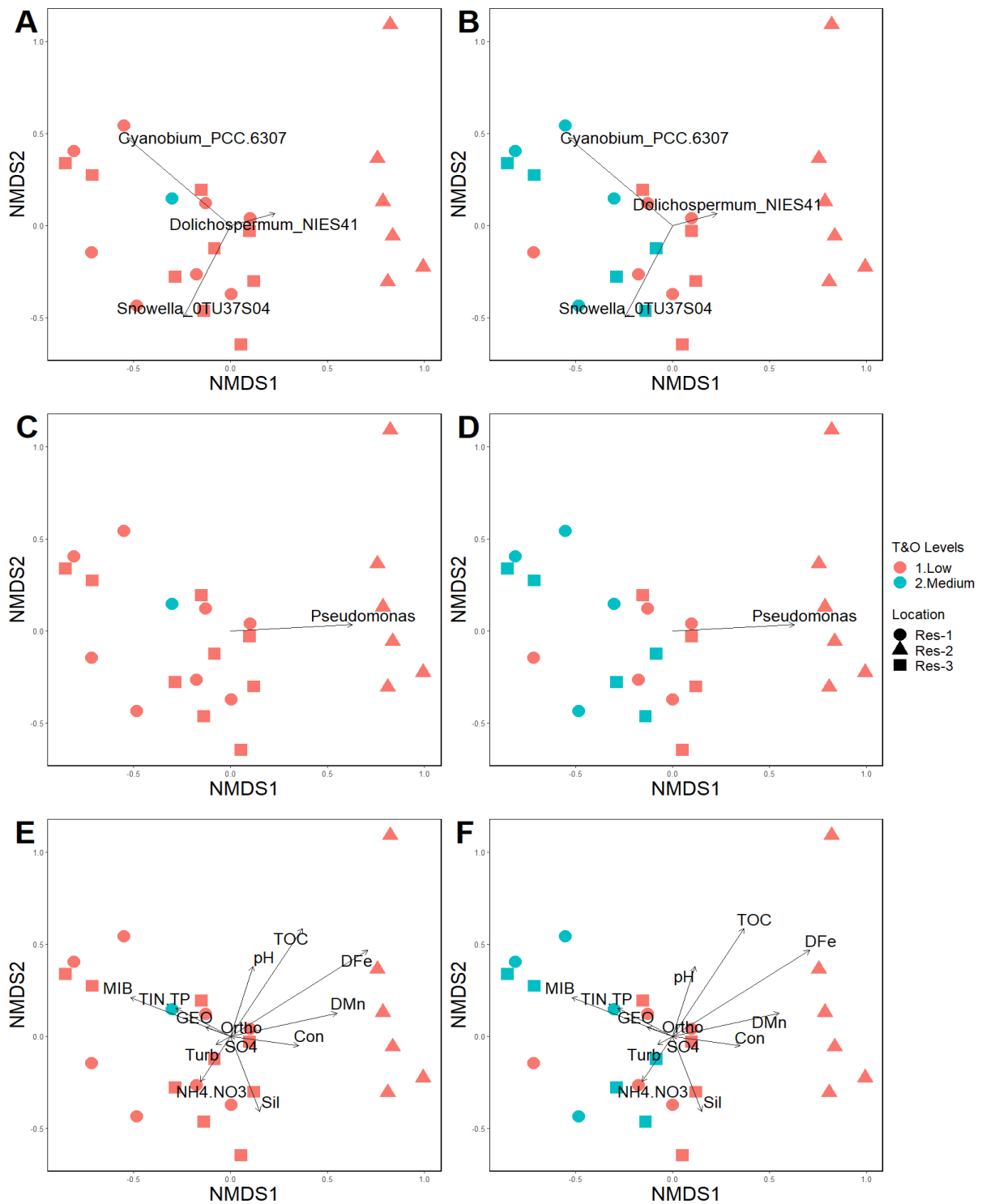


Figure 4.36: Six 16S rRNA NMDS ordination of samples from Pontsticill (n = 24) reservoir coloured by geosmin (left – A, C and E) and 2-MIB level (right – B, D and F) and shaped by sample point locations, arrows represent vectors of A+B) significant co-occurring *Cyanobacteria* from co-occurrence analysis C+D) significant co-occurring T&O degrading bacteria from co-occurrence analysis E+F) environmental variables of interest.

## 4.3.4 – NMDS analysis for *rbcL* genera

### 4.3.4.1 – Alaw

The community composition structure for *rbcL* in the Alaw reservoir did not appear to cluster according to geosmin levels (Figure 4.37). Significant *rbcL* genera as determined through NMDS analysis, revealed that most identified genera are directed towards a region of high geosmin level datapoints, except for *Asterionella* and *Tabellaria* (Figure 4.37A). *Tabellaria* appeared to be directed towards a mixture of geosmin levels, whereas *Asterionella* was directed towards a grouping of high geosmin levels. Significantly co-occurring algal genera with *Cyanobacteria*, as determined in Section 4.3.3, were directed towards a grouping of high geosmin levels (Figure 4.37B), except for *Aulacoseira*, *Thalassiosira*, and *Stephanodiscus* that were directed more towards low and medium geosmin levels. No environmental variables were deemed significant in structuring the *rbcL* community composition (Figure 4.37C).

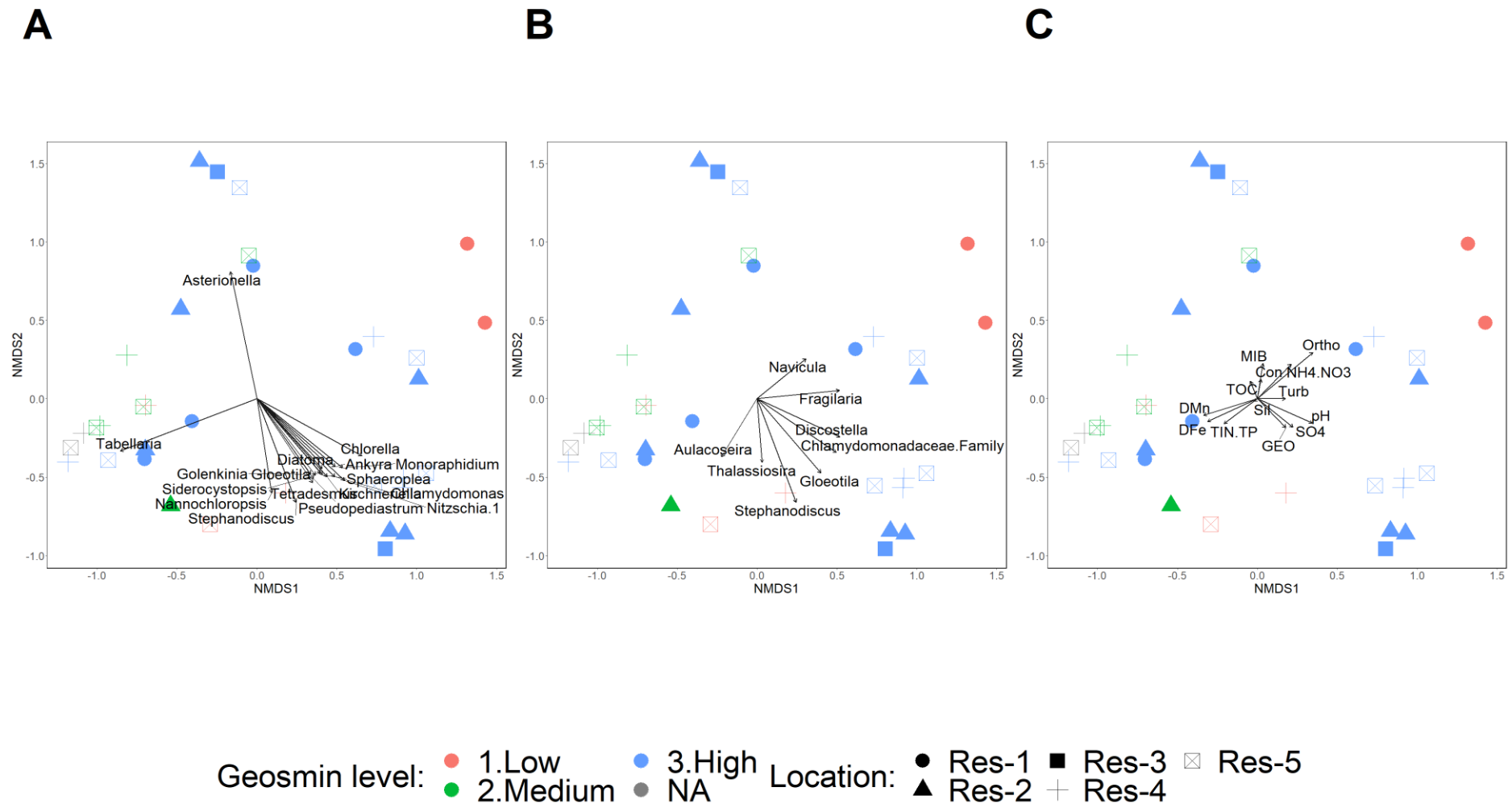


Figure 4.37: Three *rbcL* NMDS ordination of samples from Alaw (n = 34) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant *rbcL* genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring *rbcL* genera with *Cyanobacteria* from co-occurrence analysis C) environmental variables of interest.

#### 4.3.4.2 – Cefni

In the Cefni reservoir, *rbcL* genera composition did not reveal any clustering of data points according to geosmin levels (Figure 4.38).

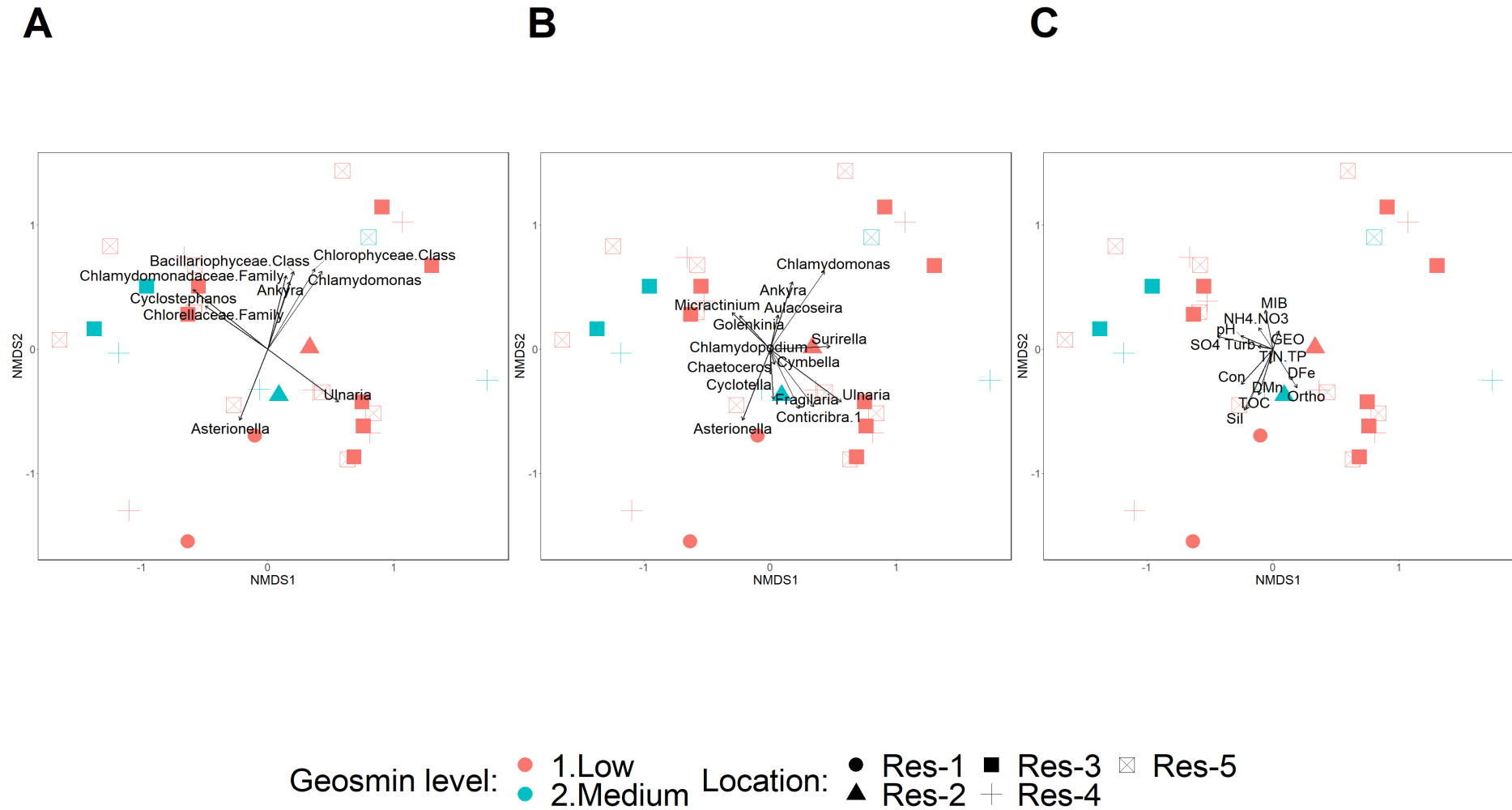


Figure 4.38: Three *rbcL* NMDS ordination of samples from Cefni (n = 32) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant *rbcL* genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring *rbcL* genera with cyanobacteria from co-occurrence analysis C) environmental variables of interest.



#### 4.3.4.3 – Dolwen

In the Dolwen reservoir, *rbcL* genera did not cluster according to geosmin concentration levels (Figure 4.39), with overlapping of data points.

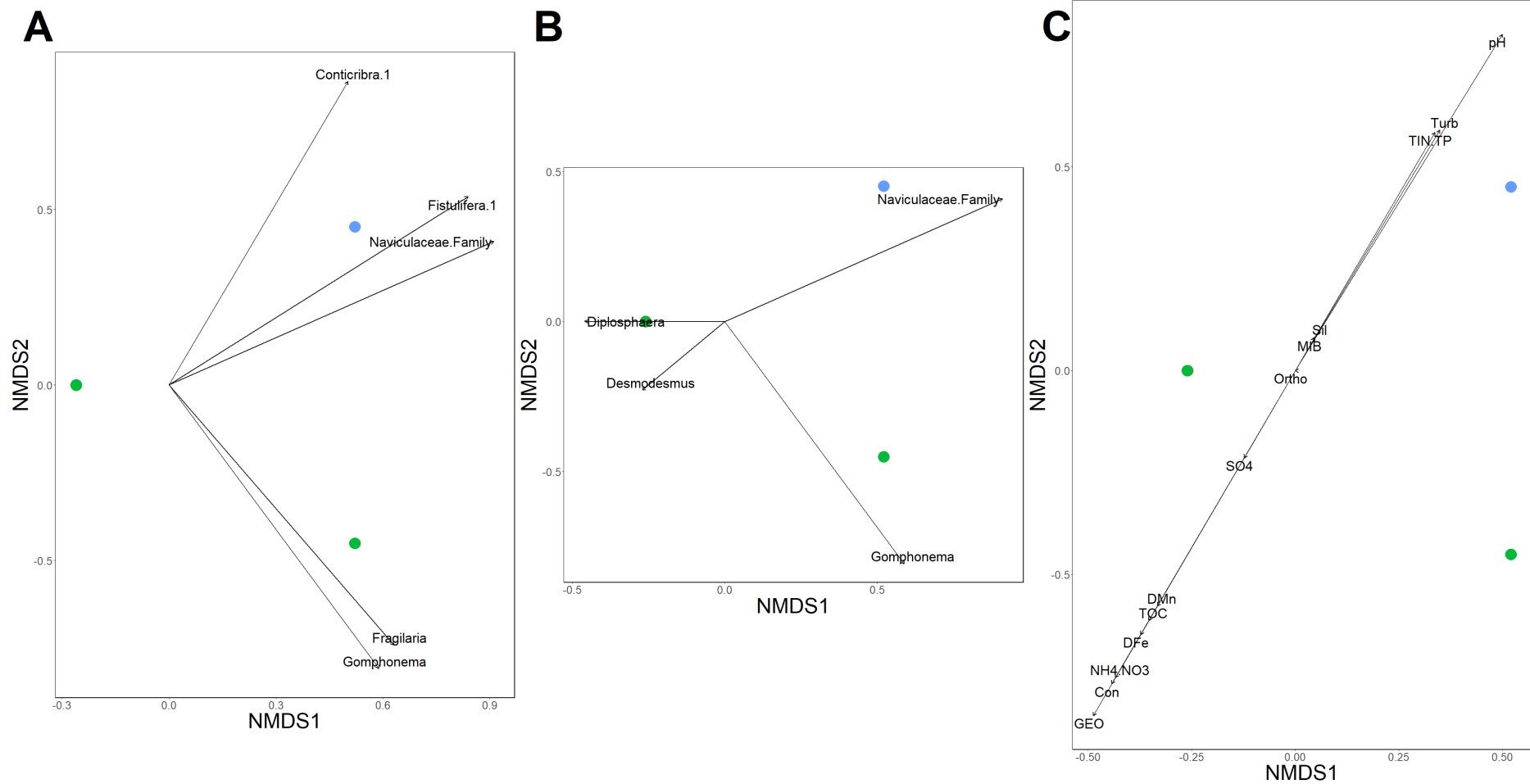


Figure 4.39: Three *rbcL* NMDS ordination of samples from Dolwen (n = 6) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant *rbcL* genera governed by NMDS analysis ( $p \leq 0.05$ ) B) significant co-occurring *rbcL* genera with *Cyanobacteria* from co-occurrence analysis C) environmental variables of interest.

#### 4.3.4.4 – Llwyn On

Distinct clustering of *rbcL* genera composition occurred according to geosmin levels in Llwyn On reservoir (Figure 4.40). Due to only one (*Asterionella*) algae significantly co-occurring with a *Cyanobacteria* (*Cyanobium* PCC-6307) in this reservoir, no plot was formed. As governed by NMDS analysis, significant genera vectors revealed *Asterionella* as the most closely related genus to medium and high geosmin levels (Figure 4.40A). All other significant genera were associated with low geosmin levels. The main significant environmental variables associated with high geosmin levels were geosmin (GEO) and pH.

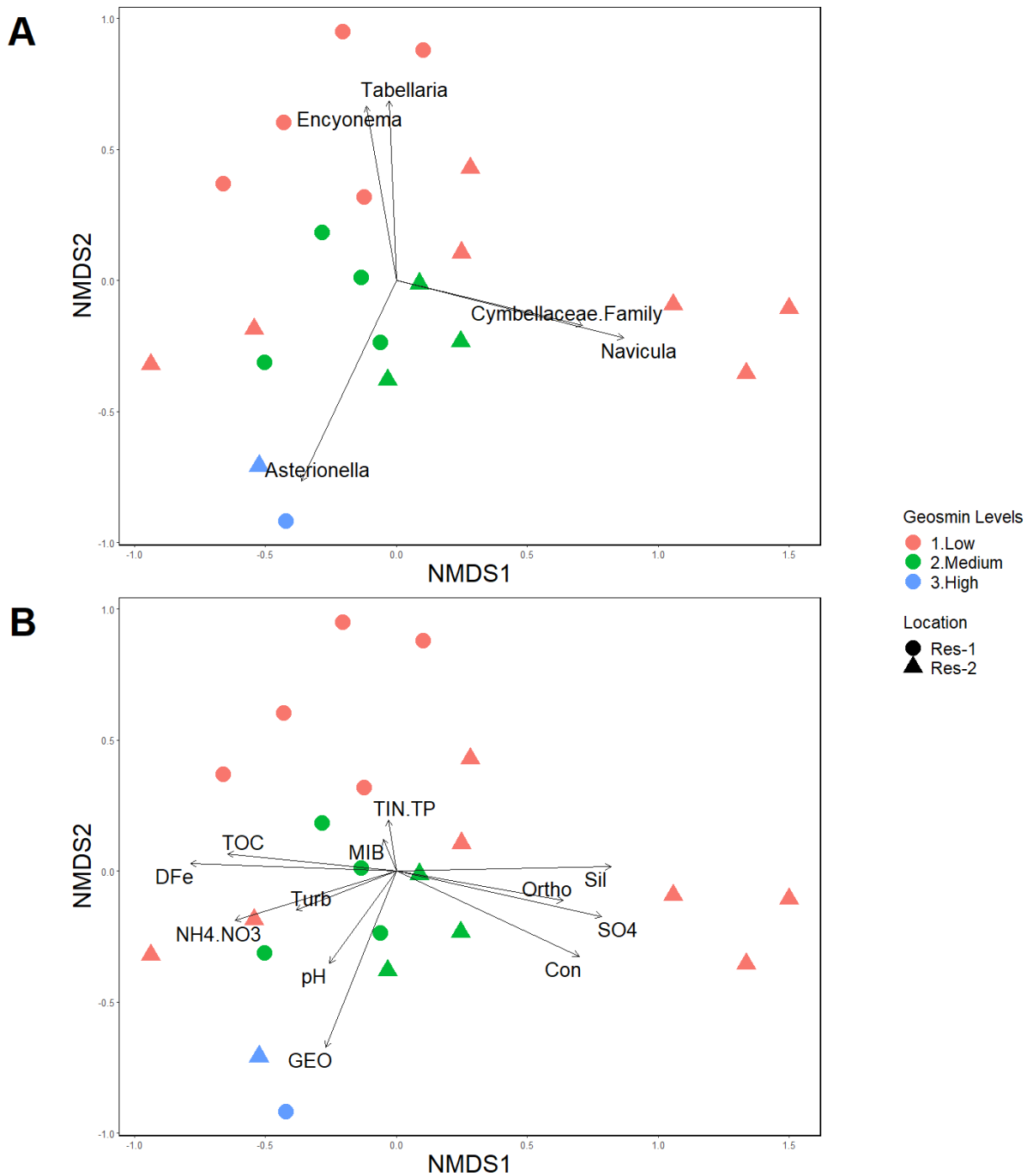
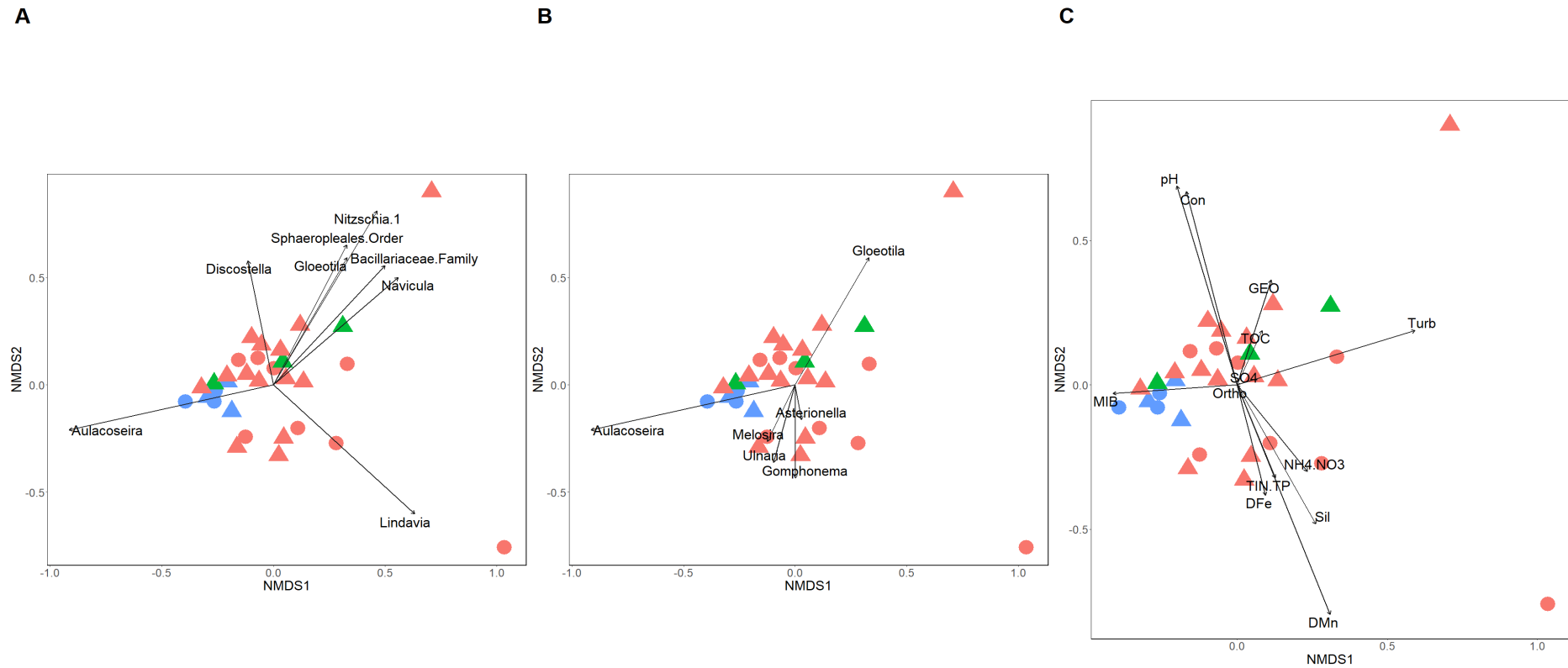


Figure 4.40: Three *rbcL* NMDS ordination of samples from Llwyn On (n = 21) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant *rbcL* genera governed by NMDS analysis ( $p \leq 0.005$ ) B) environmental variables of interest.

#### 4.3.4.5 – Pentwyn

In the Pentwyn reservoir, *rbcL* genera composition can be seen to form a cluster for data points subjected to high 2-MIB levels (Figure 4.41). From comparing significant genera as governed by NMDS analysis (Figure 4.41A) and significant co-occurring algal genera with *Cyanobacteria* determined in Section 4.3.3 (Figure 4.41B), it is evident that *Aulacoseira* is the most significantly associated genus with high 2-MIB levels. The only significant environmental variable associated with high 2-MIB levels was 2-MIB (MIB) (Figure 4.41C).



Location: ● Res-1    ▲ Res-2    2-MIB level: ● 1.Low    ● 2.Medium    ● 3.High

Figure 4.41: Three *rbcL* NMDS ordination of samples from Pentwyn (n = 31) reservoir coloured by 2-MIB level and shaped by sample point locations, arrows represent vectors of A) significant *rbcL* genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring *rbcL* genera with *Cyanobacteria* from co-occurrence analysis C) environmental variables of interest.

#### 4.3.4.6 – Plas Uchaf

*rbcL* genera composition in the Plas Uchaf reservoir did not reveal any true clustering of data points according to geosmin level (Figure 4.42), although medium geosmin levels did form two separate clusters. *Melosira* and *Asterionella* were revealed to be significantly related to medium geosmin levels as governed by NMDS analysis (Figure 4.42A). Significant co-occurring algal genera, as determined in Section 4.3.3, showed *Stephanodiscus* to be significantly related to medium geosmin levels also (Figure 4.42B). Environmental variables exPOsed pH and 2-MIB (MIB) to be associated with medium geosmin levels (Figure 4.42C).

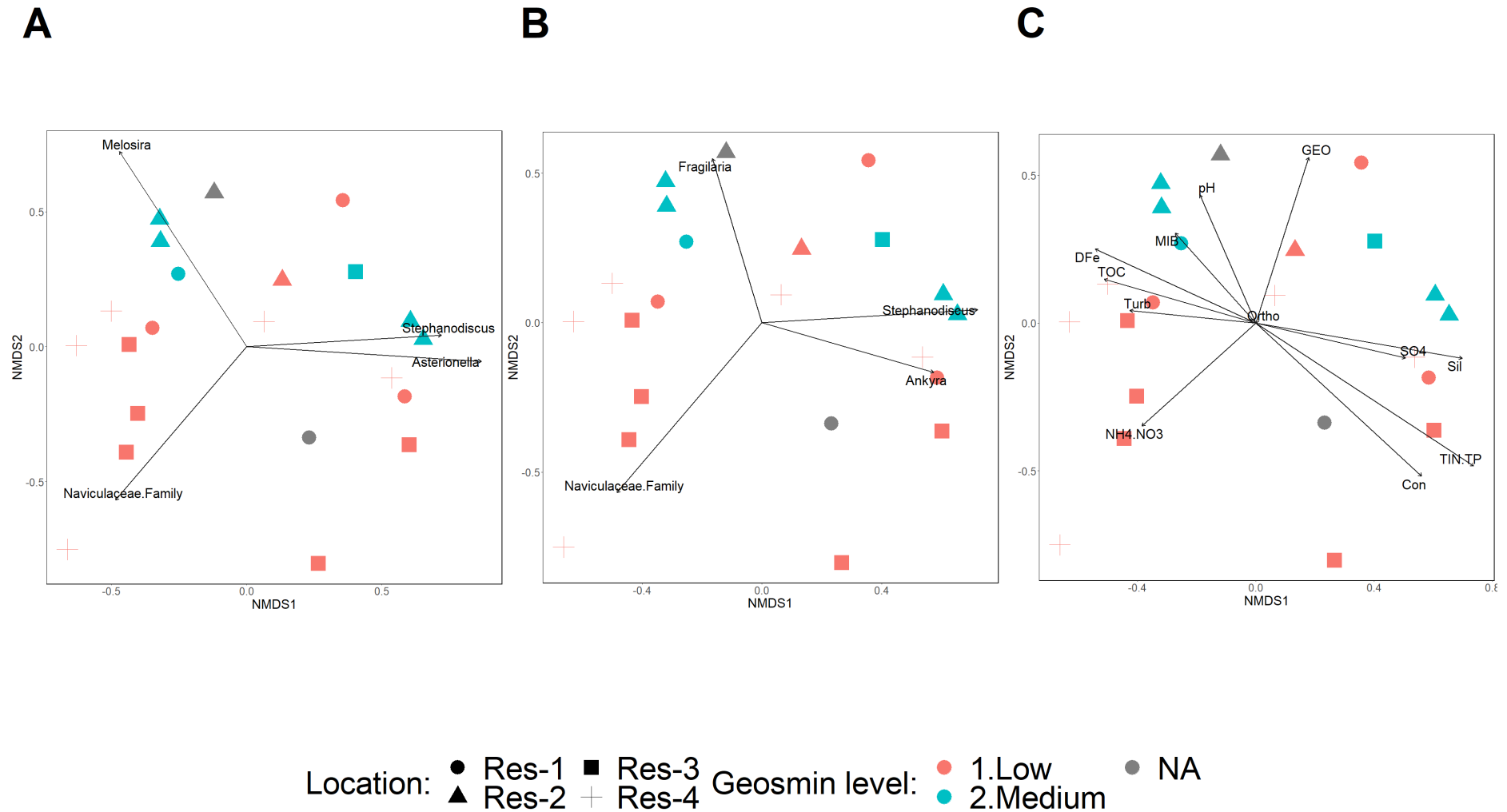


Figure 4.42: Three *rbcL* NMDS ordination of samples from Plas Uchaf (n = 22) reservoir coloured by geosmin level and shaped by sample point locations, arrows represent vectors of A) significant *rbcL* genera governed by NMDS analysis ( $p \leq 0.001$ ) B) significant co-occurring *rbcL* genera with *Cyanobacteria* from co-occurrence analysis C) environmental variables of interest.



#### 4.3.4.7 – Pontsticill

No true clustering of *rbcL* genera composition was related to geosmin levels (Figure 4.43 – left) or 2-MIB levels (Figure 4.43 – right) in the Pontsticill reservoir.

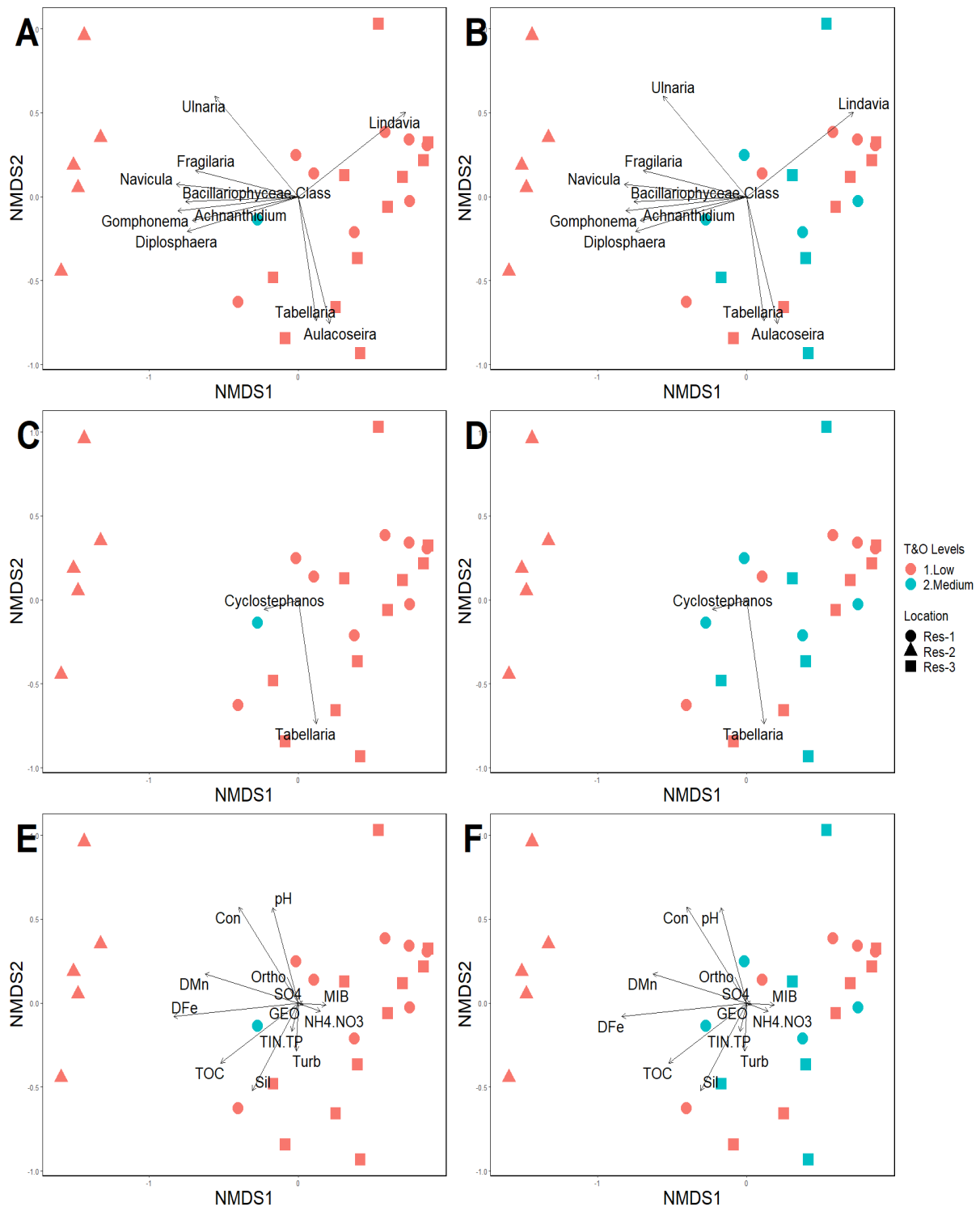


Figure 4.43: Six *rbcL* NMDS ordination of samples from Pontsticill (n = 25) reservoir coloured by geosmin (left – A, C and E) and 2-MIB level (right – B, D and F) and shaped by sample point locations, arrows represent vectors of A+B) significant *rbcL* genera governed by NMDS analysis ( $p \leq 0.001$ ) C+D) significant co-occurring *rbcL* genera with *Cyanobacteria* from co-occurrence analysis E+F) environmental variables of interest.

#### 4.3.5 – Indicative genera of a T&O ‘event’ in Alaw, Llwyn On and Pentwyn reservoirs experiencing high T&O concentrations

Random forest analysis correctly classified at least 75% of samples in the Alaw reservoir for geosmin concentration levels (low, medium, and high) and at least 66.67% of samples in the Pentwyn reservoir for 2-MIB concentrations (Table 4.11). Allocation of high and medium geosmin concentration classifications in Llwyn On reservoir was less accurate, with 0 – 50% of samples being correctly classified, respectively; this could be due to only two samples having high geosmin concentrations. However, low geosmin concentration classification was highly accurate, with 94.12% of samples being correctly classified. These random forest models were constructed using subsets of only the 20 most discriminatory taxa associated with T&O concentration levels. Random forest models performed comparably to the original random forest model in their respective classifications, indicating that these subsets captured much of the relevant information from the entire dataset. Please see Table 4.12 for a list of discriminatory taxa used for each random forest for geosmin concentration levels in Alaw and Llwyn On reservoirs and 2-MIB concentration levels in Pentwyn reservoir.

Table 4.13: Results from random forest classification of T&O status from significantly co-occurring taxa from reservoirs experiencing high T&O concentrations.

| T&O compound | Reservoir | True Class | Predicted Class |           |           | Error rate (%) |
|--------------|-----------|------------|-----------------|-----------|-----------|----------------|
|              |           |            | Low             | Medium    | High      |                |
| Geosmin      | Alaw      | Low        | <b>16</b>       | 3         | 0         | 15.79          |
|              |           | Medium     | 2               | <b>12</b> | 2         | 25.00          |
|              |           | High       | 1               | 2         | <b>17</b> | 15.00          |
| Geosmin      | Llwyn On  | Low        | <b>32</b>       | 2         | 0         | 5.88           |
|              |           | Medium     | 3               | <b>4</b>  | 1         | 50.00          |
|              |           | High       | 1               | 1         | <b>0</b>  | 100.00         |
| 2-MIB        | Pentwyn   | Low        | <b>20</b>       | 2         | 0         | 9.10           |
|              |           | Medium     | 1               | <b>2</b>  | 0         | 33.33          |
|              |           | High       | 0               | 0         | <b>6</b>  | 0.00           |

Rows represent the number of samples from each true class allocated to each predicted class (columns). The error rate is the percentage of samples misclassified from each group. The predicted class for each given sample was determined by a simple majority of votes from the trees in which that sample was out-of-bag (OOB). Correct classifications are shown in bold.

Table 4.14 cont'd: Variable importance table displaying the ranking of variables used in the random forest classification for T&O levels (low, medium, and high) using the Mean Decrease in Accuracy (MDA) showing how much removing each taxon reduces the accuracy of the model for Alaw, Llwyn On, and Pentwyn reservoirs. Cyanobacterial taxa are bolded and shaded grey.

| ALAW   |               |                   |               |               | LLWYN ON                                   |              |              |              |              | PENTWYN                          |        |       |        |        |
|--|---------------|-------------------|---------------|---------------|--|--------------|--------------|--------------|--------------|----------------------------------|--------|-------|--------|--------|
| Taxa   | Low           | Med               | High          | MDA           | Taxa                                       | Low          | Med          | High         | MDA          | Taxa                             | Low    | Med   | High   | MDA    |
| <i>Peptostreptococcaceae</i><br>Family       | 146.89        | 80.88             | 122.28        | 185.75        | <b><i>Cyanobium</i></b><br><b>PCC-6307</b> | <b>90.35</b> | <b>35.93</b> | <b>57.05</b> | <b>92.08</b> | <i>Limnobacter</i>               | 130.35 | 31.88 | 179.74 | 157.66 |
| <b><i>Aphanizomenon</i></b><br><b>NIES81</b> | <b>43.15</b>  | <b>19.75</b>      | <b>144.58</b> | <b>132.93</b> | <i>OM27 clade</i>                          | 63.88        | 48.99        | -<br>15.88   | 69.18        | <i>NS9 marine group</i>          | 61.66  | 4.64  | 56.11  | 62.92  |
| <i>Discostella</i>                           | -12.09        | 62.32             | 63.97         | 76.60         | <i>Prostheco bacter</i>                    | 65.57        | 6.41         | 20.10        | 61.87        | <i>Entomoneis</i>                | -3.42  | 74.25 | 9.22   | 44.34  |
| <i>SH3-11</i>                                | 21.71         | -0.72             | 80.14         | 70.59         | <i>Fragilaria</i>                          | 23.96        | 39.48        | 49.67        | 61.61        | <i>Limnohabitans</i>             | 46.44  | -1.21 | 25.49  | 44.11  |
| Uncultured<br><i>Pedosphaeraceae</i>         | 46.71         | 50.28             | 8.57          | 66.05         | <i>Sphaeropleales</i><br>Order             | 58.54        | 43.56        | -<br>19.60   | 58.92        | <i>CL500-3</i>                   | 31.29  | 39.03 | -4.54  | 43.73  |
| <b><i>Cyanobium</i></b> PCC-6307             | <b>59.61</b>  | -<br><b>10.21</b> | <b>50.22</b>  | <b>59.70</b>  | <i>Achnantheidium</i>                      | 57.37        | 10.29        | 15.81        | 57.64        | <i>Sporichthyaceae</i><br>Family | 28.38  | 28.21 | 16.34  | 40.96  |
| <i>Navicula</i>                              | 35.90         | 19.87             | 22.45         | 46.15         | Uncultured<br><i>Pedosphaeraceae</i>       | 35.28        | 16.95        | 21.72        | 38.94        | <i>Asterionella</i>              | 22.34  | 34.75 | 14.65  | 35.40  |
| <i>Sphingorhabdus</i>                        | 6.77          | 30.13             | 38.90         | 45.45         | <i>Chthoniobacter</i>                      | 25.92        | 29.50        | 8.68         | 35.87        | <i>Gomphonema</i>                | 21.06  | 13.49 | -2.98  | 22.19  |
| Uncultured<br>Verrucomicrobiaceae            | <b>-22.02</b> | <b>14.75</b>      | <b>58.70</b>  | <b>45.11</b>  | <i>Armatimonas</i>                         | 16.07        | 31.34        | -<br>21.50   | 25.30        | <i>Pedosphaeraceae</i><br>Family | 18.52  | 5.17  | 22.58  | 21.61  |
| <i>Fragilaria</i>                            | 0.83          | 43.71             | 19.31         | 42.90         | <i>Chlamydomonas</i>                       | 19.13        | 23.23        | 0.00         | 24.44        | <i>Urosolenia</i>                | 19.30  | 16.22 | -3.62  | 21.38  |

Table 4.14 cont'd: Variable importance table displaying the ranking of variables used in the random forest classification for T&O levels (low, medium, and high) using the Mean Decrease in Accuracy (MDA) showing how much removing each taxon reduces the accuracy of the model for Alaw, Llwyn On, and Pentwyn reservoirs. cyanobacterial taxa are bolded and shaded grey.

| ALAW                               |             |             |              |              | LLWYN ON  |       |        |       |       | PENTWYN                            |              |             |              |             |
|------------------------------------|-------------|-------------|--------------|--------------|---|-------|--------|-------|-------|------------------------------------|--------------|-------------|--------------|-------------|
| Taxa                               | Low         | Med         | High         | MDA          | Taxa  | Low   | Med    | High  | MDA   | Taxa                               | Low          | Med         | High         | MDA         |
| <i>Uncultured Planctomycetales</i> | 7.90        | 44.96       | 18.58        | 38.72        | <i>Cytophagales</i> Order                           | 18.65 | 23.09  | 0.00  | 24.13 | Uncultured <i>Gemmataceae</i>      | 12.04        | 10.12       | 14.56        | 19.71       |
| <i>Pedosphaeraceae</i>             | 27.89       | 31.07       | 1.13         | 38.63        | <i>Opitutaceae</i> Family                           | 13.22 | 15.99  | 21.84 | 24.05 | <i>Chitinophagaceae</i> Family     | 15.78        | 6.61        | 2.04         | 15.81       |
| <b><i>Snowella</i> 0TU37S04</b>    | <b>9.87</b> | <b>8.24</b> | <b>35.26</b> | <b>35.85</b> | <i>env-OPS 17</i>                                   | 13.34 | 7.23   | 26.99 | 21.37 | <i>Rhizobiales</i> Order           | 15.31        | 7.92        | 1.57         | 15.29       |
| <i>Limnohabitans</i>               | 12.83       | 22.63       | 30.23        | 34.78        | <i>Ulnaria</i>                                      | 1.04  | 19.23  | 11.12 | 19.03 | Uncultured <i>Fibrobacteraceae</i> | 5.32         | 0.00        | 11.80        | 13.07       |
| <i>Trebouxiophyceae</i> Class      | 7.23        | 1.64        | 38.20        | 32.95        | <i>Pseudarcicella</i>                               | 10.91 | 14.38  | 6.86  | 17.51 | <i>Achnantheidium</i>              | 9.45         | 7.48        | 1.00         | 10.90       |
| <i>Gloeotila</i>                   | 5.09        | 17.04       | 28.37        | 28.66        | Uncultured <i>Rhizobiales</i> <i>Incertae Sedis</i> | 12.84 | 14.48  | 0.00  | 16.61 | <i>Achnantheidium</i>              | 8.77         | 8.37        | 1.00         | 10.68       |
| <i>Bacillus</i>                    | -2.37       | 35.78       | -1.56        | 21.37        | <i>Trebouxiophyceae</i> Class                       | 32.64 | -19.86 | 4.00  | 15.52 | <i>Fragilaria</i>                  | 4.72         | 14.14       | -0.81        | 10.42       |
| <i>Chlamydomonadaceae</i> Family   | 5.39        | 7.86        | 17.49        | 17.50        | <i>Mayamaea</i>                                     | 5.28  | 14.90  | 0.00  | 13.71 | <b><i>Cyanobium</i> PCC-6307</b>   | <b>-1.56</b> | <b>5.90</b> | <b>12.02</b> | <b>9.78</b> |
| <i>Sediminibacterium</i>           | -10.58      | 12.83       | 20.11        | 17.40        | <i>Chlorophyceae</i> Class                          | 31.64 | -20.44 | 3.61  | 13.35 | <i>R7C24</i>                       | -5.74        | 5.73        | 11.48        | 8.55        |
| <i>Stephanodiscus</i>              | -1.63       | 2.97        | 24.72        | 13.90        | <i>Urosolenia</i>                                   | 22.30 | -26.17 | 17.89 | 5.96  | <i>Pseudarcicella</i>              | 8.16         | -21.42      | 14.40        | -0.84       |

#### 4.3.5.1 – Alaw

Discriminatory taxa used in Alaw's random forest displayed that the *Peptostreptococcaceae* family and *Aphanizomenon* NIES81 were most informative on community structure during low, medium, and high geosmin concentrations (Figure 4.44). *Discostella*, *SH3-11*, uncultured *Pedosphaeraceae*, and *Cyanobium* PCC-6307 were the second group of most informative taxa indicative of geosmin concentration classification. Referring to the confusion matrix for this random forest model's accuracy (Table 4.11), it can be concluded that results can be used to highlight indicative taxa associated with high geosmin concentrations. The success of this model is graphically illustrated in Figure 4.45. The ordination plot displays a cluster of high geosmin samples with a tight ellipse, although there is the exception of two high geosmin samples that fell outside of the 95% confidence interval. From exploring the overall MDA value for each taxon (Table 4.12) and the MDA values associated with that taxon for low, medium, and high geosmin concentrations, indicative genera of high geosmin classifications were determined. Here, two groups of high geosmin indicator taxa are proposed.

The first group are “true indicator taxa” (Figure 4.46) that display negative MDA values for low geosmin classifications (Figure 4.46A). Negative MDAs denote that the random permutation worked best, implying that these taxa do not have a role in the prediction of low geosmin classifications. From these “true indicator taxa”, *Discostella* was most informative regarding high geosmin classification, followed by uncultured *Verrucomicrobiaceae*. *Sediminibacterium* and *Stephanodiscus* were less informative but still represented gradual increases in MDA, showing to be more critical in the case of high geosmin concentrations. However, when plotting the relative abundance of these indicator taxa, it was clear that a decline in the genus, *Stephanodiscus*, was most indicative of high geosmin (Figure 4.46B).

The second group of indicative taxa are presented in Figure 4.47. These taxa do not display negative MDA values during low geosmin concentrations but show a gradual increase in MDA as geosmin concentrations increase from low to high (Figure 4.47A). From this group, the genus *Sphingorhabdus*, a potential geosmin degrader phylogenetically related to the known geosmin degrading genus, *Sphingopyxis* (Jogler et al., 2013), was the most promising indicator genus. The

relative abundance of each indicative taxon revealed that all taxa were predominantly at their highest abundance during high geosmin concentrations (Figure 4.47B). *Gloeotila* and *Sphingorhabdus* were the most promising indicative genera for high geosmin concentrations.

In addition to the indicative genera responsible for the cyanobacterial production of geosmin, the *Cyanobacteria* present in the model were informative in the classification of geosmin concentration levels. *Aphaznizomenon* NIES81 had the highest MDA value for high geosmin classification, showing less importance in the classification of low geosmin and even less for medium geosmin concentrations (Table 4.12). *Snowella* OUT37S04 displayed a similar relationship to *Aphaznizomenon* NIES81 but at a lower magnitude of MDA values. *Cyanobium* PCC-6307 had its highest MDA value for low geosmin classifications, a negative MDA for medium classification and its second highest for high geosmin classification.



ALAW

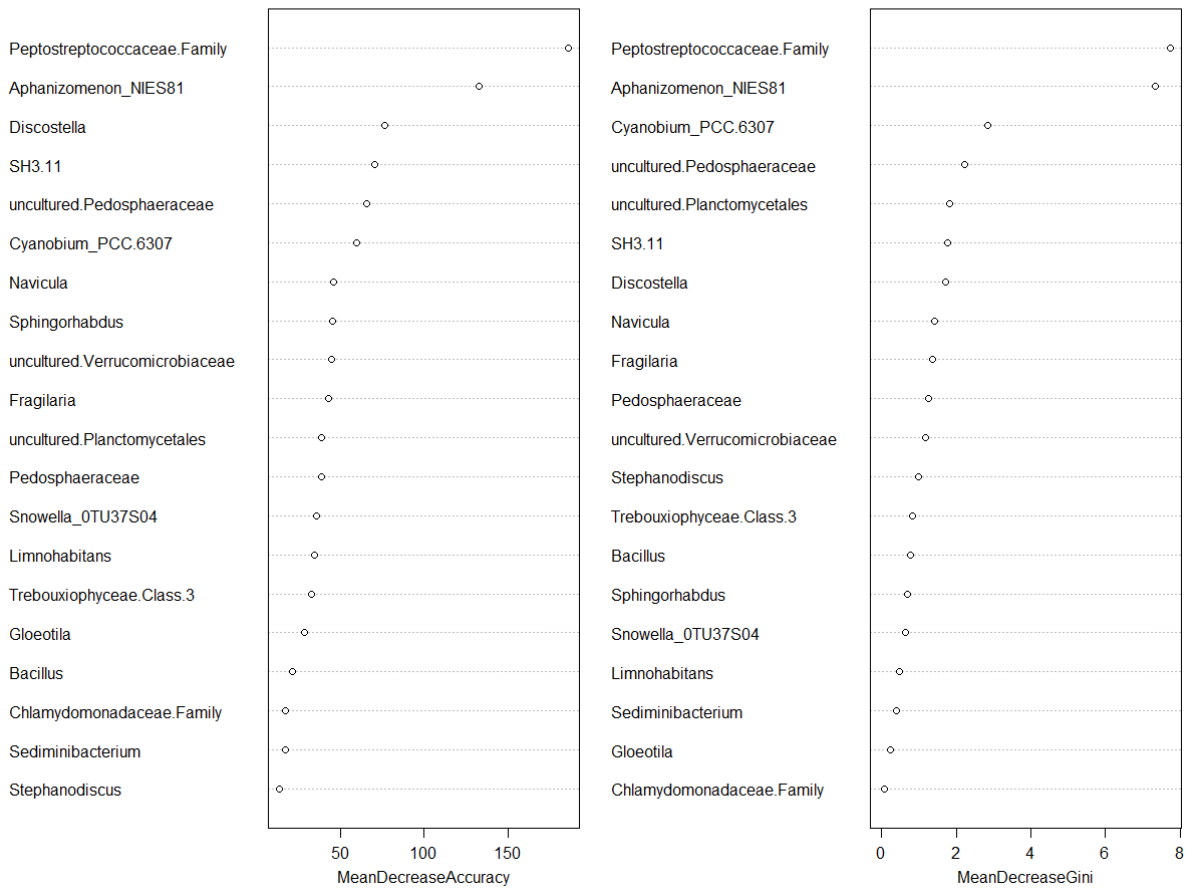


Figure 4.44: Local variable importance plot displaying the Mean Decrease in Accuracy by each taxon (x-axis) out-of-bag cross-validated prediction (left), and the mean decrease in Gini coefficient measuring how each taxon contributes to the homogeneity of the nodes and leaves in Alaw reservoirs random forest.

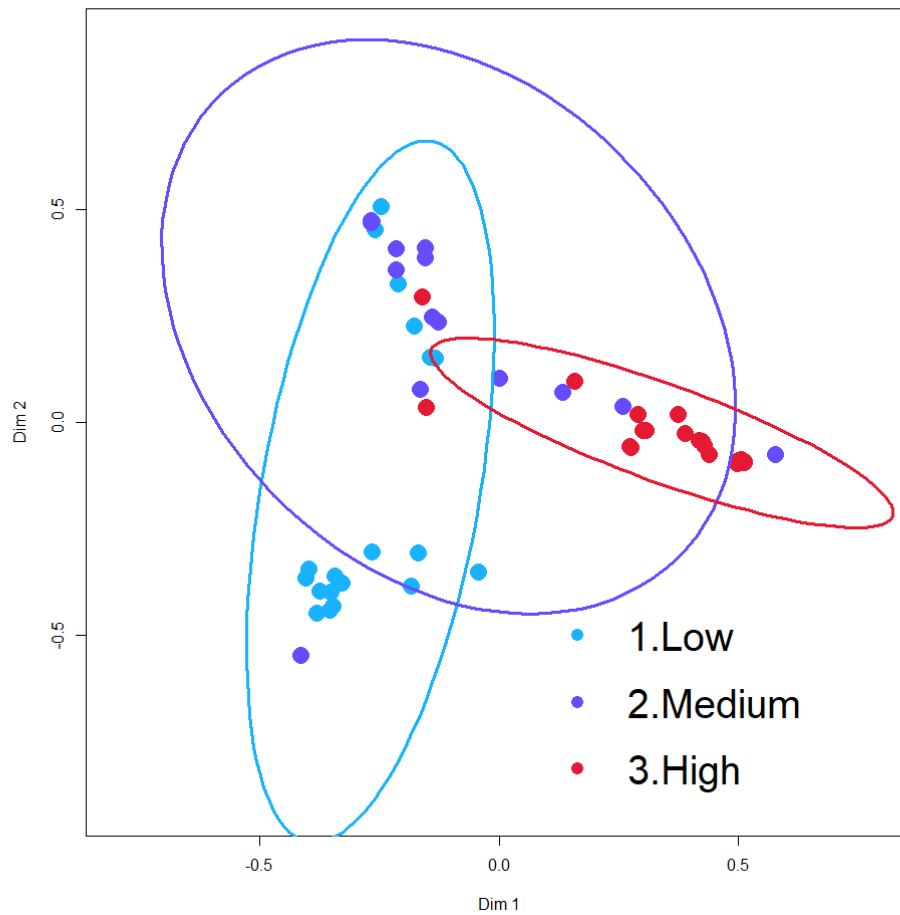


Figure 4.45: Multi-dimensional Scaling Plot of Proximity matrix from the Random Forest output for Alaw reservoir grouped by geosmin levels. Ellipses represent a 95% confidence interval.

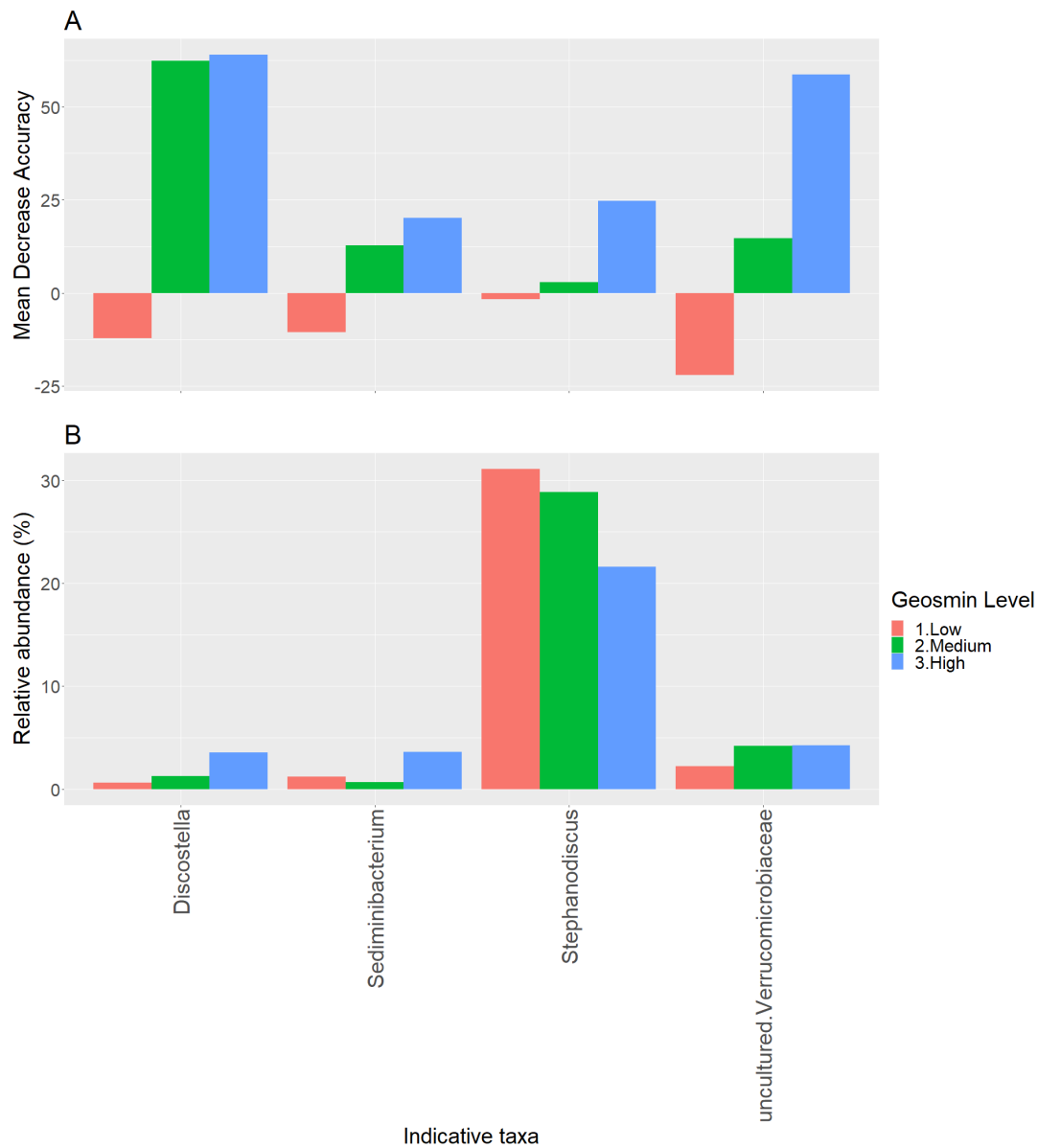


Figure 4.46: Grouped bar plots of A) Mean Decrease in Accuracy values and B) the relative abundance (%) of each proposed “true indicator taxa” coloured by geosmin concentration level classification in Alaw reservoir.

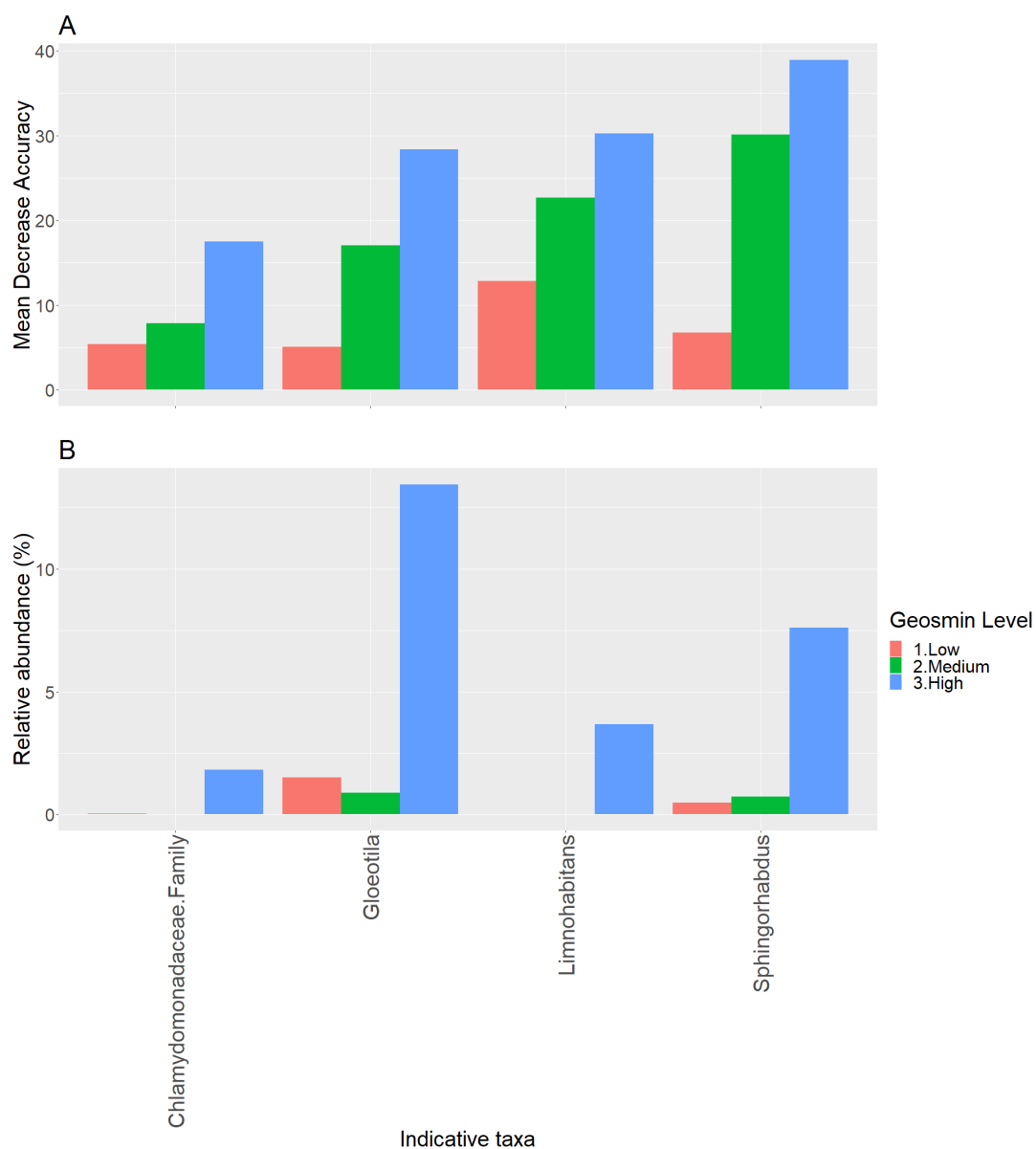


Figure 4.47: Grouped bar plots of A) Mean Decrease in Accuracy values and B) the relative abundance (%) of each proposed indicator taxa coloured by geosmin concentration level classification in Alaw reservoir.

#### 4.3.5.2 – Llwyn On

Discriminatory taxa used in Llwyn On's random forest displayed that *Cyanobium* PCC-6307 was most informative on community structure during low, medium, and high geosmin concentrations (Figure 4.48). The *OM27 clade*, *Prostheco bacter*, *Fragilaria*, *Sphaeropleales* order, and *Achnantheidium* were the second group of most informative taxa indicative of geosmin concentration classification, followed by uncultured *Pedosphaeraceae* and *Chthoniobacter*. The confusion matrix for this random forest model's accuracy (Table 4.11) highlights that caution must be taken when interpreting these results, as the model had a 50 – 100% error rate when trying to assign geosmin level classification for medium and high concentrations, respectively. Due to the model's high error rate when classifying high geosmin levels, no further analysis was done to reveal indicative taxa for this reservoir.

LLWYN\_ON

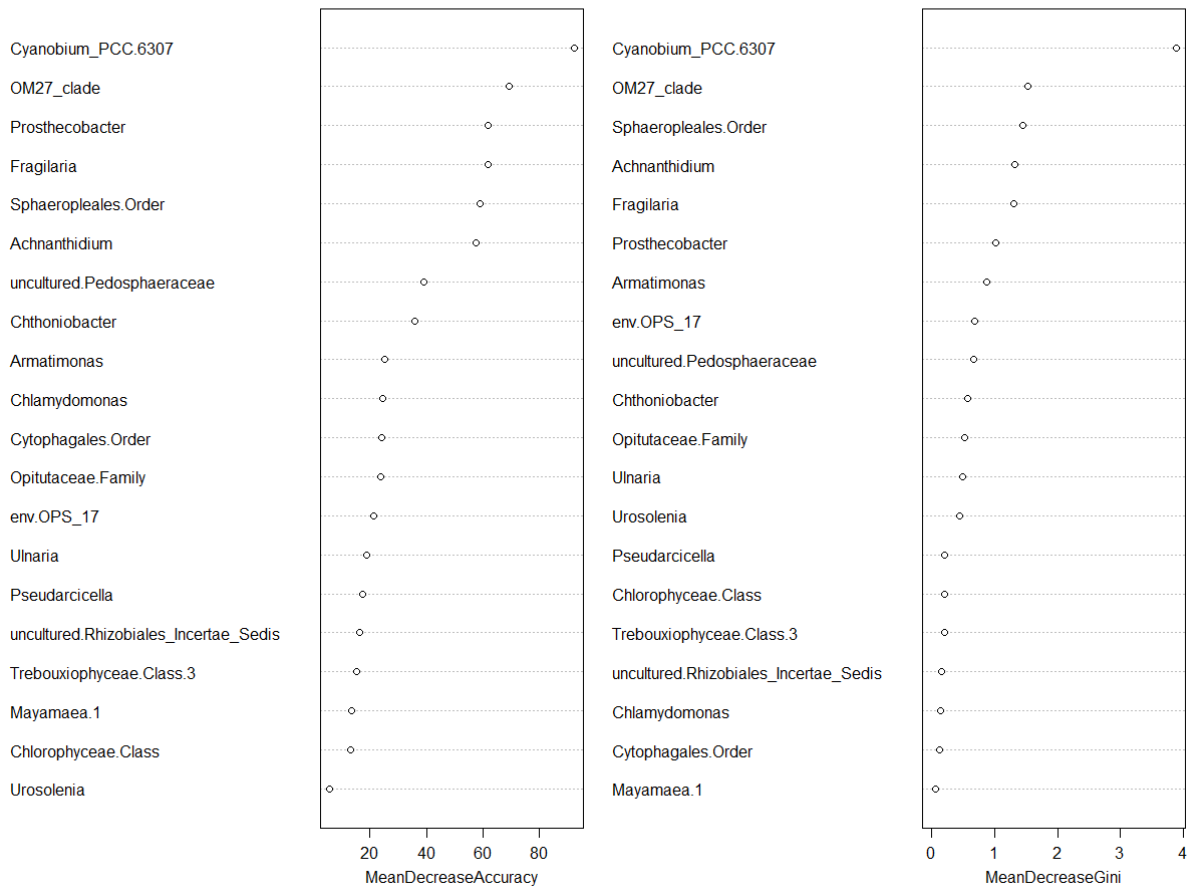


Figure 4.48: Local variable importance plot displaying the Mean Decrease in Accuracy by each taxon (x-axis) out-of-bag cross-validated prediction (left), and the mean decrease in Gini coefficient measuring how each taxon contributes to the homogeneity of the nodes and leaves in Llwyn On reservoirs random forest.

#### 4.3.5.3 – Pentwyn

Discriminatory taxa used in Pentwyn’s random forest displayed that *Limnobacter* was most informative on community structure during low, medium, and high geosmin concentrations (Figure 4.49), followed by the *NS9 marine group*. Referring to the confusion matrix for this random forest model's accuracy (Table 4.11), it can be concluded that results can be used to highlight indicative taxa associated with high 2-MIB concentrations. The success of this model is graphically illustrated in Figure 4.50, where the ordination plot displays a nice cluster of high 2-MIB samples with a tight ellipse with no overlapping 95% confidence intervals. From

exploring the overall MDA value for each taxon (Table 4.12), and the MDA values associated with that taxon for low, medium, and high 2-MIB concentrations, indicative genera of 2-MIB concentration level classification could be determined. Firstly, it was noted that there was a definite gradual increase in the importance of *Cyanobium* PCC-6307 being used in the model from low to high 2-MIB classification (Table 4.12).

The two most influential genera in the model were investigated as potential indicator genera for 2-MIB classification (Figure 4.51). Interestingly, both genera had high MDA values for classifying low 2-MIB concentrations, and both showed poor importance in the model when classifying medium 2-MIB concentrations (Figure 4.51A). *Limnobacter* was considered the most critical genus out of all 20 taxa in the classification of high 2-MIB concentrations and could potentially be used as an indicator of 2-MIB concentrations. The *NS9 marine group* had the second-highest MDA value for classifying high 2-MIB concentrations. These two taxa could help differentiate between low and high 2-MIB concentrations. By exploring the relative abundance of these taxa compared to low and high 2-MIB concentrations, *Limnobacter* was the most important genus, showing minimal abundance during low and medium 2-MIB concentrations and >8% relative abundance during high 2-MIB concentrations (Figure 4.51B). The *NS9 marine group* showed a similar trend to *Limnobacter*, however, the rise in relative abundance during high 2-MIB concentrations was smaller.

PENTWYN

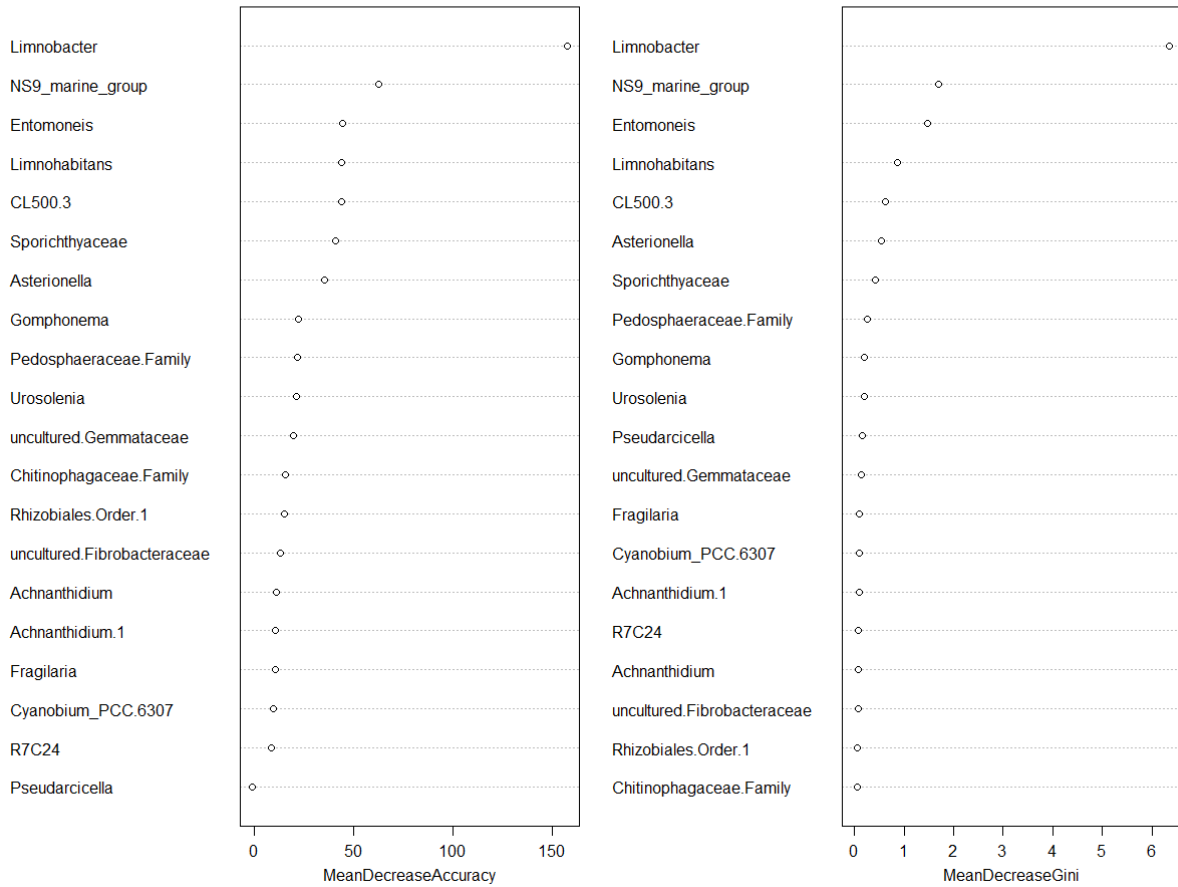


Figure 4.49: Local variable importance plot displaying the Mean Decrease in Accuracy by each taxon (x-axis) out-of-bag cross-validated prediction (left), and the mean decrease in Gini coefficient measuring how each taxon contributes to the homogeneity of the nodes and leaves in Pentwyn reservoirs random forest.



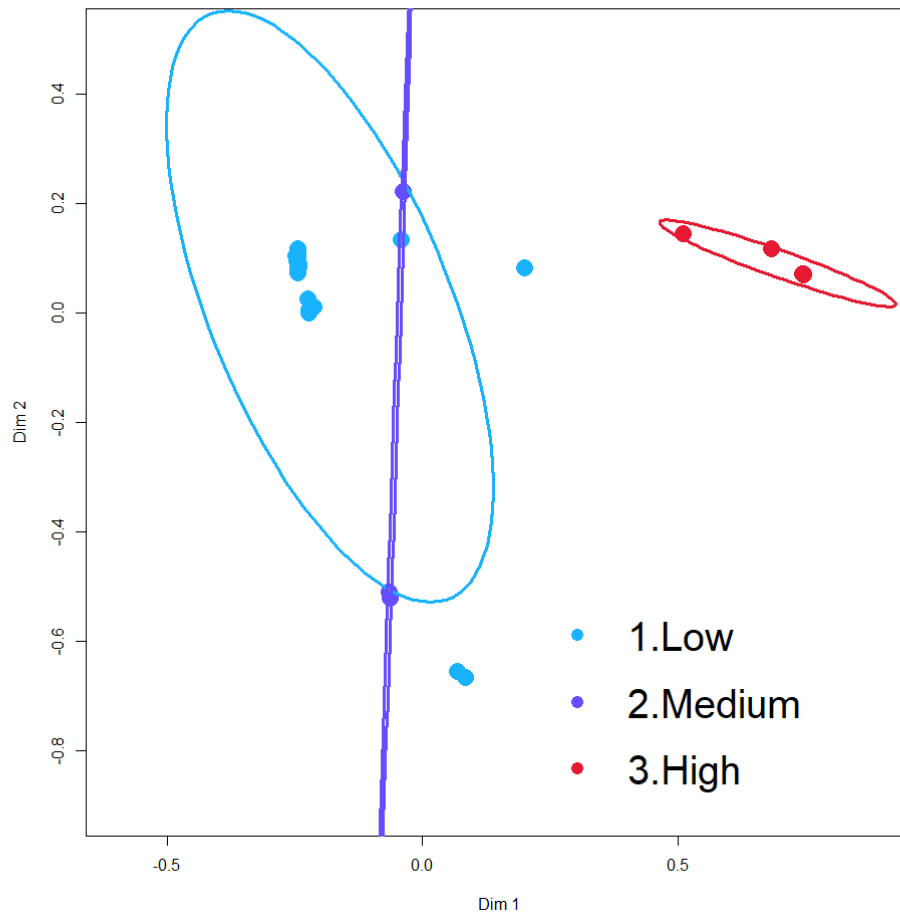


Figure 4.50: Multi-dimensional Scaling Plot of Proximity matrix from the Random Forest output for Pentwyn reservoir grouped by 2-MIB levels. Ellipses represent a 95% confidence interval.

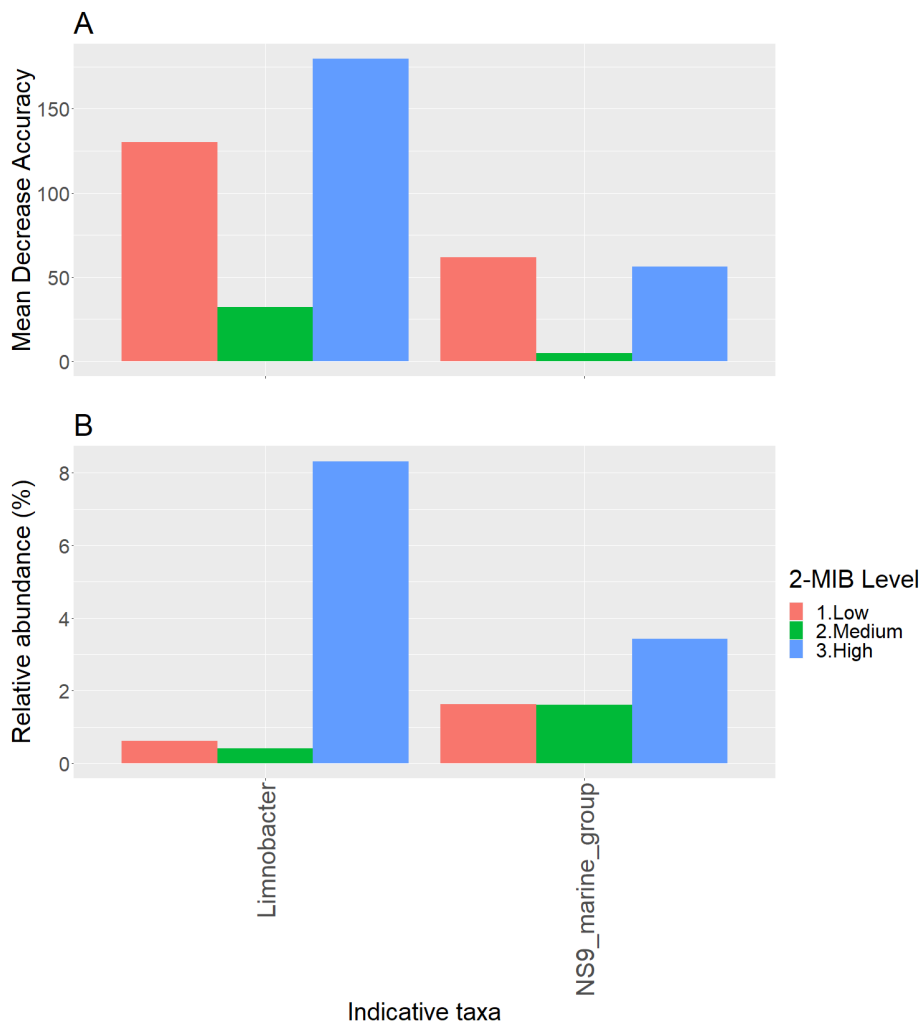


Figure 4. 51: Grouped bar plots of A) Mean Decrease in Accuracy values and B) the relative abundance (%) for the genera with the highest informative contribution to the random forest according to the overall MDA, coloured by 2-MIB concentration level classification in Pentwyn reservoir.

#### 4.4 – Discussion

The rising frequency and magnitude of T&O outbreaks in drinking water reservoirs is a relentless issue for water industries. With the consensus that *Cyanobacteria* will continue to proliferate with increasing air and water temperatures, the increased unpredictability of extreme events and climate change introduces an element of uncertainty for controlling cyanobacterial T&O events. T&O production has previously been linked to cyanobacterial biomass; however, it is evident from this study that it is productivity-driven. Utilising the biology consortium (community composition) as a response to environmental variables in the reservoirs over the

study period reflected general water quality. Further ordination analysis revealed that high T&O concentrations ( $>20 \text{ ng L}^{-1}$  – geosmin and  $>10 \text{ ng L}^{-1}$  – 2-MIB) shaped the bacterioplankton communities and, to some extent, the eukaryotic communities, confirming that the biology could be used to infer T&O outcome in reservoirs experiencing high T&O concentrations. This study identified significantly co-occurring taxa implicated in cyanobacterial T&O production from cyanospheres in each studied reservoir. Significantly co-occurring taxa in a reservoir experiencing high geosmin (Alaw) and 2-MIB (Pentwyn) concentrations enabled the distinction of indicative taxa for high T&O concentrations. Furthermore, the findings presented in this study grant the theory that negative relationships with *Cyanobacteria* can induce T&O production facilitated by symbiotic relationships within the cyanosphere.

#### 4.4.1 – Productivity driven cyanobacterial T&O events

To date, cyanobacterial blooms have been primarily associated with T&O events, being the lead cause of considerable economic impact based on drinking water quality and safety (Bruchet, 2019). These blooms have been labelled the prime cause of deteriorating water quality through toxin production and T&O events (Devi et al., 2021; Paerl et al., 2001). Generally, concentrations of geosmin and 2-MIB in waterbodies have been linked to high *Cyanobacteria* biomass (Peter et al., 2009). However, more recently, Xuwei et al. (2019) found that *Oscillatoria* were responsible for producing T&O compounds in the ‘non-blooming’ area of Lake Taihu, China. In addition, some research suggests that geosmin and 2-MIB production is greatest during high productivity (Zimmerman et al., 1995), which could explain why large 2-MIB concentrations have been reported during times of high ATP synthesis (Behr et al., 2014). The production of geosmin and 2-MIB during high productivity could also explain why some studies have failed to see correlations between cyanobacterial biomass and T&O concentrations (Graham et al., 2010). In this study, a general trend across all studied reservoirs showed maximum cyanobacterial phyla abundance to be confined to the summer and autumnal months. However, despite elevated cyanobacterial abundance being localised to the warmer months, winter geosmin events occurred in the Alaw reservoir, suggesting that the heightened relative abundance of *Cyanobacteria* is not the cause of the production of T&O compounds. Emerging evidence suggests that T&O events are not confined to

summer months, with more winter T&O events being recorded (Dzialowski et al., 2009; Jüttner and Watson, 2007; Wang et al., 2005; Watson et al., 2001). Studies have reported the importance of specific environmental variables for T&O production, e.g., temperature (Zhang et al., 2009), light intensity (Alghanmi et al., 2018) and nutrient ratios (Espinosa et al., 2021b). With key nutrient triggers being noted as ammonium (Hooper et al., 2023; Perkins et al., 2019), altering the  $\text{NH}_4^+:\text{NO}_3^-$  ratio favouring cyanobacterial productivity and hence, T&O outcome. This has led to the investigation of bacterial and algal community composition to be monitored as a response to nutrient influxes and environmental variables that can indicate T&O events in this study.

#### 4.4.2 – The “cyanosphere” in relation to T&O events

Within the cyanosphere, associated taxa are potentially involved in remineralising phytoplanktonic organic matter in exchange for vital nutrients including reduced forms of nitrogen (Bell and Mitchell, 1972). In addition, the exchange of nutrients for the organic matter could also induce short-term changes in cyanobacterial activities, i.e., the production of T&O compounds (Berg et al., 2018). Co-occurrence network analysis revealed potential taxa implicated in the cyanosphere and taxa that co-occurred at a frequency less than the observed number of co-occurrences if the taxa were distributed independently of one another (negatively co-occurring taxa). Interestingly, in the case of a reservoir experiencing extreme geosmin events (Alaw), there were only four identifiable significantly co-occurring *Cyanobacteria*, two of which were confirmed geosmin producers (*Aphanizomenon* NIES81 – Jüttner and Watson, 2007, and the *Nostocaceae* family – Kutovaya and Watson, 2014). The reservoir that contained the most cyanobacterial taxa belonged to Cefni, five of which were capable of geosmin production, yet the maximum geosmin concentration did not exceed a medium classification. In the Cefni reservoir, cyanospheres were interconnected positively or indirectly by nodes and edges, whereas *Aphanizomenon* NIES81 and the *Nostocaceae* family cyanospheres were disconnected from other cyanospheres in the Alaw reservoir. Taxa associated with the *Aphanizomenon* NIES81 cyanosphere in the Alaw reservoir tended to be negatively co-occurring. In contrast, all taxa associated with a

cyanobacterial node in the Cefni reservoir were positive, apart from the *Nostocaceae* family which had a larger number of negatively associated nodes.

In the *Aphanizomenon* NIES81 cyanosphere there was only one positively co-occurring genus, *Fluviicola*, and nine negatively co-occurring taxa. *Fluviicola* is affiliated with the family *Cryomorpaceae* which belongs to the class *Flavobacteriia* and contains many species that play an integral role in the recycling of carbon and energy in freshwater environments (Woyke et al., 2011). *Flavobacteriia* are major decomposers of high-molecular-mass organic matter in water bodies (Cottrell and Kirchman, 2000), e.g., the 2-MIB degrading genus *Flavobacterium* (Clerc et al., 2021). Guedes et al. (2018) have previously shown *Fluviicola* to have positive correlations with *Synechococcus*, which is in accordance with previous findings from natural communities (Bertos-Fortis et al., 2016; Salmaso et al., 2018). *Bacillus*, a known geosmin and 2-MIB degrader (Hsieh et al., 2010) was one negatively co-occurring genus with *Aphanizomenon* NIES81 which followed a similarly antagonistic relationship with a previously described geosmin-producing strain of *Streptomyces* (Ma et al., 2015). Intriguingly, *Aphanizomenon* NIES81 was negatively associated with iron-oxidising taxa, the *Gallionellaceae* family and *TRA3-20*; producing ferric iron ( $\text{Fe}^{3+}$ ), whereas *Cyanobacteria* require it in the ferrous iron ( $\text{Fe}^{2+}$ ) form (Molot et al., 2014). Although some *Cyanobacteria*, like *Aphanizomenon* NIES81, possess organic Fe-binding ligands (“siderophores”), they tend to be hydroxamate siderophores that are water soluble and have relatively weak Fe-binding capacity (Sorichetti et al., 2014). A negative relationship between the presence of the *Gallionellaceae* family and *TRA3-20* iron-oxidising taxa in association with *Aphanizomenon* NIES81 could be due to overwhelming the hydroxamate siderophores with oxidised forms of iron. The *Aulacoseira* genus was also negatively associated with *Aphanizomenon* NIES81, although according to Reynolds et al. (2002), this could be expected as *Aulacoseira* is sensitive to rising pH values which are associated with increases in cyanobacterial productivity (Slavin, 2020). Considering *Aphanizomenon* NIES81 is a planktonic genus (Laamanen et al., 2002) it is not surprising that they were negatively co-occurring with benthic taxa such as uncultured *Verrucomicrobiaceae*, uncultured *Pedosphaeraceae*/*Pedosphaeraceae*. Uncultured *Planctomycetales* use anaerobic denitrification as a way of accessing nitrogen to use for metabolism; with the photosynthetic capacity of *Aphanizomenon*

NIES81 surrounding water would be too oxygenated for anaerobic denitrification (Maldonado et al., 2012; Zhang, 2017). The *CL500-3* genus characteristic of oxygenated waters also had a negative association with *Aphanizomenon* NIES81. Okazaki et al. (2017) found *CL500-3* to be negatively correlated with the *Cyanobacteria*, *Nitrosospira*.

Here we have identified pairs of taxa with similar realised niches through co-occurrence networks. All cyanobacterial nodes were positively or indirectly connected in the Cefni reservoir. Tromas et al. (2018) state that if two taxa have identical ecological niches, one should competitively exclude the other unless the competition is weak due to abundant resources. In practice, closely related taxa often compete for space and resources (Cavender-Bares et al., 2009), favouring specialization to reduce overlap in niche space. Many studies have previously shown that different cyanobacterial genera can co-occur, thus sharing at least some dimensions of their realised niches (Paerl et al., 2001; Yamamoto and Tsukada, 2009). Significantly co-occurring *Cyanobacteria*, namely, *Aphanizomenon* MDT14a, *Aphanizomenon* NIES81, *Cyanobium* PCC-6307, *Gloeotrichia* PYH6, *Microcystis* PCC-7914, the *Nostocaceae* family, *Pseudanabaena* PCC-7429 and *Geitlerinema* LD9 present in Cefni reservoir implied they all shared a broadly similar physico-temporal niche. *Microcystis* has been shown to co-occur with *Dolichospermum* through partitioning niche space (Tromas et al., 2018), here *Dolichospermum* was associated with lower concentrations of dissolved nitrogen (consistent with its known ability to fix nitrogen - Andersson et al., 2015) while *Microcystis* was associated with higher concentrations of dissolved phosphorous. These genus-level traits may have existed in Cefni. Both *Aphanizomenon* genera are diazotrophic and have been shown to favour low total nitrogen concentrations as they fix atmospheric nitrogen and can thus compensate for nitrogen deficiency (Elliott and May, 2008). Diazotrophs like *Aphanizomenon* grow more slowly than other phytoplankton, requiring more iron (Sohm et al., 2011). Within the *Aphanizomenon* NIES81 cyanosphere, the *SH3-11* genus was a resident; this genus has been positively correlated with certain metabolic functions like manganese/ iron transport system (KEGG orthology entry: K09819), *feoA*; ferrous iron transport protein (KEGG orthology entry: K04758) and *feoB*; ferrous iron transport protein (KEGG orthology entry: K04759). *SH3-11* could thus supply *Aphanizomenon* NIES81 with ferrous iron

(Fe<sup>2+</sup>) in return for cyanobacterial exudates in the Cefni reservoir. Similarly, in the *Microcystis* PCC-7914 cyanosphere, *Chaetoceros* was a resident; a diatom usually forming epiphytically diatom-cyanobiont symbioses with *Calothrix* that has a high iron uptake and storage capacity (Iwade et al., 2006; Tuo et al., 2017). The association between *Microcystis* PCC-7914 and *Chaetoceros* suggests that *Chaetoceros* could be providing *Microcystis* PCC-7914 with a supply of iron. A commonly co-occurring pennate diatom shared among three cyanospheres (*Aphanizomenon* MDT14a, *Pseudanabaena* PCC-7429 and *Microcystis* PCC-7914) in the Cefni reservoir was *Asterionella*. Rapid increases in *Asterionella* have been connected to the extraction of major anions (nitrate, phosphate, and silicate), resulting in the depletion and altering of nutrient ratios (TN:TP and NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup>). It is known that a low TN:TP ratio favours *Cyanobacteria* and has been associated with increases in T&O compounds (Espinosa et al., 2021a; Harris et al., 2016). Likewise, NH<sub>4</sub><sup>+</sup> has been identified as a key nutrient trigger for the production of cyanobacterial T&O compounds (Perkins et al., 2019), especially when NH<sub>4</sub><sup>+</sup> is high in relation to NO<sub>3</sub><sup>-</sup> (Howard, 2020). Thus, *Asterionella* being contained within cyanospheres is possibly attributed to nutrient ratio alteration on a microscale which in turn is favouring the proliferation of T&O compounds. The majority of taxa contained within the cyanospheres in the Cefni reservoir have been previously identified as co-occurring with *Cyanobacteria*, *Terrimicrobium* with *Nostoc* epi-bacteria (Satjarak et al., 2021), *Candidatus Aquirestis* presence is often related to dissolved organic matter derived from *Cyanobacteria* (Farkas et al., 2020), *Pirellula* is usually found attached to filamentous algae and *Cyanobacteria* (Clum et al., 2009), *Volvocaceae* family have been previously associated with *Nostoc* as epibionts (Satjarak et al., 2021), metagenomic analysis has revealed *LD29* to be associated with cyanobacterial cells (Mohapatra et al., 2020), uncultured *Microscillaceae* has been shown to be positively correlated with *Microcystis* (Chun et al., 2020), *UKL 13-1* is associated with bloom-forming *Cyanobacteria* (Driscoll et al., 2017), and *Planothidium* cells are usually solitary and found attached to cyanobacterial cells (Stancheva, 2019). Mentioned taxa have been associated with *Cyanobacteria*, although only hypotheses about the interactions have been established, and further work needs to be done to elucidate symbiotic pathways and their implications in T&O production. Some taxa, like uncultured *Elsteraceae* involved in the *Nostocaceae* family cyanosphere, have no associated literature, which makes underpinning relationships difficult. Genera like

*Chlamydomonas* (associated with *Geitlerinema* LD9 cyanosphere) have been more widely studied in association with *Cyanobacteria*. Previous research shows that additions of *Aphanizomenon* or cylindrospermopsin induce genes in *Chlamydomonas* typically upregulated under phosphorous limitation and gives rise to extracellular APase activity that the *Cyanobacteria* can utilise for T&O production (Bar-Yosef et al., 2010).

Interestingly only a small number of co-occurring taxa with *Cyanobacteria* in the Cefni reservoir were negatively associated. Yang et al. (2019) identified *Rhizobiales Incertae Sedis* (an uncultured taxon negatively associated with *Aphanizomenon* MDT14a in the Cefni reservoir) as a bioindicator of river degradation mainly influenced by heightened  $\text{NO}_3^-$ . This function of *Rhizobiales Incertae Sedis* could explain the negative co-occurrence, as *Aphanizomenon* MDT14a would prefer the most reduced form of nitrogen ( $\text{NH}_4^+$ ) (Collos and Berges, 2003) whereas the presence of uncultured *Rhizobiales Incertae Sedis* could indicate heightened levels of  $\text{NO}_3^-$ . Although previous literature suggests *Fragilaria* has been associated with increases in cyanobacterial cell numbers (Espinosa et al., 2020), here, in association with *Gloeotrichia* PYH6 and the *Nostocaceae* family it has been reported to have a negative influence on the *Cyanobacteria*. This negative association is in accordance with the findings from Horn and Uhlmann (1995), finding high yields of *Fragilaria* caused low yields of *Cyanobacteria* and vice versa. More recent work has proven that elevated pH associated with cyanobacterial productivity negatively affected the growth rate of *Fragilaria* and diatom silica deposition (Zepernick et al., 2021). Similarly, *Ulnaria* also experienced a negative co-occurrence with *Gloeotrichia* PYH6, yet there is no supporting literature to explain reasons behind this negative relationship. Although, Kahlert et al. (2022) showed considerable overlap in geographical distribution, habitat and ecological preferences for *Ulnaria* and *Fragilaria*, which could explain why these two genera are negatively associated with the *Cyanobacteria* in question. *Armatimonas* is a mesophile (optimum 30–35°C) and a neutrophile (optimum pH 6.5) which was found to be negatively associated with *Gloeotrichia* PYH6 in Cefni reservoir (Dunfield et al., 2012). As a mesophile, *Armatimonas* would compete with *Gloeotrichia* PYH6 during warmer months, and due to its neutrophile nature, the increased productivity of the *Cyanobacteria* would create an alkaline environment in which it could not compete.



Other negatively co-occurring taxa with cyanobacterial taxa in the Cefni reservoir, like *Bacillus* and the *Gallionellaceae* family, shared the same relationships as previously discussed in the Alaw reservoir. The 2-MIB degrading bacteria, *Flavobacterium*, also shared a negative relationship with *Geitlerinema* LD9. Cai et al. (2018) said *Flavobacterium* could maintain cyanobacterial blooms by catalysing the turnover of complex organic matter released by the *Cyanobacteria* to recycle previously loaded nutrient sources. However, here in this study, the negative relationship could indicate the presence of *Flavobacterium* after *Cyanobacteria* have died off or an antagonistic relationship similar to that described by *Bacillus* with geosmin-producing strains of *Streptomyces* (Zhi et al., 2016). Previously It has been reported that dimethylsulfide evolves from dimethyl- $\beta$ -propiothetin from cultures of *Cyclotella* (Ackman et al., 1966), which could explain the negative relationship between *Cyclotella* and the *Nostocaceae* family. Sulphide can limit  $\text{Fe}^{2+}$  diffusion rates because of insoluble iron sulphide formation (Carignan and Tessier, 1988); the presence of dimethylsulfide could thus prevent cyanobacterial uptake of  $\text{Fe}^{2+}$ . Negative relationships between taxa like the *NS11-12* marine group and *Cyanobacteria* are less conclusive due to the literature not having any clear ecological implications for these associated taxa (Farkas et al., 2020).

Pentwyn reservoir was the only reservoir to experience extreme 2-MIB concentrations ( $58 \text{ ng L}^{-1}$ ), followed by the Pontsticill reservoir ( $7.9 \text{ ng L}^{-1}$ ). There were two significantly co-occurring *Cyanobacteria* in the Pentwyn reservoir capable of 2-MIB production; *Cyanobium* PCC-6307 and *Dolichospermum* NIES41. Within the *Cyanobium* PCC-6307 cyanosphere there were only three positively co-occurring residents. The *Sporichthyaceae* family's primary habitat appears to be soil (Tamura, 2014), and no literature reports any relationships between this family and *Cyanobacteria* – Suggesting a commensalism or mutualism relationship shared by *Cyanobium* PCC-6307 and *Sporichthyaceae* within the sediment of the reservoir. A proposed relationship within the sediments of the reservoir follows *Cyanobium* PCC-6307 preference for benthic growth in the sediment layer (Harland et al., 2014). Another resident of the *Cyanobium* PCC-6307 cyanosphere was *Asterionella*; as previously discussed, rapid increases in this genus leads to the extraction of major anions resulting in nutrient depletion and altering nutrient ratios to favour *Cyanobacteria* growth and productivity, and hence T&O outcome (Krivtsov et al.,

2000). The final resident of the *Cyanobium* PCC-6307 cyanosphere was *Gloeotila*; this genus is indicative of mixed water bodies and is usually found in the clear epilimnia of deep lakes (Reynolds et al., 2002). This relationship indicates vertical mixing and/ or advection of the water body in the Pentwyn reservoir, which could increase vertical nutrient transport to the epilimnion and hence increase cyanobacterial productivity (Merino-Ibarra et al., 2021). Most of the taxa co-occurring with *Cyanobium* PCC-6307 were negatively associated. The negative relationship with *Gomphonema* could result from resource partitioning through nutritional selection. *Gomphonema* is less favoured by ammonia and thrives in the presence of nitrate (Veraart et al., 2008); in contrast, *Cyanobacteria* prefer the most reduced forms of nitrogen, e.g., ammonia (Perkins et al., 2019). Another negative co-occurring genus with *Cyanobium* PCC-6307 was *Ulnaria*; as previously discussed, this genus shares an overlap in geographical distribution, habitat and ecological preferences with *Fragilaria*, which is sensitive to high pH changes caused by cyanobacterial productivity (Kahlert et al., 2022). Uncultured *Rubinisphaeraceae* was also negatively associated with *Cyanobium* PCC-6307. However, not much is known about the *Rubinisphaeraceae* family; it has been reported that members are typically associated with wetlands and bogs, and bioinformatics analysis on sequenced genomes points to the presence of extremely high glycolytic potential in these bacteria (Dedysh and Ivanova, 2019). Uncultured *Gemmataceae* has been found previously attached to diatoms and *Cyanobacteria* (Mujakić et al., 2021) and has a strong phosphorous accumulation ability (Zhang et al., 2003). However, Uncultured *Gemmataceae* was found to negatively co-occur with *Cyanobium* PCC-6307 here in the Pentwyn reservoir. Previous work found *Gemmataceae* to negatively correlate with *Melosira* (Liu et al., 2021), the only resident genus belonging to the *Dolichospermum* cyanosphere.

Only two of the nine studied reservoirs did not experience geosmin or 2-MIB concentrations over the expected event level threshold (Alwen and Llandegfedd). Although Alwen did not have a recorded T&O event, this could result from a small sample size (n = 7) missing any spikes in T&O concentrations and will not be discussed further. In the Llandegfedd reservoir, the two cyanospheres were connected, one with geosmin-producing capabilities (*Aphanizomenon* MDT14a) and the other with 2-MIB-producing capabilities (*Cyanobium* PCC-6307). These two

cyanospheres consisted mainly of positively co-occurring taxa. Many of the taxa involved in both cyanospheres have been previously linked to *Cyanobacteria*, like *Candidatus Aquirestis* (Farkas et al., 2020) and *Limnohabitans* (Luo et al., 2022; Seok Jea Youn et al., 2020). The *Sporichthyaceae* family only contains one genus, *Sporichthya* (Tamura, 2014), associated with both cyanospheres. The primary habitat of *Sporichthyaceae* strains appears to be soil (Tamura, 2014), and no literature reports any relationships between this family and *Cyanobacteria*. Although, associations between the *Sporichthyaceae* family and *Cyanobium* PCC-6307 could exist due to the benthic formation of *Cyanobium* PCC-6307 (Harland et al., 2014). Although *Aphanizomenon* MDT14a is planktonic, the *Sporichthyaceae* family could be included in the cyanosphere due to low water levels in the Llandegfedd reservoir or as a result of sediment upturning. Each cyanosphere had a negative association with only one taxon: the geosmin and 2-MIB degrading genus *Pseudomonas* in the *Aphanizomenon* MDT14a cyanosphere and *Schlesneria* in the *Cyanobium* PCC-6307 cyanosphere. *Pseudomonas* is a known geosmin degrader, with evidence of geosmin degradation in a consortium with *Sphingopyxis* and *Novosphingobium* (Hoefel et al., 2006), and independently yielding 2-ketogeosmin and 7-ketogeosmin as well as several minor products (Eaton and Sandusky, 2010). *Pseudomonas* species employ enzymes inducible by  $\gamma$ -terpinene, which are likely to be involved in  $\gamma$ -terpinene metabolism for which a pathway has not yet been established (Eaton and Sandusky, 2010). Work conducted by LewisOscar et al. (2018) revealed the antipathogenic abilities of *Cyanobacteria* against *Pseudomonas aeruginosa*. The methanolic extract was taken from the *Cyanobacteria*, *Spirulina platensis*, and was shown to inhibit quorum sensing mediated virulence factors like biofilm formation, pyoverdinin (siderophore production), pyocyanin, EPS production, reduced the cell surface hydrophobicity and motility without inhibiting *S. platensis* growth. *Aphanizomenon* MDT14a may be hindering *Pseudomonas* growth through quorum sensing inhibition. *Schlesneria* is a moderate acidophile with an optimum pH between 5 and 6, which was negatively associated with *Cyanobium* PCC-6307 (Kulichevskaya et al., 2007). At a low pH, the inorganic carbon pool contains more highly permeable  $\text{H}_2\text{CO}_3$ , necessitating a substantial expenditure of energy on transport to maintain internal inorganic carbon levels in *Cyanobacteria* (Mangan et al., 2016). An intracellular pH  $\sim 8$  reduces leakage, making the  $\text{CO}_2$  concentrating

mechanism significantly more energetically efficient. This negative relationship could highlight niche partitioning associated with different pH gradients.

#### 4.4.3 – Signature 16S rRNA and *rbcL* communities reflecting T&O concentrations in reservoirs experiencing high T&O concentrations

This is the first report to date of bacterioplankton and eukaryotic communities clustering in ordination plots according to T&O concentrations. Bacterial community profiles for samples from reservoirs experiencing geosmin concentrations exceeding 20 ng L<sup>-1</sup> (Alaw, Dolwen, and Llwyn On) showed distinct clustering in ordination plots. Similarly, bacterial community profiles were seen to be shaped by 2-MIB concentrations in a reservoir experiencing high concentrations above 10 ng L<sup>-1</sup> (Pentwyn) and in a reservoir with medium concentrations between 2.5 – 10 ng L<sup>-1</sup> (Pontsticill). Eukaryotic communities could also be seen to be shaped by geosmin concentrations in the case of Dolwen and Llwyn On reservoirs (although Dolwen's low sample size should be considered). However, there was less distinct clustering within the eukaryotic community in the case of extreme geosmin concentrations (520 ng L<sup>-1</sup>). The eukaryotic community composition in the Pentwyn reservoir was shaped by 2-MIB concentrations, although no distinct clustering was witnessed for medium 2-MIB concentrations in the Pontsticill reservoir.

It is becoming well known that persistently high T&O concentrations entering water treatment plants are enabling the colonisation of sand filters with a consortium of T&O degrading bacteria (Ho et al., 2007; Hoefel et al., 2009; Hsieh et al., 2010; Pham et al., 2015; Vandermaesen et al., 2017), shaping the community structure of the microbiome found within sand filters. Previous fractionation analyses of both T&O compounds showed that 2-MIB was found more frequently in the dissolved fraction while geosmin was mostly cell-bound (Clercic et al., 2021), hypothesised that different T&O compounds would select for different degrading bacteria. Clercic et al. (2021) found geosmin concentrations in a eutrophic reservoir related to *Novosphingobium hassiacum* and *Sphingomonas oligophenolica* and *Flavobacterium* species for 2-MIB concentrations. The roles of *Pseudomonas* and *Bacillus* were ambiguous, yet both have been identified to have the capabilities of degrading both T&O compounds (Eaton and Sandusky, 2010; Ho et al., 2007; Ma et

al., 2015). *Bacillus* in the Alaw reservoir was associated with high geosmin concentrations but not significantly, and In Dolwen reservoir *Bacillus* was not associated with high geosmin concentrations. The potential geosmin degrading genus *Sphingorhabdus* was the most significantly associated bacteria with high geosmin concentrations in Alaw and Llwyn On reservoirs. Some geosmin-degrading bacteria belonging to the *Sphingopyxis* genus (Hoefel et al., 2009) have been reclassified as *Sphingorhabdus*, and some members of this genus have been reclassified as another geosmin-degrading genus, *Novosphingobium* – highlighting the need for further investigation into the geosmin biodegradation capabilities of *Sphingorhabdus* species. Interestingly, in the Llwyn On reservoir, *Asterionella* was significantly associated with high geosmin concentrations. As previously discussed, this genus has been found in cyanospheres contained in other reservoirs, altering nutrient ratios and favouring T&O outcome (Krivtsov et al., 2000). Results from Pentwyn indicate that *Cyanobium* PCC-6307 was responsible for the production of 2-MIB. Although, neither *Novosphingobium* nor *Rhodospirillum rubrum* were significantly associated with high 2-MIB concentrations, both genera having the capabilities to degrade 2-MIB (Clerc, 2014; Clerc et al., 2021).

Results from this study highlight that both geosmin and 2-MIB concentrations may be shaping the bacterioplankton community composition and to a certain extent the eukaryotic algal community composition. Identifying significantly co-occurring organisms and possible interactions between taxa within cyanospheres in reservoirs experiencing T&O events made it possible to pinpoint indicative taxa.

#### 4.4.4 – Indicative taxa for T&O events

Indicative taxa of T&O events were identified in a reservoir experiencing extreme geosmin concentrations (Alaw) and in a reservoir with elevated 2-MIB concentrations (Pentwyn). In the Alaw reservoir, the taxon that contributed the most significant information in classifying geosmin concentrations into a low, medium, and high category was the *Peptostreptococcaceae* family, followed by *Aphanizomenon* NIES81. Interestingly this family is also an indicator of nutrient enrichment in urban lakes (Numberger et al., 2022), functionally linked to amino acid metabolism, biosynthesis of secondary metabolites, carbohydrate metabolism, energy metabolism, glycan biosynthesis and metabolism, lipid metabolism, metabolism of

cofactors and vitamins, metabolism of terpenoids and polyketides, nucleotide metabolism and xenobiotics biodegradation (Ghate et al., 2021). The *Peptostreptococcaceae* family was negatively associated with *Aphanizomenon* NIES81 in the Alaw reservoir. Both taxa are capable of nitrogen fixation; the *Peptostreptococcaceae* family are recognized as hyper-ammonia-producing bacteria capable of producing high ammonia levels from amino acids or peptides (Paster et al., 1993). The *Peptostreptococcaceae* family can be used as a proxy for nutrient enrichment in the reservoir. The absence of this family can also indicate geosmin production by *Aphanizomenon* NIES81 when the *Peptostreptococcaceae* family decline, leaving  $\text{NH}_4^+$  in excess for the *Cyanobacteria* to use..

Using the MDA values from the random forest classification for low, medium, and high T&O concentrations, it was possible to identify discriminatory taxa related to these classifications in Alaw (geosmin) and Pentwyn (2-MIB) reservoirs. True indicative taxa reflecting changes in community composition during medium and high geosmin concentrations were revealed to be *Discostella*, *Sediminibacterium*, *Stephanodiscus*, and uncultured *Verrucomicrobiaceae*. Interestingly, only uncultured *Verrucomicrobiaceae* was associated (negatively) with a geosmin producer, *Aphanizomenon* NIES81 – highlighting that geosmin production is more complex than direct interactions between taxa and that it is the general overall architecture of the community network. Uncultured *Verrucomicrobiaceae* has previously been connected to non-diazotrophic *Cyanobacteria* (Woodhouse et al., 2013) and is implicated in root carbon metabolism (Newitt, 2020). This family belongs to the *Verrucomicrobia* phylum, functions and ecophysiology of this phylum are not well understood. Although metagenomic analysis conducted by He et al. (2017) revealed this phylum to be potential polysaccharide degraders and suggested their adaptation to carbon sources of different origins in a eutrophic lake and a humic bog across multiple years. Their increased relative abundance during medium and high geosmin concentrations suggests that uncultured *Verrucomicrobiaceae* may be utilising cyanobacterial-derived polysaccharides, which inadvertently negatively affected both *Aphanizomenon* NIES81, and *Snowella* OTU37S04. Thus, increases in uncultured *Verrucomicrobiaceae* may be a good indicator for heightened cyanobacterial productivity, and the negative interactions between the *Cyanobacteria* may be inducing geosmin production as a competitive advantage. *Steophanodiscus* is an

indicator of heightened phosphorous concentrations favouring high phosphorous concentrations and a low silica to phosphorous ratio (Youn et al., 2020). Here, in Alaw reservoir *Steophanodiscus* was negatively associated with *Snowella* OTU37S04, and relative abundance of the diatom decreased from medium to high geosmin concentrations. Although *Snowella* OUT37S04 is not a known producer of geosmin, *Steophanodiscus* can be used as a proxy for phosphorous competition between the diatom and *Cyanobacteria*. Both *Discotella* and *Sediminibacterium* were positively associated with the *Cyanobium* PCC-6307 and *Snowella* OTU37S04 cyanospheres, respectively. Interestingly, these two cyanospheres in question overlapped. Both associated taxa in the cyanospheres benefitted the *Cyanobacteria*. Previous genome sequencing of cyanospheres exposed *Sediminibacterium* to be a resident (Sethuraman et al., 2022), offering protection against the colonisation of opportunistic bacteria by producing bacteriocin and toxoflavin. In return, the *Cyanobacteria* provide a stable microhabitat and nutrient supply for the bacteria within the cyanosphere. It has been established that *Sediminibacterium* can form biofilm consortia with *Cyanobacteria* (Velichko et al., 2015), and its relationship with *Cyanobacteria* in wastewater has also been previously reported (Lee et al., 2013). Thus, *Sediminibacterium* can be considered to play a role as a supplier of inorganic nutrients for cyanobacterial growth and hence productivity and T&O outcome. Higher abundances of *Discostella* have been linked to reduced mixing depths and enhanced thermal stratification (Reynolds, 1980), conditions which favour *Cyanobacteria* productivity and hence T&O outcome (Clercin and Druschel, 2019). Studies have documented a decline in epilimnetic nutrient concentrations due to greater thermal stability and reduced mixing, thus giving *Discostella* a physiological advantage when competing for limited nutrients due to the high surface area to volume ratios of these relatively small diatoms (Winder and Hunter, 2008). *Discostella* has been shown to efficiently take up nutrients, often showing a negative relationship to nutrients (Pasciak and Gavis, 1974; Rühland et al., 2015, 2003). Within the *Cyanobium* PCC-6307 cyanosphere, *Discostella* may provide the *Cyanobacteria* with nutrients during times of stratification.

General geosmin indicators present in the Alaw reservoir only indicated high concentrations (>20 ng L<sup>-1</sup>). Here, the *Chlamydomondaceae* family and *Limnohabitans* were only present during high geosmin concentrations. The

*Chlamydomondaceae* family were found to be contained within the *Nostocaceae* family cyanosphere. The *Nostocaceae* family contain members capable of producing cylindrospermopsin (Li et al., 2001), which has been shown to induce the genus *Chlamydomonas* to produce extracellular APase that *Cyanobacteria* can use for growth and productivity in a P-limited environment (Bar-Yosef et al., 2010).

*Limnohabitans* was found to be a resident of the *Snowella* OTU37S04 cyanosphere; this genus has been associated with other cyanospheres constituting 11% of the composition of the cyanosphere (Kim et al., 2020). This genus has been found to thrive along with cyanobacterial blooms (Šimek et al., 2013) and utilises cyanobacterial-derived organic material as a key substrate for growth (Paver et al., 2013). Although the benefits of *Limnohabitans* occupying a cyanosphere are known, the benefits for the *Cyanobacteria* need further investigation, especially in relation to T&O production. Additional general geosmin indicators were *Gloeotila* and *Sphingorhabdus*; these genera' relative abundance significantly increased in times of high geosmin concentrations. *Gloeotila* was negatively associated with *Snowella* OTU37S04, and its relative abundance greatly increased during times of high geosmin production, most likely due to temporal niche partitioning. *Cyanobacteria* tend to be present during the summer months, whereas *Gloeotila* is shown to be present later in the year after summer T&O events (Çelik and Sevindik, 2016). In contrast, *Sphingorhabdus* was entrained within the *Nostocaceae* family cyanosphere, and its relative abundance greatly increased during high geosmin concentrations. *Sphingorhabdus* is a suspected geosmin-degrading genus as it is contained within the *Sphingomonadaceae* family, which includes other known degraders (Glaeser and Kämpfer, 2014). In addition, many geosmin-degrading genera, like *Spingopyxis*, have been reclassified as *Spingorhabdus*, and some members of *Spingorhabdus* have been reclassified as the geosmin-degrading genus, *Novosphingopyxis* (Jogler et al., 2013; Sharma et al., 2021).

Interestingly, indicative taxa for 2-MIB events in the Pentwyn reservoir were not included directly in either cyanosphere. The two most informative taxa from the model were taken; *Limnohabitans* and *NS9 marine group*. Both taxa had negative associations with *Gloeotila* contained within the *Cyanobium* PCC-6307 cyanosphere. *Limnohabitans* and *NS9 marine group* Increases seemed to stimulate high 2-MIB concentrations, possibly through their negative interactions with *Gloeotila*. *Gloeotila*



favours stronger stratification and increased nitrogen-to-phosphorous ratio (Winder and Hunter, 2008). *Cyanobium* PCC-6307 is a non-diazotrophic *Cyanobacteria* and has been shown to contain six copies of nitrate import permease (*nrtABCD*) (Ehrenfels et al., 2021) and could favour a slightly higher nitrogen-to-phosphorous ratio like *Gloeotila*. The disruption in the symbiosis between *Cyanobium* PCC-6307 and *Gloeotila* could stress the *Cyanobacteria* and induce 2-MIB production.

## 4.5 – Conclusions

This study has shown that geosmin and 2-MIB production are productivity driven rather than biomass-related. Geosmin and 2-MIB signature communities were apparent; ordination plots revealed 16S rRNA and, to a lesser extent, *rbcL* communities to cluster according to T&O concentrations for reservoirs experiencing high T&O concentrations. Further network analysis revealed complex architectures between the bacterial and algal communities and exposed cyanospheres and interactions between taxa within the cyanospheres. Antagonistic/negative interactions between *Cyanobacteria* and implicated taxa appeared to encourage T&O production. Negative associations could result from competition between *Cyanobacteria* and implicated taxa for nutrients (e.g.,  $\text{NH}_4^+$ ) during heightened concentrations. Competition for nutrients subsequently stimulates cyanobacterial productivity if conditions are favoured towards the *Cyanobacteria*, resulting in T&O production. T&O production was facilitated by positive interactions with taxa contained within the cyanospheres through nutrient mediation, recycling of nutrients within the cyanospheres and by providing the *Cyanobacteria* with protective agents. The synergy in the cyanospheres with positively co-occurring taxa appeared to reduce the chance of high T&O concentration levels. Identifying significantly co-occurring taxa in reservoirs that exhibited high T&O concentrations (Alaw and Pentwyn) enabled indicative taxa to be exposed through RF analysis. Pinpointing indicative taxa for geosmin production enabled indicator categories to be formed based on function-specific traits shared by the taxa (Table 4.15). Discovering function-specific traits from these indicator taxa permitted the categorisation of taxa co-occurring with *Cyanobacteria* in the Alaw reservoir that possessed shared function-specific traits (Table 4.16). However, the additional indicator taxa presented in Table 4.16 did have an additional category, codon D1 – “siderophore

interference”, composed of iron-oxidising bacteria. For 2-MIB concentrations, indicative taxa for high concentrations were also identified as *Limnobacter* and the *NS9 marine group*. Nutrients, metabolites, and signalling molecules exchanged between *Cyanobacteria* and associated taxa (bacterial and algal) within the cyanosphere dictate the nature of their relationship, which can stimulate T&O production during times of environmental stress.

This is the first study to date identifying indicator taxa associated with heightened T&O risk. Combining *rbcL* and 16S rRNA community compositions revealed synergy between the two different kingdoms, highlighting the need for eDNA analysis to be incorporated into routine water monitoring. Standard phytoplankton (including *Cyanobacteria*) microscopy currently used in water industries does not account for bacterial cell counts, compromising the understanding of essential interactions between taxa emphasised in this study. Applying routine eDNA analysis to drinking water reservoirs could allow water companies to predict the onset of T&O events, especially now that T&O indicator taxa have been revealed.

Table 4.14: An overview of indicative taxa and their function-specific traits for the production of geosmin as determined by Random Forest analysis in the Alaw reservoir.

| Codon | Taxa function                          | Key taxa  | Importance for geosmin production  |
|-------|--|---|--|
| A1    | Geosmin producers                      | <i>Aphanizomenon</i> NIES81   | Codon A1 must be present and cyanobacterial productivity must be increased by influential environmental triggers, e.g., heightened NH <sub>4</sub> <sup>+</sup> and TP concentrations (as discussed in Chapter 3).   |
| A2    | Geosmin degraders                      | <i>Spingorhabdus</i>  | Taxa are indicative of heightened geosmin concentrations. Can also give rise to fluctuations in geosmin concentrations if degradation is not taking into consideration whilst monitoring production.   |
| B1    | Nutrient mediators                     | Members of the <i>Chlamydomonadaceae</i> family                           | Under P deficiency certain <i>Cyanobacteria</i> produce cyanotoxins, e.g., cylindrospermopsin. When exposed to cylindrospermopsin, genes are upregulated in species from this codon that induce rises in the production of extracellular APase for <i>Cyanobacteria</i> to exploit in times of P stress. |
| B2    | Internal loading indicators            | <i>Discostella</i> , <i>Gloeotila</i>                                     | Taxa belonging to codon B2 can be indicative of well-mixed water bodies; resuspension of sediments in well-mixed environments increases the return of nutrients from the sediments back to the water column.   |
| B3    | Nutrient enrichment indicators         | Members of the <i>Peptostreptococcaceae</i> family, <i>Stephanodiscus</i> | These taxa can be used as a proxy for nutrient enrichment. Altering NH <sub>4</sub> <sup>+</sup> :NO <sub>3</sub> <sup>-</sup> and TIN:TP ratios favouring cyanobacterial productivity and hence geosmin outcome.  |
| B4    | Recycling nutrients within cyanosphere | Members of the <i>Verrucomicrobiaceae</i> family, <i>Limnohabitans</i>    | Members of this codon are found to thrive alongside <i>Cyanobacteria</i> and have been shown to use <i>Cyanobacteria</i> -derived organic matter as key substrates for growth. Important in biogeochemical cycling within the cyanosphere.   |
| C1    | Offers protection                      | <i>Sediminibacterium</i>  | This codon contains taxa that are capable of the production of protective agents, e.g., bacteriocin and toxoflavin, offering <i>Cyanobacteria</i> with protection against opportunistic bacteria in return for a stable microhabitat.  |

Table 4.15: Additional indicator taxa taken from co-occurrence analysis for geosmin production based on shared function-specific traits with indicator taxa identified by Random Forest analysis in the Alaw reservoir.

| Codon | Taxa function                          | All example representatives   |
|-------|--|---|
| A1    | T&O producers                          | <i>Aphanizomenon</i> NIES81, members of the <i>Nostocaceae</i> family, <i>Cyanobium</i> PCC-6307                      |
| A2    | T&O degraders                          | <i>Spingorhabdus</i> , <b><i>Bacillus</i></b>   |
| B1    | Nutrient mediators                     | Members of the <i>Chlamydomonadaceae</i> family, <b>members of the <i>Planctomycetales</i> order, SH3-11</b>          |
| B2    | Internal loading indicators            | <i>Discostella</i> , <b><i>Gloeotila</i>, members of the <i>Pedosphaeraceae</i> family</b>                            |
| B3    | Nutrient enrichment indicators         | <b>Members of the <i>Peptostreptococcaceae</i> family, <i>Stephanodiscus</i>, <i>Aulacoseira</i>, <i>Navicula</i></b> |
| B4    | Recycling nutrients within cyanosphere | Members of the <i>Verrucomicrobiaceae</i> family, <i>Limnohabitans</i> , <i>Fluviicola</i>                            |
| C1    | Offers protection                      | <i>Sediminibacterium</i> , <i>Thalassiosira</i>   |
| D1    | Siderophore interference               | <b>Members of the <i>Gallionellaceae</i> family, TRA3-20</b>  |

*Bolded taxa names represent negative associations with cyanobacterial nodes.*

# Chapter 5: The importance of nutrient ratios in determining elevations in geosmin synthase (*geoA*) and 2-MIB cyclase (*mic*) resulting in Taste and Odour events



## 5.1 – Introduction

Taste and odour (T&O) are the primary sensory considerations used by customers to assess the quality of drinking water (Kehoe et al., 2015). Odorous or unpalatable T&O compounds in treated drinking water can erode customer trust in water quality and generate complaints to water companies worldwide (Webber et al., 2015). Although these compounds pose no risk to human health (Sotero-Martins et al., 2021), significant, costly treatment of drinking water is required to remove them. Adsorption by activated carbon is considered an effective measure in the removal of T&O compounds (Kim et al., 2014b). However, Rodriguez (2018) estimated that to remove 15 ng L<sup>-1</sup> of a T&O compound at a flow rate of 40 million gallons per day, 5,077 kg of Powdered Activated Carbon (PAC) would be required. With PAC having a market cost of around 1.2 – 2 \$ kg<sup>-1</sup> (Alhashimi and Aktas, 2017), this is very costly.

Geosmin (trans-1-10 dimethyl-trans-9-decalol) and 2-MIB (2-methylisoborneol) are the most common compounds associated with T&O complaints worldwide (Clercín and Druschel, 2019; Echenique et al., 2006; Hayes and Burch, 1989; Menezes et al., 2020; Perkins et al., 2019; van Rensburg et al., 2016). These compounds are produced by a variety of distantly related bacteria, including *Actinobacteria*, *Cyanobacteria* and *Proteobacteria* (Watson, 2003), but *Cyanobacteria* are considered to be the main producers of the volatile T&O compounds in aquatic environments (Suurnäkki et al., 2015). *Cyanobacteria* have also established important connections with other phytoplankton. For example, Bar-Yosef et al. (2010) revealed a close relationship between other phytoplankton capable of alkaline phosphatase (AP) production and the cyanotoxin cylindrospermopsin. This cyanotoxin stimulates algae to produce APs used to access orthophosphate from organic phosphorous, which in turn facilitates cyanobacterial growth (Bar-Yosef et al., 2010). This is supported by Olsen et al. (2017), who suggested that, specifically, diatoms influence geosmin and 2-MIB production, as demonstrated by stronger correlations between 2-MIB production and diatom abundance as compared to cyanobacterial abundance. This is consistent

with previous findings that link diatoms with the proliferation of geosmin and 2-MIB (Izaguirre and Taylor, 1998; Schrader et al., 2011; Sugiura et al., 2004, 1998).

Geosmin and 2-MIB exist as an irregular sesquiterpenoid and a monoterpene, respectively (Watson and Jüttner, 2019). Geosmin and 2-MIB are produced along the metabolic pathways for isoprenoid synthesis involved in the 2-methylerythritol-4-phosphate isoprenoid (MEP) pathway, the mevalonate pathway, and the leucine pathway (Jüttner and Watson, 2007). The molecular foundation of geosmin production is from the geosmin synthase gene (*geoA*), which encodes for a bi-functional domain enzyme in a two-step  $Mg^{2+}$  dependent reaction (Churro et al., 2020). The N-terminal part of the enzyme causes the ionization and cyclization of farnesyl diphosphate (FPP) into germacradienol, whilst the C-terminal part facilitates the protonation, cyclization, and fragmentation of the precursor germacradienol molecules into geosmin and acetone (Watson et al., 2016). For the biosynthesis of 2-MIB in Cyanobacteria, there are two metabolic steps: firstly, a S-adenosylmethionine-dependent methylation of the monoterpene precursor geranyl diphosphate (GPP) to 2-methyl-GPP catalyzed by geranyl diphosphate 2-methyltransferase (GPPMT), and secondly, further cyclization of 2-methyl-GPP to 2-MIB catalyzed by 2-MIB cyclase (*mic*) forming a putative operon (Figure 2) (Giglio et al., 2011). The Discovery of the *geoA* and *mic* genes has enabled biomolecular methods like quantitative polymerase chain reaction (qPCR) to become available to monitor their abundance (Cane and Watt, 2003; Dickschat et al., 2007; Steven Giglio et al., 2011; Gust et al., 2003; Komatsu et al., 2008; Wang and Cane, 2008). However, to the best of our knowledge, no primers currently developed target all geosmin and 2-MIB-producing *Cyanobacteria*.

Geosmin and 2-MIB are recalcitrant to conventional drinking water treatment procedures such as clarification, filtration, and oxidation using chlorine (Srinivasan and Sorial, 2011). With both compounds exhibiting extremely low odour thresholds ( $1.3 \text{ ng L}^{-1}$  for geosmin and  $6.3 \text{ ng L}^{-1}$  for 2-MIB; Young et al., 1996), it is of vital importance that water companies remove the source of these compounds through mitigative measures before proceeding to water treatment in order to avoid customer dissatisfaction and ultimately complaints. This is of particular concern considering that climate change can promote cyanobacterial bloom formation, which can lead to more T&O events worldwide (Davis et al., 2009; Taranu et al., 2015; J. Zhang et al.,



2017). Warming can encourage cyanobacterial growth selectively as they have higher optimal growth temperatures than eukaryotic algae (J. Zhang et al., 2017). Shen et al. (2022) found that warmer temperatures favoured cyanobacterial growth, leading to increased T&O compounds. However, lower temperatures (15°C) have also been shown to promote the expression levels of *geoA* and *mic* genes compared to 25°C and 35°C (Shen et al., 2022). Jeong et al. (2021) reported that *Pseudanabaena yagii* produced 2-MIB during the summer season and released 2-MIB under low-temperature conditions in the autumn.

Many studies have focused on the environmental triggers for geosmin and 2-MIB production (Saadoun et al., 2001; Journey et al., 2013; Oh et al., 2017; Clercin and Druschel, 2019). Individual studies have focused on *geoA* and *mic* copy numbers in relation to seasonal occurrences of benthic production of geosmin and 2-MIB (Gaget et al., 2020), the transcription of the genes in response to temperature (Shen et al., 2022), and developing early detection methods (Chiu et al., 2016; John et al., 2018; Suurnäkki et al., 2015). However, to the best of our knowledge, no studies to date have evaluated the combined effects of seasonality, temperature, and nutrient concentrations on *geoA* and *mic* copy numbers together. Kutovaya and Watson (2014) developed taxon-specific PCR and qPCR assays for the early detection of geosmin, however, they were unable to determine any correlations between gene expression, temperature, and nutrient concentrations. The usual nutrients implicated in the production of geosmin and 2-MIB are nitrogen and phosphorous (Harris et al., 2016). Further, Molot et al. (2014) proposed a critical role of ferrous iron and sulphate for cyanobacterial bloom formation, yet neither ferrous iron nor sulphate have been evaluated in relation to geosmin and 2-MIB production before.

Here we assess the associations between seasonality, temperature and nutrients and the production of geosmin and 2-MIB in reservoir drinking water. We employ quantitative Polymerase Chain Reaction (qPCR) to quantify the gene abundance of *geoA* and *mic* in nine reservoirs across Wales, U.K., using a newly developed reverse *mic* primer to aid the detection of 2-MIB producing *Cyanobacteria*. Findings are discussed with relevance to triggers for, and prediction of, T&O events in drinking water supply.

## 5.2 – Materials and methods

### 5.2.1 – Defining a T&O event

Classification of a defined low, medium and, high geosmin and 2-MIB “event” is detailed in Chapter 2 – Materials and Methods, Section 2.2.

### 5.2.2 – Sample locations

Samples were collected according to Chapter 2 – Materials and Methods 2.1, from sites detailed in Sections 2.1.1 – 2.1.7 with the sampling exception mentioned in Section 2.1.8. Sampled months were calculated according to Chapter 2 – Material and Methods 2.1. Seasons were assigned as spring; March 1<sup>st</sup> – May 31<sup>st</sup>, summer; June 1<sup>st</sup> – August 31<sup>st</sup>, autumn; September 1<sup>st</sup> – November 30<sup>th</sup>, and winter; December 1<sup>st</sup> – February 28<sup>th</sup>.

### 5.2.3 – Genomic eDNA extraction

Water was filtered and eDNA was extracted according to Chapter 2 – Materials and Methods, Section 2.4.

### 5.2.4 – Putative classification of standards used for 16S rRNA, *geoA* and *mic* standard curves

Cyanobacterial isolates provided by culture collections were expanded in aseptic cultures at 18°C under a 8:16 light:dark cycle with 40-60  $\mu\text{mol m}^{-2} \text{s}^{-1}$  illumination in BG11 media (Sigma-Aldrich) supplemented with Z8 algal trace metal (see Rippka, 1988). Cells were harvested during log phase growth by centrifugation of 10 mL of culture at 16000 RCF for 10 mins. The cell pellet was resuspended by vortexing in 500  $\mu\text{L}$  ATL buffer (QIAGEN, Germany) with 50  $\mu\text{L}$  of proteinase K (20  $\text{mg mL}^{-1}$  – QIAGEN, Germany) and 20  $\mu\text{L}$  of Lysozyme (25  $\text{mg mL}^{-1}$ ) and incubated at 59°C for 60 mins. Post-incubation 100  $\mu\text{L}$  of the solution was transferred to a fresh tube and DNA was purified as described for the material removed from the Sterivex filters (Section 5.2.3). DNA was extracted from mid-log *Streptomyces* strain S.

*coelicolor* (provided by Dr G. Webster – Cardiff University) using MP Biomedicals™ FastDNA Spin Kit for Soil (MP Biomedicals, USA) as per manufacturer's instructions.

Before sequencing, RNA was removed by the addition of 2  $\mu\text{L}$  of RNase solution (4 mg  $\mu\text{L}^{-1}$ ) to 50  $\mu\text{L}$  of extract sample. Subsequently, DNA was purified using SPRI beads (Beckman Coulter Ltd) using a 0.6:1, Bead:DNA ratio. DNA was quantified using qubit and quality assessment using a TapeStation (Agilent).

DNA concentrations were normalised to 0.25 ng  $\mu\text{L}^{-1}$ , and libraries were generated and indexed using the Nextera XT workflow kit (as per manufacturer's instructions) with conditions optimised to generate an insert size of 200-250 bp. MiSeq v2 cartridge was used to generate 150 bp paired-end data of approx. 100x coverage of the individual genomes. De-multiplex data were trimmed using Trimmomatic (v 0.36) (Bolger et al., 2014) and assembled and polished using Unicycler (v 0.4.7) (Wick et al., 2017). Prokka (v 1.13.3) (Seemann, 2014) was used to annotate the draft assemblies exploiting Barrnap (v 0.9) to predict 16S ribosomal RNAs (<https://github.com/tseemann/barrnap>). Species identification was performed by blast analysis of the 16S ribosomal RNA against the NCBI 16S ribosomal database.

The acquisition of strains used for the generation of each gene qPCR standard curve (*geoA*, *mic*, 16S rRNA) are displayed in Table 5.1, along with the genomic results of the putative 16S rRNA gene classification of the strains used. 16S rRNA gene numbers were used to normalise copy numbers for *geoA* and *mic* samples to account for biomass.

Table 5.1: List of strains used for standard curves in all qPCR reactions with their origin and the putative 16S rRNA classification after genomic classification with corresponding percentage identity.

| Culture strain                             | Culture collection  | 16S rRNA classification (identity %)   |
|--|---|--|
| <i>Anabaena</i> sp. 1446/1c                | Culture Collection of Algae and Protozoa (CCAP)                         | <i>Microcystis aeruginosa</i> NIES 843 (99.66%)  |
| <i>Anabaena flos-aquae</i> 30.87           | The collection of algae cultures at the University of Göttingen (SAG)   | <i>Anabaena cylindrica</i> PCC 7122 (95.33%)<br><i>Massilia agri</i> (98.95%)<br><i>Methanospirillum psychrodurum</i> (86.14%)<br><i>Gemmatimonas aurantiaca</i> T-27 (90.32%) |
| <i>Cylindrospermopsis raciborskii</i> 1.97 | The collection of algae cultures at the University of Göttingen (SAG)   | <i>Anabaena cylindrica</i> PCC 7122 (95.43%)<br><i>Peteryoungia desertarenae</i> (98.51%)<br><i>Fluviicola taffensis</i> (95.90%)  |
| <i>Oscillatoria</i> sp. UHCC 0327          | HAMBI Microbial Culture Collection, University of Helsinki (HAMBI/UHCC) | <i>Oscillatoria nigro-viridis</i> (95.27%)<br><i>Roseomonas mucosa</i> (99.93%)<br><i>Massilia putida</i> (98.51%)   |
| <i>Tolypothrix</i> sp. UHCC 0328           | HAMBI Microbial Culture Collection, University of Helsinki (HAMBI/UHCC) | <i>Oscillatoria nigro-viridis</i> (99.40%)   |
| <i>Planktothrix agardhii</i> NIVA-CYA 126  | The Norwegian Culture Collection of Algae (NORCCA)                      | <i>Planktothrix agardhii</i> NIES-204 (100%)<br><i>Hydrogenophaga palleronii</i> (98.75%)<br><i>Cypionkella psychrotolerans</i> (97.84%)                                       |
| <i>Planktothrix</i> sp. 18                 | HAMBI Microbial Culture Collection, University of Helsinki (HAMBI/UHCC) | <i>Planktothrix agardhii</i> NIES-204 (99.93%)   |
| <i>Streptomyces coelicolor</i> M145        | Cardiff University  | <i>Streptomyces anthocyanicus</i> (100%)   |

## 5.2.5 – qPCR 16S rRNA, *geoA* and *mic* standards preparation

A summary of isolated strains utilised for the preparation of all three gene templates for the establishment of standard curves is detailed in Table 5.2. Standard curve template formation for all three qPCR reactions was performed through PCR using a SimpliAmp™ Thermal Cycler (Thermo Fisher Scientific, USA) using the primer pairs detailed in Table 5.3. PCR's were carried out with the thermocycling conditions listed in Table 5.5 for 16S rRNA, Table 5.6 for *geoA* and Table 5.7 for *mic*.

Table 5.2: Isolated bacterial strains used for the qPCR template preparation for 16S rRNA, *geoA* and *mic* used in the standard curve set-up.

| Strain                                     | 16S rRNA | <i>geoA</i> | <i>mic</i> |
|--|----------|-------------|------------|
| <i>Anabaena flos-aquae</i> 1446/1c         | x        | x           | x          |
| <i>Cylindrospermopsis raciborskii</i> 1.97 | x        |             | x          |
| <i>Anabaena flos-aquae</i> 30.87           | x        | x           | x          |
| <i>Planktothrix</i> sp. 18                 | x        | x           | x          |
| <i>Streptomyces coelicolor</i> M145        | x        |             |            |
| <i>Planktothrix aegardhii</i> cya 126      |          |             | x          |
| <i>Oscillatoria</i> sp. UHCC 0327/2        |          |             | x          |
| <i>Tolypothrix</i> sp. UHCC 0328           |          |             | x          |

Table 5.3: Primer pairs used to amplify 16S rRNA, *geoA* and *mic* gene templates during the initial PCR when generating qPCR standards.

| Target gene      | Primer   | Sequences 5' → 3'         | Product length (bp) | Reference               |
|------------------|----------|---------------------------|---------------------|-------------------------|
| 16S rRNA         | 27F      | AGAGTTTGATCMTGGCTCAG      | 1490                | DeLong (1992)           |
|                  | 1492R    | GGTTACCTTGTTACGACTT       |                     |                         |
| Geosmin synthase | geo78F   | GCATTCCAAAGCCTGGGCTT<br>A | 905                 | Suurnäkki et al. (2015) |
|                  | geo982R  | TCGCATGTGCCACTCGTGAC      |                     |                         |
| 2-MIB cyclase    | MIBS02 F | ACCTGTTACGCCACCTTCT       | 307                 | Chiu et al. (2016)      |
|                  | MIBS02 R | CCGCAATCTGTAGCACCATG      |                     |                         |

Table 5.4: PCR master mix volumes for a 50  $\mu\text{L}$  reaction volume for primer pairs 27F and 1492R, *geo*78F and *geo*982R and MIBS02F and MIBS02R respectively.

|                                | 16S rRNA single reaction / $\mu\text{L}$ | <i>geoA</i> single reaction / $\mu\text{L}$ | <i>mic</i> single reaction / $\mu\text{L}$ |
|--------------------------------|--|---|--|
| Nuclease free H <sub>2</sub> O | 35                                       | 31.50                                       | 31.50                                      |
| AllTaq Mastermix (4x)          | 12.50                                    | 12.50                                       | 12.50                                      |
| Forward primer                 | 0.75                                     | 2.50  | 2.50                                       |
| Reverse primer                 | 0.75                                     | 2.50  | 2.50                                       |
|                                |  |   |  |
| Volume of template DNA         | 1  | 1   | 1  |
| Final volume of PCR reaction   | 50                                       | 50  | 50   |
| Final primer concentration     | 0.30 pmol $\mu\text{L}^{-1}$             | 0.50 pmol $\mu\text{L}^{-1}$                | 0.50 pmol $\mu\text{L}^{-1}$               |

Table 5.5: PCR conditions used for 16S rRNA primer pair 27F and 1492R.

|           | PCR stages           | Time (minutes)                            | Temperature $^{\circ}\text{C}$ |
|-----------|----------------------|---|--------------------------------|
|           | Initial denaturation | 2:00                                      | 95                             |
| 36 Cycles | Denaturation         | 0:30                                      | 94                             |
|           | Annealing            | 0:30                                      | 52                             |
|           | Extension            | 1:30<br>Increased by 1 second every cycle | 72                             |
|           | Final Extension      | 5:00                                      | 72                             |
|           | Hold                 | $\infty$                                  | 4                              |

Table 5.6: The PCR conditions used for *geoA* primer pair 78F and 982R.

|           | PCR stages           | Time (minutes) | Temperature $^{\circ}\text{C}$ |
|-----------|----------------------|----------------|--------------------------------|
|           | Initial denaturation | 2:00           | 95                             |
| 40 Cycles | Denaturation         | 0:30           | 95                             |
|           | Annealing            | 0:30           | 58                             |
|           | Extension            | 0:45           | 72                             |
|           | Final Extension      | 5:00           | 72                             |
|           | Hold                 | $\infty$       | 4                              |

Table 5.7: The PCR conditions used for *mic* primer pair MIBS02F and MIBS02R.

|           | PCR stages           | Time (minutes) | Temperature °C |
|-----------|----------------------|----------------|----------------|
|           | Initial denaturation | 5:00           | 95             |
| 40 Cycles | Denaturation         | 0:10           | 95             |
|           | Annealing/ Extension | 0:20           | 60             |
|           | Final Extension      | 5:00           | 60             |
|           | Hold                 | ∞              | 4              |

All amplicon templates were cleaned and purified with the QIAquick PCR purification kit (QIAGEN, Germany) as per the manufacturer's instructions. Additionally, before use, standards were quality control checked using the QIAxcel (QIAGEN, Germany) to ensure amplicons were of the expected size; 905 bp, 307 bp and 1490 bp for *geoA*, *mic* and 16S rRNA, respectively. Six replicates of the stock template for 16S rRNA, *geoA* and *mic* can be seen in Appendices 5.1 – 5.2.

### 5.2.6 – Reverse primer design for *mic* gene

For the *mic* gene, a new reverse primer was designed using Geneious 9.1.8 for the qPCR. This software aligned genes involved in 2-MIB production from different organisms and identified a conserved region of the genes to allow a reverse primer to be formed. A phylogenetic representation of the organisms used for the construction of the reverse primer can be found in Figure 5.1. The forward primer used in conjunction with the newly developed reverse primer was MIBS02F from Chiu et al. (2016). The newly designed reverse primer was named MIBAHR and can be seen in Figure 5.2. All organisms listed in Figure 5.1 were able to detect both MIBS02F and MIBAHR. *In silico* testing of the two primers was also performed in the statistical software R using the package 'primerTree' (Cannon et al., 2016), which identified 30 species that can be detected using this primer set. A list of all species that can be identified using the primers is shown in Supplementary Information 5.3. Using the reverse primer from Chiu et al. (2016) in conjunction with MIBS02F, only 23 species were identified. When using primers Mtcf and Mtrc from Wang et al. (2011), only 13 species were identified; using primers MIB3324F MIB4050R and MIB3313F and MIB4226R from Suurnäkki et al. (2015), only 17 and 7 species were able to be identified. All species identified by the primers listed are presented in

Appendix 5.4 for comparison. Meaning MIBAHR may be more selective to the *mic* gene, as indicated by the *in silico* testing.

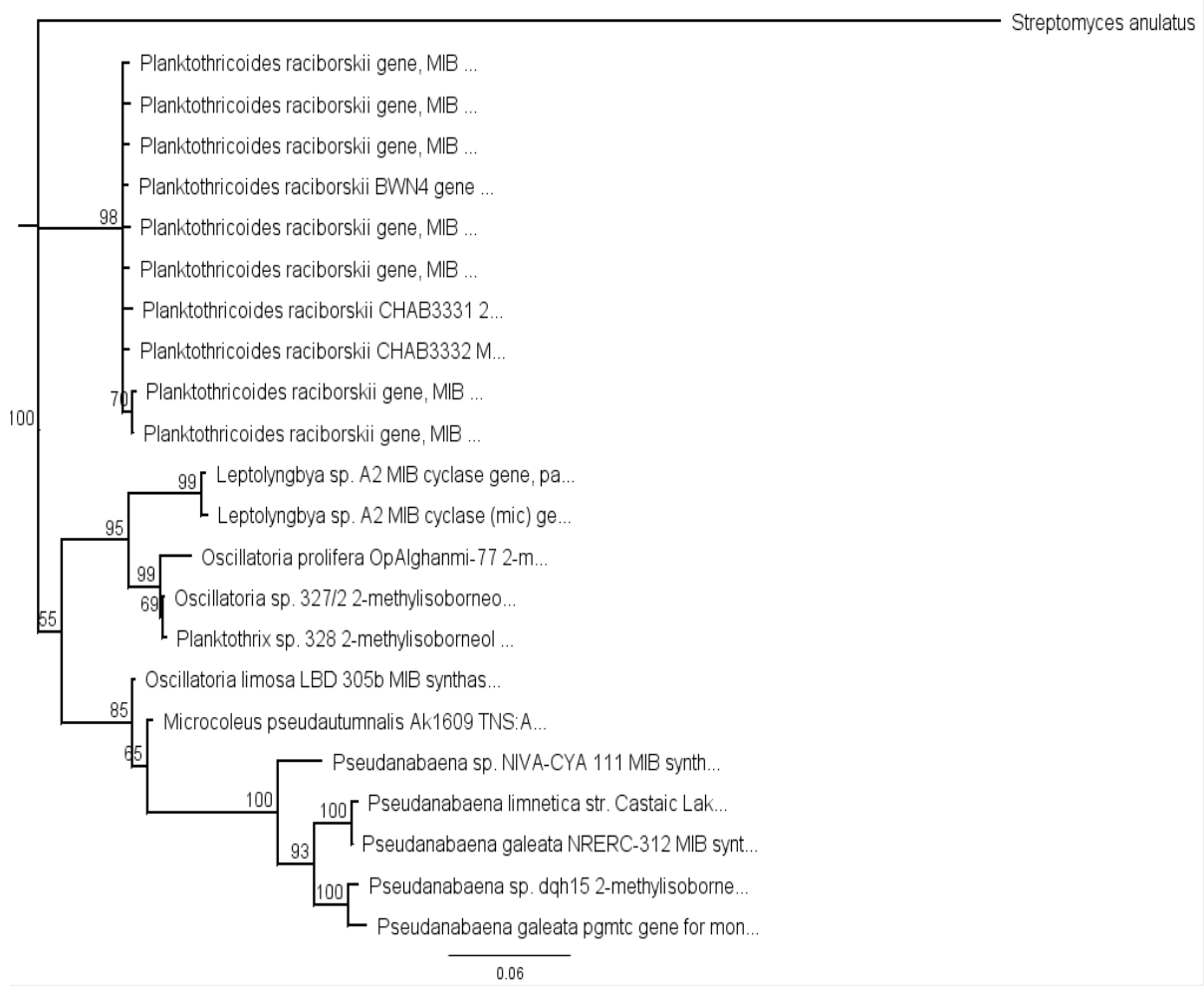


Figure 5.1: Neighbour-joining (NJ) phylogenetic tree of 2-MIB associated genes using Tamura-Nei genetic distance model, with *Streptomyces anulatus* as an outgroup. Accession numbers: *Planktothrix* sp. PD12 (MT515744.1) *Pseudanabaena* sp. PD8 (MT3602661.1) *Pseudanabaena yagii* NIES-4238 (LC507459.1) *Pseudanabaena* sp. 1 (LC507462.1) *Pseudanabaena yagii* NIES-4237 (LC507458.1) *Pseudanabaena* sp. NIVA-CYA 1119 (HQ630887.1) *Pseudanabaena* sp. 2 (HQ830028.1) *Pseudanabaena galeata* NIES-512 (AB826230.1) *Pseudanabaena galeata* (MN167115.1) *Pseudanabaena limnetica* (HQ630883.1) *Pseudanabaena cinerea* (LC507461.1) *Pseudanabaena cinerea* NIES-4062 (LC507460.1) *Microcoleus pseudautumnalis* (LC486303.1) *Oscillatoria limosa* (HQ630885.1) *Leptolyngbya* sp. (KP013063.1) *Oscillatoria prolifera* (MK759878.1) *Planktothrix* sp. 328 (KJ658378.1) *Oscillatoria prolifera* (MK759878.1) *Planktothricoides raciborskii* 1 (LC512931.1) *Planktothricoides raciborskii* 2 (LC157992.1) *Planktothricoides raciborskii* 3 (LC157991.1) *Planktothricoides raciborskii* 4 (LC157990.1) *Planktothricoides raciborskii* 5 (LC157989.1) *Planktothricoides raciborskii* 6 (LC157988.1) *Planktothricoides raciborskii* 7 (LC157987.1) *Planktothricoides raciborskii* 8 (LC157986.1) *Planktothricoides raciborskii* 9 (HQ830029.1). Genbank Accession numbers are provided with associated version identifiers.





Figure 5.2: Sequence view of the newly developed primer MIBAHR (24 bp).

### 5.2.7 – qPCR reactions for 16S rRNA, *geoA* and *mic* genes

The gene copy numbers for all targeting genes (16S rRNA, *geoA* and *mic*) were calculated using New England Biolabs (2020) calculator using the DNA PCR fragment lengths (1490 bp, 905 bp and 307 bp respectively) and average Qubit readings (Table 5.8). These were then used to generate dilutions required for the standards to create standard curves for the qPCR reactions. The DNA copy number obtained for the 16S rRNA gene was  $6.416 \times 10^{10}$  copies  $\mu\text{L}^{-1}$ , the DNA copy number obtained for the *geoA* gene was  $9.38 \times 10^{10}$  copies  $\mu\text{L}^{-1}$  and the DNA copy number for *mic* was  $9.65 \times 10^{10}$  copies  $\mu\text{L}^{-1}$ . All copy numbers were multiplied by 2.5 to account for the DNA sample volume added in the qPCR reaction (2.5  $\mu\text{L}$ ). The dynamic linear range for 16S rRNA was  $1.60 \times 10^1$  to  $1.60 \times 10^7$  copies  $\mu\text{L}^{-1}$ , the dynamic linear range for *geoA* was  $2.34 \times 10^1$  to  $2.34 \times 10^7$  copies  $\mu\text{L}^{-1}$  and the dynamic linear range for *mic* was  $2.41 \times 10^1$  to  $2.41 \times 10^7$  copies  $\mu\text{L}^{-1}$ . Both genes' sample copy numbers were then extrapolated from the standard curves generated from the prepared standards.

Table 5.8: Triplicate Qubit readings for both 16S rRNA, *geoA* and *mic* concentrations and an average in ng  $\mu\text{L}^{-1}$ .

| Replicate | Qubit readings                 |                                   |                                  |
|-----------|--------------------------------|-----------------------------------|----------------------------------|
|           | 16S rRNA ng $\mu\text{L}^{-1}$ | <i>geoA</i> ng $\mu\text{L}^{-1}$ | <i>mic</i> ng $\mu\text{L}^{-1}$ |
| 1         | 102                            | 92.8                              | 31.6                             |
| 2         | 96.0                           | 81.6                              | 31.0                             |
| 3         | 96.4                           | 86.8                              | 28.6                             |
| Average   | 98.1                           | 87.1                              | 30.4                             |

## 5.2.8 – qPCR set-up on the Eppendorf epMotion robot

Before the qPCR reaction set-up, the Eppendorf epMotion® 5075 (Eppendorf AG, Hamburg, Germany) robot was programmed to carry out 1:100 and 1:10 dilutions of the samples to be used for the qPCR DNA templates for all genes, in 96-well plates. Dilutions were used to dilute any PCR inhibitors contained within the extracted DNA.

Stock reagents for the master mix and gene standards were set up in an additional 96-well plate that the robot used to dispense into the according wells in a 384-well plate. The diluted samples were then aliquoted out into triplicate replicates into the same 384-well plate (Figure 5.3) (1:10 dilutions for the *geoA* gene and *mic* gene, and 1:100 dilutions for the 16S rRNA gene).

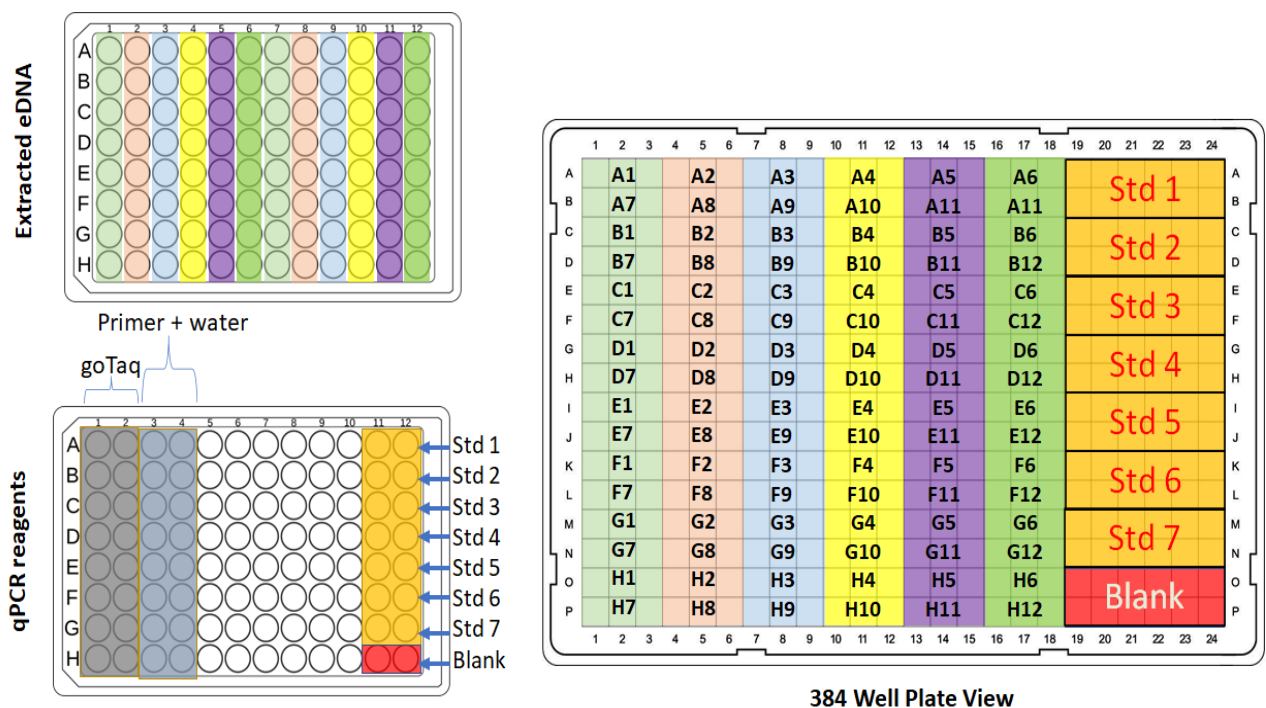


Figure 5.3: An overview of plate setups for the Eppendorf epMotion® 5075 robot. One 96-well plate setup for diluted samples (1:100 16S rRNA, 1:10 *geoA* and *mic*) and another 96-well plate for the reagent stock plates and standards. A general map for the qPCR reaction in a 384-well plate.

## 5.2.9 – qPCR reactions

The primers used for the qPCR are displayed in Table 5.9, and diagrams illustrating the binding sites of both *geoA* and *mic* primers are shown in Figures 5.4 and 5.5. GoTaq® qPCR Master Mix (Promega A6001; BSA; Promega, Southampton, UK) was used to carry out these qPCR reactions. This master mix includes a novel fluorescent DNA-binding dye like SYBR® Green I and is optimised for fast and reproducible quantitative PCR assays. The composition and concentrations for the qPCR reactions are displayed in Table 5.10. qPCR conditions (Table 5.11, 5.12 and 5.13) have been adapted from the manufacturer's instruction manual and current protocols for the qPCR master mix and the bacterial 16S rRNA primers (Kille et al., 2019; Webster et al., 2015). All qPCR reactions were carried on a QuantStudio™ 7 Flex Real-Time PCR System, 384-well (ThermoFisher Scientific, Massachusetts, United States).

Table 5.9: Primer pairs for 16S rRNA, *geoA* and *mic* used for the qPCR.

| Target gene      | Primer  | Sequences 5' → 3'        | Product length (bp) | Reference               |
|------------------|---------|--------------------------|---------------------|-------------------------|
| 16S rRNA         | 534F    | GCCAGCAGCCGCGGTAAT       | 400                 | Muyzer et al. (1993)    |
|                  | 907R    | CCGTCAATTCCTTTGAGTTT     |                     | Muyzer (1998)           |
| Geosmin synthase | geo799F | GCCGCTAACCTCACTAACGA     | 184                 | John et al. (2018)      |
|                  | geo982R | ATCGCATGTGCCACTCGTGAC    |                     | Suurnäkki et al. (2015) |
| 2-MIB cyclase    | MIBS02F | ACCTGTTACGCCACCTTCT      | 131                 | Chiu et al. (2016)      |
|                  | MIBAHR  | GTCATGGAGGTGTAGAAGCTGTCG |                     | This study              |

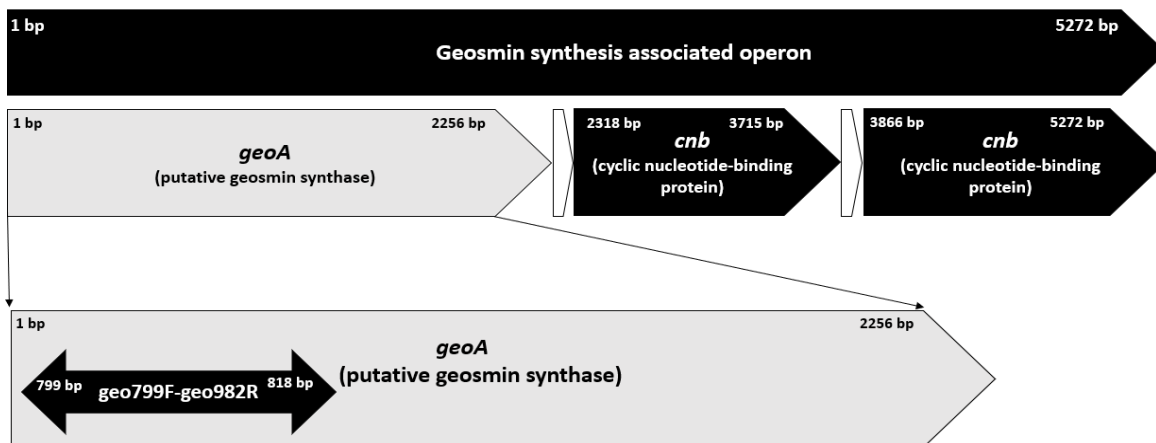


Figure 5.4: Primer positions and alignment of the amplicon on *Anabaena ucrainica* CHAB2155 geosmin synthesis associated operon, complete sequence (HQ404996).

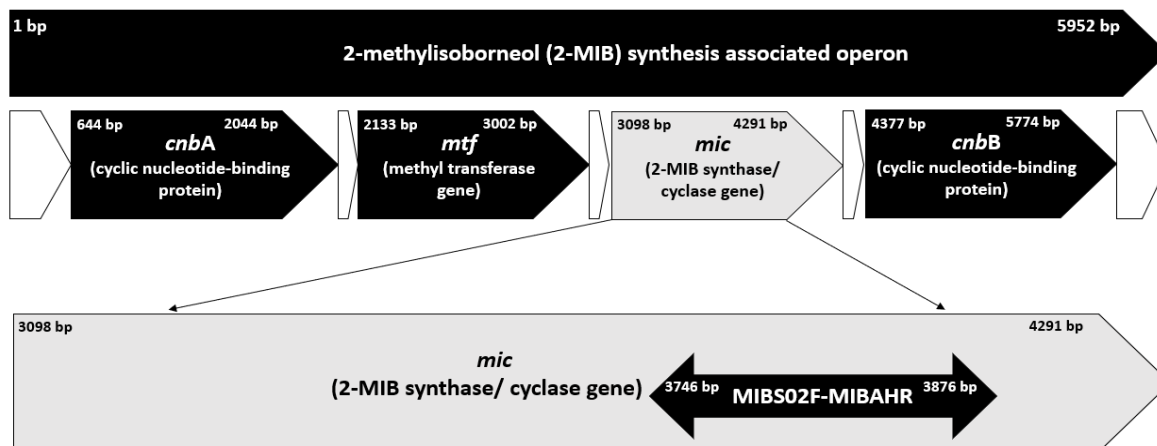


Figure 5.5: Primer positions and alignment of the amplicon on *Pseudoanabaena* sp. dqp15 2-methylisoborneol (2-MIB) synthesis associated operon, complete sequence (HQ830028.1).

Table 5.10: qPCR master mix volumes for a 10  $\mu\text{L}$  reaction volume for primer pairs 534F, 907R and *geo*799F *geo*982R and MIBS02F and MIBAHR respectively.

|  | 16S rRNA single reaction / $\mu\text{L}$ | <i>geoA</i> single reaction / $\mu\text{L}$ | <i>mic</i> single reaction / $\mu\text{L}$ |
|--|--|---|--|
| Nuclease free H <sub>2</sub> O               | 2.1                                      | 1.7   | 1.7  |
| GoTaq qPCR Master Mix (2x)                   | 5  | 5   | 5  |
| Forward primer (10 pmol $\mu\text{L}^{-1}$ ) | 0.2                                      | 0.4   | 0.4  |
| Reverse primer (10 pmol $\mu\text{L}^{-1}$ ) | 0.2                                      | 0.4   | 0.4  |
|  |  |   |  |
| Volume of template DNA                       | 2.5                                      | 2.5   | 2.5  |
| Final volume of qPCR reaction                | 10                                       | 10  | 10   |
| Final primer concentration                   | 0.20 pmol $\mu\text{L}^{-1}$             | 0.40 pmol $\mu\text{L}^{-1}$                | 0.40 pmol $\mu\text{L}^{-1}$               |

Table 5.11: The qPCR conditions used for 16S rRNA primer pair 534F and 907R.

|           | qPCR Stages            | Time (minutes) | Temperature ( $^{\circ}\text{C}$ ) |
|-----------|------------------------|----------------|------------------------------------|
|           | Hot-start activation   | 0:20           | 95                                 |
| 40 cycles | Denaturation           | 0:01           | 95                                 |
|           | Annealing              | 0:01           | 55                                 |
|           | Extension              | 0:30           | 72                                 |
|           | Dissociation (melting) |                | 95                                 |

Table 5.12: The qPCR conditions used for *geoA* primer pair *geo*799F and *geo*982R.

|           | qPCR Stages            | Time (minutes) | Temperature ( $^{\circ}\text{C}$ ) |
|-----------|------------------------|----------------|------------------------------------|
|           | Hot-start activation   | 0:20           | 95                                 |
| 40 cycles | Denaturation           | 0:01           | 95                                 |
|           | Annealing/ Extension   | 0:20           | 60                                 |
|           | Dissociation (melting) |                | 95                                 |

Table 5.13: The qPCR conditions used for *mic* primer pair MIBS02F and MIBAHR.

|           | qPCR Stages            | Time (minutes) | Temperature ( $^{\circ}\text{C}$ ) |
|-----------|------------------------|----------------|------------------------------------|
|           | Hot-start activation   | 0:20           | 95                                 |
| 40 cycles | Denaturation           | 0:01           | 95                                 |
|           | Annealing/ extension   | 0:20           | 65                                 |
|           | Dissociation (melting) |                | 95                                 |

## 5.2.10 – Cross validation of 16S rRNA and *geoA* standards

At the start of this project, 310 samples had been previously processed using the methods stated above (5.2.7) for 16S rRNA and *geoA* qPCR (data kindly provided by Ms I. Elfferich). A different standard curve was used with a dynamic linear range of  $5.12 \times 10^8 - 5.12 \times 10^2$  copies  $\mu\text{L}^{-1}$  for 16S rRNA and  $4.85 \times 10^9$  to  $4.85 \times 10^3$  copies  $\mu\text{L}^{-1}$  for *geoA* for this previous work. To account for differences in results due to using standard curves being prepared at different times by different operators for both 16S rRNA and *geoA* and to allow datasets to be comparable, a cross-calibration of the standards was performed. The standards used for the first set of samples (310) underwent a qPCR using materials and methods described in Section 5.2.7 in parallel with the standards detailed in Section 5.2.5 to establish a cross-calibration coefficient. Both sets of standards for 16S rRNA and *geoA* were performed over six replicates.

All replicates for each standard dilution were averaged, and then the gradient (*m*) and intercept (*c*) of the standard curve for standards detailed in Section 5.2.7 were calculated using the least squares method. This was done by taking the average *Ct* of standards and the known log copy number  $\mu\text{L}$  of said standards. The average *Ct* of the first set of standards used in this project was supplemented into the following equation using the ‘*c*’ and ‘*m*’ generated from the standard curve detailed in Section 5.2.5:

$$\text{Average } Ct - c/m$$

From this equation, an antilog of the result was taken, and a ratio was determined by dividing this by the known concentration of the first set of standards used. An average of all standard ratios was then calculated, and a cross-calibration coefficient was determined. For 16S rRNA, the coefficient determined was 4.74, and for *geoA*, the coefficient was 4.05. For the first 310 samples, established results for determined copies  $\text{mL}^{-1}$  were multiplied by the derived cross-calibration coefficient. For 16S rRNA, a histogram illustrating the frequency of samples copies  $\text{mL}^{-1}$  before and after the calibration coefficient has been added is displayed in Figure 5.6.

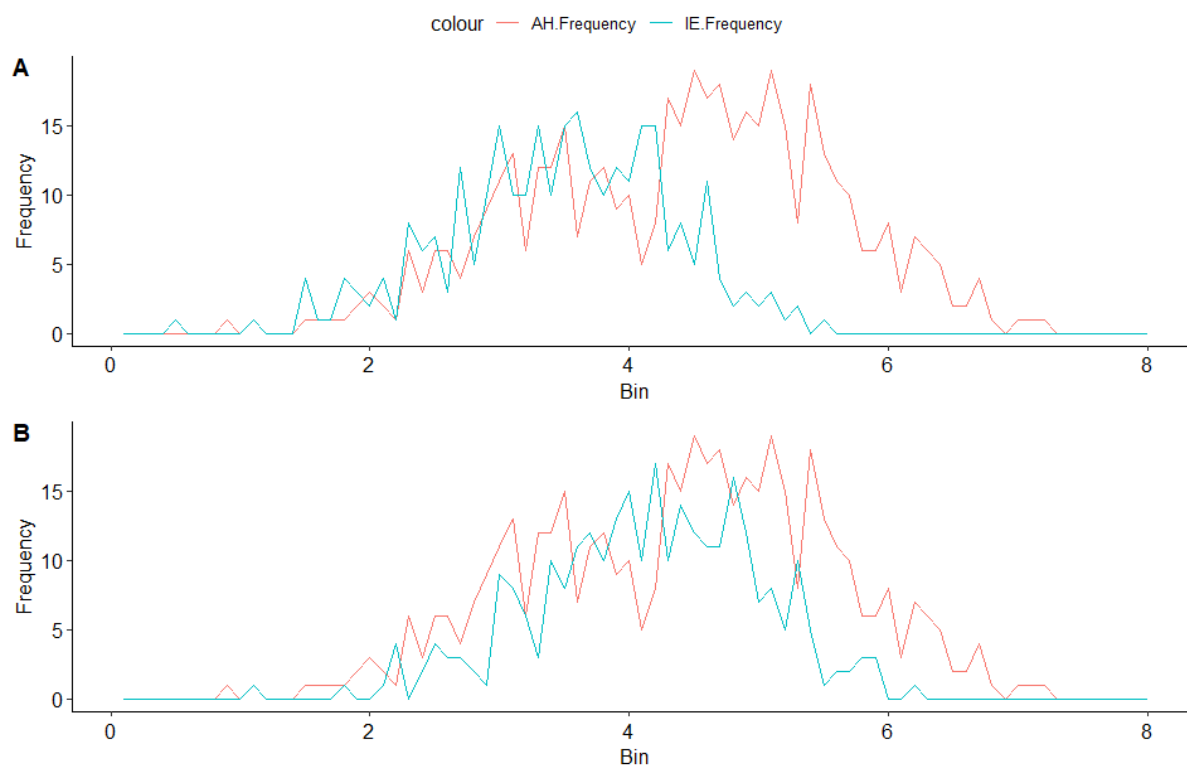


Figure 5.6: Histograms of (A) the frequency of 16S rRNA copy numbers from this study (red – AH.Frequency) and the frequency of copy numbers of 16S rRNA processed previously (blue – IE.Frequency) (B) frequency of 16S rRNA copy numbers from this study (red – AH.Frequency) and the frequency of copy numbers of 16S rRNA processed previously (blue – IE.Frequency) after the calibration coefficient had been added.

## 5.3 – Results

### 5.3.1 – Temporal changes in gene copy numbers and T&O concentrations by year and month

Strong positive linear correlations were observed between geosmin and  $\log_{10}$  *geoA*:16S copies  $\text{mL}^{-1}$  during the months where geosmin concentrations were above  $100 \text{ ng L}^{-1}$  (July 2019,  $p < 0.0001$ ; September 2019,  $p < 0.0001$ ; July 2020,  $p < 0.0001$ ; August 2020,  $p < 0.0001$ ) (Figure 5.7). The significance of relationships decreased when geosmin fell below  $100 \text{ ng L}^{-1}$  (August 2019,  $p < 0.05$ ; October 2019,  $p < 0.05$ ; December 2019,  $p < 0.01$ ). No significant relationship was determined for November 2019 ( $p = 0.482$ ).

No consistent relationships could be deduced from  $\log_{10}$  *mic*:16S copy numbers  $\text{mL}^{-1}$  and  $\log_{10}$  2-MIB concentrations  $\text{ng L}^{-1}$  (Figure 5.8). The maximum 2-MIB  $\text{ng L}^{-1}$  concentration ( $58 \text{ ng L}^{-1}$ ) was seen in September 2019, which coincided with a high *mic*:16S copy numbers  $\text{mL}^{-1}$  value ( $5.12 \text{ mic}:16\text{S copy number mL}^{-1}$ ). A slight positive relationship was observed in August 2020, although this was not significant ( $p = 0.688$ ), and a significantly negative relationship was witnessed in October 2019 ( $p < 0.01$ ).

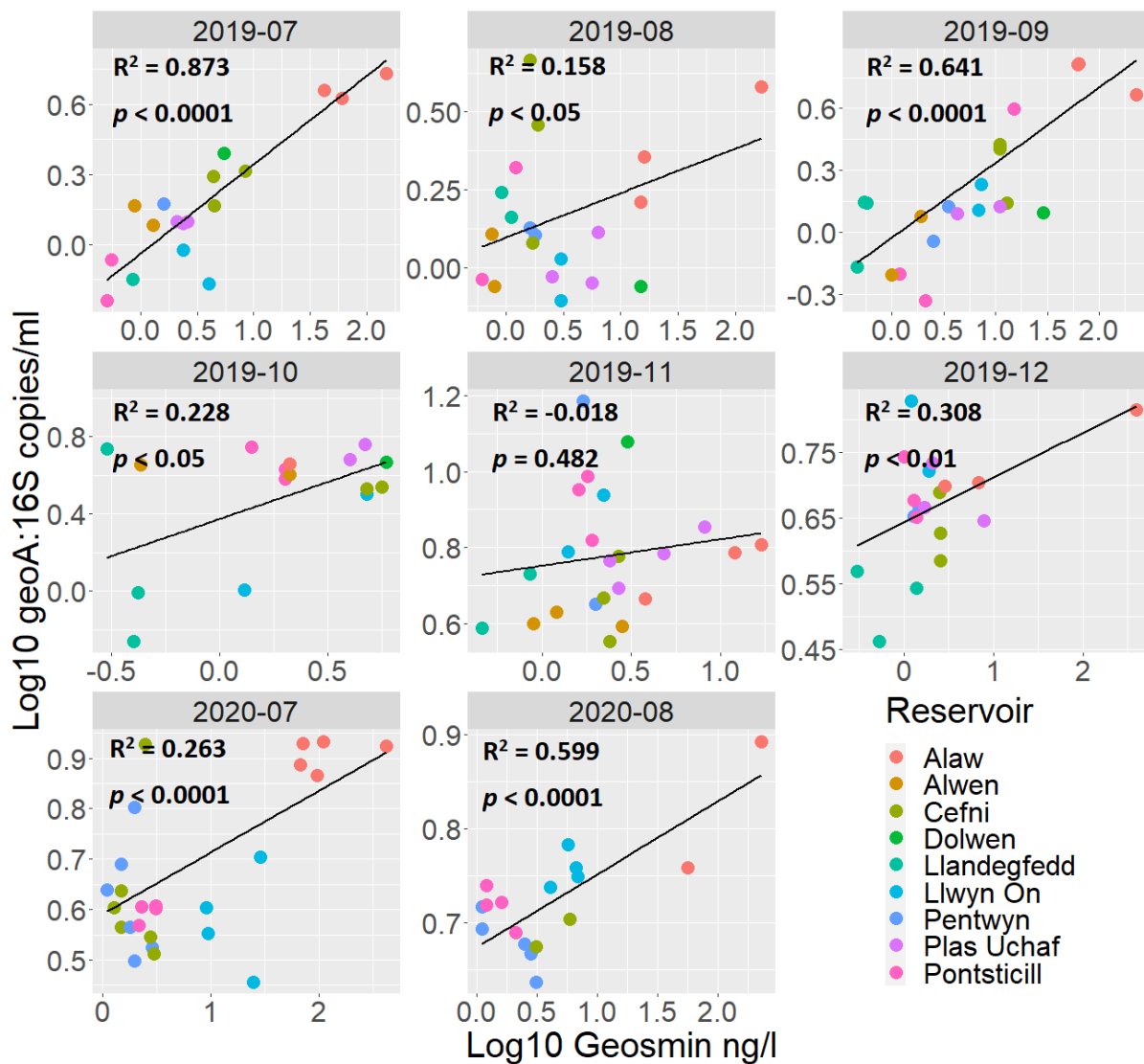


Figure 5.7: Scatterplots of  $\log_{10}$  concentrations of geosmin ( $\text{ng L}^{-1}$ ) and *geoA*:16S (copy numbers  $\text{mL}^{-1}$ ) from all reservoirs. Individual points are coloured corresponding to the reservoir and facet wrapped according to sampling month and year. Each scatterplot includes a linear regression line of best fit with the  $R^2$  value and associated significance assigned by  $p$  values.



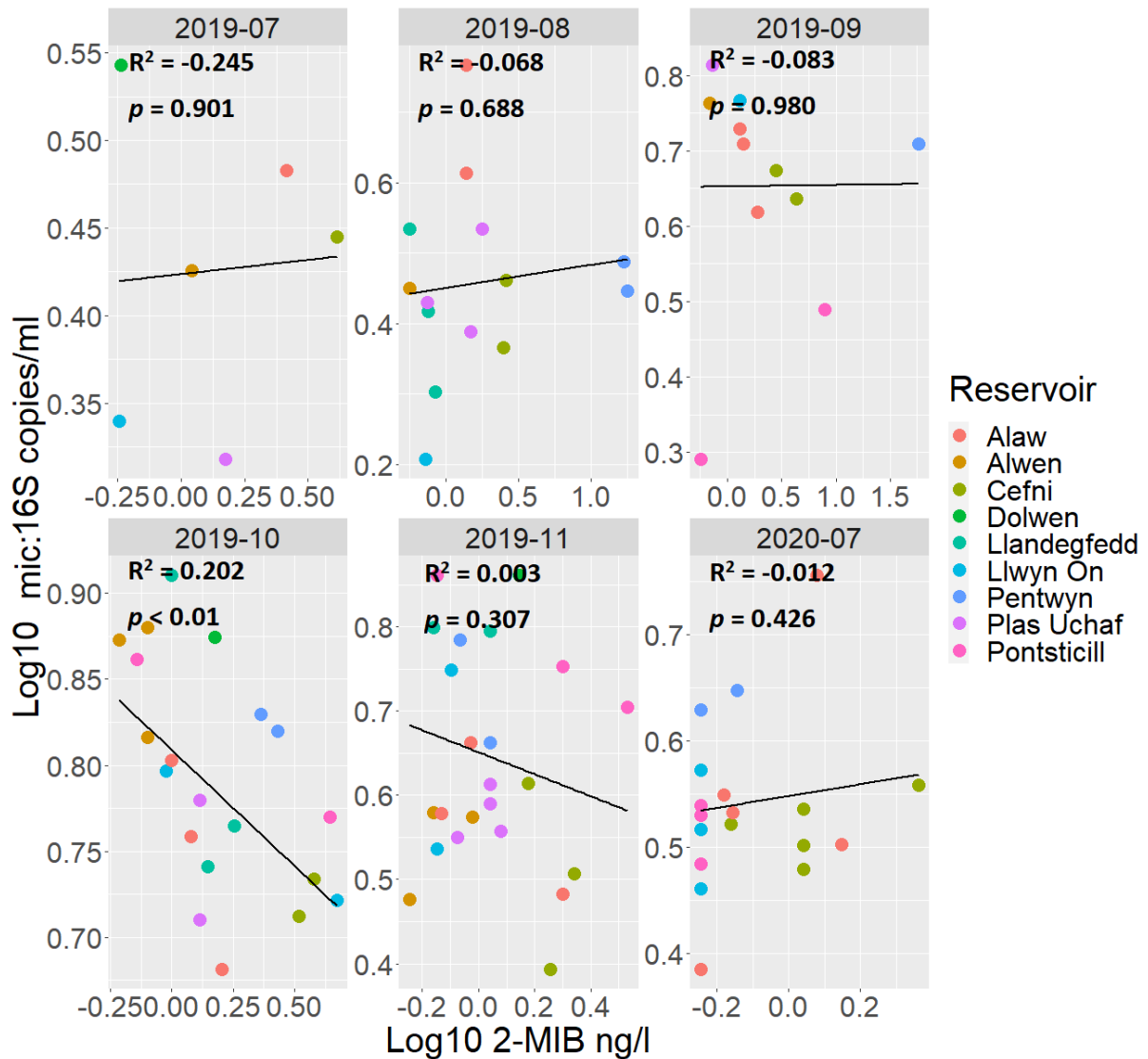


Figure 5.8: Scatterplots of  $\text{log}_{10}$  concentrations of 2-MIB (ng L<sup>-1</sup>) and *mic:16S* (copy numbers mL<sup>-1</sup>) from all reservoirs. Individual points are coloured corresponding to the reservoir and facet wrapped according to sampling month and year. Each scatterplot includes a linear regression line of best fit with the  $R^2$  value and associated significance assigned by  $p$  values.

### 5.3.2 – Temporal changes in gene copy numbers and T&O concentrations by season and year

Highly significant positive relationships were observed for geosmin and *geoA:16S* gene copy numbers (Figure 5.9) for summer 2019 ( $p < 0.0001$ ) and summer 2020 ( $p < 0.0001$ ). Significant relationships were apparent for winter 2019 ( $p < 0.01$ ) and Autumn 2019 ( $p < 0.05$ ). When geosmin concentrations fell below 3.16

ng L<sup>-1</sup>, a slightly negative non-significant correlation was detected, as seen in winter 2020 ( $p = 0.305$ ).

For 2-MIB and *mic*:16S no significant relationships could be determined. In both summer 2019 and 2020, a weak positive relationship can be seen in Appendix C.2 ( $p = 0.611$  and  $0.151$ , respectively). In autumn 2019 and winter 2020, non-significant negative relationships are displayed ( $p = 0.437$  and  $0.127$ ). Winter 2019 was removed from the log<sub>10</sub> 2-MIB and log<sub>10</sub> *mic*:16S copy number plot (Figure 5.10) due to no variation seen in log<sub>10</sub> 2-MIB concentrations.

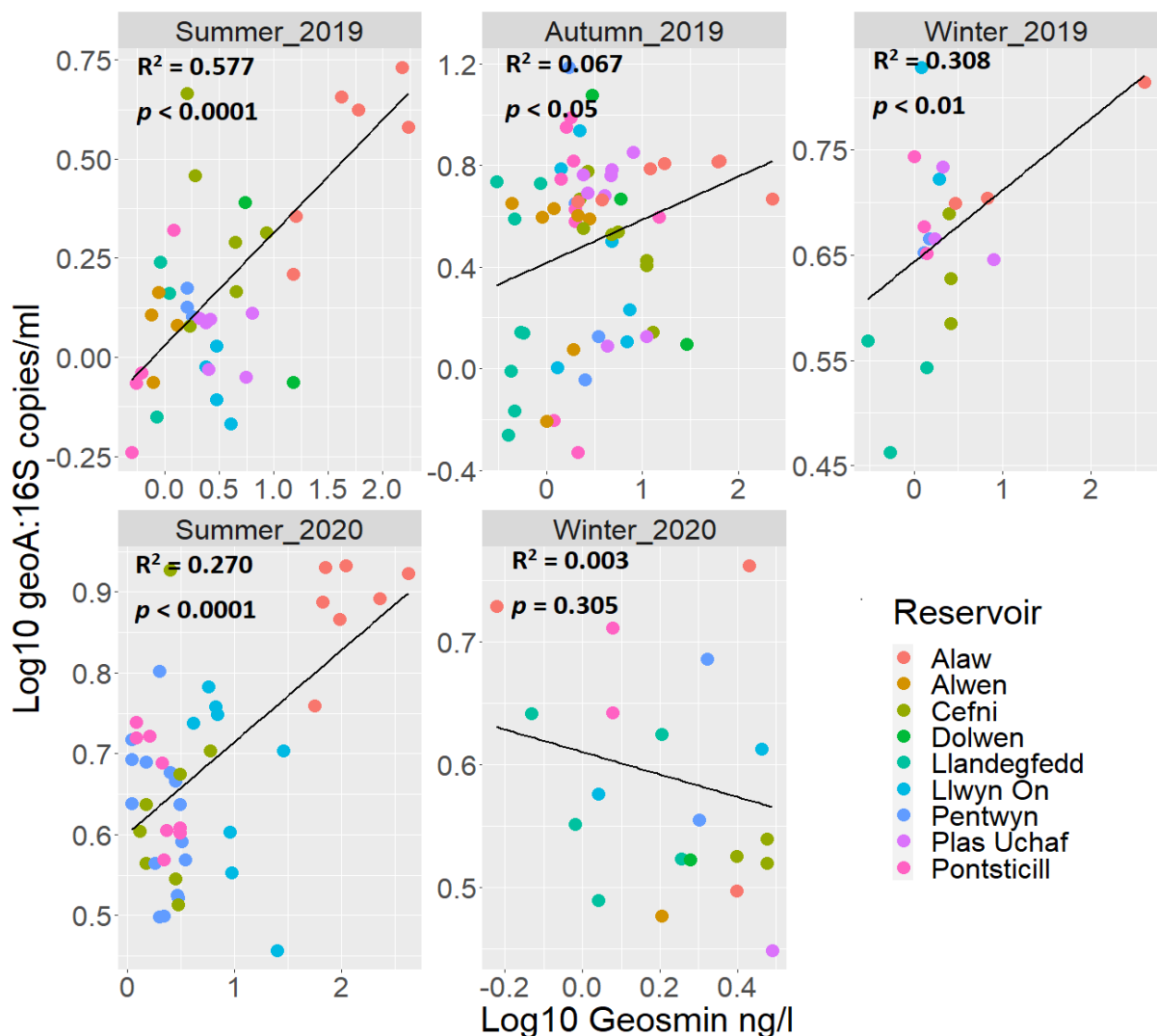


Figure 5.9: Scatterplots of log<sub>10</sub> concentrations of geosmin (ng L<sup>-1</sup>) and *geoA*:16S (copy numbers mL<sup>-1</sup>) from all reservoirs. Individual points are coloured corresponding to the reservoir and facet wrapped according to sampling season and year. Each scatterplot includes a linear regression line of best fit with the R<sup>2</sup> result and associated significance assigned by  $p$  values.

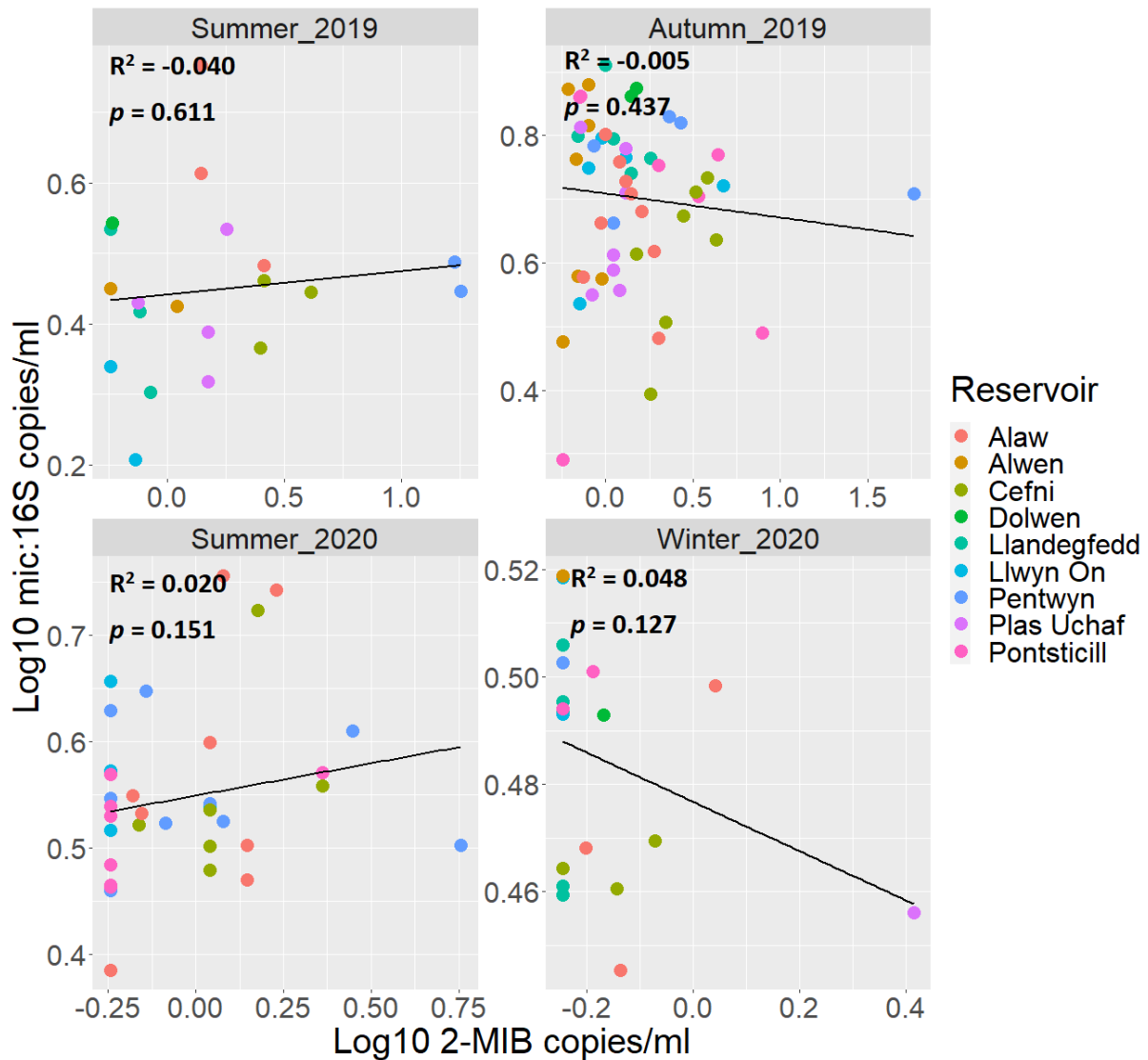


Figure 5.10: Scatterplots of  $\log_{10}$  concentrations of 2-MIB ( $\text{ng L}^{-1}$ ) and *mic:16S* (copy numbers  $\text{mL}^{-1}$ ) from all reservoirs. Individual points are coloured corresponding to the reservoir and facet wrapped according to sampling season and year. Each scatterplot includes a linear regression line of best fit with the  $R^2$  result and associated significance assigned by  $p$  values.

### 5.3.3 – Modelling a T&O event

For *geoA:16S* copy numbers  $\text{mL}^{-1}$ , two GAMs were employed. The first GAM used data only from the Alaw reservoir (deviance = 97.7%,  $R^2 = 0.87$ ,  $n = 41$ : Appendix 5.5). The second GAM incorporated data from all other reservoirs excluding Alaw (deviance = 67.20%,  $R^2 = 0.33$ ,  $n = 183$ : Appendix 5.6). The *mic:16S*

copy numbers mL<sup>-1</sup> GAM incorporated data from all reservoirs (deviance = 58.70%, R<sup>2</sup> = 0.47, n = 172: Appendix 5.7).

### 5.3.3.1 - *geoA*:16S abundance in a reservoir with major geosmin events (Alaw)

All linear environmental covariates were significantly associated with *geoA*:16S copy numbers mL<sup>-1</sup> for the Alaw reservoir; except autumn 2019 and winter 2020 which showed no significant difference in *geoA*:16S abundance compared with summer 2019 (Table 5.14). *geoA*:16S during winter 2019 was significantly lower and summer 2020 was significantly higher than summer 2019 ( $p < 0.001$  for both). The mean temperature had a significant negative impact on *geoA*:16S copy numbers mL<sup>-1</sup> ( $p < 0.001$ ). Sulphate and dissolved iron were positively associated with *geoA*:16S copy numbers mL<sup>-1</sup> ( $p < 0.0001$ ), whereas dissolved reactive silicate was negatively associated with *geoA*:16S copy numbers mL<sup>-1</sup> ( $p < 0.0001$ ). When geosmin levels were between 50 – 200 ng L<sup>-1</sup>, *geoA*:16S copy numbers mL<sup>-1</sup> rapidly increased from a partial effect size of around -0.5 to 1.5 ( $p < 0.001$ ) (Figure 5.11 top). Increased levels of geosmin (> 400 ng L<sup>-1</sup>) were negatively associated with *geoA*:16S copy numbers mL<sup>-1</sup>. NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> had a significant non-linear relationship with *geoA*:16S copy numbers mL<sup>-1</sup> ( $p < 0.001$ ): With the exception of two datapoints (<0.025), heightened NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> ratios (>0.1) showed a strong positive relationship with the abundance of *geoA*:16S copy numbers mL<sup>-1</sup> (Figure 5.11 middle). NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> ratios <0.1 reduced the average response values of *geoA*:16S copy numbers mL<sup>-1</sup>. TIN:TP also had a significantly non-linear relationship with *geoA*:16S copy numbers mL<sup>-1</sup> ( $p < 0.0001$ ). Low levels of TIN:TP (0:1 – 24:1) had the greatest partial effect size on *geoA*:16S copy numbers mL<sup>-1</sup> (Figure 5.11 bottom). TIN:TP levels between 24:1 – 36:1 appeared to reduce the response below the average levels of *geoA*:16S copy numbers mL<sup>-1</sup>. This inhibitory effect was reduced when TIN:TP levels were above 36:1.

Table 5.14: GAM model results for Alaw reservoir with *geoA*:16S copy numbers mL<sup>-1</sup> as the response variable. Using summer 2019 as the reference level for seasonal comparison.

| <b>Parametric coefficients</b>                             | Estimate | standard error | <i>p</i> -value |
|--|----------|----------------|-----------------|
| Autumn 2019  | 0.000    | 0.499          | 0.999           |
| Winter 2019  | -3.733   | 0.960          | < 0.001 ***     |
| Summer 2020  | 2.073    | 0.461          | < 0.001 ***     |
| Winter 2020  | 0.732    | 0.687          | 0.297           |
| Mean Temperature   | -0.405   | 0.093          | < 0.001 ***     |
| Sulphate   | 1.338    | 0.162          | < 0.001 ***     |
| Dissolved Reactive Silicate                                | -1.465   | 3.128          | < 0.001 ***     |
| Dissolved Iron   | 19.325   | 3.128          | < 0.001 ***     |
| <b>Smooth terms</b>  | edf      | <i>F</i>       | <i>p</i> -value |
| Geosmin  | 2.009    | 39.040         | < 0.001 ***     |
| NH <sub>4</sub> <sup>+</sup> :NO <sub>3</sub> <sup>-</sup> | 3.000    | 11.240         | < 0.001 ***     |
| TIN:TP   | 2.315    | 22.140         | < 0.001 ***     |

Note: Variables with significant influences are indicated by: .  $p < 0.1$ ,

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

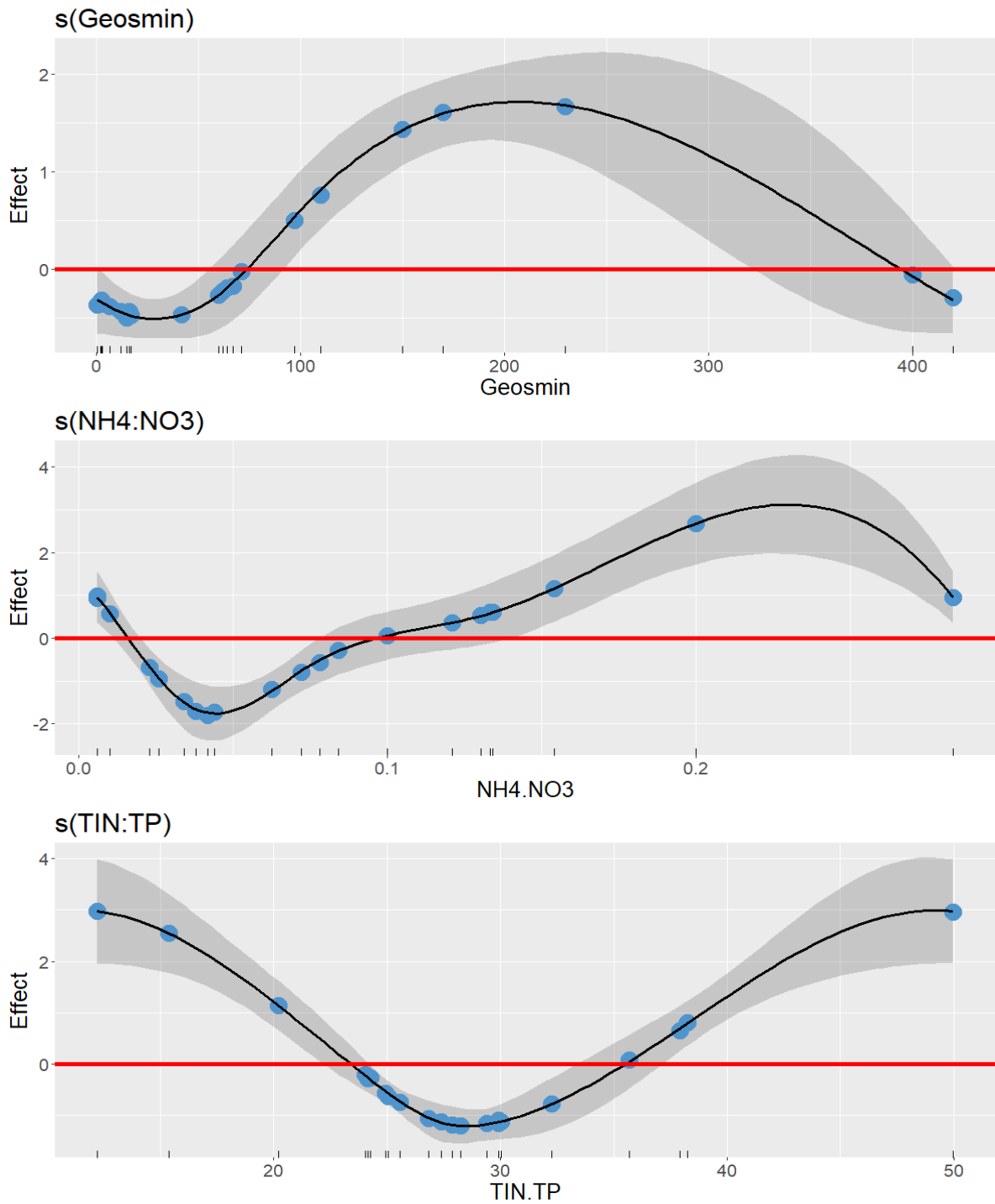


Figure 5.11: Smooth function plots for predictor variables in the Alaw *geoA*:16S copy numbers  $\text{mL}^{-1}$  GAM. Estimated smooth functions (solid lines) with 95% confidence intervals (grey shaded area) are shown for each smoothed predictor: (top) Geosmin concentrations  $\text{ng L}^{-1}$  (middle)  $\text{NO}_3\text{:NH}_4^+$  ratio and (bottom) TIN:TP ratio. The y-axis denotes the partial effect size, the comb on the x-axis shows where the value of predictor data points lie and the points are the residuals. The horizontal red line at  $y = 0$  intercept indicates the overall mean of the response (*geoA* abundance).

### 5.3.3.2 - *geoA*:16S abundance in reservoirs with minor geosmin concentrations

For all reservoirs except Alaw, *geoA*:16S copy numbers mL<sup>-1</sup> GAM, summer 2020 and summer 2019 were found to be positively significantly different to winter 2019 ( $p < 0.001$  and  $p < 0.05$ , respectively) (Table 5.15). In addition, *geoA*:16S in all reservoirs were found to be significantly different to *geoA*:16S in Llandegfedd, apart from Pontsticill. Both mean temperature and dissolved reactive silicate were significantly negatively associated with *geoA* ( $p < 0.001$ , for both). Geosmin concentrations were significantly non-linearly associated *geoA*:16S abundance ( $p < 0.001$ ), having the greatest positive influence on concentrations between 4 – 35 ngL<sup>-1</sup> (Figure 5.12 top). NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> was also significantly non-linearly associated with *geoA*:16S copy number mL<sup>-1</sup> ( $p < 0.001$ ), with the greatest partial effect when NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> ratios were high (0.02 – 0.17) (Figure 5.12 middle); except for one spurious data point at a ratio of 0.20. TIN:TP was also significantly non-linearly associated with *geoA*:16S copy numbers mL<sup>-1</sup> ( $p < 0.001$ ), with the greatest positive influence at a ratio <25:1 (Figure 5.12 bottom). A TIN:TP ratio between 25:1 – 310:1 had a negative effect on the abundance of *geoA*:16S, although the abundance of *geoA*:16S increased at a ratio >310:1.

Table 5.15: GAM model results for all reservoirs except Alaw, with *geoA*:16S copy numbers mL<sup>-1</sup> as the response variable. Using winter 2019 as the reference level for seasonal comparison and Llandegfedd as the reference level for reservoir comparison.

| <b>Parametric coefficients</b>                             | Estimate | standard error | <i>p</i> -value |
|--|----------|----------------|-----------------|
| Summer 2019  | 1.554    | 0.738          | 0.037 *         |
| Autumn 2019  | 0.45     | 0.69           | 0.514           |
| Summer 2020  | 2.967    | 0.772          | < 0.001 ***     |
| Winter 2020  | -0.4     | 0.716          | 0.578           |
| Mean Temperature   | -0.569   | 0.051          | < 0.001 ***     |
| Sulphate   | 0.129    | 0.08           | 0.111           |
| Dissolved Reactive Silicate                                | -0.249   | 0.055          | < 0.001 ***     |
| Dissolved Iron   | 0.023    | 0.158          | 0.884           |
| Alwen  | -1.494   | 0.388          | < 0.001 ***     |
| Cefni  | -1.238   | 0.697          | 0.078 .         |
| Dolwen   | -1.556   | 0.672          | 0.022 *         |
| Llwyn On   | -1.105   | 0.39           | 0.005 **        |
| Pentwyn  | 0.562    | 0.325          | 0.086 .         |
| Plas Uchaf   | -1.214   | 0.324          | < 0.001 ***     |
| Pontsticill  | -0.01    | 0.284          | 0.973           |
| <b>Smooth terms</b>  | edf      | <i>F</i>       | <i>p</i> -value |
| Geosmin  | 2.692    | 14.72          | < 0.001 ***     |
| NH <sub>4</sub> <sup>+</sup> :NO <sub>3</sub> <sup>-</sup> | 2.956    | 16.95          | < 0.001 ***     |
| TIN:TP   | 1.339    | 4.7            | < 0.001 ***     |

Note: Variables with significant influences are indicated by: .  $p < 0.1$ ,

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .



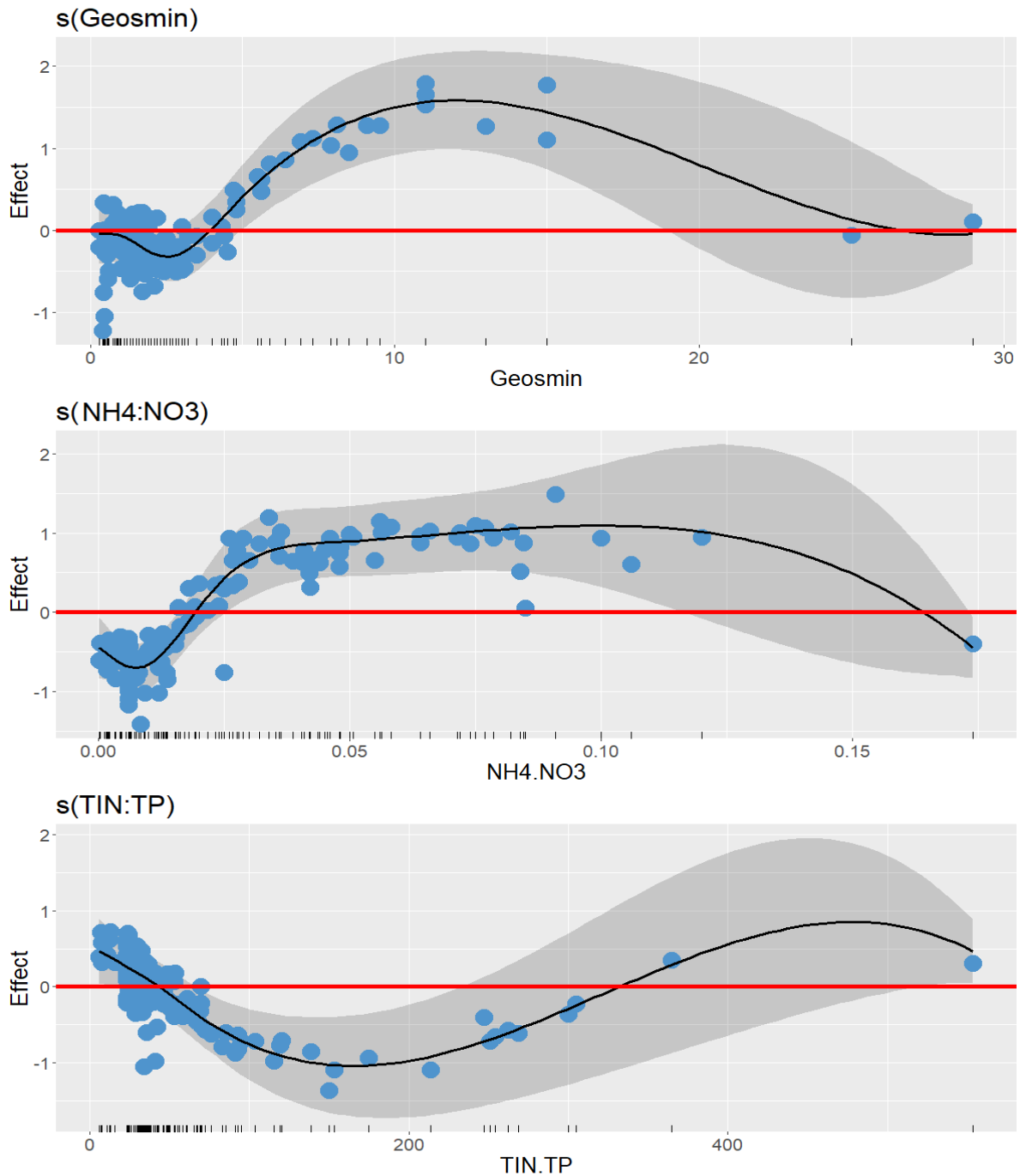


Figure 5.12: Smooth function plots for predictor variables in the control *geoA*:16S copy numbers mL<sup>-1</sup> GAM. Estimated smooth functions (solid lines) with 95% confidence intervals (grey shaded area) are shown for each smoothed predictor: (top) Geosmin concentrations ng L<sup>-1</sup> (middle) NO<sub>3</sub><sup>-</sup>:NH<sub>4</sub><sup>+</sup> ratio and (bottom) TIN:TP ratio. The y-axis denotes the partial effect size, the comb on the x-axis shows where the value of predictor data points lie and the points are the residuals. The horizontal red line at y = 0 intercept indicates the overall mean of the response (*geoA* abundance).

### 5.3.3.3 – *mic:16S* abundance in all reservoirs

Autumn and summer 2019 were significantly positively and negatively different from winter 2019 ( $p < 0.001$ , for both) for *mic:16S* copy numbers mL<sup>-1</sup> (Table 5.16). Dissolved reactive silicate also had a slightly significant negative relationship with *mic:16S* copy numbers mL<sup>-1</sup> ( $p < 0.1$ ). 2-MIB concentrations were not found to be significantly non-linearly associated with *mic:16S* copy numbers mL<sup>-1</sup> up to concentrations of 8 ng L<sup>-1</sup> as depicted by the horizontal estimated smooth function in Figure 5.13 top. NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> had a significantly non-linear relationship with *mic:16S* copy numbers mL<sup>-1</sup> ( $p < 0.1$ ), showing a positive trend with low ratios of NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> between ~0.00 – 0.07 and at higher ratios (> 0.18) (Figure 5.13 middle). An inhibitory effect of NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> on *mic:16S* copy numbers mL<sup>-1</sup> was observed between ~0.07 – 0.18.

Table 5.16: GAM model results with *mic*:16S copy numbers mL<sup>-1</sup> as the response variable. Using winter 2019 as the reference level for seasonal comparison, and Llandegfedd as the reference level for reservoir comparison.

| <b>Parametric coefficients</b>                             | Estimate | standard error | <i>p</i> -value |
|--|----------|----------------|-----------------|
| Summer 2019  | -0.670   | 0.470          | < 0.001 ***     |
| Autumn 2019  | 1.871    | 0.281          | < 0.001 ***     |
| Summer 2020  | 0.174    | 0.380          | 0.647           |
| Winter 2020  | -0.319   | 0.206          | 0.125           |
| Mean Temperature   | -0.020   | 0.041          | 0.630           |
| Sulphate   | 0.048    | 0.043          | 0.268           |
| Dissolved Reactive Silicate                                | -0.099   | 0.056          | 0.077 .         |
| Dissolved Iron   | 0.261    | 0.456          | 0.567           |
| Alaw   | 0.019    | 0.322          | 0.953           |
| Alwen  | -0.016   | 0.649          | 0.980           |
| Cefni  | -0.122   | 0.415          | 0.769           |
| Dolwen   | 0.488    | 0.598          | 0.416           |
| Llwyn On   | -0.364   | 0.267          | 0.175           |
| Pentwyn  | 0.454    | 0.318          | 0.156           |
| Plas Uchaf   | -0.233   | 0.352          | 0.510           |
| Pontsticill  | -0.347   | 0.324          | 0.285           |
| <b>Smooth terms</b>  | edf      | <i>F</i>       | <i>p</i> -value |
| 2-MIB  | 0.000    | 0.000          | 0.341           |
| NH <sub>4</sub> <sup>+</sup> :NO <sub>3</sub> <sup>-</sup> | 1.020    | 1.023          | 0.052 .         |
| TIN:TP   | 0.000    | 0.000          | 0.901           |

Note: Variables with significant influences are indicated by: .  $p < 0.1$ ,

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

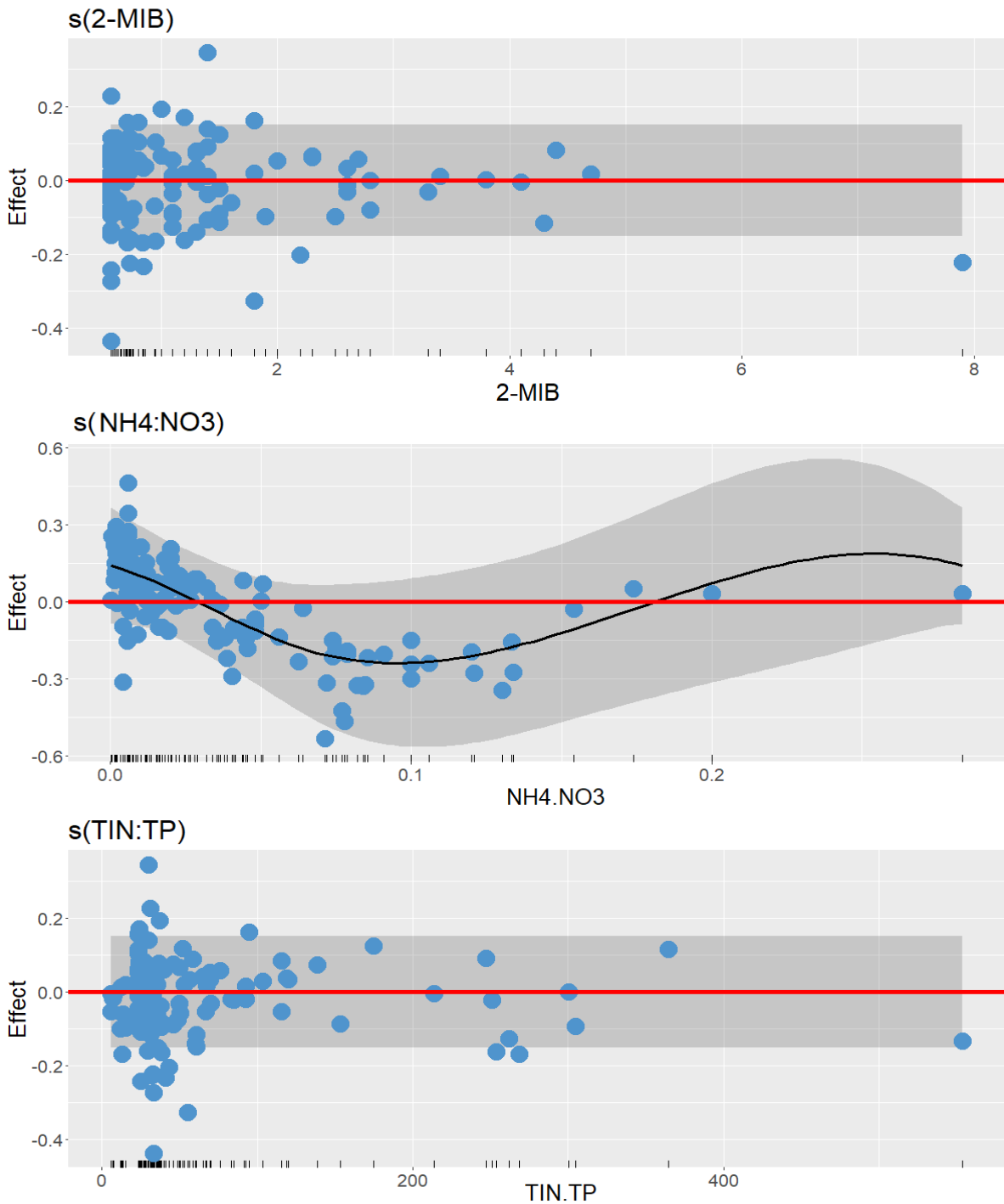


Figure 5.13: Smooth function plots for predictor variables in the *mic*:16S copy numbers  $\text{mL}^{-1}$  GAM. Estimated smooth functions (solid lines) with 95% confidence intervals (grey shaded area) are shown for each smoothed predictor: (top) 2-MIB concentrations  $\text{ng L}^{-1}$  (middle)  $\text{NO}_3:\text{NH}_4^+$  ratio (bottom) TIN:TP ratio and (C). The y-axis denotes the partial effect size, the comb on the x-axis shows where the value of predictor data points lie and the points are the residuals. The horizontal red line at  $y = 0$  intercept indicates the overall mean of the response (*mic* abundance).

## 5.4 – Discussion

This is the first study to our knowledge to demonstrate variations in *geoA* and *mic* abundance between different seasons. This is also the first study to report the associations between iron, sulphate, and dissolved reactive silicate with *geoA* and *mic* copy numbers in relation to T&O levels. Significant correlations were determined between *geoA* and geosmin concentrations by month and by season, whereas no relationships could be deduced from correlations observed between *mic* and 2-MIB concentrations by month or by season. *geoA* was deemed to be a suitable indicator of geosmin concentrations, especially when geosmin concentrations were elevated ( $>100 \text{ ng L}^{-1}$ ), as seen in the Alaw reservoir from this study but also at concentrations  $>4 \text{ ng L}^{-1}$  from the model analysis. Indicators of heightened *geoA* abundance between reservoirs experiencing elevated geosmin concentrations and non-elevated concentrations were negative linear relationships with mean temperature and dissolved reactive silicate. Both nutrient ratios (TIN:TP and  $\text{NH}_4^+:\text{NO}_3^-$ ) were significantly associated with the abundance of *geoA*. TIN:TP generally had the greatest effect on *geoA* abundance at low ratios, with inhibitory effects witnessed at intermediate levels for TIN:TP ratios which were suppressed at higher TIN:TP ratios. The model analysis also revealed that when  $\text{NH}_4^+:\text{NO}_3^-$  ratios were high, *geoA* abundance was also high. In addition, the positive linear relationships between sulphate and dissolved iron with *geoA* in a reservoir experiencing severe geosmin “events” should not be ignored. Although no correlations between *mic* and 2-MIB concentrations could be determined, the model analysis revealed the significance of a negative linear relationship with *mic* and dissolved reactive silicate and a smoothed relationship with  $\text{NH}_4^+:\text{NO}_3^-$ .

A limitation of this study was the frequency of sampling (monthly) which was a constraint of the water industry project. Paerl et al. (2022) demonstrated that a large change in geosmin concentration can occur on a week to week basis during spring – summer, and Pochiraju et al. (2021) reported that geosmin concentration may decline by 12% within a week. Hence monthly monitoring may miss significant spikes in *geoA* and geosmin concentrations. Similarly, monthly monitoring of *mic* and 2-MIB may miss spikes in 2-MIB concentrations or the resuspension of underlying sediment

containing cyanobacterial species containing the *mic* gene. It is therefore suggested that weekly or at least biweekly monitoring should be implemented by water companies, to facilitate suitable accuracy in predictive capacity.

#### 5.4.1 – Triggers of geosmin “events”

Of the nine reservoirs studied, Alaw showed the most elevated geosmin concentrations considered to be significant “events”, and hence was chosen for modelling triggers of *geoA* prevalence relating to geosmin concentrations. Findings from this study identified *geoA* to be a suitable indicator of geosmin concentrations, although direct significant correlations were only apparent when geosmin concentrations had a large range with high maximum concentrations ( $\leq 420$  ng L<sup>-1</sup>). In accordance with this, previous studies that have reported significant correlations between *geoA* and geosmin also had large ranges of concentrations with elevated maximum concentrations ( $10^5$  ng L<sup>-1</sup>, Su et al., 2013;  $10^2$  ng L<sup>-1</sup>, Tsao et al., 2014;  $10^3$  ng L<sup>-1</sup>, Otten et al., 2016). In addition, Jørgensen et al. (2016) were unable to detect *geoA* from surface and bottom waters that had geosmin concentrations of 1.4 and 5.8 ng L<sup>-1</sup>, respectively. In contrast, Gaget et al. (2020) found low correlations between *geoA* and geosmin with geosmin concentrations up to 18 ng L<sup>-1</sup>. This highlights the need for physical and chemical parameters to be measured in parallel with qPCR to facilitate predictive models (as seen in this study) to determine significant trigger thresholds at geosmin concentrations as low as 4 ng L<sup>-1</sup>.

For reservoirs experiencing mild geosmin concentrations ( $\leq 15$  ng L<sup>-1</sup>) and elevated geosmin concentrations ( $\leq 420$  ng L<sup>-1</sup>) negative relationships existed between mean temperatures, dissolved reactive silicate and *geoA* abundance. Similarly, Shen et al. (2022) found *geoA* gene expression to be higher at 15°C than at 25°C and 35°C. This was consistent with findings from Zhang et al. (2009) that showed geosmin production by *Lyngbya kuetzingii* was maximal at a low temperature (10°C), while Saadoun, Schrader and Blevins (2001) suggested that at low temperatures, more geosmin was synthesized by *Anabaena* sp. Negative relationships with *geoA* and dissolved reactive silicate could be used as a gauge for diatom formation, as depletion of dissolved reactive silicate is usually an indicator of diatom production of silicified cell walls containing amorphous silica (frustules)

(Shimizu et al., 2001). Negative relationships between dissolved reactive silicate and *geoA* point towards a potential mutualistic symbiotic relationship between the two phytoplankton. Olsen, Chislock and Wilson (2016) found that T&O production throughout their study may have been linked to *Synedra* sp. being used as a substrate for cyanobacterial growth and the proliferation of T&O compounds. In addition, to mean temperature and dissolved reactive silicate, geosmin and both nutrient ratios (TIN:TP and  $\text{NH}_4^+:\text{NO}_3^-$ ) ratios were good indicators for elevated *geoA* levels in a reservoir experiencing geosmin “events” and reservoirs with mild geosmin concentrations. This coincides with the findings of Howard (2020) who suggested that low TN:TP favoured the growth and dominance of *Cyanobacteria*, whilst low  $\text{NO}_3^-:\text{NH}_4^+$  promoted the production of T&O compounds. From this study, both low TIN:TP and high  $\text{NH}_4^+:\text{NO}_3^-$  ratios were shown to be significant in relation to *geoA* abundance. Interestingly, at intermediate ratios of TIN:TP the abundance of *geoA* was reduced below the average value of expected *geoA*. TIN:TP ratios revealed that low levels of TIN:TP favoured *geoA* abundance in reservoirs experiencing extreme and mild geosmin concentrations. However, at intermediate ratios the response of *geoA* was inhibited below the average value and when TIN:TP was high the response of *geoA* was regained. Youn et al. (2020) found that cyanobacterial community composition affected geosmin levels, when nitrogen concentrations were high (changing to a high TN:TP) non-nitrogen fixing *Cyanobacteria* dominated. Regained *geoA* levels in this study after an intermediate inhibitory effect of heightened TIN:TP may reveal a transition from nitrogen-fixing *Cyanobacteria* to non-nitrogen-fixing *Cyanobacteria*. Non-nitrogen-fixing *Cyanobacteria* typically prefer a high TN:TP ratio, whereas nitrogen-fixing *Cyanobacteria* are more commonly observed in water columns experiencing low TN:TP (Elliott and May, 2008; Vrede et al., 2009). *Cyanobacteria* have also been reported to assimilate  $\text{NH}_4^+$  more efficiently than  $\text{NO}_3^-$  (Hampel et al., 2018), and  $\text{NO}_3^-$  has been shown to have inhibitory effects on the production of T&O compounds, for example, geosmin in *Dolichospermum* (Saadoun et al., 2001). Here, we support these findings and identify high  $\text{NH}_4^+$  in proportion to low  $\text{NO}_3^-$  to be a key trigger in causing elevations in reservoirs experiencing mild and extreme geosmin concentrations. Therefore, the most useful water chemistry parameters were the ratio of ammonium to nitrate which was previously found in analysing drinking water reservoirs in Wales and England (Perkins et al., 2019).

Dissolved iron and sulphate were both significantly positively associated with *geoA* for a reservoir experiencing geosmin “events”. Molot et al. (2014) proposed that the availability of ferrous iron ( $\text{Fe}^{2+}$ ) regulates the ability of *Cyanobacteria* to compete with other phytoplankton counterparts to assert dominance. *Cyanobacteria* also possess siderophores to readily convert ferric iron ( $\text{Fe}^{3+}$ ) to usable  $\text{Fe}^{2+}$  forms in Fe-limited environments (Wilhelm and Trick, 1994). In combination with dissolved iron, sulphate reduction to sulphide can limit  $\text{Fe}^{2+}$  diffusion rates from anoxic sediments due to insoluble iron sulphide formation (Molot et al., 2014). The increase of sulphate concentrations can thus promote the availability of  $\text{Fe}^{2+}$  for cyanobacterial dominance assertion. However, a negative significant relationship with dissolved reactive silicate may pose as a better early indicator for elevated *geoA* shown in this study.

#### 5.4.2 – *mic* and 2-MIB concentrations

In this study, no relationship between *mic* and 2-MIB concentrations could be determined by month or by season; however, this was likely due to the low concentrations of 2-MIB detected throughout most of this study period (0.57 – 58 ng L<sup>-1</sup>). Chiu et al. (2016) found that *mic* gene levels in some open water samples were below the limit of detection despite 2-MIB being detected. They proposed that this was likely a result of 2-MIB production not being indigenous to the pelagic region, e.g., originating from benthic *Cyanobacteria*, which diffused 2-MIB to the open water sampling site hence why no *mic* genes were detected. Low concentrations of 2-MIB observed throughout this study, despite high levels of *mic* detected, could be due to sediment resuspension, suspending benthic species containing *mic*. Another possible reason for the poor correlation is the periodicity of sampling, as 2-MIB is lost more readily from the water column compared to geosmin owing to higher volatility and biodegradation (Cho, 2007; Li et al., 2012). Both T&O compounds are associated with the thylakoid and cytoplasmic membrane proteins, although 2-MIB is less closely bound and more easily excreted than geosmin (Wu and Juttner, 1988).

For *mic* the  $\text{NH}_4^+:\text{NO}_3^-$  ratio was considered a better indicator of elevated *mic* gene levels in the water column than 2-MIB concentrations, along with a negative linear relationship with dissolved reactive silicate. Both low and high  $\text{NH}_4^+:\text{NO}_3^-$  ratios



revealed the greatest partial effect seen on *mic* levels, with intermediate ratios inhibiting the levels of *mic*. This could be explained by the preference of *Cyanobacteria* for  $\text{NH}_4^+$ ; when *Cyanobacteria* absorb  $\text{NH}_4^+$  they immediately incorporate it into amino acids, whereas they require enzymatic reduction to use  $\text{NO}_3^-$  (Kim et al., 2017). Thus, *Cyanobacteria* that use  $\text{NH}_4^+$  prior to  $\text{NO}_3^-$  may experience inhibition of  $\text{NO}_3^-$  uptake (Dortch, 1990) before being able to produce the enzymes capable of reducing  $\text{NO}_3^-$  to  $\text{NO}_2^-$  then finally to  $\text{NH}_4^+$ . Thus, *mic* abundance would favour a high  $\text{NH}_4^+:\text{NO}_3^-$  ratio but also have the capability of increasing *mic* abundance at low  $\text{NH}_4^+:\text{NO}_3^-$  ratios, leading to potential elevations in 2-MIB concentrations. Dissolved reactive silicate was also a significant proxy for elevated *mic*. This supports additional studies that have linked diatoms to 2-MIB production (Izaguirre and Taylor, 1998; Schrader et al., 2011; Sugiura et al., 2004, 1998). Although many studies have only been able to identify a correlation between 2-MIB and diatoms (Olsen et al., 2016), additional research is required to understand the relationship.

#### 5.4.3 – qPCR primer specificity for *mic* and *geoA*

*In silico* testing, comparing established primer sets (MIBS02F and MIBS02R; (Chiu et al., 2017), Mtcf and MtcR; (Wang et al., 2011), MIB3324F and MIB4050R; (Suurnäkki et al., 2015), MIB3313F and MIB4226R; (Suurnäkki et al., 2015) for *mic* quantification revealed a lack of universality (Appendix A.3). The designed reverse primer MIBHR from this study in combination with MIBS02F (Chiu et al., 2017), allowed for the detection of 30 cyanobacterial strains that possess *mic* (Appendix A.3), seven more strains compared to the original primer set (Appendix A.3; MIBS02F and MIBS02R; (Chiu et al., 2016)). Capturing a larger proportion of *Cyanobacteria* that possess the *mic* gene enabled us to better quantify *mic* present in the water body from this study. However, Wang et al. (2011) stated that more than 40 *Cyanobacteria* species have been identified to produce 2-MIB. This would imply that the *mic*:16S copy number  $\text{mL}^{-1}$  recorded in this study may be underestimated. Thus, the lack of universality in primers used for *mic* detection would indicate that the data reported here is not fully representational of all *Cyanobacteria* that possess the *mic* gene in the water column.

Likewise, the *geoA* qPCR primers used in this study, namely, *geo799F* (John et al., 2018) and *geo982R* (Suurnäkki et al., 2015) were not universal. For future reference, forward primer *geo799F* should be used in conjunction with the reverse primer *geo927R* (John et al., 2018). If taxon-specificity is required to see which producers are present and how much *geoA* they contribute to, this would require multiple sets of taxon-specific primers with differing protocols (Devi et al., 2021).

#### 5.4.4 – Seasonal influence of T&O compounds

In accordance with Oh et al. (2017), geosmin was predicted to have the potential of causing drinking water problems in all seasons, this can be reaffirmed by the winter 2019 results from this study. When geosmin concentrations were low during winter 2020 ( $\leq 0.4 \log_{10}(\text{ng L}^{-1})$ ), an uncoupling of the relationship occurred, illustrated by a slight negative association. Model analysis on a reservoir experiencing geosmin “events” revealed significant differences between *geoA* levels during all seasons when compared to summer 2019. Dzialowski et al. (2009) found that elevated geosmin concentrations were not necessarily confined to summer months, and heightened concentrations of geosmin were found during the winter in some studied reservoirs like this study. For *mic* levels autumn 2019 was the only significant season when compared to winter 2019; this was likely due to the 2-MIB “event” (58 ng L<sup>-1</sup>) witnessed in Pentwyn reservoir during this time.

#### 5.5 – Conclusions

This study demonstrates that *geoA* copy numbers can be implemented as a suitable direct proxy for geosmin concentrations during periods of elevated geosmin concentrations. Through modelling the response of *geoA* in parallel with physical and chemical water parameters, it can be concluded that *geoA* also has suitable predictive applications for geosmin concentrations  $\geq 4 \text{ ng L}^{-1}$ . From these data, nutrient ratios (TIN:TP, and  $\text{NH}_4^+:\text{NO}_3^-$ ) were better predictors of 2-MIB events than 2-MIB concentrations. The  $\text{NH}_4^+:\text{NO}_3^-$  ratio was significantly non-linearly associated with heightened *geoA* abundance when the ratio was high in all studied reservoirs, with the Alaw reservoir having a trigger threshold  $>0.1$ . Negative linear relationships between *geoA* and dissolved reactive silicate and mean temperature should also be

considered as variables for inclusion in modelling T&O event prediction. Therefore, a combined molecular dataset with water chemistry and physical data provides a powerful predictor of geosmin-based T&O events. However, the frequency of samples taken needs to be at least bi-weekly to ensure that fluctuations seen in nutrients, gene levels and T&O concentrations are detected and fully represent the water body. Sample type concerning 2-MIB-producing species (i.e., sediment samples) should also be considered to ensure benthic communities are assessed for *mic* abundance. To evaluate *geoA* response to high and low extremes in TIN:TP ratios, metabarcoding should be implemented to assess species composition in relation to N-fixing and non-N-fixing *Cyanobacteria*.

# Chapter 6: General discussion



The findings presented in this thesis provide substantial knowledge highlighting important abiotic and biotic triggers linked to the production of geosmin and 2-MIB in drinking water reservoirs. Since their identification in the early 1960s, these earthy-muddy-smelling metabolites have been the focus of considerable research. Nevertheless, despite this extensive body of knowledge, geosmin and 2-MIB remain poorly understood throughout much of the water industry, and misconceptions of their proliferation are proceeding to impede the prediction, treatment and control of T&O compounds. This thesis demonstrates a holistic approach to understanding T&O by combining water's chemical, biological, and physical properties. This approach has illustrated the predictive capacity of T&O events when collating datasets which water industries can use to aid water management and pre-emptively schedule treatments, such as PAC in the treatment works. In summary, key nutrient triggers for heightened T&O concentrations were determined (Chapter 3), bacterial and algal community networks were explored – revealing indicative taxa for heightened T&O risk allowing function-specific trait classifications to be formed (Chapter 4), and predictive models were conducted, revealing directly and non-linearly associated variables causing heightened *geoA* and *mic* concentrations (Chapter 5). The first half of this general discussion will summarise and tie Chapters together through common themes shared in the results. The second half will focus on the potential biological and ecological functions of T&O compounds, finishing with a conceptual theory explaining how drinking water reservoirs can transition from a stable (little to no geosmin production) to an unstable state (increased geosmin concentrations).

## 6.1 – T&O events are productivity driven and not always biomass-related

Historically, T&O events have been linked to cyanobacterial blooms (Harris and Graham, 2017; John et al., 2018; Lee et al., 2020). However, in juxtaposition, findings from this thesis support the notion that T&O production is not solely influenced by cyanobacterial biomass. This notion was supported by evidence of no correlations existing between changes in cyanobacterial abundance with geosmin or 2-MIB concentrations (Figure 6.1 – using analysis from Chapter 4). This is following previous research failing to establish correlations between cyanobacterial biomass and *geoA* (Jørgensen et al., 2016; Koltsidou, 2019; Otten et al., 2016) and *mic* gene

copy numbers (Gaget et al., 2020; Koltzidou, 2019; Wang et al., 2016). Equally important, the geosmin event in the Alaw reservoir during winter 2019 supports that T&O production is productivity driven as winter 2019 was accompanied by the lowest Chl *a* TSI indicator value (Chapter 3). Since Chl *a* is typically used as a gauge for phytoplankton biomass (Caputo et al., 2008), it seems unusual that this would be low when *Cyanobacteria* are meant to be in abundance correlating with geosmin concentrations. However, during winter 2019, the Alaw reservoir had the second-highest recorded TP TSI indicator value, suggesting that heightened TP stimulated geosmin. Increases in TP would alter the TN:TP ratio favouring cyanobacterial productivity (Espinosa et al., 2021b; Olsen et al., 2016). Harris et al. (2016) discovered that low TN:TP ratios (<30:1) favoured geosmin events in reservoirs. Similarly, here in this study, heightened *geoA* copy numbers were apparent at a TIN:TP ratio of <24:1 in the Alaw reservoir and <25:1 in all other studied reservoirs (Hooper et al., 2023) (Chapter 5). However, this study hypothesised a shift in communities between diazotrophic and non-diazotrophic *Cyanobacteria* during low and high TIN:TP ratios, respectively. A shift in cyanobacterial community structure could have been owed to the intermediate inhibitory ratios of TN:TP effects on *geoA* copy numbers, which is supported by non-diazotrophic favouring higher TN:TP ratios (Ehrenfels et al., 2021).

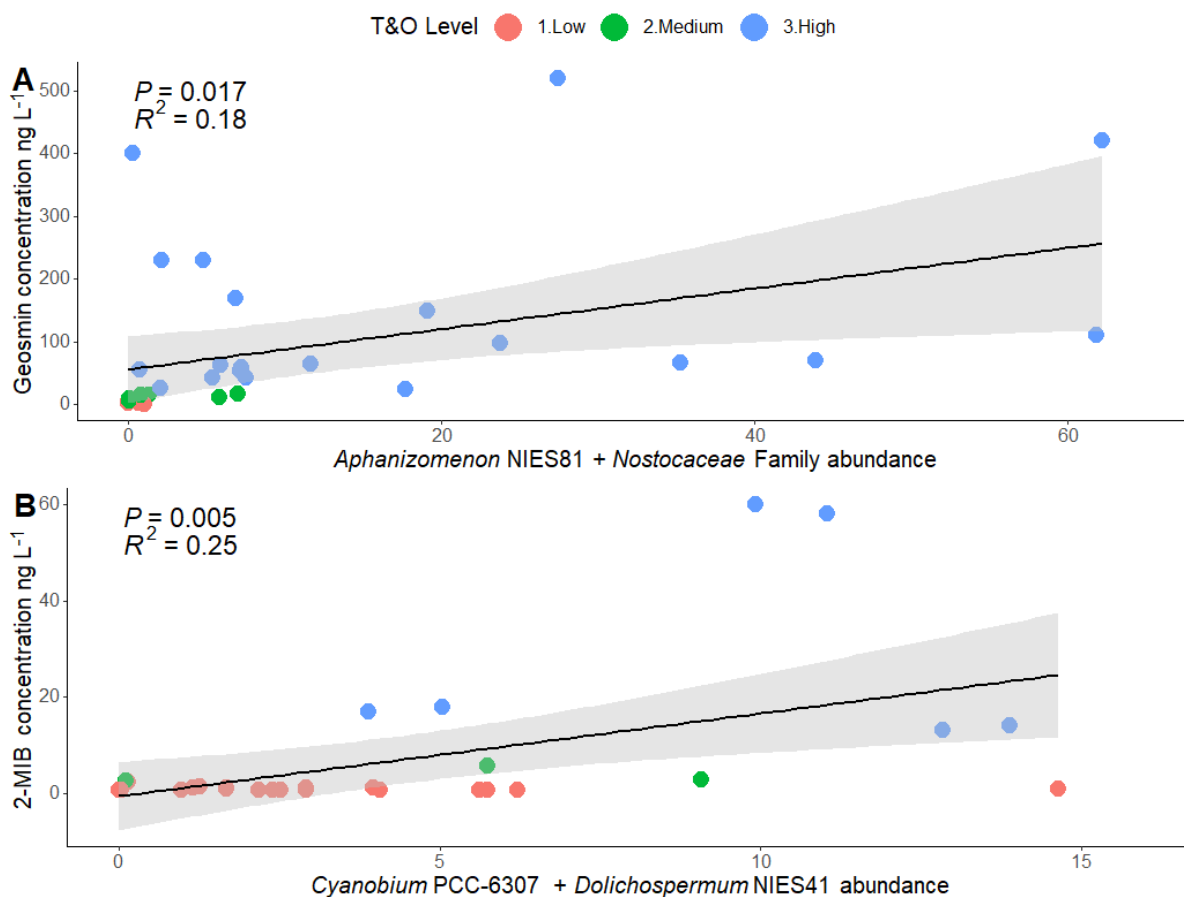


Figure 6.1: Illustrations of linear regression models conducted on the combined relative abundances of significantly co-occurring T&O-producing *Cyanobacteria* against geosmin concentrations in A) the Alaw reservoir and 2-MIB concentrations in B) the Pentwyn reservoir. Associated  $p$  and  $r^2$  values are located in the top left corner.

### 6.3 –T&O triggers and indicators

Preference for the most reduced forms of nitrogen ( $\text{NH}_3$  and  $\text{NH}_4^+$ ) for cyanobacterial growth and productivity has been previously associated with T&O production (Perkins et al., 2019). In Chapter 3, a significant time-lapse existed between the monthly changes in geosmin concentrations and monthly changes in  $\text{NH}_4^+$  concentrations. Identifying heightened changes in  $\text{NH}_4^+$  as a key trigger for monthly lagged changes in geosmin production along with monthly changes in TP, orthophosphate, and non-significant negative effects of  $\text{NO}_3^-$ . Increases in TP would alter the TN:TP ratio, which has been shown to affect geosmin production directly (Dzialowski et al., 2009). Further, a study conducted by Shen et al. (2021) concluded that elevated nitrogen concentrations with phosphorus concentrations promoted Chl



a production, whereas geosmin synthesis decreased in *Dolichospermum smithii* NIES-824. In addition to TP, monthly changes in orthophosphate were also significantly associated with lagged monthly changes in geosmin concentrations. Orthophosphate is the most readily utilisable form of phosphorous for *Cyanobacteria* (Dae-Kyun et al., 2001), and adjustments in orthophosphate concentrations would alter the TN:TP ratio. As well as using available orthophosphate, it has been revealed that *Cyanobacteria* capable of producing the cyanotoxin, cylindrospermopsin, can induce phytoplankton, namely species of *Chlamydomonas*, to produce alkaline phosphatases (APases) to access orthophosphate from organic phosphorous (Bar-Yosef et al., 2010). The *Chlamydomonadaceae* family include the genus *Chlamydomonas*, a positively associated resident in the *Nostocaceae* family cyanosphere in the Alaw reservoir, as demonstrated in Chapter 4. Furthermore, the *Chlamydomonadaceae* family was revealed to indicate high geosmin concentrations by random forest analysis, with the relative abundance of the *Chlamydomonadaceae* family absent during low and medium geosmin concentrations. Although no reports suggest that the *Nostocaceae* family can produce cylindrospermopsin, Kust et al. (2018) stated that *Nostocaceae* are a rich source of unknown secondary metabolites that requires further investigation. These metabolites could have a similar effect to cylindrospermopsin on the *Chlamydomonadaceae* family, inducing the algae to produce APases for the *Cyanobacteria* to use.

$\text{NH}_4^+:\text{NO}_3^-$  ratios were explored alongside TIN:TP ratios with increases in *geoA* and *mic* copy numbers (Chapter 5). Here, heightened  $\text{NH}_4^+:\text{NO}_3^-$  ratios in the Alaw reservoir ( $>0.1$ ) and all other reservoirs except Alaw (0.02 – 0.17) were shown to increase *geoA* abundance significantly and hence potentially increase geosmin production. Similarly,  $\text{NH}_4^+:\text{NO}_3^-$  ratios were also significantly non-linearly associated with *mic* copy numbers, favouring higher ratios. When *Cyanobacteria* absorb  $\text{NH}_4^+$ , they immediately incorporate it into amino acids; however, *Cyanobacteria* can only use  $\text{NO}_3^-$  after enzymatic reduction to  $\text{NO}_2^-$  and  $\text{NH}_4^+$ , and these enzymatic reactions require cellular energy and thereby affect cell growth and productivity (Flores et al., 2005; Flores and Herrero, 2005). Thus, *Cyanobacteria* that utilise  $\text{NH}_4^+$  before  $\text{NO}_3^-$  may experience T&O inhibition from  $\text{NO}_3^-$  uptake (Dortch, 1990; Takamura et al., 1985). For example, Saadoun et al. (2001) reported geosmin production in *Dolichospermum* (previously, *Anabaena*) to be inhibited by increased  $\text{NO}_3^-$

concentrations and promoted by  $\text{NH}_4^+$ . In agreement, Harris et al. (2016) reported increased geosmin and 2-MIB concentrations during high  $\text{NH}_4^+$  concentrations relative to  $\text{NO}_3^-$  concentrations. Therefore, geosmin and 2-MIB production is likely triggered when  $\text{NH}_4^+$  concentrations are high relative to  $\text{NO}_3^-$  concentrations, as it stimulates more rapid growth and hence productivity by *Cyanobacteria*, identifying  $\text{NH}_4^+$  as a critical trigger for T&O production. Investigating nutrient ratios determined that  $\text{NH}_4^+:\text{NO}_3^-$  ratios were better predictors of high *geoA* and *mic* copy numbers than TIN:TP ratios.  $\text{NH}_4^+:\text{NO}_3^-$  ratios were better predictors as TIN:TP had no significant non-linear effects on *mic* copy numbers and low and high TIN:TP ratios had similar effects on *geoA* copy numbers, with intermediate TIN:TP ratios having inhibitory effects. These findings coincide with findings from Howard (2020), who suggested that low TN:TP ratios favoured cyanobacterial growth and dominance, whilst low  $\text{NO}_3^-:\text{NH}_4^+$  ratios promoted the production of T&O compounds. Low and high TIN:TP ratios from this study were hypothesised to reflect a cyanobacterial community shift from diazotrophic *Cyanobacteria* to non-diazotrophic *Cyanobacteria*, respectively.

Chapter 5 revealed *geoA* to be a suitable indicator for geosmin concentrations, although only when geosmin concentrations were elevated ( $>100 \text{ ng L}^{-1}$ ), as seen in the Alaw reservoir. Following this, previous studies that reported significant correlations between *geoA* and geosmin concentrations also had elevated concentrations (Otten et al., 2016; Su et al., 2013; Tsao et al., 2014). In contrast, Jørgensen et al. (2016) could not detect *geoA* in surface and bottom waters with geosmin concentrations of 1.4 and 5.8  $\text{ng L}^{-1}$ , respectively. In addition, Gaget et al. (2020) found low correlations between geosmin concentrations ( $\leq 18 \text{ ng L}^{-1}$ ) and *geoA* copy numbers. Lack of correlations between *geoA* and geosmin concentrations when geosmin concentrations were low emphasised the need for physical and chemical parameters to be measured simultaneously with qPCR data to better monitor changes in T&O biosynthetic gene abundance in the absence of elevated T&O concentrations. Through modelling *geoA* and *mic* copy numbers in all studied reservoirs, triggers causing elevations in the T&O biosynthetic genes were identified. One trigger that exposed a direct negative relationship with both T&O biosynthetic genes was dissolved reactive silicate. A direct inverse relationship was identified in Chapter 3 between monthly changes in geosmin concentrations and monthly

changes in dissolved reactive silicate. Considering dissolved reactive silicate can be used as a gauge for diatom formation, an inverse relationship could reveal a mutualistic symbiosis between *Cyanobacteria* and diatoms. One relationship that was exposed in Chapter 4 existed between *Asterionella* and T&O-producing *Cyanobacteria* in a reservoir experiencing mild geosmin concentrations (Cefni  $\leq 13$  ng L<sup>-1</sup>) and another experiencing 2-MIB events (Pentwyn  $\leq 58$  ng L<sup>-1</sup>). *Asterionella* species are known to extract major anions like NO<sub>3</sub><sup>-</sup>, and incorporation of this diatom within a T&O-producing cyanosphere would dramatically alter the NH<sub>4</sub><sup>+</sup>:NO<sub>3</sub><sup>-</sup> ratio and favour cyanobacterial productivity, hence T&O production.

The roles of dissolved iron and sulphate were less conclusive than previously mentioned nutrients. Direct monthly changes in iron and sulphate described in Chapter 3 did not show any relationships with lagged or regular monthly changes in geosmin and 2-MIB concentrations. In contrast, Chapter 5 revealed *geoA* copy numbers to be positively influenced by sulphate and dissolved iron concentrations in the Alaw reservoir, having the potential to cause geosmin events. No studies have evaluated iron and sulphate concentrations' potential influences on T&O production. Although the critical roles of ferrous iron and sulphate for cyanobacterial bloom formation have been proposed (Molot et al., 2014), highlighting the importance of ferrous iron in regulating the ability of *Cyanobacteria* to compete with other phytoplankton. Combined with dissolved iron, sulphate reduction to sulphide can limit ferrous iron diffusion rates from anoxic sediments due to insoluble iron sulphide formation (Molot et al., 2014). Previous studies have suggested that *Cyanobacteria* have higher iron requirements than eukaryotic algae (Brand, 1991; Morton and Lee, 1974; Parr and Smith, 1976), with diazotrophic *Cyanobacteria* imposing an even higher demand (Glass et al., 2009; Murphy et al., 1976). Indirect evidence for high iron requirements in *Cyanobacteria* comes from the production of siderophores (Molot et al., 2010). The production of cyanobacterial siderophores to aid iron scavenging has been postulated to obtain iron and deprive competitors of iron (Brown and Trick, 1992; Murphy et al., 1976; Wilhelm et al., 1996). Iron requirements have been shown to limit the growth of *Cyanobacteria*; for example, Molot et al. (2010) showed the growth of *Synechococcus* sp. and *Anabaena flos-aquae* to be inhibited in the presence of a strong iron chelator, oxine (8-hydroxyquinoline,

C<sub>9</sub>H<sub>7</sub>NO). Further work must be done to address the potential roles of iron and sulphate on cyanobacterial productivity and hence T&O production.

Out of all weather data associated with samples, the mean temperature was the only variable showing significance with *geoA* copy numbers in reservoirs experiencing mild (all reservoirs except Alaw) and severe (Alaw) geosmin concentrations in Chapter 5. The mean temperature had a direct negative influence on the *geoA* copy numbers; this highlights that unfavourable conditions like low temperatures may induce T&O production. Unfavourable conditions (e.g., low temperature) favouring T&O production are supported by findings from Shen et al. (2022) that revealed low temperatures (15°C) inhibited the growth of *Dolichospermum smithii*, but the expression of the *geoA* gene significantly increased. Inhibition of cyanobacterial growth at low temperatures is not surprising but heightened *geoA* expression further strengthens the case that T&O production is an adaptation in response to environmental stress. Eukaryotic algae have adapted specific physiological responses to environmental stress, e.g., increases in photoprotective pigments during times of high levels of ultraviolet radiation (Alderkamp et al., 2012; Cassar et al., 2011), reduction in photosynthetic rate and inactivation of specific enzymes during low light exposure (Villafane et al., 2003) and evolution of cold shock and antifreeze proteins and photosynthetic electron transport chain adaptations in times of low temperature (Morgan-Kiss et al., 2006). In addition, eukaryotic algae also express molecular changes in response to environmental stress, e.g., up-regulation of cytoprotective genes, down-regulation of genes related to photosynthesis and increases in fucoxanthin Chl *a/c*-binding proteins during periods of increased temperature (Hwang et al., 2008). The ecological functions of geosmin and 2-MIB are still unknown, although the molecular response of increasing *geoA* during unfavourable low temperatures suggests that T&O compounds may serve as protective agents during environmental stress.

T&O production due to environmental stress is mainly highlighted by interpreting the bacterial and algal community compositions discussed in Chapter 4. T&O is known to reflect the environmental and nutrient conditions of the water column, which is informative to the microbiome of a specific niche; thus, T&O production could be a consequence of unfavourable ecological interactions. Unfavourable ecological interactions between T&O-producing *Cyanobacteria* and

specific taxa were identified through co-occurrence network analysis. In this analysis, stable cyanospheres (cyanospheres with more positive than negative significant co-occurring taxa) appeared to be associated with mild to minimum geosmin and 2-MIB concentrations. In juxtaposition, in the Alaw reservoir, there were only two *Cyanobacteria* capable of geosmin production (*Aphanizomenon* NIES81 and the *Nostocaceae* family), although geosmin production likely originated from *Aphanizomenon* NIES81 as this *Cyanobacteria* contained only one resident in the cyanosphere with nine negatively co-occurring associated taxa. Production of geosmin concentrations in the Alaw reservoir by *Aphanizomenon* NIES81 is further supported by this genus having the highest MDA value in categorising geosmin concentration levels from the RF analysis (Chapter 4). Similarly, two *Cyanobacteria* present in the Pentwyn reservoir could produce 2-MIB (*Cyanobium* PCC-6307 and *Dolichospermum* NIES41), although heightened 2-MIB concentrations were likely associated only with *Cyanobium* PCC-6307. *Cyanobium* PCC-6307 had more negatively co-occurring taxa associated than positively co-occurring taxa occupying the cyanosphere. The production of 2-MIB by *Cyanobium* PCC-6307 is further supported by this taxon being within the top 20 MDA values used to categorise 2-MIB concentration levels, whereas *Dolichospermum* NIES41 was absent. In contrast, the Llandegfedd reservoir experienced no geosmin ( $\leq 1.80 \text{ ng L}^{-1}$ ) or 2-MIB ( $\leq 1.80 \text{ ng L}^{-1}$ ) events above the stated event level threshold. However, the Llandegfedd reservoir contained a significantly co-occurring geosmin producer (*Aphanizomenon* MDT14a) and a 2-MIB producer (*Cyanobium* PCC-6307). Llandegfedd's cyanospheres were directly linked and contained many positively co-occurring taxa, with only one taxon per *Cyanobacteria* negatively co-occurring. *Aphanizomenon* MDT14a in the Llandegfedd reservoir was negatively co-occurring with *Pseudomonas*, a known geosmin-degrading genus (Eaton and Sandusky, 2010). This negative association with *Aphanizomenon* MDT14a could be due to *Pseudomonas* only being present in times of heightened geosmin concentrations, which were not present in the Llandegfedd reservoir. Interestingly, Churro et al. (2020) identified three *Pseudomonas* strains with the *geoA* gene, although they could not determine if they are effective geosmin producers. Most studies related to the biodegradation of T&O have identified T&O-degrading bacteria associated with bacterial consortiums in specific niches (Guttman and van Rijn, 2012; Hoefel et al., 2009, 2006; Zhou et al., 2011). Numerous bacteria have evolved to rely on less

abundant and less explored carbon sources as their primary carbon source, like volatile organic compounds (Marmulla and Harder, 2014), which could subsequently result in losing the ability to synthesise terpenoids like geosmin and 2-MIB in natural environments.

Using the biology consortium in Chapter 4 as a response to the environmental pressures subjected to the water column over the study period reflected the water quality. For example, the most informative taxon for classifying geosmin concentrations in the Alaw reservoir was the *Peptostreptococcaceae* family. Increases in the *Peptostreptococcaceae* family typically indicate nutrient enrichment in the water, conditions favouring cyanobacterial T&O production. Interestingly, the community composition of bacteria and, to a certain extent, algae illustrated the existence of signature communities according to geosmin and 2-MIB concentration levels (low, medium, and high) in reservoirs that experienced “high” geosmin and both “high” and “medium” 2-MIB concentrations. This is the first report to date that has shown communities clustering in response to T&O concentrations. Rooney-Varga et al. (2005) suggested that changes in phytoplankton community composition may influence the composition of bacterial communities that function as part of the microbial loop. Differences in the quality of organic matter produced by different phytoplankton communities (including *Cyanobacteria*) have been said to influence the bacterioplankton communities (Van Hannen et al., 1999). Niu et al. (2011) found that the biomass of *Cyanophyta* and *Bacillariophyta* were the most influential factors of bacterioplankton community composition in Lake Taihu, China. Although Niu et al. (2011) did not assess community compositional changes with cyanobacterial T&O, it could be hypothesised that organic matter produced in the form of geosmin and/or 2-MIB selects for the bacterial and algal communities. *Sphingorhabdus* associated with the *Nostocaceae* family cyanosphere in the Alaw reservoir could result from a positive feedback loop where the *Cyanobacteria* is supplying *Sphingorhabdus* with geosmin to degrade. *Sphingorhabdus* was also proven to be a good indicator of high geosmin concentrations according to the RF analysis, significantly increasing in abundance when geosmin concentrations were high (>20 ng L<sup>-1</sup>). Although no reports indicate that *Sphingorhabdus* is a definite geosmin degrader, it possesses genotypic and physiological features similar to the geosmin-degrading genus *Sphingopyxis* (Sharma et al., 2021). Further, some species of the geosmin-

degrading genus *Sphingopyxis* have been reclassified as *Sphingorhabdus* (Subhash et al., 2014; Yang et al., 2017); thus, *Sphingorhabdus* may have the potential to degrade geosmin. However, all other T&O degraders in the analysis were negatively associated with *Cyanobacteria*. Negative associations of T&O degrading taxa with *Cyanobacteria* are hypothesised to be due to antagonistic interactions (like, *Bacillus* and geosmin-producing strains of *Streptomyces* (Zhi et al., 2016)) or because the T&O degrader is only present after cyanobacterial cell lysis. A potential 2-MIB degrader identified in this study that needs further investigation is *Limnobacter*. In the Pentwyn reservoir, *Limnobacter* was significantly the most informative genus in categorising 2-MIB concentrations into low, medium, and high levels, with heightened relative abundance in *Limnobacter* during times of high 2-MIB concentrations. It has been postulated that *Limnobacter* species can degrade a vast array of aromatic compounds (Pérez-Pantoja et al., 2012; Vedler et al., 2013), although there are no reports to date of this genus degrading T&O compounds. In addition, this is the first study to date that reports indicative taxa for increasing levels of geosmin and 2-MIB concentrations, providing function-specific trait categories with example taxa. These categories consisted of geosmin producers, geosmin degraders, nutrient mediators, internal loading indicators, nutrient enrichment indicators, taxa capable of recycling nutrients within the cyanosphere, taxa that protect *Cyanobacteria*, and taxa causing siderophore interference.

## 6.1 – Seasonal occurrences of T&O concentrations

Assumptions of T&O production being confined to warmer months (summer and autumn) can no longer be considered (Li et al., 2016). Here, in Chapter 4, it was revealed that most T&O production was apparent during summer and autumn when cyanobacterial abundance was highest. However, geosmin and 2-MIB concentrations did not correlate with significantly co-occurring T&O-producing cyanobacterial abundance in the Alaw and Pentwyn reservoirs, respectively (Figure 6.1). Heightened T&O concentrations were likely owed to a combination of elevated  $\text{NH}_4^+$  (Chapter 3) altering the  $\text{NH}_4^+:\text{NO}_3^-$  ratios (Chapter 5), lower and higher TIN:TP ratios (depending on nitrogen-fixing capabilities of present *Cyanobacteria*) (Chapter 5), the influence of positive and negative relationships on T&O-producing cyanospheres (Chapter 4) and decreasing temperatures (Chapter 5). Chapter 4

demonstrated that overall cyanobacterial abundances were highest in all reservoirs during the summer and autumn, coinciding with increases in all other taxa. Increases in other taxa abundances may have caused increases in competition for nutrients and induced T&O production, conditions that mainly exist in warmer months. Although during the winter, when there are pulses of key nutrients (e.g.,  $\text{NH}_4^+$ , TP, orthophosphate – highlighted in Chapter 3) from external loading, competition would arise between *Cyanobacteria* and other taxa to attain nutrients for growth and productivity. Competition for nutrients, like TP in the Alaw reservoir during winter 2019 (TP TSI – Chapter 3), further implies that T&O production is productivity-driven rather than cyanobacterial biomass-related. T&O production being productivity driven is supported by findings from Watson et al. (2007) that could not distinguish any relationship between late summer peaks of geosmin (5 – 200 ng L<sup>-1</sup>) and *Cyanobacteria*. In addition, a study conducted by the Oklahoma Water Resources Board (2002) could not establish a connection between geosmin concentrations ( $\leq$  2000 ng L<sup>-1</sup>) and the dominant cyanobacterium, *Anabaena circinalis*.

Generally, Carlson Trophic State Index (CTSI), to determine eutrophication levels proposed by Carlson (1977), was not a good indicator of T&O outcome when assessed by season and year in Chapter 3. Reservoirs considered eutrophic during seasons typically experienced T&O production of either geosmin or 2-MIB over the “event” level threshold. However, two reservoirs (Alwen and Llandgedd) gained eutrophic status during the monitored seasons and did not experience T&O concentrations over the “event” level threshold. However, individual Trophic State Indicators (TSI) values served as better indicators for total phosphorus (TP) in the Alaw reservoir during the winter of 2019 when a geosmin event was apparent. Increases in the monthly changes in TP have also been shown to be a key indicator for increases in lagged geosmin concentrations. Periods of heightened TP would alter the total nitrogen to total phosphorus (TN:TP) ratio selecting for increased cyanobacterial productivity (Wu et al., 2021) and increases in *geoA* abundance and hence geosmin production (Hooper et al., 2023).

Significant variability arose between T&O biosynthetic gene abundances and season and year for geosmin (*geoA*) and 2-MIB (*mic*), as indicated by GAM analysis in Chapter 5. Interestingly, in the Alaw reservoir, there was no significant statistical difference in *geoA* abundance between summer 2019 and winter 2020.



Nevertheless, winter 2019 and summer 2020 significantly differed from *geoA* abundances during summer 2019. Similarly, *geoA* abundances in all other reservoirs except Alaw showed similar seasonal relationships. Indicating that under the right environmental conditions, *geoA*, and to a certain extent, *mic* can increase despite the season, giving the potential to cause T&O events during all seasons. In addition to variability in *geoA* abundances between seasons and years, the mean temperature was significantly inversely proportional to *geoA* abundances in all studied reservoirs. Low temperatures promoting increases in T&O biosynthetic gene abundances are supported by previous findings showing that expression levels of *geoA*, and two 2-MIB synthase genes (*mtf* and *mic*) at 15°C were significantly higher than those at 25°C and 35°C (Shen et al., 2022). Evidence suggests that T&O production is favoured during periods of lower temperatures despite *Cyanobacteria* having an optimum temperature of around 29°C (Lüring et al., 2013). T&O production in times of suboptimal conditions further implies that T&O production may serve as a defence mechanism and somehow geosmin and 2-MIB protect *Cyanobacteria* during environmental stress. T&O production in times of stress seems logical, given that the geosmin gene cluster has two global transcription regulator genes known to modulate cellular signals associated with responses to environmental stress (Zhou et al., 2012). The high conservation of the arrangement of the *geoA* gene cluster in *Cyanobacteria* could therefore indicate the integral role of geosmin production in controlling environmental adaptation and its implication in defence/offence mechanisms towards other taxa (Asquith et al., 2013; Churro et al., 2020; Wang et al., 2013).

## 6.6 – Potential biological and ecological functions of T&O

Eukaryotic algae can produce volatile organic carbon (VOC) compounds that enhance resistance to abiotic stresses, transfer information between algae, play allelopathic roles and protect against predators (Zuo, 2019). However, *Cyanobacteria* are the only known phytoplankton to produce geosmin and 2-MIB (Kwon et al., 2021). Since eukaryotic algae and *Cyanobacteria* inhabit the same niches, these differences in synthesised VOC/T&O compounds may bring competitive superiority to both of them in the phytoplankton community (Koksharova and Safronov, 2022). Martín-Sánchez et al. (2019) suggested that geosmin may

have an important ecological function as a chemical signal or as a specialised protective metabolite against biotic and abiotic stresses, similar to the roles played by terpenoids in plants. Martín-Sánchez et al. (2019) also discovered that the absence of *geoA* in a *Streptomyces* strain was likely due to it being an endophytic plant growth-promoting bacterium. It was hypothesised that the absence of *geoA* may have been complemented by geosmin-like metabolites by the plant host. Further, *geoA* and *mic* have likely been lost through reductive genomic evolution, and now geosmin and 2-MIB producers and non-producers positively co-occur together to share the benefits of these T&O compounds. Gene loss is generally said to leave an organism dependent on co-occurring taxa for lost metabolic functions; it can provide a selective advantage by conserving organisms limiting resources, providing the gene's function is dispensable (Morris et al., 2012). Genetic functions like geosmin (*geoA*) and 2-MIB (*mic*) production are expelled from the cyanobacterial cell and made available to the entire community. Thus geosmin and 2-MIB are dispensable for *Cyanobacteria*, provided they are not lost entirely from the community. This theory could help explain the variability of T&O and non-T&O-producing *Cyanobacteria* and why most cyanospheres in Chapter 4 were interconnected with all *Cyanobacteria* positively co-occurring. Bacteria and their associations with cyanobacterial hosts (forming cyanospheres) from natural environments are profoundly understudied (Sethuraman et al., 2022), with algal associations further understudied. The absence of literature regarding T&O production and cyanospheres highlights the need for further research into the overarching communities of both eukaryotic algae and bacteria that can be associated with T&O risk.

Previous experiments with labelled precursors indicated that geosmin and 2-MIB synthesis are not waste “overflow” products of pigment biosynthesis, as earlier thought (Naes et al., 1989; Zimmerman et al., 1995); alternatively, their production is coupled to more fundamental cell energetics (Zimba et al., 1999). Although these links to fundamental cell energetics are unresolved (Watson, 2003). The Release of T&O compounds is affected by environmental and nutritional conditions, which is informative of the microbiome of a specific niche and T&O production could be a consequence of ecological interactions (as demonstrated in Chapter 4). The synthesis of cyanobacterial T&O compounds is an energy-consuming process; thus,

*Cyanobacteria* must use these compounds as adaptive tools in response to environmental changes (Koksharova, 2020). Geosmin and 2-MIB have the potential to act as “infochemicals” and/or “allelochemicals”, having a significant impact on the physiological functions of prokaryotic and eukaryotic organisms. The T&O compounds have been hypothesised to be allelopathically active compounds, interfering with the settlement and growth of competitors and having negative or positive biological effects on different organisms (Li and Li, 2011). Ikawa et al. (2001) indicated that geosmin and other T&O compounds might act as allelopathic chemicals as geosmin inhibited the growth of *Chlorella pyrenoidosa* by diffusion at a 2-5 mg mL<sup>-1</sup> concentration range. Since the biosynthesis of both T&O compounds is energy consuming, it seems energetically more beneficial for cells to biosynthesise the metabolites upon stress, e.g., by different abiotic and biotic factors (light, temperature, pH, availability of nutrients, grazers, and competitors’ presence). The highest recorded geosmin concentration was from the Alaw reservoir, and RF analysis implied *Aphanizomenon* NIES81 to be responsible for the geosmin production. The *Aphanizomenon* NIES81 cyanosphere had many negatively associated taxa with only one resident (*Fluviicola*). Similarly, RF analysis revealed *Cyanobium* PCC-6307 to be the likely culprit of 2-MIB production in the Pentwyn reservoir; the *Cyanobium* PCC-6307 cyanosphere had a larger number of negatively co-occurring taxa than positive residents. Insinuating that T&O production may have served as a defence mechanism against competitors to assert dominance in the water column if nutrient conditions were suitable for the *Cyanobacteria*. This theory has evolved from studies in higher plants capable of producing T&O compounds and is implicated in defence mechanisms, pollination, and competition (Kanchiswamy et al., 2015; Lu et al., 2003; Ogura et al., 2000; Schlumpberger et al., 2004).

## 6.7 – Cyanobacterial response to environmental stress – the cyanosphere T&O stable state theory

The stable state theory was first proposed in the 1960s (Lewontin, 1969) and described mathematically within ecological communities in the 1970s (May, 1977; Noy-Meir, 1975). This contemporary view states that all ecosystems (cyanospheres in this example) are exposed to gradual changes in climate, nutrients or biotic interactions, usually assumed to respond smoothly (Figure 6.2a and b) (Scheffer et

al., 2001). However, this smooth change can be interrupted by sudden drastic switches induced by environmental stress to a contrasting state (Tilman et al., 2001; Vitousek et al., 1997); ecosystems (cyanospheres in this example) may be inert over specific ranges of stressors, responding more drastically when stressors approach critical levels (Figure 6.2c) (e.g., competition for  $\text{NH}_4^+$  when  $\text{NH}_4^+:\text{NO}_3^-$  ratio is high,  $>0.1$  – Chapter 5). This theory has been adopted in this research and transposed into the “cyanosphere T&O stable state theory”. This theory states that cyanobacterial productivity exists in a stable cyanosphere state which can transition into an unstable cyanosphere state with increased cyanobacterial productivity induced by environmental stress, causing heightened T&O concentrations. Drawing distinctions for T&O risk has previously been difficult without specifying what a T&O event is or having set categories (low, medium, and high) to aid the modelling and forecasting of T&O concentrations into risk categories. In Chapter 2 (Section 2.2), threshold concentrations for a T&O event for geosmin and 2-MIB were defined, and T&O concentration level categories were formed. The importance of defining T&O concentrations into categories is highlighted in Chapter 4 by identifying indicator species via RF analysis using T&O categories. Water companies worldwide have required an applicable conceptual framework based on the underlying environmental dynamics (abiotic and biotic), which can identify T&O risk and predict T&O outcomes. Here, this thesis demonstrates the power of combining the biological (eDNA analysis), chemical (e.g., nutrients) and physical (weather data) properties characteristic of and/or influencing T&O concentrations to provide an adaptable framework for water companies to tackle T&O events. This methodological framework described in this thesis can be adapted and used for forecasting and predicting T&O events in water industries by combining eDNA, chemical and physical water datasets. Here, the “cyanosphere stable state theory” is detailed to help better understand and explain the driving factors behind T&O events, proposing an adaptable evidence-based framework to explain T&O events. It is theorised here that although environmental stress can trigger such shifts in states, loss of resilience may also pave the way for a switch to a prolonged alternative state (Scheffer et al., 2001), similar to the prolonged unstable states causing heightened geosmin concentrations in the Alaw reservoir (Chapter 3). A loss of cyanosphere resilience in conjunction with environmental stressors (e.g., reduced temperatures – Chapter 5, negative associations/competition between *Cyanobacteria* and taxa – Chapter 4)

and/or favourable environmental conditions (e.g., heightened  $\text{NH}_4^+:\text{NO}_3^-$  ratios – Chapter 5, increased phosphorous – Chapter 3, and positive influences in the cyanosphere – Chapter 4) are conditions that existed in the Alaw reservoir. Here, the cyanosphere T&O stable state theory is presented using the Alaw reservoir, which experienced prolonged heightened geosmin concentrations.

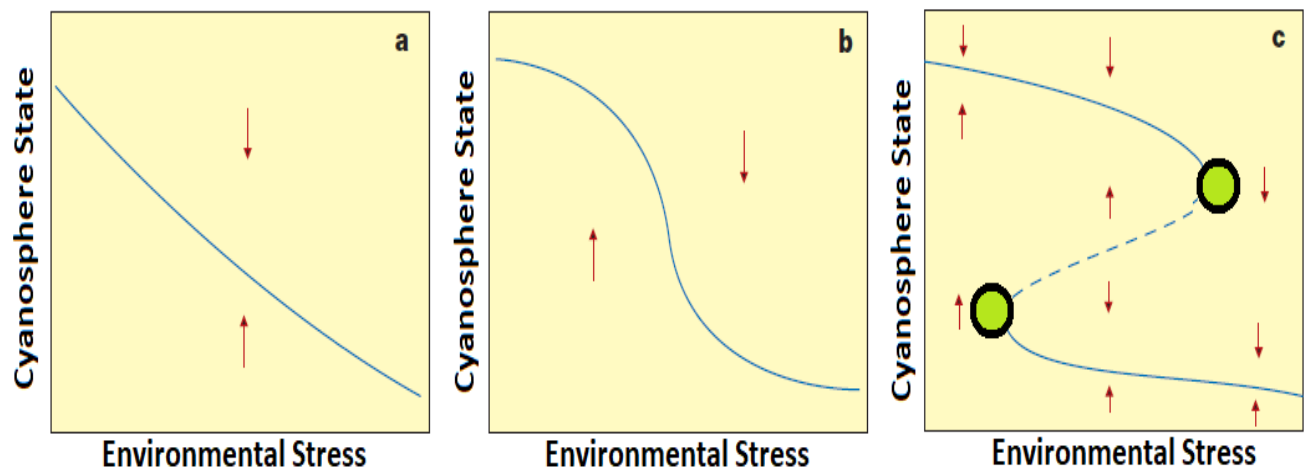


Figure 6.2: Example of ways a cyanobacterial equilibrium state can vary with environmental stressors that create suboptimal conditions for *Cyanobacteria* and lead to T&O events. In a) and b), only one equilibrium exists for each environmental stressor. However, if the equilibrium curve is folded backwards (c), three equilibria can exist for a given environmental stressor. It can be seen from the red arrows indicating the direction of change that, in this case, equilibria on the dashed middle section (between green balls) are unstable and represent the transitional state between the two states.

Chapter 5 revealed *geoA* copy numbers to be a suitable direct proxy for geosmin concentrations (Hooper et al., 2023), so it can be assumed that *geoA* will reflect geosmin concentrations in this theory. In the stable state, it is presumed that positive associations in the cyanosphere stabilise *Cyanobacteria*, and cyanobacterial *geoA* copy numbers remain consistently low and steady; hence geosmin is sustained at consistently low concentrations (Figure 6.3A). In this stable state, cyanospheres are stabilised by positively co-occurring taxa and competition between *Cyanobacteria* and other taxa for resources (e.g.,  $\text{NH}_4^+$ , TP, and orthophosphate – Chapter 3) is limited. When cyanospheres are exposed to environmental stress/suboptimal conditions (e.g., low temperatures – Chapter 5, competition for nutrients

and negative associations between *Cyanobacteria* and taxa – Chapters 3 and 4), this can cause *geoA* copy numbers to enter a transition into the unstable state (Figure 6.3B). In the unstable state, *geoA* copy numbers increase in response to environmental stress (negative associations between *Cyanobacteria* and taxa occur – Chapter 4, decreasing temperature – Chapter 5). When *geoA* increases, there is a risk of geosmin events occurring as *geoA* has the potential to be transcribed and translated to catalyse the production of geosmin. Both states can exist in the cyanosphere stable state theory, stabilised by several buffer mechanisms through positive and/or negative feedbacks. These buffer mechanisms maintain that state by either influencing T&O production through a bottom-up approach (e.g., geosmin producers and nutrient mediators – Chapter 4 and critical nutrients like  $\text{NH}_4^+$  – Chapter 3) or minimising T&O concentrations through a top-down approach (e.g., geosmin degraders – Chapter 4 and increased  $\text{NO}_3^-$  – Chapter 3). Long-standing evidence indicates that shifts between alternative stable states occur in lakes (Carpenter et al., 1999; Scheffer et al., 1997), with one of the best-studied state shifts being the loss of transparency and vegetation observed in shallow lakes subject to human-induced eutrophication (Abell et al., 2022; Moi et al., 2021). In this example, plants increase water clarity through a top-down buffering approach, causing the clear state to be a self-stabilising alternative to the heightened turbidity and phytoplankton blooms characteristic of the unstable state. In juxtaposition, fish are central in sustaining the turbid unstable state through a bottom-up buffering approach (by resuspending sediment) controlling *Daphnia* abundance in the absence of plants. In the cyanosphere T&O stable state theory, both top-down and bottom-up mechanisms help stabilise both states of the cyanosphere, which in turn controls the geosmin production. A detailed summary of factors affecting and contributing to the transition from stable to unstable states can be found below in Table 6.1.

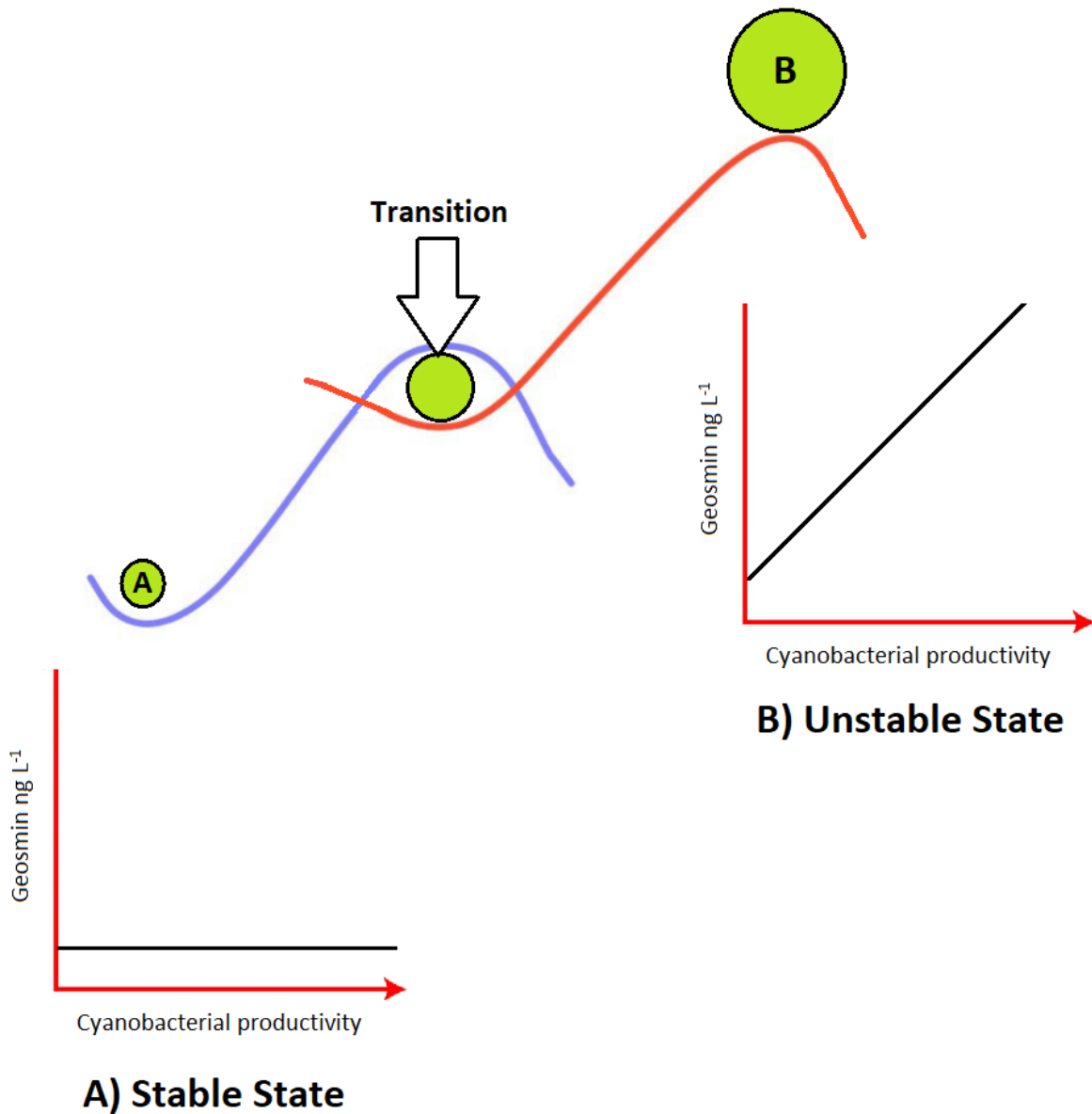


Figure 6.3: Diagram illustrating the “cyanosphere stable state theory”, displaying A) the stable cyanosphere state where *geoA* abundance remains at a steady level with stable cyanobacterial productivity and hence low geosmin production and B) the unstable cyanosphere state where *geoA* abundance increases directly proportional to increases in cyanobacterial productivity leading to increased geosmin production. The size of the green balls is representative of *geoA* abundance.

Table 6.1: Buffer mechanisms maintaining stable and unstable states through bottom-up and top-down approaches involved in the cyanosphere stable state theory.

| Buffer Mechanism | Buffer Representative | Control over T&O influence   |
|------------------|-----------------------|--|
| Bottom-up        | Geosmin producers     | Responsible for the production of geosmin during times of environmental stress (e.g., low temperature - Chapter 5, and negative associations with taxa - Chapter 4) and when $\text{NH}_4^+:\text{NO}_3^-$ ratios are high enough to stimulate heightened cyanobacterial productivity causing elevated <i>geoA</i> abundance (Chapter 5).  |
|                  | Nutrient mediators    | Positive relationships within the cyanosphere with these taxa can supply <i>Cyanobacteria</i> with essential nutrients for growth and productivity and provide the <i>Cyanobacteria</i> with Apase during times of P deficiency (Chapter 4).   |
|                  | $\text{NH}_4^+$       | Increases in the concentration of the most reduced form of nitrogen (relative to nitrate concentrations) will cause rapid increases in cyanobacterial productivity giving rise to elevated geosmin concentrations (Chapter 3). Elevated concentrations of $\text{NH}_4^+$ could also stimulate competition between <i>Cyanobacteria</i> and other taxa causing increased negative associations with the cyanosphere (Chapter 4). |
|                  | TP and orthophosphate | Rises in TP and orthophosphate were shown to influence geosmin production in the Alaw reservoir (Chapter 3). Elevations in these nutrients will also alter TIN:TP ratios causing fluctuations in <i>geoA</i> abundance and hence potential geosmin production (Chapter 5).   |
| Top-down         | Geosmin degraders     | These taxa can be residents within the cyanosphere decreasing and sustaining low geosmin concentrations by geosmin biodegradation. Some geosmin degraders were found to be negatively associated with <i>Cyanobacteria</i> ; the presence of geosmin degraders in the absence of <i>Cyanobacteria</i> could be due to degraders only being present after geosmin is released after cell death/lysis (Chapter 4).                 |
|                  | $\text{NO}_3^-$       | Increases in this oxidised form of nitrogen will inhibit geosmin production (Chapter 3) causing decreases in the $\text{NH}_4^+:\text{NO}_3^-$ ratio which is shown to decrease <i>geoA</i> abundance (Chapter 5).   |

The Alaw reservoir was the only reservoir in this study to experience severe geosmin events ( $\leq 520 \text{ ng L}^{-1}$ ). Results presented in Chapter 3 illustrated that Alaw had substantially higher geosmin concentrations than all other reservoirs over the entire study. In addition, the average geosmin concentration in the Alaw reservoir ( $87.32 \pm 118.12 \text{ ng L}^{-1}$ ) suggests that events were frequent and perhaps prolonged. Prolonged heightened geosmin concentrations could indicate a loss of resilience within the *Aphanizomenon* NIES81 cyanosphere. This loss of resilience is reflected by the lack of stability within the *Aphanizomenon* NIES81 cyanosphere, having more



negatively co-occurring taxa (10 taxa) than positively associated taxa contained within the cyanosphere (one taxon – *Fluviicola*). *Fluviicola* was the only resident of the *Aphanizomenon* NIES81 cyanosphere and has been identified as an indicator taxon belonging to Codon B4 in Chapter 4 and is responsible for the recycling of nutrients within cyanospheres (Farkas et al., 2020; Guedes et al., 2018). However, the magnitude of negatively associated taxa impeded the positive influence of *Fluviicola* in maintaining cyanobacterial stability. Negative relationships between the *Cyanobacteria* and other taxa were most likely due to competition for resources, although amensalism and prey-predator relationships should also not be ignored (Faust and Raes, 2012). In addition, negative relationships may have also been time-lagged; for example, the T&O degrading genus, *Bacillus*, was negatively associated with *Aphanizomenon* NIES81. Intracellular geosmin release after cyanobacterial cell death could therefore be a discriminatory carbon resource selecting for taxa, like taxa belonging to the codon A2 (detailed in Chapter 4). Interestingly, NMDS analysis of both 16S rRNA and *rbcL* communities appeared to cluster according to categorised geosmin concentration levels in reservoirs experiencing high concentrations ( $>20 \text{ ng L}^{-1}$ ), exposing existing T&O signature communities (Chapter 4). In agreement, Watson (2003) stated that although T&O compounds may represent waste by-products of cyanobacterial metabolism, it is highly feasible that these compounds actively modify ecosystem processes and hence community compositions.

Here, in the Alaw reservoir, it is proposed that the loss of resilience in the *Aphanizomenon* NIES81 cyanosphere created a fragile system leading to easy transitions from a stable to an unstable state, especially when subjected to additional environmental stress (e.g., suboptimal temperatures increasing *geoA* abundance – Chapter 5). This fragile system in the Alaw reservoir also had a delayed top-down buffering mechanism (Table 6.1) due to negative associations between *Aphanizomenon* NIES81 and *Bacillus*. This delayed buffering mechanism further exacerbated and prolonged the heightened geosmin concentrations because production and degradation must co-exist to stabilise a state. Thus, due to the time lag in the production (*Aphanizomenon* NIES81) and degradation (*Bacillus*) of geosmin, the cyanosphere could not transition from the unstable state (heightened cyanobacterial productivity, *geoA* and geosmin concentrations) to the stable state

(sustained cyanobacterial productivity, consistently low *geoA* and geosmin concentrations). In contrast to the *Aphanizomenon* NIES81 cyanosphere, members of the *Nostocaceae* family in the Alaw reservoir were positively associated with the geosmin-degrading genus *Sphingorhabdus*. In the *Nostocaceae* cyanosphere, there were no negatively associated taxa, and the production and degradation of geosmin simultaneously occurred. Here, the *Nostocaceae* family sustained a stable state and maintained positive associations in the cyanosphere with a taxon belonging to codons B1 (nutrient mediators - *Chlamydomonadaceae* family) and C1 (offers protection - *Thalassiosira*) (Chapter 4 – Tables 4.15 – 4.16).

In combination with abiotic (e.g., low temperature) and biotic (e.g., negatively co-occurring taxa with *Cyanobacteria*) pressures influencing cyanosphere resilience, anthropogenic forces reduce resilience further by altering the physicochemical environment (e.g., through additions of fertilisers and pesticides). The Alaw reservoir was artificially constructed in the 1960s on a lowland fen (Hatton-ellis, 2016). Fens are wetlands that are underlain by peat and have become of strategic agricultural importance due to their highly fertile soils (Johnson and Land, 2019). Therefore, the Alaw reservoir had a nutrient-rich basin on the lowland; lowlands generally exhibit a low water table, enabling groundwater infiltration into the water column. The catchment area surrounding the Alaw reservoir consists mainly of intensive agriculture (91%) (Edwards, 2016). Agricultural practices enable pesticide residues and nutrient surpluses to accumulate in the environment from using pesticides and fertilisers. Via leaching, run-off, drainage, erosion, volatilisation, and deposition, these agrochemicals enter groundwater that permeates drinking water reservoirs (Siedt et al., 2021). Nitrogen and phosphorous are two macronutrients present in these fertilisers, which can alter important nutrient ratios that can give rise to heightened *geoA* copy numbers leading to increases in geosmin concentrations (Hooper et al., 2023). Pesticides mainly consist of herbicides, insecticides and fungicides, which can positively and negatively influence bacterial and algal communities (Wijewardene et al., 2021). Direct effects of pesticides often result in adverse negative impacts on bacterial and algal communities through inhibition of growth, photosynthesis, and reproduction (Rico-Martínez et al., 2012). The indirect effects of pesticides like insecticides can alter top-down selection pressures, such as predation and competition (Fleeger et al., 2003). In drinking water reservoirs, these

emerging agrochemicals can influence the cyanosphere stable state by inducing changes in communities associated with the cyanosphere, causing competition between more resistant taxa and mortality of the most sensitive taxa (Corcoll et al., 2015).

## Bibliography

- Abd El-Hack, M.E., El-Saadony, M.T., Elbestawy, A.R., Ellakany, H.F., Abaza, S.S., Geneedy, A.M., Salem, H.M., Taha, A.E., Swelum, A.A., Omer, F.A., AbuQamar, S.F., El-Tarabily, K.A., 2022. Undesirable odour substances (geosmin and 2-methylisoborneol) in water environment: Sources, impacts and removal strategies. *Mar. Pollut. Bull.* 178, 1–16.  
<https://doi.org/10.1016/j.marpolbul.2022.113579>
- Abell, J.M., Özkundakci, D., Hamilton, D.P., Reeves, P., 2022. Restoring shallow lakes impaired by eutrophication: Approaches, outcomes, and challenges. *Crit. Rev. Environ. Sci. Technol.* 52, 1199–1246.  
<https://doi.org/10.1080/10643389.2020.1854564>
- Ackman, R.G., Tocher, C.S., McLachlan, J., 1966. Occurrence of Dimethyl- $\beta$ -Propiothetin in Marine Phytoplankton. *J. Fish. Board Canada* 23, 357–364.
- Adam, B., Klawonn, I., Svedén, J.B., Bergkvist, J., Nahar, N., Walve, J., Littmann, S., Whitehouse, M.J., Lavik, G., Kuypers, M.M.M., Ploug, H., 2016. N<sub>2</sub>-fixation, ammonium release and N-transfer to the microbial and classical food web within a plankton community. *ISME J.* 10, 450–459.  
<https://doi.org/10.1038/ismej.2015.126>
- Adams, H., Southard, M., Reeder, S., Buerkens, F., Hallford, R.L., Ikehata, K., Nix, D.K., 2021. Successfully Detecting and Mitigating Algal Blooms and Taste and Odor Compounds. *J. Am. Water Works Assoc.* 113, 10–19.  
<https://doi.org/10.1002/awwa.1743>
- Al-Saud, S., Florea, K.M., Webb, E.A., Thrash, J.C., 2020. Metagenome-Assembled Genome Sequence of Kapabacteriales. *Microbiol. Resour. Announc.* 9, 118–120. <https://doi.org/doi.org/10.1128/MRA.01118-20>
- Alderkamp, A.C., Kulk, G., Buma, A.G.J., Visser, R.J.W., Van Dijken, G.L., Mills, M.M., Arrigo, K.R., 2012. The effect of iron limitation on the photophysiology of *phaeocystis antarctica* (prymnesiophyceae) and *fragilariopsis cylindrus* (bacillariophyceae) under dynamic irradiance. *J. Phycol.* 48, 45–59.  
<https://doi.org/10.1111/j.1529-8817.2011.01098.x>

- Alghanmi, H.A., Alkam, F.M., Al-tae, M.M., 2018. Effect of light and temperature on new cyanobacteria producers for geosmin and 2-methylisoborneol. *J. Appl. Phycol.* 30, 319–328. <https://doi.org/10.1007/s10811-017-1233-0>
- Alhashimi, H.A., Aktas, C.B., 2017. Life cycle environmental and economic performance of biochar compared with activated carbon: A meta-analysis. *Resour. Conserv. Recycl.* 118, 13–26. <https://doi.org/10.1016/j.resconrec.2016.11.016>
- Allahverdiyeva, Y., Suorsa, M., Tikkanen, M., Aro, E., 2015. Photoprotection of photosystems in fluctuating light intensities. *J. Exp. Bot.* 66, 2427–2436. <https://doi.org/10.1093/jxb/eru463>
- Almende, B., Thieurmel, B., Robert, T., 2019. visNetwork: network visualization using 'vis. js' Library. R Package Version 2.0.9.
- Amé, M.V., Wunderlin, D.A., 2005. Effects of iron, ammonium and temperature on microcystin content by a natural concentrated *Microcystis aeruginosa* population. *Water. Air. Soil Pollut.* 168, 235–248. <https://doi.org/10.1007/s11270-005-1774-8>
- An, K.G., Park, S.S., 2003. Influence of seasonal monsoon on the trophic state deviation in an Asian reservoir. *Water. Air. Soil Pollut.* 145, 267–287. <https://doi.org/10.1023/A:1023688819724>
- Andersson, A., Högländer, H., Karlsson, C., Huseby, S., 2015. Key role of phosphorus and nitrogen in regulating cyanobacterial community composition in the northern Baltic Sea. *Estuar. Coast. Shelf Sci.* 164, 161–171. <https://doi.org/10.1016/j.ecss.2015.07.013>
- Asquith, E.A., Evans, C.A., Geary, P.M., Dunstan, R.H., Cole, B., 2013. The role of Actinobacteria in taste and odour episodes involving geosmin and 2-methylisoborneol in aquatic environments. *J. Water Supply Res. Technol. - AQUA* 62, 452–467. <https://doi.org/10.2166/aqua.2013.055>
- Aubriot, L., Bonilla, S., 2012. Rapid regulation of phosphate uptake in freshwater cyanobacterial blooms. *Aquat. Microb. Ecol.* 67, 251–263.
- Azpiazu-muniozguren, M., Garc, M., Laorden, L., Martinez-malaxetxebarria, I., Seoane, S., Bikandi, J., Garaizar, J., Mart, I., 2022. *Anianabacter salinae* gen.

- nov., sp. nov. ASV31, a Facultative Alkaliphilic and Extremely Halotolerant Bacterium Isolated from Brine of a Millennial Continental Saltern. *Diversity* 14, 1–17.
- Baek, D., Lim, J., Cho, Y., Ahn, Y.-T., Lee, H., Park, D., Jung, D., Kim, T.-U., 2015. Investigation of geosmin removal efficiency by microorganism isolated from biological activated carbon. *J. Korean Soc. Water Wastewater* 29, 47–55. <https://doi.org/10.11001/jksww.2015.29.1.047>
- Bai, X., Zhang, T., Wang, C., Zong, D., Li, H., Yang, Z., 2017. Occurrence and distribution of taste and odor compounds in subtropical water supply reservoirs and their fates in water treatment plants. *Environ. Sci. Pollut. Res.* 24, 2904–2913. <https://doi.org/10.1007/s11356-016-7966-5>
- Baker, P., 1992. Identification of Common Noxious Cyanobacteria Part II – Chroococales Oscillatoriales. Australia.
- Baker, P., 1991. Identification of common noxious cyanobacteria part I-Nostocales. Australia.
- Balls, P.W., Macdonald, A., Pugh, K., Edwards, A.C., 1995. Long-term nutrient enrichment of an estuarine system: Ythan, Scotland (1958-1993). *Environ. Pollut.* 90, 311–321. [https://doi.org/10.1016/0269-7491\(95\)00025-M](https://doi.org/10.1016/0269-7491(95)00025-M)
- Bar-Yosef, Y., Sukenik, A., Hadas, O., Viner-Mozzini, Y., Kaplan, A., 2010. Enslavement in the water body by toxic aphanizomenon ovalisporum, inducing alkaline phosphatase in phytoplanktons. *Curr. Biol.* 20, 1557–1561. <https://doi.org/10.1016/j.cub.2010.07.032>
- Barberán, A., Bates, S.T., Casamayor, E.O., Fierer, N., 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6, 343–351. <https://doi.org/10.1038/ismej.2011.119>
- Barry, C., Williamson, J., Pickard, A., Gunn, I., Spears, B., Feuchtmayr, H., Elliot, A., Montheith, D., Evans, C., Grant, A., Olszewska, J., White, S., 2021. Understanding the influence of reservoir dissolved organic matter dynamics on key customer acceptability metrics: DOM Disinfection by-product precursors. Bangor.
- Behr, M., Serchi, T., Cocco, E., Guignard, C., Sergeant, K., Renaut, J., Evers, D.,

2014. Description of the mechanisms underlying geosmin production in *Penicillium expansum* using proteomics. *J. Proteomics* 96, 13–28.  
<https://doi.org/10.1016/j.jprot.2013.10.034>
- Bell, W., Mitchell, R., 1972. Chemotactic and Growth Responses of Marine Bacteria to Algal Extracellular Products. *Biol. Bull.* 143, 265–277.  
<https://doi.org/https://doi.org/10.2307/1540052>
- Ben Hania, W., Joseph, M., Bunk, B., Spröer, C., Klenk, H.P., Fardeau, M.L., Spring, S., 2017. Characterization of the first cultured representative of a Bacteroidetes clade specialized on the scavenging of cyanobacteria. *Environ. Microbiol.* 19, 1134–1148. <https://doi.org/10.1111/1462-2920.13639>
- Bennett, A., Bogorad, L., 1973. Complementary chromatic adaptation in a filamentous blue-green alga. *J. Cell Biol.* 58, 419–435.  
<https://doi.org/10.1083/jcb.58.2.419>
- Bentley, R., Meganathan, R., 1981. Geosmin and methylisoborneol biosynthesis in streptomycetes. Evidence for an isoprenoid pathway and its absence in non-differentiating isolates. *FEBS Lett.* 125, 220–222. [https://doi.org/10.1016/0014-5793\(81\)80723-5](https://doi.org/10.1016/0014-5793(81)80723-5)
- Berg, C., Dupont, C.L., Asplund-Samuelsson, J., Celepli, N.A., Eiler, A., Allen, A.E., Ekman, M., Bergman, B., Ininbergs, K., 2018. Dissection of Microbial Community Functions during a Cyanobacterial Bloom in the Baltic Sea via Metatranscriptomics. *Front. Mar. Sci.* 5, 1–12.  
<https://doi.org/10.3389/fmars.2018.00055>
- Berg, K.A., Lyra, C., Sivonen, K., Paulin, L., Suomalainen, S., Tuomi, P., Rapala, J., 2009. High diversity of cultivable heterotrophic bacteria in association with cyanobacterial water blooms. *ISME J.* 3, 314–325.  
<https://doi.org/10.1038/ismej.2008.110>
- Berry, M.A., Davis, T.W., Cory, R.M., Duhaime, M.B., Johengen, T.H., Kling, G.W., Marino, J.A., Den Uyl, P.A., Gossiaux, D., Dick, G.J., Deneff, V.J., 2017. Cyanobacterial harmful algal blooms are a biological disturbance to Western Lake Erie bacterial communities. *Environ. Microbiol.* 19, 1149–1162.  
<https://doi.org/10.1111/1462-2920.13640>

- Bertone, E., O'halloran, K., 2016. Analysis and modelling of taste and odour events in a shallow subtropical reservoir. *Environ. - MDPI* 3, 1–14.  
<https://doi.org/10.3390/environments3030022>
- Bertos-Fortis, M., Farnelid, H.M., Lindh, M. V., Casini, M., Andersson, A., Pinhassi, J., Legrand, C., 2016. Unscrambling cyanobacteria community dynamics related to environmental factors. *Front. Microbiol.* 7, 1–13.  
<https://doi.org/10.3389/fmicb.2016.00625>
- Bodilis, J., Nsique-Meilo, S., Besaury, L., Quillet, L., 2012. Variable copy number, intra-genomic heterogeneities and lateral transfers of the 16S rRNA gene in *Pseudomonas*. *PLoS One* 7, 1–14.  
<https://doi.org/10.1371/journal.pone.0035647>
- Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciulek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X., Lofffield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A. V, Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Priesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ull-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. <https://doi.org/10.1038/s41587-019-0209-9>



- Bordowitz, J.R., Montgomery, B.L., 2008. Photoregulation of cellular morphology during complementary chromatic adaptation requires sensor-kinase-class protein RcaE in *Fremyella diplosiphon*. *J. Bacteriol.* 190, 4069–4074.  
<https://doi.org/10.1128/JB.00018-08>
- Bostrom, B., Andersen, J.M., Fleischer, S., Jansson, M., 1988. Exchange of phosphorus across the sediment-water interface. *Hydrobiologia* 170, 229–244.
- Bowmer, K.H., Padovan, A., Oliver, R.L., Korth, W., Ganf, G.G., 1992. Physiology of geosmin production by *Anabaena circinalis* isolated from the Murrumbidgee River, Australia. *Water Sci. Technol.* 25, 259–267.
- Bowtell, H.D., Hill, G., 2006. Reservoir Builders of South Wales: Dam Builders in the Age of Steam: Book 6, First Edit. ed. Industrial Locomotive Society.
- Brand, L., 1991. Minimum iron requirements of marine phytoplankton and the implications for the biogeochemical control of new production. *Limnol. Oceanogr.* 36, 1756–1771.
- Brauer, V.S., Stomp, M., Bouvier, T., Fouilland, E., Leboulanger, C., Confurius-Guns, V., Weissing, F.J., Stal, L.J., Huisman, J., 2014. Competition and facilitation between the marine nitrogen-fixing cyanobacterium *Cyanothece* and its associated bacterial community. *Front. Microbiol.* 5, 1–14.  
<https://doi.org/10.3389/fmicb.2014.00795>
- Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5–32.
- Brown, C., Trick, C.G., 1992. Response of the cyanobacterium, *Oscillatoria tenuis*, to low iron environments: the effect on growth rate and evidence for siderophore production. *Arch. Microbiol.* 157, 349–354.
- Bruchet, A., 2019. Chemical analytical techniques for taste and odour compounds, in: Lin, T.-F., Watson, S., Dietrich, A.M., Suffet, I.H. (Eds.), *Taste and Odour in Source and Drinking Water: Causes, Controls, and Consequences*. IWA Publishing, London, UK, pp. 113–138.  
<https://doi.org/doi.org/10.2166/9781780406664>
- Bruchet, A., 1999. Solved and unsolved cases of taste and odor episodes in the files of inspector Cluzeau. *Water Sci. Technol.* 40, 15–21.  
[https://doi.org/10.1016/S0273-1223\(99\)00532-6](https://doi.org/10.1016/S0273-1223(99)00532-6)

- Bruder, S., Babbar-Sebens, M., Tedesco, L., Soyeux, E., 2014. Use of fuzzy logic models for prediction of taste and odor compounds in algal bloom-affected inland water bodies. *Environ. Monit. Assess.* 186, 1525–1545.  
<https://doi.org/10.1007/s10661-013-3471-1>
- Bukin, Y.S., Galachyants, Y.P., Morozov, I. V., Bukin, S. V., Zakharenko, A.S., Zemskaya, T.I., 2019. The effect of 16s rRNA region choice on bacterial community metabarcoding results. *Sci. Data* 6, 1–14.  
<https://doi.org/10.1038/sdata.2019.7>
- Burlingame, G.A., Dann, R.M., Brock, G.L., 1986. A Case Study of Geosmin in Philadelphia's Water. *Am. Water Work. Assoc.* 78, 56–61.
- Cai, H., Jiang, H., Krumholz, L.R., Yang, Z., 2014. Bacterial Community Composition of Size-Fractionated Aggregates within the Phycosphere of Cyanobacterial Blooms in a Eutrophic Freshwater Lake. *PLoS One* 9.  
<https://doi.org/10.1371/journal.pone.0102879>
- Cai, H., Zeng, Y., Wang, Y., Cui, H., Jiang, H., 2018. *Flavobacterium cyanobacterium* sp. nov., isolated from cyanobacterial aggregates in a eutrophic lake. *Int. J. Syst. Evol. Microbiol.* 68, 1279–1284.  
<https://doi.org/10.1099/ijsem.0.002664>
- Çako, V., Baci, S., Shena, M., 2013. Water Turbidity as One of the Trophic State Indices in Butrinti Lake. *J. Water Resour. Prot.* 05, 1144–1148.  
<https://doi.org/10.4236/jwarp.2013.512120>
- Callahan, B.J., 2018. Silva taxonomic training data formatted for DADA2 (Silva version 132) [WWW Document]. [dataset].  
<https://doi.org/10.5281/zenodo.1172783>
- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. <https://doi.org/10.1038/nmeth.3869>
- Cane, D.E., Watt, R.M., 2003. Expression and mechanistic analysis of a germacadienol synthase from *Streptomyces coelicolor* implicated in geosmin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1547–1551.  
<https://doi.org/10.1073/pnas.0337625100>

- Cannon, M. V., Hester, J., Shalkhauser, A., Chan, E.R., Logue, K., Small, S.T., Serre, D., 2016. In silico assessment of primers for eDNA studies using PrimerTree and application to characterize the biodiversity surrounding the Cuyahoga River. *Sci. Rep.* 6, 1–11. <https://doi.org/10.1038/srep22908>
- Caputo, L., Naselli-Flores, L., Ordoñez, J., Armengol, J., 2008. Phytoplankton distribution along trophic gradients within and among reservoirs in Catalonia (Spain). *Freshw. Biol.* 53, 2543–2556. <https://doi.org/10.1111/j.1365-2427.2008.02082.x>
- Carey, C.C., Haney, J.F., Cottingham, K.L., 2007. First record of the ground beetle *Trechoblemus postilenatus* (Coleoptera, Carabidae) in Primorskii krai. *Environ. Toxicol.* 22, 337–339. <https://doi.org/doi:10.1002/tox.20245>
- Carey, C.C., Ibelings, B.W., Hoffmann, E.P., Hamilton, D.P., Brookes, J.D., 2011. Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. *Water Res.* 46, 1394–1407. <https://doi.org/10.1016/j.watres.2011.12.016>
- Carignan, R., Tessier, A., 1988. The co-diagenesis of sulfur and iron in acid lake sediments of southwestern Québec. *Geochim. Cosmochim. Acta* 52, 1179–1188. [https://doi.org/10.1016/0016-7037\(88\)90271-2](https://doi.org/10.1016/0016-7037(88)90271-2)
- Carlson, R.E., 1977. A trophic state index for lakes. *Limnol. Oceanogr.* 22, 361–369. <https://doi.org/10.4319/lo.1977.22.2.0361>
- Carlson, R.E., Simpson, J., 1996. *A Coordinator's Guide to Volunteer Lake Monitoring Methods*, 96th ed. North American Lake Management Society.
- Carpenter, S.R., Ludwig, D., Brock, W.A., 1999. Management of Eutrophication for Lakes Subject to Potentially Irreversible Change. *Ecol. Appl.* 9, 751–771.
- Cassar, N., Difiore, P.J., Barnett, B.A., Bender, M.L., Bowie, A.R., Tilbrook, B., Petrou, K., Westwood, K.J., Wright, S.W., Lefevre, D., 2011. The influence of iron and light on net community production in the Subantarctic and Polar Frontal Zones. *Biogeosciences* 8, 227–237. <https://doi.org/10.5194/bg-8-227-2011>
- Cattaneo, C.R., Rodríguez, Y., Rene, E.R., García-Depraect, O., Muñoz, R., 2022. Biogas bioconversion into poly(3-hydroxybutyrate) by a mixed microbial culture in a novel Taylor flow bioreactor. *Waste Manag.* 150, 364–372.

<https://doi.org/10.1016/j.wasman.2022.07.017>

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A., Kembel, S.W., 2009. The merging of community ecology and phylogenetic biology. *Ecol. Lett.* 12, 693–715.

<https://doi.org/10.1111/j.1461-0248.2009.01314.x>

Çelik, K., Sevindik, T.O., 2016. Seasonal variations of phytoplankton community in relation to some physical and chemical parameters in a temperate eutrophic reservoir, Turkey. *Jordan J. Biol. Sci.* 9, 249–260.

Chapuis-Lardy, L., Temminghoff, E.J.M., De Goede, R.G.M., 2003. Effects of different treatments of cattle slurry manure on water-extractable phosphorus. *NJAS - Wageningen J. Life Sci.* 51, 91–102. [https://doi.org/10.1016/S1573-5214\(03\)80028-3](https://doi.org/10.1016/S1573-5214(03)80028-3)

Chen, J., Liu, J., Rühland, K.M., Smol, J.P., Zhang, X., Zhang, Z., Zhou, A., Shen, Z., Chen, F., 2021. Aquatic ecosystem responses to environmental and climatic changes in NE China since the last deglaciation (~17, 500 cal yr BP) tracked by diatom assemblages from Lake Moon. *Quat. Sci. Rev.* 272, 1–12.

<https://doi.org/10.1016/j.quascirev.2021.107218>

Chen, L., Weng, D., Du, C., Wang, J., Cao, S., 2019. Contribution of frustules and mucilage trails to the mobility of diatom *Navicula* sp. *Sci. Rep.* 9, 1–12.

<https://doi.org/10.1038/s41598-019-43663-z>

Chen, Y., Hobson, P., Burch, D., Lin, T., 2010. In situ measurement of odor compound production by benthic cyanobacteria. *J. Environ. Monit.* 12, 769–775.

<https://doi.org/10.1039/b918487b>

Cheng, F., Cheng, Z., 2015. Research progress on the use of plant allelopathy in agriculture and the physiological and ecological mechanisms of allelopathy.

*Front. Plant Sci.* 6, 1–16. <https://doi.org/10.3389/fpls.2015.01020>

Cheng, J., Karambelkar, B., Xie, Y., 2017. leaflet: Create Interactive Web Maps with the JavaScript “Leaflet” Library.

Chiovitti, A., Higgins, M.J., Harper, R.E., Wetherbee, R., Bacic, A., 2003. The complex polysaccharides of the raphid diatom *Pinnularia viridis*

(Bacillariophyceae). *J. Phycol.* 39, 543–554. <https://doi.org/10.1046/j.1529-8817.2003.02162.x>

- Chiu, Y.T., Chen, Y.H., Wang, T.S., Yen, H.K., Lin, T.F., 2017. A qPCR-based tool to diagnose the presence of harmful cyanobacteria and cyanotoxins in drinking water sources. *Int. J. Environ. Res. Public Health* 14.  
<https://doi.org/10.3390/ijerph14050547>
- Chiu, Y.T., Yen, H.K., Lin, T.F., 2016. An alternative method to quantify 2-MIB producing cyanobacteria in drinking water reservoirs: Method development and field applications. *Environ. Res.* 151, 618–627.  
<https://doi.org/10.1016/j.envres.2016.08.034>
- Cho, H., 2007. 2-Methylisoborneol and natural organic matter adsorption by powdered activated carbon. University of Colorado.
- Choi, Y., Lee, S.J., Kim, H.S., Eom, J.S., Jo, S.U., Guan, L.L., Park, T., Seo, J., Lee, Y., Bae, D., Lee, S.S., 2022. Red seaweed extracts reduce methane production by altering rumen fermentation and microbial composition in vitro. *Front. Vet. Sci.* 9, 1–17. <https://doi.org/10.3389/fvets.2022.985824>
- Chong, S., Lee, H., An, K.-G., 2018. Predicting Taste and Odor Compounds in a Shallow Reservoir Using a Three – Dimensional Hydrodynamic. *Water* 10, 1–19.  
<https://doi.org/10.3390/w10101396>
- Christensen, V.G., Milligan, C.R., Pope, L.M., Zeigler, A.C., 2006. Water Quality and Relation to Taste-and-Odor Compounds in the North Fork Ninnescah River and Cheney Reservoir, South-Central Kansas, 1997–2003. Virginia.  
<https://doi.org/10.3133/sir20065095>
- Chun, S.J., Cui, Y., Lee, J.J., Choi, I.C., Oh, H.M., Ahn, C.Y., 2020. Network analysis reveals succession of *Microcystis* genotypes accompanying distinctive microbial modules with recurrent patterns. *Water Res.* 170.  
<https://doi.org/10.1016/j.watres.2019.115326>
- Churro, C., Semedo-Aguiar, A.P., Silva, A.D., Pereira-Leal, J.B., Leite, R.B., 2020. A novel cyanobacterial geosmin producer, revising GeoA distribution and dispersion patterns in Bacteria. *Sci. Rep.* 10, 1–18.  
<https://doi.org/10.1038/s41598-020-64774-y>
- Clercín, N.A., 2019. Origin and fate of odorous metabolites, 2-methylisoborneol and geosmin, in a eutrophic reservoir. Indiana University.

- Clercin, N.A., 2014. Identification of Taste-and-Odor-Producing Bacteria, in: Using Metagenomics To Identify Taste-and-Odor Producing and Degrading Bacteria in Eagle Creek Reservoir, Central Indiana. IUPUI, Boston, MA, pp. 1–30.
- Clercin, N.A., Druschel, G.K., 2019. Influence of Environmental Factors on the Production of MIB and Geosmin Metabolites by Bacteria in a Eutrophic Reservoir. *Water Resour. Res.* 55, 5413–5430.  
<https://doi.org/10.1029/2018WR023651>
- Clercin, N.A., Druschel, G.K., Gray, M., 2021. Occurrences of 2-methylisoborneol and geosmin –degrading bacteria in a eutrophic reservoir and the role of cell-bound versus dissolved fractions. *J. Environ. Manage.* 297, 1–9.  
<https://doi.org/10.1016/j.jenvman.2021.113304>
- Clercin, N.A., Koltsidou, I., Picard, C.J., Druschel, G.K., 2022. Prevalence of Actinobacteria in the production of 2-methylisoborneol and geosmin, over Cyanobacteria in a temperate eutrophic reservoir. *Chem. Eng. J. Adv.* 9, 1–14.  
<https://doi.org/10.1016/j.cej.2021.100226>
- Clum, A., Tindall, B.J., Sikorski, J., Ivanova, N., Mavrommatis, K., Lucas, S., del Rio, T.G., Nolan, M., Chen, F., Tice, H., Pitluck, S., Cheng, J.F., Chertkov, O., Brettin, T., Han, C., Detter, J.C., Kuske, C., Bruce, A., Goodwin, L., Ovchinnikova, G., Pati, A., Mikhailova, N., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y.J., Jeffries, C.D., Chain, P., Rohde, M., Göker, M., Bristow, J., Eisen, J.A., Markowitz, V., Hugenholtz, H., Kyrpides, N.C., Klenk, H.P., Lapidus, A., 2009. Complete genome sequence of *Pirellula staleyi* type strain (ATCC 27377 T). *Stand. Genomic Sci.* 1, 308–316.  
<https://doi.org/10.4056/sigs.51657>
- Collos, Y., Berges, J., 2003. Nitrogen metabolism in phytoplankton, in: Duarte, C.M., Lot Helgueras, A. (Eds.), *Marine Ecology. Encyclopaedia of Life Support Systems (EOLSS)*, pp. 262–280.
- Cook, D., Newcombe, G., Sztajn bok, P., 2001. THE APPLICATION OF POWDERED ACTIVATED CARBON FOR MIB AND GEOSMIN REMOVAL : PREDICTING PAC DOSES IN FOUR RAW WATERS. *Water Res.* 35, 1325–1333.
- Corcoll, N., Casellas, M., Huerta, B., Guasch, H., Acuña, V., Rodríguez-Mozaz, S., Serra-Compte, A., Barceló, D., Sabater, S., 2015. Effects of flow intermittency

and pharmaceutical exposure on the structure and metabolism of stream biofilms. *Sci. Total Environ.* 503–504, 159–170.  
<https://doi.org/10.1016/j.scitotenv.2014.06.093>

Cottrell, M.T., Kirchman, D.L., 2000. Natural assemblages of marine proteobacteria and members of the Cytophaga-flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* 66, 1692–1697. <https://doi.org/10.1128/AEM.66.4.1692-1697.2000>

Couto, C., Prentice, C., Rui, H., Alves, C.M., Carlos, S., Borges, R., Roselet, F., Cesar, P., 2021. Optimization of a low- cost fertilizer-based medium for large-scale cultivation of the coastal diatom *Conticribra weissflogii* using response surface methodology and its effects on biomass composition. *J. Appl. Phycol.* 33, 2767–2781. <https://doi.org/10.1007/s10811-021-02519-8>

Curren, E., Yoshida, T., Kuwahara, V.S., Leong, S.C.Y., 2019. Rapid profiling of tropical marine cyanobacterial communities. *Reg. Stud. Mar. Sci.* 25.  
<https://doi.org/10.1016/j.rsma.2018.100485>

da Silva Brito, M.T., Duarte-Neto, P.J., Reis Molica, R.J., 2018. *Cylindrospermopsis raciborskii* and *Microcystis aeruginosa* competing under different conditions of pH and inorganic carbon. *Hydrobiologia* 815, 253–266.  
<https://doi.org/10.1007/s10750-018-3567-2>

Dadi, T., Rinke, K., Friese, K., 2020. Trajectories of sediment-water interactions in reservoirs as a result of temperature and oxygen conditions. *Water (Switzerland)* 12, 1–19. <https://doi.org/10.3390/W12041065>

Dae-Kyun, P., Maeng, J., Ahn, C.-Y., Chung, A.-S., Lee, J.-H., Oh, H.M., 2001. Geosmin Concentration and Its Relation to Environmental Factors in Daechung Reservoir , Korea. *Korean J. Limnol.* 34, 319–326.

Dantas, Ê.W., Moura, A.N., Bittencourt-oliveira, M.D.O.C., 2011. Dantas 2011 Cyanobacterial blooms in stratified and destratified eutrophic reservoirs.pdf. *Ann. Brazilian Acad. Sci.* 83, 1327–1338.

Davis, T.W., Berry, D.L., Boyer, G.L., Gobler, C.J., 2009. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. *Harmful Algae* 8, 715–725.

<https://doi.org/10.1016/j.hal.2009.02.004>

Dedysh, S.N., Ivanova, A.A., 2019. Planctomycetes in boreal and subarctic wetlands: Diversity patterns and potential ecological functions. *FEMS Microbiol. Ecol.* 95, 1–10. <https://doi.org/10.1093/femsec/fiy227>

Devi, A., Chiu, Y.T., Hsueh, H.T., Lin, T.F., 2021. Quantitative PCR based detection system for cyanobacterial geosmin/2-methylisoborneol (2-MIB) events in drinking water sources: Current status and challenges. *Water Res.* 188, 116478. <https://doi.org/10.1016/j.watres.2020.116478>

Dickschat, J.S., Nawrath, T., Thiel, V., Kunze, B., Müller, R., Schulz, S., 2007. Biosynthesis of the off-flavor 2-methylisoborneol by the myxobacterium *Nannocystis exedens*. *Angew. Chemie - Int. Ed.* 46, 8287–8290. <https://doi.org/10.1002/anie.200702496>

Domingues, R.B., Barbosa, A.B., Sommer, U., Galvão, H.M., 2011. Ammonium, nitrate and phytoplankton interactions in a freshwater tidal estuarine zone: Potential effects of cultural eutrophication. *Aquat. Sci.* 73, 331–343. <https://doi.org/10.1007/s00027-011-0180-0>

Dortch, Q., 1990. The interaction between ammonium and nitrate uptake in phytoplankton. *Mar. Ecol. Prog. Ser.* 61, 183–201. <https://doi.org/10.3354/meps061183>

Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., Stetsenko, P., Short, T., Lianoglou, S., Antonyan, E., Bonsch, M., Parsonage, H., Ritchie, S., 2019. Package 'data.table.' Ext. 'data.frame.'

Downing, J.A., Watson, S.B., McCauley, E., 2001. Predicting Cyanobacteria dominance in lakes. *Can. J. Fish. Aquat. Sci.* 58, 1905–1908. <https://doi.org/10.1139/cjfas-58-10-1905>

Driscoll, C.B., Otten, T.G., Brown, N.M., Dreher, T.W., 2017. Towards long-read metagenomics: Complete assembly of three novel genomes from bacteria dependent on a diazotrophic cyanobacterium in a freshwater lake co-culture. *Stand. Genomic Sci.* 12, 1–16. <https://doi.org/10.1186/s40793-017-0224-8>

Du, K., Zhou, B., Yuan, R., 2017. Biodegradation of 2-methylisoborneol by single bacterium in culture media and river water environment. *Int. J. Environ. Stud.*



74, 399–411. <https://doi.org/10.1080/00207233.2017.1308158>

- Dunck, B., Nogueira, I. de S., Machado, M. das G., 2012. Planktonic diatoms in lotic and lentic environments in the Lago dos Tigres hydrologic system (Britânia, Goiás, Brazil): Coscinodiscophyceae and Fragilariophyceae. *Brazilian J. Bot.* 35, 181–193. <https://doi.org/10.1590/s1806-99592012000200007>
- Dunfield, P.F., Tamas, I., Lee, K.C., Morgan, X.C., McDonald, I.R., Stott, M.B., 2012. Electing a candidate: A speculative history of the bacterial phylum OP10. *Environ. Microbiol.* 14, 3069–3080. <https://doi.org/10.1111/j.1462-2920.2012.02742.x>
- Durrer, M., Zimmermann, U., Jüttner, F., 1999. Dissolved and particle-bound geosmin in a mesotrophic lake (Lake Zurich): spatial and seasonal distribution and the effect of grazers. *Water Res.* 33, 3628–3636.
- Dziallas, C., Grossart, H.P., 2011. Temperature and biotic factors influence bacterial communities associated with the cyanobacterium *Microcystis* sp. *Environ. Microbiol.* 13, 1632–1641. <https://doi.org/10.1111/j.1462-2920.2011.02479.x>
- Dzialowski, A.R., Smith, V.H., Huggins, D.G., DeNoyelles, F., Lim, N.C., Baker, D.S., Beury, J.H., 2009. Development of predictive models for geosmin-related taste and odor in Kansas, USA, drinking water reservoirs. *Water Res.* 43, 2829–2840. <https://doi.org/10.1016/j.watres.2009.04.001>
- Eaton, R.W., Sandusky, P., 2010. Biotransformations of (+/-)-geosmin by terpene-degrading bacteria. *Biodegradation* 21, 71–79. <https://doi.org/10.1007/s10532-009-9282-y>
- Echenique, R., Giannuzzi, L., Ferrari, L., 2006. Drinking water: problems related to water supply in Bahía Blanca, Argentina. *Acta toxicológica argentina* 14, 23–30.
- Edwards, P., 2016. Nitrogen source apportionment for candidate lake NVZ designations in Wales. NRW Tech Memo TMW16\_01.
- Ehrenfels, B., Bartosiewicz, M., Mbonde, A.S., Baumann, K.B.L., Dinkel, C., Junker, J., Kamulali, T.M., Kimirei, I.A., Niederdorfer, R., Odermatt, D., Pomati, F., Sweke, E.A., Wehrli, B., 2021. Diazotrophic Cyanobacteria are Associated With a Low Nitrate Resupply to Surface Waters in Lake Tanganyika. *Front. Environ. Sci.* 9, 1–16. <https://doi.org/10.3389/fenvs.2021.716765>

- Eiler, A., Bertilsson, S., 2004. Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environ. Microbiol.* 6, 1228–1243. <https://doi.org/10.1111/j.1462-2920.2004.00657.x>
- Eiler, A., Heinrich, F., Bertilsson, S., 2012. Coherent dynamics and association networks among lake bacterioplankton taxa. *ISME J.* 6, 330–342. <https://doi.org/10.1038/ismej.2011.113>
- Elliott, J.A., May, L., 2008. The sensitivity of phytoplankton in Loch Leven (U.K.) to changes in nutrient load and water temperature. *Freshw. Biol.* 53, 32–41. <https://doi.org/10.1111/j.1365-2427.2007.01865.x>
- Elser, J.J., Bracken, M.E.S., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B., Smith, J.E., 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol. Lett.* 10, 1135–1142. <https://doi.org/10.1111/j.1461-0248.2007.01113.x>
- Emerson, J.W., Green, W.A., Schloerke, B., Crowley, J., Cook, D., Hofmann, H., Wickham, H., 2013. The generalized pairs plot. *J. Comput. Graph. Stat.* 22, 79–91. <https://doi.org/10.1080/10618600.2012.694762>
- Enzingmüller-bleyl, T.C., Boden, J.S., Herrmann, A.J., Ebel, K.W., 2021. Lack of Fe (II) transporters in basal Cyanobacteria complicates iron uptake in ferruginous Archean oceans. *bioRxiv* 1–47. <https://doi.org/https://doi.org/10.1101/2021.11.08.467730>
- Espinosa, C., Abril, M., Bretxa, È., Jutglar, M., Ponsá, S., Sellarès, N., Vendrell-Puigmitjà, L., Llenas, L., Ordeix, M., Proia, L., 2021a. Driving Factors of Geosmin Appearance in a Mediterranean River Basin: The Ter River Case. *Front. Microbiol.* 12, 1–14. <https://doi.org/10.3389/fmicb.2021.741750>
- Espinosa, C., Abril, M., Guasch, H., Pou, N., Proia, L., Ricart, M., Ordeix, M., Llenas, L., 2020. Water Flow and Light Availability Influence on Intracellular Geosmin Production in River Biofilms. *Front. Microbiol.* 10, 1–13. <https://doi.org/10.3389/fmicb.2019.03002>
- Espinosa, C., Abril, M., Ponsá, S., Ricart, M., Vendrell-Puigmitjà, L., Ordeix, M., Llenas, L., Proia, L., 2021b. Effects of the interaction between nutrient

- concentration and DIN:SRP ratio on geosmin production by freshwater biofilms. *Sci. Total Environ.* 768, 1–11. <https://doi.org/10.1016/j.scitotenv.2020.144473>
- Farkas, M., Kaszab, E., Radó, J., Háhn, J., Tóth, G., Harkai, P., Ferincz, Á., Lovász, Z., Táncsics, A., Vörös, L., Kriszt, B., Szoboszlay, S., 2020. Planktonic and Benthic Bacterial Communities of the Largest Central European Shallow Lake, Lake Balaton and Its Main Inflow Zala River. *Curr. Microbiol.* 77, 4016–4028. <https://doi.org/10.1007/s00284-020-02241-7>
- Faust, K., Raes, J., 2012. Microbial interactions: From networks to models. *Nat. Rev. Microbiol.* 10, 538–550. <https://doi.org/10.1038/nrmicro2832>
- Fawley, M.W., Fawley, K.P., 2004. A simple and rapid technique for the isolation of DNA from microalgae. *J. Phycol.* 40, 223–225. <https://doi.org/10.1111/j.0022-3646.2004.03-081.x>
- Fernandes, U.L., Castro, M.S., Giroldo, D., 2021. Morphological features of the *Geminella protogenita* (Kützing) West (Chlorophyta, Trebouxiophyceae): a rare freshwater green alga found in southern of Brazil. *Biologia (Bratisl.)* 76, 2889–2894. <https://doi.org/10.1007/s11756-021-00861-3>
- Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S.G., González, J.M., Pedrós-Alió, C., 2013. Ecology of marine bacteroidetes: A comparative genomics approach. *ISME J.* 7, 1026–1037. <https://doi.org/10.1038/ismej.2012.169>
- Fink, P., 2007. Ecological functions of volatile organic compounds in aquatic systems. *Mar. Freshw. Behav. Physiol.* 40, 155–168. <https://doi.org/10.1080/10236240701602218>
- Fink, P., Elert, E. Von, Jüttner, F., 2006a. Volatile Foraging Kairomones in the Littoral Zone: Attraction of an Herbivorous Freshwater Gastropod to Algal Odors. *J. Chem. Ecol.* 32, 1867–1881. <https://doi.org/10.1007/s10886-006-9115-y>
- Fink, P., Von Elert, E., Juttner, F., 2006b. Oxylipins from freshwater diatoms act as attractants for a benthic herbivore. *Arch. für Hydrobiol.* 167, 561–574. <https://doi.org/10.1127/0003-9136/2006/0167-0561>
- Fleeger, J.W., Carman, K.R., Nisbet, R.M., 2003. Indirect effects of contaminants in

- aquatic ecosystems. *Sci. Total Environ.* 317, 207–233.  
[https://doi.org/10.1016/S0048-9697\(03\)00141-4](https://doi.org/10.1016/S0048-9697(03)00141-4)
- Flores, E., Frías, J.E., Rubio, L.M., Herrero, A., 2005. Photosynthetic nitrate assimilation in cyanobacteria. *Photosynth. Res.* 83, 117–133.  
<https://doi.org/10.1007/s11120-004-5830-9>
- Flores, E., Herrero, A., 2010. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat. Rev. Microbiol.* 8, 39–50.  
<https://doi.org/10.1038/nrmicro2242>
- Flores, E., Herrero, A., 2005. Nitrogen assimilation and nitrogen control in cyanobacteria. *Biochem. Soc. Trans.* 33, 164–168.
- Foster, R.A., Kuypers, M.M.M., Vagner, T., Paerl, R.W., Musat, N., Zehr, J.P., 2011. Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses. *ISME J.* 5, 1484–1493. <https://doi.org/10.1038/ismej.2011.26>
- Gaget, V., Hobson, P., Keulen, A., Newton, K., Monis, P., Humpage, A.R., Weyrich, L.S., Brookes, J.D., 2020. Toolbox for the sampling and monitoring of benthic cyanobacteria. *Water Res.* 169, 1–13.  
<https://doi.org/10.1016/j.watres.2019.115222>
- Gerber, N., 1969. A volatile metabolite of *Streptomyces*, 2-methylisoborneol. *J. Antibiot. (Tokyo)*. 22, 508–509. <https://doi.org/10.7164/antibiotics.22.508>
- Gerber, N., Brunswick, N., 1965. Geosmin, Earthy-Smelling Substance Isolated from *Actinomycetes*. *Society* 13, 935–938.
- Ghate, S.D., Shastry, R.P., Arun, A.B., Rekha, P.D., 2021. Unraveling the bacterial community composition across aquatic sediments in the Southwestern coast of India by employing high-throughput 16S rRNA gene sequencing. *Reg. Stud. Mar. Sci.* 46, 1–9. <https://doi.org/10.1016/j.rsma.2021.101890>
- Giglio, S., Chou, W.K.W., Ikeda, H., Cane, D.E., Monis, P.T., 2011. Biosynthesis of 2-Methylisoborneol in Cyanobacteria. *The Phototrophic Prokaryotes* 45, 992–998.  
[https://doi.org/10.1007/978-1-4615-4827-0\\_10](https://doi.org/10.1007/978-1-4615-4827-0_10)
- Giglio, S., Jiang, J., Saint, C.P., Cane, D., Monis, P.T., 2008. Isolation and Characterization of the Gene Associated with Geosmin Production in Cyanobacteria. *Environ. Sci. Technol.* 42, 8027–8032.

- Giglio, Steven, Saint, C.P., Monis, P.T., 2011. Expression of the geosmin synthase gene in the cyanobacterium *Anabaena circinalis* AWQC318. *J. Phycol.* 47, 1338–1343. <https://doi.org/10.1111/j.1529-8817.2011.01061.x>
- Giovannoni, S.J., Cameron, T.J., Temperton, B., 2014. Implications of streamlining theory for microbial ecology. *ISME J.* 8, 1553–1565. <https://doi.org/10.1038/ismej.2014.60>
- Glaeser, S.P., Kämpfer, P., 2014. The Family Sphingomonadaceae, in: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F. (Eds.), *The Prokaryotes*. Springer, Berlin, Heidelberg, pp. 641–707.
- Glass, J.B., Wolfe-Simon, F.W., Anbar, A.D., 2009. Coevolution of metal availability and nitrogen assimilation in cyanobacteria and metal. *Geobiology* 7, 100–123.
- Glibert, P.M., Wilkerson, F.P., Dugdale, R.C., Raven, J.A., Dupont, C.L., Leavitt, P.R., Parker, A.E., Burkholder, J.M., Kana, T.M., 2016. Pluses and minuses of ammonium and nitrate uptake and assimilation by phytoplankton and implications for productivity and community composition, with emphasis on nitrogen-enriched conditions. *Limnol. Oceanogr.* 61, 165–197. <https://doi.org/10.1002/lno.10203>
- Graham, J.L., Loftin, K.A., Meyer, M.T., Ziegler, A.C., 2010. Cyanotoxin mixtures and taste-and-odor compounds in cyanobacterial blooms from the midwestern United States. *Environ. Sci. Technol.* 44, 7361–7368. <https://doi.org/10.1021/es1008938>
- Griffith, D.M., Veech, J.A., Marsh, C.J., 2016. Cooccur: Probabilistic species co-occurrence analysis in R. *J. Stat. Softw.* 69, 1–17. <https://doi.org/10.18637/jss.v069.c02>
- Grover, J.P., Scott, J.T., Roelke, D.L., Brooks, B.W., 2019. Dynamics of nitrogen-fixing cyanobacteria with heterocysts: a stoichiometric model. *Mar. Freshw. Res.* 71, 644–658. <https://doi.org/10.1071/MF18361>
- Guedes, I.A., Rachid, C.T.C.C., Rangel, L.M., Silva, L.H.S., Bisch, P.M., Azevedo, S.M.F.O., Pacheco, A.B.F., 2018. Close link between harmful cyanobacterial dominance and associated bacterioplankton in a tropical eutrophic reservoir. *Front. Microbiol.* 9. <https://doi.org/10.3389/fmicb.2018.00424>

- Guo, D., Liang, J., Chen, W., Wang, J., Ji, B., Luo, S., 2021. Bacterial community analysis of two neighboring freshwater lakes originating from one lake. *Polish J. Environ. Stud.* 30, 111–117. <https://doi.org/10.15244/pjoes/119094>
- Gupta, R.S., Patel, S., 2020. Robust Demarcation of the Family Caryophanaceae (Planococcaceae) and Its Different Genera Including Three Novel Genera Based on Phylogenomics and Highly Specific Molecular Signatures. *Front. Microbiol.* 10, 1–28. <https://doi.org/10.3389/fmicb.2019.02821>
- Gurnell, A.M., Downward, S.R., Jones, R., 1994. Channel planform change on the river dee meanders, 1876–1992. *Regul. Rivers Res. Manag.* 9, 187–204. <https://doi.org/10.1002/rrr.3450090402>
- Gurney Environmental, 2016. Pontsticill Reservoir (2016) [WWW Document]. Install. ResMix 'Source Manag. Syst. to Prev. Tast. odour issues water Treat. Work. URL [https://waterprojectsonline.com/wp-content/uploads/case\\_studies/2016/DCWW\\_Pontsticill\\_2016.pdf](https://waterprojectsonline.com/wp-content/uploads/case_studies/2016/DCWW_Pontsticill_2016.pdf) (accessed 8.9.22).
- Gust, B., Challis, G.L., Fowler, K., Kieser, T., F, C.K., 2003. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *PNAS* 100, 1–30. <https://doi.org/10.1073/pnas.0337542100>
- Guttman, L., van Rijn, J., 2012. Isolation of bacteria capable of growth with 2-methylisoborneol and geosmin as the sole carbon and energy sources. *Appl. Environ. Microbiol.* 78, 363–70. <https://doi.org/10.1128/AEM.06333-11>
- Gutu, A., Kehoe, D.M., 2012. Emerging Perspectives on the Mechanisms , Regulation , and Distribution of Light Color Acclimation in Cyanobacteria. *Mol. Plant* 5, 1–13. <https://doi.org/10.1093/mp/ssr054>
- Halac, S.R., Bazán, R. V., Larrosa, N.B., Nadal, A.F., Ruibal-Conti, A.L., Rodríguez, M.I., Ruiz, M.A., López, A.G., 2019. First report on negative association between cyanobacteria and fecal indicator bacteria at San Roque reservoir (Argentina): impact of environmental factors. *J. Freshw. Ecol.* 34, 273–291. <https://doi.org/10.1080/02705060.2019.1595752>
- Hallbeck, L., Pederson, K., 2014. The Family Gallionellaceae, in: Rosenberg, E.,

- DeLong, E.F., Lory, S.S., Thompson, F. (Eds.), *The Prokaryotes*. Springer, Berlin, Heidelberg, pp. 853–858. [https://doi.org/https://doi.org/10.1007/978-3-642-30197-1\\_398](https://doi.org/https://doi.org/10.1007/978-3-642-30197-1_398)
- Hampel, J.J., McCarthy, M.J., Gardner, W.S., Zhang, L., Xu, H., Zhu, G., Newell, S.E., 2018. Nitrification and ammonium dynamics in Taihu Lake, China: Seasonal competition for ammonium between nitrifiers and cyanobacteria. *Biogeosciences* 15, 733–748. <https://doi.org/10.5194/bg-15-733-2018>
- Harland, F.M.J., Wood, S.A.O., Broady, P.A., Gaw, S., Williamson, W.M., 2014. Polyphasic studies of cyanobacterial strains isolated from benthic freshwater mats in Canterbury, New Zealand. *New Zeal. J. Bot.* 52, 116–135. <https://doi.org/10.1080/0028825X.2013.846266>
- Harris, T.D., Graham, J.L., 2017. Predicting cyanobacterial abundance, microcystin, and geosmin in a eutrophic drinking-water reservoir using a 14-year dataset. *Lake Reserv. Manag.* 33, 32–48. <https://doi.org/10.1080/10402381.2016.1263694>
- Harris, T.D., Smith, V.H., Graham, J.L., Waa, D.B. Van de, Tedesco, L.P., Clercin, N., 2016. Combined effects of nitrogen to phosphorus and nitrate to ammonia ratios on cyanobacterial metabolite concentrations in eutrophic Midwestern USA reservoirs. *Int. Waters* 6, 199–210. <https://doi.org/10.5268/IW-6.2.938>
- Hatton-ellis, T., 2016. Evidence Review of Lake Eutrophication in Evidence Review of Lake Eutrophication in Wales. *NRW Evid. Rep. No.* 135 135. <https://doi.org/10.13140/RG.2.2.35448.44806>
- Havens, K.E., 2008. *Cyanobacteria blooms: Effects on aquatic ecosystems*, Cyanobacte. ed. Springer, New York. [https://doi.org/10.1007/978-0-387-75865-7\\_33](https://doi.org/10.1007/978-0-387-75865-7_33)
- Havens, K.E., Ji, G., Beaver, J.R., Fulton, R.S., Teacher, C.E., 2019. Dynamics of cyanobacteria blooms are linked to the hydrology of shallow Florida lakes and provide insight into possible impacts of climate change. *Hydrobiologia* 829, 43–59. <https://doi.org/10.1007/s10750-017-3425-7>
- Hayes, K.P., Burch, M.D., 1989. Odorous compounds associated with algal blooms in South Australian water. *Water Res.* 23, 115–121.

- He, S., Stevens, S.L.R., Chan, L.-K., Bertilsson, S., Glavina del Rio, T., Tringe, S.G., Malmstrom, R.R., McMahon, K.D., 2017. Ecophysiology of Freshwater Verrucomicrobia Inferred from. *mSphere* 2, e00277-17.
- Herdman, M., Janvier, M., Rippka, R., Stanier, R.Y., 1979. Genome size of cyanobacteria. *J. Gen. Microbiol.* 111, 73–85. <https://doi.org/10.1099/00221287-111-1-73>
- Heudre, D., Wetzel, C.E., Moreau, L., Van de Vijver, B., Ector, L., 2019. On the identity of the rare fragilaria subconstricta (fragilariaceae), with fragilaria species forming ribbon-like colonies shortly reconsidered. *Plant Ecol. Evol.* 152, 327–339. <https://doi.org/10.5091/plecevo.2019.1619>
- Hietanen, S., Moisander, P.H., Kuparinen, J., Tuominen, L., 2002. No sign of denitrification in a Baltic Sea cyanobacterial bloom. *Mar. Ecol. Prog. Ser.* 242, 73–82. <https://doi.org/10.3354/meps242073>
- Hilton, J.A., Foster, R.A., James Tripp, H., Carter, B.J., Zehr, J.P., Villareal, T.A., 2013. Genomic deletions disrupt nitrogen metabolism pathways of a cyanobacterial diatom symbiont. *Nat. Commun.* 4, 1–7. <https://doi.org/10.1038/ncomms2748>
- Hmelo, L.R., Van Mooy, B.A.S., Mincer, T.J., 2012. Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. *Aquat. Microb. Ecol.* 67, 1–14. <https://doi.org/10.3354/ame01571>
- Ho, L., Hoefel, D., Bock, F., Saint, C.P., Newcombe, G., 2007. Biodegradation rates of 2-methylisoborneol (MIB) and geosmin through sand filters and in bioreactors. *Chemosphere* 66, 2210–2218. <https://doi.org/10.1016/j.chemosphere.2006.08.016>
- Ho, M., Soulier, N.T., Canniffe, D.P., Shen, G., Bryant, D.A., 2017. Light regulation of pigment and photosystem biosynthesis in cyanobacteria. *Curr. Opin. Plant Biol.* 37, 24–33. <https://doi.org/10.1016/j.pbi.2017.03.006>
- Hoefel, D., Ho, L., Aunkofer, W., Monis, P.T., Keegan, A., Newcombe, G., Saint, C.P., 2006. Cooperative biodegradation of geosmin by a consortium comprising three gram-negative bacteria isolated from the biofilm of a sand filter column. *Lett. Appl. Microbiol.* 43, 417–423.



- Hoefel, D., Ho, L., Monis, P.T., Newcombe, G., Saint, C.P., 2009. Biodegradation of geosmin by a novel Gram-negative bacterium; isolation, phylogenetic characterisation and degradation rate determination. *Water Res.* 43, 2927–2935. <https://doi.org/10.1016/j.watres.2009.04.005>
- Hojun, L., Depuydt, S., Choi, S., Kim, G., Pandey, L.K., Hader, D., Han, T., Park, J., 2021. Potential use of nuisance cyanobacteria as a source of anticancer agents, in: Sinha, R.P., Hader, D. (Eds.), *Natural Bioactive Compounds*. Academic Press, pp. 203–231.
- Holland, A., Kinnear, S., 2013. Interpreting the Possible Ecological Role(s) of Cyanotoxins: Compounds for Competitive Advantage and/or Physiological Aide? *Mar. Drugs* 11, 2239–2258. <https://doi.org/10.3390/md11072239>
- Hong, J.W., Jo, S.W., Cho, H.W., Nam, S.W., Shin, W., Park, K.M., Lee, K.I., Yoon, H.S., 2015. Phylogeny, morphology, and physiology of *Micractinium* strains isolated from shallow ephemeral freshwater in Antarctica. *Phycol. Res.* 63, 212–218. <https://doi.org/10.1111/pre.12097>
- Hooper, A.S., Kille, P., Watson, S.E., Christofides, S.R., Perkins, R.G., 2023. The importance of nutrient ratios in determining elevations in geosmin synthase (*geoA*) and 2-MIB cyclase (*mic*) resulting in taste and odour events. *Water Res.* 232, 1–13. <https://doi.org/10.1016/j.watres.2023.119693>
- Hopkins, J., 2018. Dee regulation reservoirs. *Dams Reserv.* 28, 94–101. <https://doi.org/10.1680/jdare.18.00035>
- Hopkinson, B.M., Morel, F.M.M., 2009. The role of siderophores in iron acquisition by photosynthetic marine microorganisms. *BioMetals* 22, 659–669. <https://doi.org/10.1007/s10534-009-9235-2>
- Horn, H., Uhlmann, D., 1995. Competitive growth of blue-greens and diatoms (*Fragilaria*) in the Saidenbach Reservoir, saxony. *Water Sci. Technol.* 32, 77–88.
- Howard, C.S., 2020. Taste and odor event dynamics of a midwestern freshwater reservoir. Indiana University.
- Hsieh, S.T., Lin, T.F., Wang, G.S., 2010. Biodegradation of MIB and geosmin with slow sand filters. *J. Environ. Sci. Heal. - Part A Toxic/Hazardous Subst. Environ.*

- Eng. 45, 951–957. <https://doi.org/10.1080/10934521003772352>
- Huang, Z., Jiang, C., Xu, S., Zheng, X., Lv, P., Wang, C., Wang, D., Zhuang, X., 2022. Spatiotemporal changes of bacterial communities during a cyanobacterial bloom in a subtropical water source reservoir ecosystem in China. *Sci. Rep.* 12, 1–29. <https://doi.org/10.1038/s41598-022-17788-7>
- Huisman, J., Sharples, J., Stroom, J.M., Visser, P.M., Kardinaal, W.E.A., Verspagen, J.M.H., Sommeijer, B., 2004. Changes in turbulent mixing shift competition for light between phytoplankton species. *Ecology* 85, 2960–2970. <https://doi.org/10.1890/03-0763>
- Humbert, J.F., Barbe, V., Latifi, A., Gugger, M., Calteau, A., Coursin, T., Lajus, A., Castelli, V., Oztas, S., Samson, G., Longin, C., Medigue, C., de Marsac, N.T., 2013. A Tribute to Disorder in the Genome of the Bloom-Forming Freshwater Cyanobacterium *Microcystis aeruginosa*. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0070747>
- Hwang, Y. sic, Jung, G., Jin, E.S., 2008. Transcriptome analysis of acclimatory responses to thermal stress in Antarctic algae. *Biochem. Biophys. Res. Commun.* 367, 635–641. <https://doi.org/10.1016/j.bbrc.2007.12.176>
- Ikawa, M., Sasner, J.J., Haney, J.F., 2001. Activity of cyanobacterial and algal odor compounds found in lake waters on green alga *Chlorella pyrenoidosa* growth. *Hydrobiologia* 443, 19–22. <https://doi.org/10.1023/A:1017535801766>
- Ishida, H., Miyaji, Y., 1992. Biodegradation of 2-methylisoborneol by oligotrophic bacterium isolated from a eutrophied lake. *Water Sci. Technol.* 25, 269–276. <https://doi.org/10.2166/wst.1992.0061>
- Ivanova, A.A., Naumoff, D.G., Miroshnikov, K.K., Liesack, W., Dedysh, S.N., 2017. Comparative genomics of four isosphaeraceae planctomycetes: A common pool of plasmids and glycoside hydrolase genes shared by *Paludisphaera borealis* PX4, *isosphaera pallida* IS1B, *Singulisphaera acidiphila* DSM 18658, and strain SH-PL62. *Front. Microbiol.* 8, 1–14. <https://doi.org/10.3389/fmicb.2017.00412>
- Iwade, S., Kuma, K., Isoda, Y., Yoshida, M., Kudo, I., Nishioka, J., Suzuki, K., 2006. Effect of high iron concentrations on iron uptake and growth of a coastal diatom *Chaetoceros sociale*. *Aquat. Microb. Ecol.* 43, 177–191.

<https://doi.org/10.3354/ame043177>

- Izaguirre, G., Hwang, C.J., Krasner, S.W., McGuire, M.J., 1982. Geosmin and 2-methylisoborneol from cyanobacteria in three water supply systems. *Appl. Environ. Microbiol.* 43, 708–714. <https://doi.org/10.1128/aem.43.3.708-714.1982>
- Izaguirre, G., Taylor, W., 1998. A *Pseudanabaena* species from Castaic Lake, California that produces 2-methylisoborneol. *Water Res.* 32, 3192. [https://doi.org/10.1016/S0043-1354\(98\)00326-1](https://doi.org/10.1016/S0043-1354(98)00326-1)
- Izaguirre, G., Taylor, W.D., 2007. Geosmin and MIB events in a new reservoir in southern California. *Water Sci. Technol.* 55, 9–14. <https://doi.org/10.2166/wst.2007.156>
- Jahnichen, S., Jaschke, K., Wieland, F., Packroff, G., Benndorf, J., 2011. Spatio-temporal distribution of cell-bound and dissolved geosmin in Wahnbach Reservoir: Causes and potential odour nuisances in raw water. *Water Res.* 45, 4973–4982. <https://doi.org/10.1016/j.watres.2011.06.043>
- James, R.T., Havens, K., Zhu, G., Qin, B., 2009. Comparative analysis of nutrients, chlorophyll and transparency in two large shallow lakes (Lake Taihu, P.R. China and Lake Okeechobee, USA). *Hydrobiologia* 627, 211–231. <https://doi.org/10.1007/s10750-009-9729-5>
- Jarosiewicz, A., Ficek, D., Zapadka, T., 2012. Eutrophication parameters and Carlson-type trophic state indices in selected Pomeranian lakes. *Limnol. Rev.* 11, 15–23. <https://doi.org/10.2478/v10194-011-0023-3>
- Jasti, S., Sieracki, M.E., Poulton, N.J., Giewat, M.W., 2005. Phylogenetic Diversity and Specificity of Bacteria Closely Associated with. *Society* 71, 3483–3494. <https://doi.org/10.1128/AEM.71.7.3483>
- Jeong, J.Y., Lee, S.H., Yun, M.R., Oh, S.E., Lee, K.H., Park, H.D., 2021. 2-Methylisoborneol (2-Mib) Excretion By *Pseudanabaena Yagii* Under Low Temperature. *Microorganisms* 9. <https://doi.org/10.3390/microorganisms9122486>
- Jeppesen, E., Søndergaard, M., Jensen, J.P., Havens, K.E., Anneville, O., Carvalho, L., Coveney, M.F., Deneke, R., Dokulil, M.T., Foy, B., Gerdeaux, D., Hampton, S.E., Hilt, S., Kangur, K., Köhler, J., Lammens, E.H.H.R., Lauridsen, T.L.,

- Manca, M., Miracle, M.R., Moss, B., Nõges, P., Persson, G., Phillips, G., Portielje, R., Romo, S., Schelske, C.L., Straile, D., Tatrai, I., Willén, E., Winder, M., 2005. Lake responses to reduced nutrient loading - An analysis of contemporary long-term data from 35 case studies. *Freshw. Biol.* 50, 1747–1771. <https://doi.org/10.1111/j.1365-2427.2005.01415.x>
- Jiang, J., He, X., Cane, D.E., 2007. Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme. *Natl. Chem. Biol.* 3, 711–715. <https://doi.org/10.1038/nchembio.2007.29>
- Jiang, Y., Cheng, B., Liu, M., Nie, Y., 2016. Spatial and Temporal Variations of Taste and Odor Compounds in Surface Water, Overlying Water and Sediment of the Western Lake Chaohu, China. *Bull. Environ. Contam. Toxicol.* 96, 186–191. <https://doi.org/10.1007/s00128-015-1698-y>
- Jogler, M., Chen, H., Simon, J., Rohde, M., Busse, H.J., Klenk, H.P., Tindall, B.J., Overmann, J., 2013. Description of *Sphingorhabdus planktonica* gen. nov., sp. nov. and reclassification of three related members of the genus *Sphingopyxis* in the genus *Sphingorhabdus* gen. nov. *Int. J. Syst. Evol. Microbiol.* 63, 1342–1349. <https://doi.org/10.1099/ijs.0.043133-0>
- John, N., Koehler, A. V., Ansell, B.R.E., Baker, L., Crosbie, N.D., Jex, A.R., 2018. An improved method for PCR-based detection and routine monitoring of geosmin-producing cyanobacterial blooms. *Water Res.* <https://doi.org/10.1016/j.watres.2018.02.041>
- Jöhnk, K.D., Huisman, J., Sharples, J., Sommeijer, B., Visser, P.M., Stroom, J.M., 2008. Summer heatwaves promote blooms of harmful cyanobacteria. *Glob. Chang. Biol.* 14, 495–512. <https://doi.org/10.1111/j.1365-2486.2007.01510.x>
- Johnson, S., Land, D., 2019. Productive lowland peatlands.
- Jørgensen, N.O.G., Podduturi, R., Burford, M.A., 2016. Relations between abundance of potential geosmin- and 2-MIB-producing organisms and concentrations of these compounds in water from three Australian reservoirs. *J. Water Supply Res. Technol. - AQUA* 65, 504–513. <https://doi.org/10.2166/aqua.2016.001>
- Journey, C.A., Beaulieu, K.M., Bradley, P.M., 2013. Environmental Factors that

- Influence Cyanobacteria and Geosmin Occurrence in Reservoirs, in: INTECH. IntechOpen, pp. 137–164. <https://doi.org/10.5772/32009>
- Ju, F., Zhang, T., 2015. Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant. *ISME J.* 9, 683–695. <https://doi.org/10.1038/ismej.2014.162>
- Jung, S.W., Baek, K.H., Yu, M.J., 2004. Treatment of taste and odor material by oxidation adsorption. *Water Sci. Technol.* 49, 289–295. <https://doi.org/10.2166/wst.2004.0588>
- Jüttner, F., Watson, S.B., 2007. Biochemical and ecological control of geosmin and 2-methylisoborneol in source waters. *Appl. Environ. Microbiol.* 73, 4395–4406. <https://doi.org/10.1128/AEM.02250-06>
- Kahlert, M., Maaria Karjalainen, S., Keck, F., Kelly, M., Ramon, M., Rimet, F., Schneider, S., Tapolczai, K., Zimmermann, J., 2022. Co-occurrence, ecological profiles and geographical distribution based on unique molecular identifiers of the common freshwater diatoms *Fragilaria* and *Ulnaria*. *Ecol. Indic.* 141. <https://doi.org/10.1016/j.ecolind.2022.109114>
- Kakimoto, M., Ishikawa, T., Miyagi, A., Saito, K., Miyazaki, M., Asaeda, T., Yamaguchi, M., Uchimiya, H., Kawai-Yamada, M., 2014. Culture temperature affects gene expression and metabolic pathways in the 2-methylisoborneol-producing cyanobacterium *Pseudanabaena galeata*. *J. Plant Physiol.* 171, 292–300. <https://doi.org/10.1016/j.jplph.2013.09.005>
- Kanchiswamy, C.N., Malnoy, M., Maffei, M.E., 2015. Chemical diversity of microbial volatiles and their potential for plant growth and productivity. *Front. Plant Sci.* 6, 1–23. <https://doi.org/10.3389/fpls.2015.00151>
- Karsten, U., Friedl, T., Schumann, R., Hoyer, K., Lembcke, S., 2005. Mycosporine-like amino acids and phylogenies in green algae: *Prasiola* and its relatives from the Trebouxiophyceae (Chlorophyta). *J. Phycol.* 41, 557–566. <https://doi.org/10.1111/j.1529-8817.2005.00081.x>
- Kathuria, S., Martiny, A.C., 2011. Prevalence of a calcium-based alkaline phosphatase associated with the marine cyanobacterium *Prochlorococcus* and other ocean bacteria. *Environ. Microbiol.* 13, 74–83.

<https://doi.org/10.1111/j.1462-2920.2010.02310.x>

Keene, B., 2002. Bay of Quinte remedial action plan update on impaired beneficial use No. 9: restriction on drinking water consumption, taste and odour. Belleville, Ontario, Canada.

Kehoe, M.J., Chun, K.P., Baulch, H.M., 2015. Who Smells? Forecasting Taste and Odor in a Drinking Water Reservoir. *Environ. Sci. Technol.* 49, 10984–10992. <https://doi.org/10.1021/acs.est.5b00979>

Keshri, J., Pradeep Ram, A.S., Nana, P.A., Sime-Ngando, T., 2018. Taxonomical Resolution and Distribution of Bacterioplankton Along the Vertical Gradient Reveals Pronounced Spatiotemporal Patterns in Contrasted Temperate Freshwater Lakes. *Microb. Ecol.* 76, 372–386. <https://doi.org/10.1007/s00248-018-1143-y>

Khairunisa, B.H., Loganathan, U., Ogejo, J.A., Mukhopadhyay, B., 2022. Nitrogen Transformation Processes in Manure Microbiomes of Earthen Pit and Concrete Storages on Commercial Dairy Farms. *Res. Sq. PREPRINT*, 1–24.

Khanh Tran, H.N., Kim, J.A., Youn, U.J., Kim, S., Woo, M.H., Min, B.S., 2019. Investigation of chemical compounds from *Chlamydomonas* sp. KSF108 (*Chlamydomonadaceae*). *Biochem. Syst. Ecol.* 83, 4–6. <https://doi.org/10.1016/j.bse.2018.12.009>

Kille, P., Svendsen, C., Procházková, P., 2019. Assessing the impact of NPs on invertebrate immune cells : A system Toxicology Approach 43–46.

Kim, C., Il, S., Hwang, S., Cho, M., Kim, H., Hong, S., 2014a. Removal of geosmin and 2-methylisoboneol ( 2-MIB ) by membrane system combined with powdered activated carbon ( PAC ) for drinking water treatment. *J. Water Process Eng.* 4, 91–98. <https://doi.org/10.1016/j.jwpe.2014.09.006>

Kim, C., Lee, S. Il, Hwang, S., Cho, M., Kim, H.S., Noh, S.H., 2014b. Removal of geosmin and 2-methylisoboneol (2-MIB) by membrane system combined with powdered activated carbon (PAC) for drinking water treatment. *J. Water Process Eng.* 4, 91–98. <https://doi.org/10.1016/j.jwpe.2014.09.006>

Kim, H., Jo, B.Y., Kim, H.S., 2017. Effect of different concentrations and ratios of ammonium, nitrate, and phosphate on growth of the blue-green alga

- (Cyanobacterium) *microcystis aeruginosa* isolated from the Nakdong River, Korea. *Algae* 32, 275–284. <https://doi.org/10.4490/algae.2017.32.10.23>
- Kim, K., Park, C., Yoon, Y., Hwang, S.J., 2018. Harmful cyanobacterial material production in the north han river (South Korea): Genetic potential and temperature-dependent properties. *Int. J. Environ. Res. Public Health* 15. <https://doi.org/10.3390/ijerph15030444>
- Kim, K.Y., Khan, J.B., Choi, I.C., Hong, S.H., Lee, J.B., Lee, S.H., J, L.J., 2015. Temporal and Spatial Distribution of Geosmin and 2-MIB in the Daecheong Reservoir. *Korean J. Environ. Agric.* 34, 14–20.
- Kim, M., Lee, J., Yang, D., Park, H.Y., Park, W., 2020. Seasonal dynamics of the bacterial communities associated with cyanobacterial blooms in the Han River. *Environ. Pollut.* 266, 115198. <https://doi.org/10.1016/j.envpol.2020.115198>
- Kim, T.K., Moon, B.R., Kim, T., Kim, M.K., Zoh, K.D., 2016. Degradation mechanisms of geosmin and 2-MIB during UV photolysis and UV/chlorine reactions. *Chemosphere* 162, 157–164. <https://doi.org/10.1016/j.chemosphere.2016.07.079>
- Kimmel, B.L., Groeger, A.W., 2009. Lake and Reservoir Management FACTORS CONTROLLING PRIMARY PRODUCTION IN LAKES AND RESERVOIRS : A PERSPECTIVE. *Lake Reserv. Manag.* 2381, 277–281. <https://doi.org/10.1080/07438148409354524>
- Kingston, J.C., 2003. Araphid and monoraphid diatoms, in: Wehr, J.D., Sheath, R.G. (Eds.), *Freshwater Algae of North America - Ecology and Classification*. Academic Press, Ely, pp. 595–636.
- Koksharova, O.A., 2020. Cyanobacterial VOCs as Allelopathic Tools, in: Ryu, C.-M., Weisskopf, L., Piechulla, B. (Eds.), *Bacterial Volatile Compounds as Mediators of Airborne Interactions*. Springer Singapore, Singapore, pp. 257–280. [https://doi.org/10.1007/978-981-15-7293-7\\_11](https://doi.org/10.1007/978-981-15-7293-7_11)
- Koksharova, O.A., Safronov, N.A., 2022. The effects of secondary bacterial metabolites on photosynthesis in microalgae cells. *Biophys. Rev.* 14, 843–856. <https://doi.org/10.1007/s12551-022-00981-3>
- Koltsidou, I., 2019. Detection and Quantification of Taste and Odor Producing

Bacteria in Eagle Creek Reservoir. Purdue University.

- Komárek, J., Anagnostidis, K., 1989. Modern Approach to the Classification System of Cyanophytes 4. Nostocales. Arch. für Hydrobiol. 82, 247–435.
- Komatsu, M., Tsuda, M., Omura, S., Oikawa, H., Ikeda, H., 2008. Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. Proc. Natl. Acad. Sci. U. S. A. 105, 7422–7427.  
<https://doi.org/10.1073/pnas.0802312105>
- Kong, P., Richardson, P., Hong, C., 2019. Seasonal dynamics of cyanobacteria and eukaryotic phytoplankton in a multiple- reservoir recycling irrigation system. Ecol. Process. 8, 1–11.
- Körner, H., Sofia, H.J., Zumft, W.G., 2003. Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs. FEMS Microbiol. Rev. 27, 559–592.  
[https://doi.org/10.1016/S0168-6445\(03\)00066-4](https://doi.org/10.1016/S0168-6445(03)00066-4)
- Kranzler, C., Rudolf, M., Keren, N., Schleiff, E., 2013. Iron in Cyanobacteria, Advances in Botanical Research. <https://doi.org/10.1016/B978-0-12-394313-2.00003-2>
- Krasner, S.W., Hwang, C.J., McGuire, M.J., 1983. A Standard Method for Quantification of Earthy-Musty Odorants in Water, Sediments, and Algal Cultures. Water Sci. Technol. 15, 127–138.
- Krivtsov, V., Bellinger, E.G., Sigeo, D.C., 2000. Changes in the elemental composition of *Asterionella formosa* during tim diatom spring bloom. J. Plankton Res. 22, 169–184.
- Kulichevskaya, I.S., Ivanova, A.A., Naumoff, D.G., Beletsky, A. V, Rijpstra, W.I.C., Sinninghe Damsté, J.S., Mardanov, A. V, Ravin, N. V, Dedysh, S.N., 2020. *Frigoriglobus tundricola* gen. nov., sp. nov., a psychrotolerant cellulolytic planctomycete of the family Gemmataceae from a littoral tundra wetland. Syst. Appl. Microbiol. 43, 1–10. <https://doi.org/10.1016/j.syapm.2020.126129>
- Kulichevskaya, I.S., Ivanova, A.O., Belova, S.E., Baulina, O.I., Bodelier, P.L.E., Rijpstra, W.I.C., Sinninghe Damsté, J.S., Zavarzin, G.A., Dedysh, S.N., 2007. *Schlesneria paludicola* gen. nov., sp. nov., the first acidophilic member of the



- order Planctomycetales, from Sphagnum-dominated boreal wetlands. *Int. J. Syst. Evol. Microbiol.* 57, 2680–2687. <https://doi.org/10.1099/ijs.0.65157-0>
- Kust, A., Urajová, P., Hrouzek, P., Vu, D.L., Čapková, K., Štenclová, L., Řeháková, K., Kozlíková-Zapomělová, E., Lepšová-Skáclová, O., Lukešová, A., Mareš, J., 2018. A new microcystin producing Nostoc strain discovered in broad toxicological screening of non-planktic Nostocaceae (cyanobacteria). *Toxicon* 150, 66–73. <https://doi.org/10.1016/j.toxicon.2018.05.007>
- Kutovaya, O.A., Watson, S.B., 2014. Development and application of a molecular assay to detect and monitor geosmin-producing cyanobacteria and actinomycetes in the Great Lakes. *J. Great Lakes Res.* 40, 404–414. <https://doi.org/https://doi.org/10.1016/j.jglr.2014.03.016>
- Kwon, Y.S., Cho, I.H., Kim, H.K., Byun, J.H., Bae, M.J., Kim, B.H., 2021. Prediction of geosmin at different depths of lake using machine learning techniques. *Int. J. Environ. Res. Public Health* 18, 1–13. <https://doi.org/10.3390/ijerph181910303>
- Laamanen, M.J., Forsström, L., Sivonen, K., 2002. Diversity of Aphanizomenon flos-aquae (cyanobacterium) populations along a Baltic Sea salinity gradient. *Appl. Environ. Microbiol.* 68, 5296–5303. <https://doi.org/10.1128/AEM.68.11.5296-5303.2002>
- Lan, H., Yang, D., Wang, X., Qi, S., Zhang, Hao, Zhang, P., Wang, C., Liu, J., Zhang, Heng, 2020. Microbiological evaluation of nano-Fe<sub>3</sub>O<sub>4</sub>/GO enhanced the micro-aerobic activate sludge system for the treatment of mid-stage pulping effluent. *Appl. Nanosci.* 10, 1969–1980. <https://doi.org/10.1007/s13204-020-01314-0>
- Larson, C.A., Mirza, B., Rodrigues, J.L.M., Passy, S.I., 2018. Iron limitation effects on nitrogen-fixing organisms with possible implications for cyanobacterial blooms. *FEMS Microbiol. Ecol.* 94, 1–8. <https://doi.org/10.1093/femsec/fiy046>
- Lee, J., Rai, P.K., Jeon, Y.J., Kim, K.H., Kwon, E.E., 2017. The role of algae and cyanobacteria in the production and release of odorants in water. *Environ. Pollut.* 227, 252–262. <https://doi.org/10.1016/j.envpol.2017.04.058>
- Lee, J.E., Youn, S.J., Byeon, M., Yu, S.J., 2020. Occurrence of cyanobacteria, actinomycetes, and geosmin in drinking water reservoir in Korea: A case study

- from an algal bloom in 2012. *Water Sci. Technol. Water Supply* 20, 1862–1870. <https://doi.org/10.2166/ws.2020.102>
- Lee, Jangho, Lee, Juyoun, Lee, T.K., Woo, S.G., Baek, G.S., Park, J., 2013. In-depth characterization of wastewater bacterial community in response to algal growth using pyrosequencing. *J. Microbiol. Biotechnol.* 23, 1472–1477. <https://doi.org/10.4014/jmb.1303.03022>
- Lewington-Pearce, L., Narwani, A., Thomas, M.K., Kremer, C.T., Vogler, H., Kratina, P., 2019. Temperature-dependence of minimum resource requirements alters competitive hierarchies in phytoplankton. *Oikos* 128, 1194–1205. <https://doi.org/10.1111/oik.06060>
- LewisOscar, F., Nithya, C., Alharbi, S.A., Alharbi, N.S., Thajuddin, N., 2018. In vitro and in silico attenuation of quorum sensing mediated pathogenicity in *Pseudomonas aeruginosa* using *Spirulina platensis*. *Microb. Pathog.* 116, 246–256. <https://doi.org/10.1016/j.micpath.2018.01.046>
- Lewontin, R.C., 1969. The meaning of stability. *Brookhaven Symp. Biol.* 22, 13–24.
- Li, J., Dittrich, M., 2018. Dynamic polyphosphate metabolism in cyanobacteria responding to phosphorus availability. *Environ. Microbiol.* 21. <https://doi.org/10.1111/1462-2920.14488>
- Li, R., Carmichael, W.W., Brittain, S., Eaglesham, G.K., Shaw, G.R., Watanabe, M.M., 2001. First report of the cyanotoxins cylindrospermopsin and deoxycylindrospermopsin from *Raphidiopsis curvata* (Cyanobacteria). *J. Phycol.* 37, 1121–1126.
- Li, X., Yu, J., Guo, Q., Su, M., Liu, T., Yang, M., Zhao, Y., 2016. Source-water odor during winter in the Yellow River area of China: Occurrence and diagnosis. *Environ. Pollut.* 218, 252–258. <https://doi.org/10.1016/j.envpol.2016.06.069>
- Li, Y., Li, D., 2011. Competition between toxic *Microcystis aeruginosa* and nontoxic *Microcystis wesenbergii* with *Anabaena PCC7120*. *J. Appl. Phycol.* 24, 69–78. <https://doi.org/10.1007/s10811-010-9648-x>
- Li, Z., Hobson, P., An, W., Burch, M.D., House, J., Yang, M.Y., 2012. Earthy odor compounds production and loss in three cyanobacterial cultures. *Water Res.* 46, 5165–5173. <https://doi.org/10.1016/j.watres.2012.06.008>

- Liato, V., Aïder, M., 2017. Geosmin as a source of the earthy-musty smell in fruits, vegetables and water: Origins, impact on foods and water, and review of the removing techniques. *Chemosphere* 181, 9–18.  
<https://doi.org/10.1016/j.chemosphere.2017.04.039>
- Liu, S., Zhao, F., Fang, X., 2021. The relationship between the community structure and function of bacterioplankton and the environmental response in qingcaosha reservoir. *Water (Switzerland)* 13, 1–19. <https://doi.org/10.3390/w13223155>
- Loar, S.N., 2009. Seasonal Variation in Lake Erie Picoplankton. University of Tennessee, Knoxville.
- Louati, I., Pascault, N., Debroas, D., Bernard, C., Humbert, J.F., Leloup, J., 2015. Structural diversity of bacterial communities associated with bloom-forming freshwater cyanobacteria differs according to the cyanobacterial genus. *PLoS One* 10. <https://doi.org/10.1371/journal.pone.0140614>
- Lu, G., Edwards, C.G., Fellman, J.K., Scott Mattinson, D., Navazio, J., 2003. Biosynthetic origin of geosmin in red beets (*Beta vulgaris* L.). *J. Agric. Food Chem.* 51, 1026–1029. <https://doi.org/10.1021/jf020905r>
- Lu, J., Zhu, B., Struewing, I., Xu, N., Duan, S., 2019. Nitrogen–phosphorus-associated metabolic activities during the development of a cyanobacterial bloom revealed by metatranscriptomics. *Sci. Rep.* 9, 1–11.  
<https://doi.org/10.1038/s41598-019-38481-2>
- Lu, K.Y., Chiu, Y.T., Burch, M., Senoro, D., Lin, T.F., 2019. A molecular-based method to estimate the risk associated with cyanotoxins and odor compounds in drinking water sources. *Water Res.* 164, 1–9.  
<https://doi.org/10.1016/j.watres.2019.114938>
- Ludwig, F., Medger, A., Börnick, H., Opitz, M., Lang, K., Göttfert, M., Röske, I., 2007. Identification and expression analyses of putative sesquiterpene synthase genes in *Phormidium* sp. and prevalence of *geoA*-like genes in a drinking water reservoir. *Appl. Environ. Microbiol.* 73, 6988–6993.  
<https://doi.org/10.1128/AEM.01197-07>
- Lukassen, Mie B., de Jonge, N., Bjerregaard, S.M., Podduturi, R., Jørgensen, N.O.G., Petersen, M.A., David, G.S., da Silva, R.J., Nielsen, J.L., 2019.

- Microbial production of the off-flavor geosmin in tilapia production in Brazilian water reservoirs: Importance of bacteria in the intestine and other fish-associated environments. *Front. Microbiol.* 10, 1–12.  
<https://doi.org/10.3389/fmicb.2019.02447>
- Lukassen, Mie Bech, Podduturi, R., Rohaan, B., Jørgensen, N.O.G., Nielsen, J.L., 2019. Dynamics of geosmin-producing bacteria in a full-scale saltwater recirculated aquaculture system. *Aquaculture* 500, 170–177.  
<https://doi.org/10.1016/j.aquaculture.2018.10.008>
- Luo, F., Chen, H., Wu, X., Liu, L., Chen, Y., Wang, Z., 2022. Insights into the Seasonal Olfactory Mechanism of Geosmin in Raw Water of Huangpu River. *Toxics* 10, 1–12. <https://doi.org/10.3390/toxics10080485>
- Luo, F., You, Q., Yu, P., Pang, W., Wang, Q., 2019. Eunotia (Bacillariophyta) biodiversity from high altitude, freshwater habitats in the Mugecuo Scenic Area, Sichuan Province, China. *Phytotaxa* 394, 133–147.  
<https://doi.org/10.11646/phytotaxa.394.2.2>
- Lüring, M., Eshetu, F., Faassen, E.J., Kosten, S., Huszar, V.L.M., 2013. Comparison of cyanobacterial and green algal growth rates at different temperatures. *Freshw. Biol.* 58, 552–559. <https://doi.org/10.1111/j.1365-2427.2012.02866.x>
- Ma, N.N., Luo, G.Z., Tan, H.X., Yao, M.L., Wange, X.Y., 2015. Kinetic Characteristics of Degradation of Geosmin and 2-Methylisoborneol by *Bacillus subtilis*. *Huan Jing Ke Xue* 36, 1379–1384.
- Maldonado, M., Ribes, M., van Duyl, F.C., 2012. Nutrient Fluxes Through Sponges: Biology, Budgets, and Ecological Implications, in: Becerro, M.A., Uriz, M.J., Maldonado, M., Turon, X. (Eds.), *Advances in Marine Biology*. Academic Press, pp. 113–182. <https://doi.org/https://doi.org/10.1016/B978-0-12-394283-8.00003-5>
- Mallott, E.K., Garber, P.A., Malhi, R.S., 2018. Trnl outperforms rbcl as a DNA metabarcoding marker when compared with the observed plant component of the diet of wild white-faced capuchins (*cebus capucinus*, primates). *PLoS One* 13, 1–16. <https://doi.org/10.1371/journal.pone.0199556>
- Manage, P.M., Kawabata, Z., Nakano, S. ichi, 2000. Algicidal effect of the bacterium

- Alcaligenes denitrificans on Microcystis spp. *Aquat. Microb. Ecol.* 22, 111–117.  
<https://doi.org/10.3354/ame022111>
- Mangan, N.M., Flamholz, A., Hood, R.D., Milo, R., Savage, D.F., 2016. PH determines the energetic efficiency of the cyanobacterial CO<sub>2</sub> concentrating mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 113, E5354–E5362.  
<https://doi.org/10.1073/pnas.1525145113>
- Marmulla, R., Harder, J., 2014. Microbial monoterpene transformations-a review. *Front. Microbiol.* 5, 1–14. <https://doi.org/10.3389/fmicb.2014.00346>
- Maruyama, T., Kato, K., Yokoyama, A., Tanaka, T., Hiraishi, A., Park, H.D., 2003. Dynamics of microcystin-degrading bacteria in mucilage of *Microcystis*. *Microb. Ecol.* 46, 279–288. <https://doi.org/10.1007/s00248-002-3007-7>
- Matthews, R., Hilles, M., Pelletier, G., 2002. Determining trophic state in Lake Whatcom, Washington (USA), a soft water lake exhibiting seasonal nitrogen limitation. *Hydrobiologia* 468, 107–121.  
<https://doi.org/10.1023/A:1015288519122>
- Mau, D.P., Ziegler, A.C., Porter, S.D., Pope, L.M., 2004. Surface-water-quality conditions and relation to taste-and-odor occurrences in the Lake Olathe watershed, Northeast Kansas, 2000–02. Virginia.  
<https://doi.org/https://doi.org/10.3133/sir20045047>
- May, R.M., 1977. Thresholds and breakpoints in ecosystems with a multiplicity of states. *Nature* 269, 471–477.
- McDowall, B., Ho, L., Saint, C., Newcombe, G., 2007. Removal of geosmin and 2-methylisoborneol through biologically active sand filters. *Environ. Waste Manag.* 1, 311–320.
- Melo, N., Wolff, G.H., Costa-da-Silva, A.L., Arribas, R., Triana, M.F., Gugger, M., Riffell, J.A., DeGennaro, M., Stensmyr, M.C., 2019. Geosmin attracts *Aedes aegypti* mosquitoes to oviposition sites. *bioRxiv*. <https://doi.org/10.1101/598698>
- Menezes, C., Valério, E., Botelho, M.J., Dias, E., 2020. Isolation and Characterization of *Cylindrospermopsis raciborskii* Strains from Finished Drinking Water. *Toxins (Basel)*. 12, 1–13.  
<https://doi.org/10.3390/toxins12010040>

- Merino-Ibarra, M., Ramírez-Zierold, J.A., Valdespino-Castillo, P.M., Castillo-Sandoval, F.S., Guzmán-Arias, A.P., Barjau-Aguilar, M., Monroy-Ríos, E., López-Gómez, L.M., Sacristán-Ramírez, A., Quintanilla-Terminel, J.G., Zayas, R.G. De, Jimenez-Contreras, J., Valeriano-Riveros, M.E., Vilaclara-Fatjó, G., Sánchez-Carrillo, S., 2021. Vertical boundary mixing events during stratification govern heat and nutrient dynamics in a windy tropical reservoir lake with important water-level fluctuations: A long-term (2001–2021) study. *Water (Switzerland)* 13, 1–19. <https://doi.org/10.3390/w13213011>
- Millar, E.N., Kidd, K.A., Surette, M.G., Bennett, C.J., Salerno, J., Gillis, P.L., 2022. Effects of municipal wastewater effluents on the digestive gland microbiome of wild freshwater mussels (*Lasmigona costata*). *Ecotoxicol. Environ. Saf.* 241, 1–12. <https://doi.org/10.1016/j.ecoenv.2022.113774>
- Mohapatra, M., Behera, P., Kim, J.Y., Rastogi, G., 2020. Seasonal and spatial dynamics of bacterioplankton communities in a brackish water coastal lagoon. *Sci. Total Environ.* 705, 1–13. <https://doi.org/10.1016/j.scitotenv.2019.134729>
- Moi, D.A., Alves, D.C., Antigueira, P.A.P., Thomaz, S.M., Teixeira de Mello, F., Bonecker, C.C., Rodrigues, L.C., García-Ríos, R., Mormul, R.P., 2021. Ecosystem Shift from Submerged to Floating Plants Simplifying the Food Web in a Tropical Shallow Lake. *Ecosystems* 24, 628–639. <https://doi.org/10.1007/s10021-020-00539-y>
- Molot, L.A., Li, G., Findlay, D.L., Watson, S.B., 2010. Iron-mediated suppression of bloom-forming cyanobacteria by oxine in a eutrophic lake. *Freshw. Biol.* 55, 1102–1117. <https://doi.org/10.1111/j.1365-2427.2009.02384.x>
- Molot, L.A., Watson, S.B., Creed, I.F., Trick, C.G., McCabe, S.K., Verschoor, M.J., Sorichetti, R.J., Powe, C., Venkiteswaran, J.J., Schiff, S.L., 2014. A novel model for cyanobacteria bloom formation: The critical role of anoxia and ferrous iron. *Freshw. Biol.* 59, 1323–1340. <https://doi.org/10.1111/fwb.12334>
- Montgomery, B.L., 2014. The regulation of light sensing and light-harvesting impacts the use of cyanobacteria as biotechnology platforms. *Front. Bioeng. Biotechnol.* 2, 1–7. <https://doi.org/10.3389/fbioe.2014.00022>
- Moore, L.R., Huang, T., Ostrowski, M., Mazard, S., Kumar, S.S., Gamage, H.K.A.H., Brown, M. V, Messer, L.F., Seymour, J.R., Paulsen, I.T., 2019. Unicellular

- cyanobacteria are important components of phytoplankton communities in Australia's northern oceanic ecoregions. *Front. Microbiol.* 10, 1–16.  
<https://doi.org/10.3389/fmicb.2018.03356>
- Moradinejad, S., Trigui, H., Guerra Maldonado, J.F., Shapiro, J., Terrat, Y., Zamyadi, A., Dorner, S., Prévost, M., 2020. Diversity Assessment of Toxic Cyanobacterial Blooms during Oxidation. *Toxins (Basel)*. 12.  
<https://doi.org/10.3390/toxins12110728>
- Morgan-Kiss, R.M., Priscu, J.C., Pockock, T., Gudynaite-Savitch, L., Huner, N.P.A., 2006. Adaptation and Acclimation of Photosynthetic Microorganisms to Permanently Cold Environments. *Microbiol. Mol. Biol. Rev.* 70, 222–252.  
<https://doi.org/10.1128/mubr.70.1.222-252.2006>
- Morris, J.J., Johnson, Z.I., Szul, M.J., Keller, M., Zinser, E.R., 2011. Dependence of the cyanobacterium *Prochlorococcus* on hydrogen peroxide scavenging microbes for growth at the ocean's surface. *PLoS One* 6.  
<https://doi.org/10.1371/journal.pone.0016805>
- Morris, J.J., Lenski, R.E., Zinser, E.R., 2012. The black queen hypothesis: Evolution of dependencies through adaptive gene loss. *MBio* 3, 1–7.  
<https://doi.org/10.1128/mBio.00036-12>
- Morrison, J.A.E., 1997. Welsh Water's partnership approach to beating the drought in north Wales. *Water Environ. J.* 11, 377–380.
- Morton, S., Lee, T., 1974. Algal blooms – possible effects of iron. *Environ. Sci. Technol.* 8, 673–674.
- Mu, D.S., Wang, S., Liang, Q.Y., Du, Z.Z., Tian, R., Ouyang, Y., Wang, X.P., Zhou, A., Gong, Y., Chen, G.J., Van Nostrand, J., Yang, Y., Zhou, J., Du, Z.J., 2020. Bradymonabacteria, a novel bacterial predator group with versatile survival strategies in saline environments. *Microbiome* 8, 1–15.  
<https://doi.org/10.1186/s40168-020-00902-0>
- Mujakić, I., Andrei, A.-Ş., Shabarova, T., Fecskeová, L.K., Salcher, M.M., Piewosz, K., Ghai, R., Koblížek, M., 2021. Common Presence of Phototrophic Gemmatimonadota in Temperate Freshwater Lakes. *mSystems* 6, 1–19.  
<https://doi.org/10.1128/msystems.01241-20>

- Murphy, T.P., Lean, D.R.S., Nalewajko, C., 1976. Blue-green algae: their excretion of iron-selective chelators enables them to dominate other algae. *Science* (80-). 192, 900–902.
- Mustapha, S., Tijani, J.O., Ndamitso, M., Abdulkareem, A.S., Shuaib, D.T., Mohammed, A.K., 2021. A critical review on geosmin and 2-methylisoborneol in water: sources, effects, detection, and removal techniques. *Environ. Monit. Assess.* 193, 1–34. <https://doi.org/10.1007/s10661-021-08980-9>
- Muyzer, G., 1998. Denaturing gradient gel electrophoresis (DGGE) in microbial ecology, in: *Microbial Ecology Manual*. 3.4.4. Kluwer Academic Publishers.
- Muyzer, G., de Waal, E.C., Uitterlinden, A.G., 1993. Rapid Lipid Biomarker Analysis for Quantitative Assessment of Microbial Community Composition and Activity. *Appl. Environ. Microbiol.* 59, 695–700.
- Naes, H., Utkilen, H.C., Post, A.F., 1989. Geosmin production in the cyanobacterium *Oscillatoria brevis*. *Arch. Microbiol.* 151, 407–410. <https://doi.org/10.1007/BF00416598>
- Nagarajan, V., Tsai, H.C., Chen, J.S., Hussain, B., Fan, C.W., Asif, A., Hsu, B.M., 2022. The Evaluation of Bacterial Abundance and Functional Potentials in the Three Major Watersheds, Located in the Hot Spring Zone of the Tatun Volcano Group Basin, Taiwan. *Microorganisms* 10, 1–19. <https://doi.org/10.3390/microorganisms10030500>
- Narayan, L. V, Nunez, W.J., 1974. Biological control: isolation and bacterial oxidation of the taste and odour compound geosmin. *J. Am. Water Work. Assoc.* 66, 532–536.
- Natural Resources Wales, 2014. Dee River Basin District: Consultation on the draft Flood Risk Management Plan, Environment Agency. Bangor.
- New England Biolabs, 2020. NEBioCalculator [WWW Document]. dsDNA Mass to/from Moles Convert. URL <https://nebiocalculator.neb.com/#!/dsdnaamt> (accessed 12.9.20).
- Newitt, J.T., 2020. Roles and recruitment of *Streptomyces* species in the wheat root microbiome. University of East Anglia.
- Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., Bertilsson, S., Yamada, Y.,



- Kuzuyama, T., Komatsu, M., Shin-ya, K., Omura, S., Cane, D.E., Ikeda, H., Woodhouse, J.N., Ziegler, J., Grossart, H.P., Neilan, B.A., Adam, B., Klawonn, I., Svedén, J.B., Bergkvist, J., Nahar, N., Walve, J., Littmann, S., Whitehouse, M.J., Lavik, G., Kuypers, M.M.M., Ploug, H., Murchie, E.H., Lawson, T., Pattanaik, B., Lindberg, P., Woodhouse, J.N., Kinsela, A.S., Collins, R.N., Bowling, L.C., Honeyman, G.L., Holliday, J.K., Neilan, B.A., Flores, E., Herrero, A., Louati, I., Pascault, N., Debroas, D., Bernard, C., Humbert, J.F., Leloup, J., Parmar, A., Singh, N.K., Pandey, A., Gnansounou, E., Madamwar, D., Seymour, J.R., Amin, S.A., Raina, J.-B., Stocker, R., WHO, Zhang, J., Li, L., Qiu, L., Wang, X., Meng, X., You, Y., Yu, J., Ma, W., Journey, C.A., Beaulieu, K.M., Bradley, P.M., Salomon, P.S., Janson, S., Granéli, E., Dickschat, J.S., Bode, H.B., Mahmud, T., Müller, R., Schulz, S., Guttman, L., Rijn, J. Van, Liato, V., Aïder, M., Bell, W., Mitchell, R., Suurnäkki, S., Gomez-Saez, G. V., Rantala-Ylinen, A., Jokela, J., Fewer, D.P., Sivonen, K., 2015. Chemotactic and Growth Responses of Marine Bacteria to Algal Extracellular Products. *PLoS One* 10, 265–277. <https://doi.org/10.1371/journal.pone.0140614>
- Nguyen, N.H.A., Špánek, R., Kasalický, V., Ribas, D., Vlková, D., Řeháková, H., Kejzlar, P., Ševců, A., 2018. Different effects of nano-scale and micro-scale zero-valent iron particles on planktonic microorganisms from natural reservoir water. *Environ. Sci. Nano* 5, 1117–1129. <https://doi.org/10.1039/c7en01120b>
- Niu, Y., Shen, H., Chen, J., Xie, P., Yang, X., Tao, M., Ma, Z., Qi, M., 2011. Phytoplankton community succession shaping bacterioplankton community composition in Lake Taihu, China. *Water Res.* 45, 4169–4182. <https://doi.org/10.1016/j.watres.2011.05.022>
- Noy-Meir, I., 1975. Stability of grazing systems: an application of predator-prey graphs. *J. Ecol.* 63, 459–481.
- Numberger, D., Zoccarato, L., Woodhouse, J., Ganzert, L., Sauer, S., Márquez, J.R.G., Domisch, S., Grossart, H.P., Greenwood, A.D., 2022. Urbanization promotes specific bacteria in freshwater microbiomes including potential pathogens. *Sci. Total Environ.* 845, 1–13. <https://doi.org/10.1016/j.scitotenv.2022.157321>
- Nurhasanah, S., Papuangan, N., 2019. Amplification and Analysis of Rbcl Gene

(Ribulose-1,5-Bisphosphate Carboxylase) of Clove in Ternate Island, in: IOP Conference Series: Earth and Environmental Science.

<https://doi.org/10.1088/1755-1315/276/1/012061>

Ogura, T., Sunairi, M., Nakajima, M., 2000. 2-Methylisoborneol and Geosmin, the Main Sources of Soil Odor, Inhibit the Germination of Brassicaceae Seeds. *Soil Sci. Plant Nutr.* 46, 217–227. <https://doi.org/10.1080/00380768.2000.10408777>

Oh, H.S., Lee, C.S., Srivastava, A., Oh, H.M., Ahn, C.Y., 2017. Effects of environmental factors on cyanobacterial production of odorous compounds: Geosmin and 2-methylisoborneol. *J. Microbiol. Biotechnol.* <https://doi.org/10.4014/jmb.1702.02069>

Okazaki, Y., Fujinaga, S., Tanaka, A., Kohzu, A., Oyagi, H., Nakano, S.I., 2017. Ubiquity and quantitative significance of bacterioplankton lineages inhabiting the oxygenated hypolimnion of deep freshwater lakes. *ISME J.* 11, 2279–2293. <https://doi.org/10.1038/ismej.2017.89>

Oklahoma Water Resources Board, 2005. Justification for chlorophyll-a criteria to protect the public and private water supply beneficial use of sensitive water supplies.

Oklahoma Water Resources Board, 2002. Water Quality Evaluation of the Eucha/Spavinaw Lake System. Oklahoma city, OK.

Oksanen, J., 2013. Vegan : ecological diversity. *R Proj.* 368, 1–11.

Olapade, O.A., 2018. Community Composition and Diversity of Coastal Bacterioplankton Assemblages in Lakes Michigan, Erie, and Huron. *Microb. Ecol.* 75, 598–608. <https://doi.org/10.1007/s00248-017-1081-0>

Olsen, B.K., Chislock, M.F., Rebelein, A., Wilson, A.E., 2017. Nutrient enrichment and vertical mixing mediate 2-methylisoborneol and geosmin concentrations in a drinking water reservoir. *Water Sci. Technol. Water Supply* 17, 500–507. <https://doi.org/10.2166/ws.2016.159>

Olsen, B.K., Chislock, M.F., Wilson, A.E., 2016. Eutrophication mediates a common off-flavor compound, 2-methylisoborneol, in a drinking water reservoir. *Water Res.* <https://doi.org/10.1016/j.watres.2016.01.058>

Ömür-Özbek, P., Little, J.C., Dietrich, A.M., 2007. Ability of humans to smell

- geosmin, 2-MIB and nonadienal in indoor air when using contaminated drinking water. *Water Sci. Technol.* 55, 249–256. <https://doi.org/10.2166/wst.2007.186>
- Örnólfsson, E.B., Lumsden, E., Pinckney, J.L., 2004. Phytoplankton community growth-rate response to nutrient pulses in a shallow turbid estuary, Galveston Bay, Texas. *J. Plankton Res.* 26, 325–339. <https://doi.org/10.1093/plankt/fbh035>
- Otten, T.G., Graham, J.L., Harris, T.D., Dreher, T.W., 2016. Elucidation of taste and odor-producing bacteria and toxigenic cyanobacteria in a Midwestern drinking water supply reservoir by shotgun metagenomic analysis. *Appl. Environ. Microbiol.* 82, 5410–5420. <https://doi.org/10.1128/AEM.01334-16>
- Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J.A., McQuisten, K.A., Almabrazi, H.G., Pedersen, K.F., Lin, Y., Garretson, J., McEntaggart, N.O., Sailor, C.A., Dawson, R.B., Peek, A.S., 2008. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.* 36, 163–169. <https://doi.org/10.1093/nar/gkn198>
- Paerl, H., Fulton, R., Moisander, P., Dyble, J., 2001. Harmful Freshwater Algal Blooms, With an Emphasis on Cyanobacteria. *The Scientific World* 1:76-113. *ScientificWorldJournal.* 1, 76–113. <https://doi.org/10.1100/tsw.2001.16>
- Paerl, H.W., Fulton, R.S., 2006. Ecology of Harmful Cyanobacteria, in: Granéli, E., Turner, J.T. (Eds.), *Ecology of Harmful Algae*. Springer Berlin Heidelberg, pp. 95–109. [https://doi.org/10.1007/978-3-540-32210-8\\_8](https://doi.org/10.1007/978-3-540-32210-8_8)
- Paerl, H.W., Kellar, P.E., 1978. Significance of bacterial (cyanophyceae) Anabaena associations with respect to N<sub>2</sub> fixation in freshwater. *J. Phycol.* 14, 254–260.
- Paerl, H.W., Otten, T.G., 2016. Duelling “CyanoHABs”: Unravelling the environmental drivers controlling dominance and succession among diazotrophic and non-N<sub>2</sub>-fixing harmful cyanobacteria. *Environ. Microbiol.* <https://doi.org/10.1111/1462-2920.13035>
- Paerl, H.W., Otten, T.G., 2013. Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls. *Microb. Ecol.* 65, 995–1010. <https://doi.org/10.1007/s00248-012-0159-y>
- Paerl, H.W., Pinckney, J.L., 1996. A mini-review of microbial consortia: Their roles in aquatic production and biogeochemical cycling. *Microb. Ecol.* 31, 225–247.

<https://doi.org/10.1007/BF00171569>

- Paerl, R., Huang, H., Ehrlich, L., 2022. Investigating the microbial culprits of Taste/Odour issues in city of Durham drinking water reservoir lake michie and algicidal mitigation tactics. North Carolina.
- Páez, R., Ruiz, G., Márquez, R., Soto, L.M., Montiel, M., López, C., 2001. Limnological studies on a shallow reservoir in Western Venezuela (Tulé Reservoir). *Limnologica* 31, 139–145. [https://doi.org/10.1016/S0075-9511\(01\)80008-6](https://doi.org/10.1016/S0075-9511(01)80008-6)
- Parinet, J., Rodriguez, M.J., Sérodes, J., 2010. Influence of water quality on the presence of off-flavour compounds (geosmin and 2-methylisoborneol). *Water Res.* 44, 5847–5856. <https://doi.org/10.1016/j.watres.2010.06.070>
- Parmar, A., Singh, N.K., Pandey, A., Gnansounou, E., Madamwar, D., 2011. Cyanobacteria and microalgae: A positive prospect for biofuels. *Bioresour. Technol.* 102, 10163–10172. <https://doi.org/10.1016/j.biortech.2011.08.030>
- Parr, M., Smith, R., 1976. The identification of phosphorus as a growth-limiting nutrient in Lough Neagh using bioassays. *Water Res.* 10, 1151–1154.
- Parveen, B., Ravet, V., Djediat, I.M., Quiblier, C., Debroas, D., Humbert, J.F., 2013. Bacterial communities associated with *Microcystis* colonies differ from free-living communities living in the same ecosystem. *Environ. Microbiol. Rep.* 5, 716–724. <https://doi.org/doi:10.1111/1758-2229.12071>
- Pasciak, W.J., Gavis, J., 1974. Transport limitation of nutrient uptake in phytoplankton. *Science* (80-. ). 19, 881–886.
- Paster, B.J., Russell, J.B., Yang, C.M.J., Chow, J.M., 1993. Phylogeny of the Ammonia-Producing Ruminant Bacteria. *Int. J. Syst. Bacteriol.* 43, 107–110.
- Pattanaik, B., Lindberg, P., 2015. Terpenoids and Their Biosynthesis in Cyanobacteria. *Life* 5, 269–293. <https://doi.org/10.1109/MCOM.2002.995859>
- Pattanaik, B., Whitaker, M.J., Montgomery, B.L., 2012. Light quantity affects the regulation of cell shape in *Fremyella diplosiphon*. *Front. Microbiol.* 3, 1–6. <https://doi.org/10.3389/fmicb.2012.00170>
- Paver, S.F., Hayek, K.R., Gano, K.A., Fagen, J.R., Brown, C.T., Davis-Richardson,

- A.G., Crabb, D.B., Rosario-Passapera, R., Giongo, A., Triplett, E.W., Kent, A.D., 2013. Interactions between specific phytoplankton and bacteria affect lake bacterial community succession. *Environ. Microbiol.* 15, 2489–2504.  
<https://doi.org/10.1111/1462-2920.12131>
- Pérez-Pantoja, D., Donoso, R., Agulló, L., Córdova, M., Seeger, M., Pieper, D.H., González, B., 2012. Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales. *Environ. Microbiol.* 14, 1091–1117.  
<https://doi.org/10.1111/j.1462-2920.2011.02613.x>
- Pérez, A.A., Rodionov, D.A., Bryant, D.A., 2016. Identification and regulation of genes for cobalamin transport in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *J. Bacteriol.* 198, 2753–2761. <https://doi.org/10.1128/JB.00476-16>
- Perkins, R.G., Slavin, E.I., Andrade, T.M.C., Blenkinsopp, C., Pearson, P., Froggatt, T., Godwin, G., Parslow, J., Hurley, S., Luckwell, R., Wain, D.J., 2019. Managing taste and odour metabolite production in drinking water reservoirs: The importance of ammonium as a key nutrient trigger. *J. Environ. Manage.* 244, 276–284. <https://doi.org/10.1016/j.jenvman.2019.04.123>
- Peter, A., Köster, O., Schildknecht, A., von Gunten, U., 2009. Occurrence of dissolved and particle-bound taste and odor compounds in Swiss lake waters. *Water Res.* 43, 2191–2200. <https://doi.org/10.1016/j.watres.2009.02.016>
- Pham, N.D., Lee, E.H., Chae, S.H., Cho, Y., Shin, H., Son, A., 2015. Bacterial community structure shifted by geosmin in granular activated carbon system of water treatment plants. *J. Microbiol. Biotechnol.* 26, 80–88.  
<https://doi.org/10.4014/jmb.1506.06033>
- Plant Working Group, C., 2009. A DNA barcode for land plants, *Proceedings of the National Academy of Sciences, USA*.
- Ploug, H., Adam, B., Musat, N., Kalvelage, T., Lavik, G., Wolf-Gladrow, D., Kuypers, M.M.M., 2011. Carbon, nitrogen and O<sub>2</sub> fluxes associated with the cyanobacterium *Nodularia spumigena* in the Baltic Sea. *ISME J.* 5, 1549–1558.  
<https://doi.org/10.1038/ismej.2011.20>
- Pochiraju, S., Hoppe-Jones, C., Adams, C., Weinrich, L., 2021. Development and optimization of analytical methods for the detection of 18 taste and odor

- compounds in drinking water utilities. *Water Res.* X 11, 1–9.  
<https://doi.org/10.1016/J.WROA.2021.100099>
- Prasad, A.G.D., Siddaraju, P., 2012. Carlson's Trophic State Index for the assessment of trophic status of two Lakes in Mandya district. *Adv. Appl. Sci. Res.* 3, 2992–2996.
- Pushpakumara, B.L.D.U., Tandon, K., Willis, A., Verbruggen, H., 2022. Unravelling microalgal-bacterial interactions in aquatic ecosystems through 16S co-occurrence networks. *bioRxiv*.
- Qiu, Y.L., Kuang, X.Z., Shi, X.S., Yuan, X.Z., Guo, R.B., 2014. *Terrimicrobium sacchariphilum* gen. nov., sp. nov., an anaerobic bacterium of the class "Spartobacteria" in the phylum Verrucomicrobia, isolated from a rice paddy field. *Int. J. Syst. Evol. Microbiol.* 64, 1718–1723. <https://doi.org/10.1099/ijs.0.060244-0>
- Rajaneesh, K.M., Naik, R.K., Roy, R., Costa, P.M.D., 2020. Cyanobacteria in tropical and subtropical marine environments: bloom formation and ecological role, in: *Advances in Cyanobacterial Biology*. Academic Press, pp. 35–46.
- Rajaniemi-Wacklin, P., Rantala, A., Mugnai, M.A., Turicchia, S., Ventura, S., Komárková, J., Lepistö, L., Sivonen, K., 2006. Correspondence between phylogeny and morphology of *Snowella* spp. and *Woronichinia naegeliana*, cyanobacteria commonly occurring in lakes. *J. Phycol.* 42, 226–232.  
<https://doi.org/10.1111/j.1529-8817.2006.00179.x>
- Ramanan, R., Kim, B.H., Cho, D.H., Oh, H.M., Kim, H.S., 2016. Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnol. Adv.* 34, 14–29. <https://doi.org/10.1016/j.biotechadv.2015.12.003>
- Raven, J.A., 2010. Cyanotoxins: A poison that frees phosphate. *Curr. Biol.* 20, R850–R852. <https://doi.org/10.1016/j.cub.2010.08.012>
- Redfield, A.C., 1958. The biological control of chemical factors in the environment. *Am. Sci.* 16, 1–221.
- Reynolds, C.S., 1980. Phytoplankton assemblages and their periodicity in stratifying lake systems. *Ecography (Cop.)*. 3, 141–159. <https://doi.org/10.1111/j.1600-0587.1980.tb00721.x>

- Reynolds, C.S., Huszar, V., Kruk, C., Naselli-Flores, L., Melo, S., 2002. Towards a functional classification of the freshwater phytoplankton. *J. Plankton Res.* 24, 417–428. <https://doi.org/10.1093/plankt/24.5.417>
- Rico-Martínez, R., Arias-Almeida, J.C., Pérez-Legaspi, I.A., Alvarado-Flores, J., Retes-Pruneda, J.L., 2012. Adverse Effects of Herbicides on Freshwater Zooplankton, in: Hasaneen, M.N. (Ed.), *Herbicides - Properties, Synthesis and Control of Weeds*. IntechOpen, pp. 405–437. <https://doi.org/10.5772/2511>
- Rippka, R., 1988. Isolation and purification of cyanobacteria. *Methods Enzymol.* 167, 3–27. [https://doi.org/10.1016/0076-6879\(88\)67004-2](https://doi.org/10.1016/0076-6879(88)67004-2)
- Robarts, R.D., Zohary, T., 1987. Temperature effects on photosynthetic capacity , respiration , and growth rates of bloom - forming cyanobacteria. *New Zeal. J. Mar. Freshw. Res.* 21, 391–399. <https://doi.org/10.1080/00288330.1987.9516235>
- Rodriguez, A.J., 2018. Evaluation Of Powdered Activated Carbon For Control Of Taste And Odors Caused By Geosmin. University of Texas at El Paso.
- Rooney-Varga, J.N., Giewat, M.W., Savin, M.C., Sood, S., Legresley, M., Martin, J.L., 2005. Links between phytoplankton and bacterial community dynamics in a coastal marine environment. *Microb. Ecol.* 49, 163–175. <https://doi.org/10.1007/s00248-003-1057-0>
- Rosen, B.H., MacLeod, B.W., Simpson, M.R., 1992. Accumulation and release of geosmin during the growth phases of *Anabaena circinalis* (Kutz.) Rabenhorst. *Water Sci. Technol.* 25, 185–190.
- Rühland, K.M., Paterson, A.M., Smol, J.P., 2015. Lake diatom responses to warming: reviewing the evidence. *J. Paleolimnol.* 54, 1–35. <https://doi.org/10.1007/s10933-015-9837-3>
- Rühland, K.M., Smol, J.P., Reinhard, P., 2003. Ecology and spatial distributions of surface-sediment diatoms from 77 lakes in the subarctic Canadian treeline region. *Can. J. Bot.* 81, 57–73. <https://doi.org/10.1139/b03-005>
- Ruppert, K.M., Kline, R.J., Rahman, M.S., 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review

- in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* 17, 1–29. <https://doi.org/10.1016/j.gecco.2019.e00547>
- Saadoun, I.M.K., Schrader, K.K., Blevins, W.T., 2001. Environmental and nutritional factors affecting geosmin synthesis by *Anabaena* sp. *Water Res.* 35, 1209–1218. [https://doi.org/10.1016/S0043-1354\(00\)00381-X](https://doi.org/10.1016/S0043-1354(00)00381-X)
- Sabater, S., Vilalta, E., Gaudes, A., Guasch, H., Muñoz, I., Romani, A., 2003. Ecological implications of mass growth of benthic. *Aquat. Microb. Ecol.* 32, 175–184.
- Saini, J.S., Hassler, C., Cable, R., Fourquez, M., Danza, F., Roman, S., Tonolla, M., Storelli, N., Jacquet, S., Zdobnov, E.M., Duhaime, M.B., 2022. Bacterial, Phytoplankton, and Viral Distributions and Their Biogeochemical Contexts in Meromictic Lake Cadagno Offer Insights into the Proterozoic Ocean Microbial Loop. *MBio* 13, 1–20. <https://doi.org/10.1128/mbio.00052-22>
- Salmaso, N., Albanese, D., Capelli, C., Boscaini, A., Pindo, M., Donati, C., 2018. Diversity and Cyclical Seasonal Transitions in the Bacterial Community in a Large and Deep Perialpine Lake. *Microb. Ecol.* 76, 125–143. <https://doi.org/10.1007/s00248-017-1120-x>
- Salomon, P.S., Janson, S., Granéli, E., 2003. Molecular identification of bacteria associated with filaments of *Nodularia spumigena* and their effect on the cyanobacterial growth. *Harmful Algae* 2, 261–272. [https://doi.org/10.1016/S1568-9883\(03\)00045-3](https://doi.org/10.1016/S1568-9883(03)00045-3)
- Santi, I., Kasapidis, P., Karakassis, I., Pitta, P., 2021. A comparison of DNA metabarcoding and microscopy methodologies for the study of aquatic microbial eukaryotes. *Diversity* 13, 1–12. <https://doi.org/10.3390/d13050180>
- Santoferrara, L.F., 2019. Current practice in plankton metabarcoding: Optimization and error management. *J. Plankton Res.* 41, 571–582. <https://doi.org/10.1093/plankt/fbz041>
- Satjarak, A., Graham, L.E., Piotrowski, M.J., Trest, M.T., Wilcox, L.W., Cook, M.E., Knack, J.J., Arancibia-Avila, P., 2021. Shotgun metagenomics and microscopy indicate diverse cyanophytes, other bacteria, and microeukaryotes in the epimicrobiota of a northern Chilean wetland *Nostoc* (Cyanobacteria). *J. Phycol.*



- 57, 39–50. <https://doi.org/10.1111/jpy.13084>
- Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., 2001. Catastrophic shifts in ecosystems. *Nature* 413, 591–596. <https://doi.org/10.1038/35098000>
- Scheffer, M., Rinaldi, S., Gragnani, A., Mur, L.R., Nes, E.H. Van, Nes, E.H.V.A.N., 1997. On the Dominance of Filamentous Cyanobacteria in Shallow , Turbid Lakes. *Ecology* 78, 272–282.
- Schindler, D.W., 2012. The dilemma of controlling cultural eutrophication of lakes. *Proc. R. Soc. B Biol. Sci.* 279, 4322–4333. <https://doi.org/10.1098/rspb.2012.1032>
- Schindler, D.W., 1977. Evolution of phosphorus limitation in lakes. *Science* (80-. ). 195, 260–262. <https://doi.org/10.1126/science.195.4275.260>
- Schindler, D.W., 1974. Eutrophication and recovery in experimental lakes: Implications for lake management. *Science* (80-. ). 184, 897–899. <https://doi.org/10.1126/science.184.4139.897>
- Schindler, D.W., Hecky, R.E., Findlay, M.P., Stainton, M.P., Parker, B.R., Paterson, K.G., Beaty, K.G., Lyng, M., Kasian, S.E.M., 2008. Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. *PNAS* 105, 11254–11258. <https://doi.org/https://doi.org/10.1073/pnas.0805108105>
- Schlumpberger, B.O., Jux, A., Kunert, M., Boland, W., Wittmann, D., 2004. Musty-earthy scent in cactus flowers: Characteristics of floral scent production in dehydrogeosmin-producing cacti. *Int. J. Plant Sci.* 165, 1007–1015. <https://doi.org/10.1086/423878>
- Schmidt, M.L., White, J.D., Deneff, V.J., 2016. Phylogenetic conservation of freshwater lake habitat preference varies between abundant bacterioplankton phyla. *Environ. Microbiol.* 18, 1212–1226. <https://doi.org/10.1111/1462-2920.13143>
- Schrader, K.K., Davidson, J.W., Summerfelt, S.T., 2013. Aquacultural Engineering Evaluation of the impact of nitrate-nitrogen levels in recirculating aquaculture systems on concentrations of the off-flavor compounds geosmin and 2-methylisoborneol in water and rainbow trout (*Oncorhynchus mykiss*). *Aquac.*

- Eng. 57, 126–130. <https://doi.org/10.1016/j.aquaeng.2013.07.002>
- Schrader, K.K., Green, B.W., Perschbacher, P.W., 2011. Development of phytoplankton communities and common off-flavors in a biofloc technology system used for the culture of channel catfish (*Ictalurus punctatus*). *Aquac. Eng.* 45, 118–126. <https://doi.org/10.1016/j.aquaeng.2011.08.004>
- Sechi, G.M., Sulis, A., 2007. Multi-reservoir system optimization using chlorophyll-a trophic indexes. *Water Resour. Manag.* 21, 849–860. <https://doi.org/10.1007/s11269-006-9114-3>
- Sethuraman, A., Stancheva, R., Sanders, C., Caceres, L., Castro, D., Hausknecht-Buss, H., Henry, S., Johansen, H., Kasler, A., Lastor, S., Massaro, I., Mekuria, I., Moron-Solano, A., Read, N., Vengerova, G., Zhang, A., Zhang, X., Read, B., 2022. Genome of a novel *Sediminibacterium* discovered in association with two species of freshwater cyanobacteria from streams in Southern California. *G3 Genes, Genomes, Genet.* 12, 1–8. <https://doi.org/10.1093/g3journal/jkac123>
- Seto, H., Watanabe, H., Kazuo, F., 1996. Simultaneous Operation of the Mevalonate and Non-mevalonate. *Tetrahedron Lett.* 37, 7979–7982.
- Shao, X., Du, K., 2020. Biodegradation of 2-methylisoborneol by enzyme separated from *Pseudomonas mandelii*. *Water Sci. Technol. Water Supply* 20, 2096–2105. <https://doi.org/10.2166/ws.2020.100>
- Sharma, M., Khurana, H., Singh, D.N., Negi, R.K., 2021. The genus *Sphingopyxis*: Systematics, ecology, and bioremediation potential - A review. *J. Environ. Manage.* 280, 111744. <https://doi.org/10.1016/j.jenvman.2020.111744>
- Sharma, M.P., Kumar, A., Rajvanshi, S., 2010. Case of Mansi Ganga Lake in India. *Hydro Nepal J. Water, Energy Environ.* 6, 65–72.
- Shatwell, T., Köhler, J., 2019. Decreased nitrogen loading controls summer cyanobacterial blooms without promoting nitrogen-fixing taxa: Long-term response of a shallow lake. *Limnol. Oceanogr.* 64, S166–S178. <https://doi.org/10.1002/lno.11002>
- Sheela, A.M., Letha, J., Joseph, S., 2011. Environmental status of a tropical lake system. *Environ. Monit. Assess.* 180, 427–449. <https://doi.org/10.1007/s10661-010-1797-5>

- Shen, H., Niu, Y., Xie, P., Tao, M., Yang, X., 2011. Morphological and physiological changes in *Microcystis aeruginosa* as a result of interactions with heterotrophic bacteria. *Freshw. Biol.* 56, 1065–1080. <https://doi.org/10.1111/j.1365-2427.2010.02551.x>
- Shen, Q., Shimizu, K., Miao, H., Tsukino, S., Utsumi, M., Lei, Z., Zhang, Z., Nishimura, O., Asada, Y., Fujimoto, N., Takanashi, H., Akiba, M., 2021. Effects of elevated nitrogen on the growth and geosmin productivity of *Dolichospermum smithii*. *Environ. Sci. Pollut. Res.* 28, 177–184. <https://doi.org/10.1007/s11356-020-10429-4>
- Shen, Q., Wang, Q., Miao, H., Shimada, M., Utsumi, M., Lei, Z., Zhang, Z., Nishimura, O., Asada, Y., Fujimoto, N., Takanashi, H., Akiba, M., Shimizu, K., 2022. Temperature affects growth, geosmin/2-methylisoborneol production, and gene expression in two cyanobacterial species. *Environ. Sci. Pollut. Res.* 29, 12017–12026. <https://doi.org/10.1007/s11356-021-16593-5>
- Shi, L., Cai, Y., Kong, F., Yu, Y., 2012. Specific association between bacteria and buoyant *Microcystis* colonies compared with other bulk bacterial communities in the eutrophic Lake Taihu, China. *Environ. Microbiol. Rep.* 4, 669–678. <https://doi.org/10.1111/1758-2229.12001>
- Shi, L., Cai, Y., Shi, X., Zhang, M., Zeng, Q., Kong, F., Xu, P., 2022. Community structure of aerobic anoxygenic phototrophic bacteria in algae- and macrophyte-dominated areas in Taihu Lake, China. *J. Oceanol. Limnol.* 40, 1855–1867. <https://doi.org/10.1007/s00343-022-1348-2>
- Shi, L., Cai, Y., Yang, H., Xing, P., Li, P., Kong, L., Kong, F., 2009. Phylogenetic diversity and specificity of bacteria associated with *Microcystis aeruginosa* and other cyanobacteria. *J. Environ. Sci.* 21, 1581–1590. [https://doi.org/10.1016/S1001-0742\(08\)62459-6](https://doi.org/10.1016/S1001-0742(08)62459-6)
- Shimizu, K., Del Amo, Y., Brzezinski, M.A., Stucky, G.D., Morse, D.E., 2001. A novel fluorescent silica tracer for biological silicification studies. *Chem. Biol.* 8, 1051–1060. [https://doi.org/10.1016/S1074-5521\(01\)00072-2](https://doi.org/10.1016/S1074-5521(01)00072-2)
- Siedt, M., Schäffer, A., Smith, K.E.C., Nabel, M., Roß-Nickoll, M., van Dongen, J.T., 2021. Comparing straw, compost, and biochar regarding their suitability as agricultural soil amendments to affect soil structure, nutrient leaching, microbial

- communities, and the fate of pesticides. *Sci. Total Environ.* 751, 1–19.  
<https://doi.org/10.1016/j.scitotenv.2020.141607>
- Šimek, K., Kasalický, V., Jezbera, J., Horňák, K., Nedoma, J., Hahn, M.W., Bass, D., Jost, S., Boenigk, J., 2013. Differential freshwater flagellate community response to bacterial food quality with a focus on Limnohabitans bacteria. *ISME J.* 7, 1519–1530. <https://doi.org/10.1038/ismej.2013.57>
- Simiyu, B.M., Kurmayer, R., 2022. Response of planktonic diatoms to eutrophication in Nyanza Gulf of Lake Victoria, Kenya. *Limnologica* 93, 1–12.  
<https://doi.org/10.1016/j.limno.2022.125958>
- Simon, M., Grossart, H., Schweitzer, B., Ploug, H., 2002. Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat. Microb. Ecol.* 28, 175–211.
- Simpson, M.R., MacLeod, B.W., 1991. Comparison of various powdered activated carbons for the removal of geosmin and 2-methylisoborneol in selected water conditions, in: *Proceedings of American Water Works Association Annual Conference*. Philadelphia PA.
- Sinang, S.C., Reichwaldt, E.S., Ghadouani, A., 2015. Local nutrient regimes determine site-specific environmental triggers of cyanobacterial and microcystin variability in urban lakes. *Hydrol. Earth Syst. Sci.* 19, 2179–2195.  
<https://doi.org/10.5194/hess-19-2179-2015>
- Sisman-Aydin, G., Simsek, K., 2022. Investigation of the Phycoremediation Potential of Freshwater Green Algae *Golenkinia radiata* for Municipal Wastewater. *Sustainability* 14. <https://doi.org/10.3390/su142315705>
- Slavin, E., 2020. Using artificial circulation for in-reservoir management of cyanobacteria and taste and odour metabolite production. University of Bath.
- Śliwińska-Wilczewska, S., Maculewicz, J., Felpeto, A.B., Latała, A., 2018. Allelopathic and bloom-forming picocyanobacteria in a changing world. *Toxins (Basel)*. 10, 1–20. <https://doi.org/10.3390/toxins10010048>
- Smith, V.H., 1983. Low nitrogen to phosphorus ratios favor dominance by blue-green algae in lake phytoplankton. *Science (80-. )*. 221, 669–671.
- Smith, V.H., Sieber-Denlinger, J., DeNoyelles, F., Campbell, S., Pan, S., Randtke, S.J., Blain, G.T., Strasser, V.A., 2002. Managing taste and odor problems in a

- eutrophic drinking water reservoir. *Lake Reserv. Manag.* 18, 319–323.  
<https://doi.org/10.1080/07438140209353938>
- Sohm, J.A., Webb, E.A., Capone, D.G., 2011. Emerging patterns of marine nitrogen fixation. *Nat. Rev. Microbiol.* 9, 499–508. <https://doi.org/10.1038/nrmicro2594>
- Søndergaard, M., 2009. Redox Potential. *Encycl. Inl. Waters* 852–859.  
<https://doi.org/10.1016/B978-012370626-3.00115-0>
- Sorichetti, R.J., Creed, I.F., Trick, C.G., 2014. The influence of iron, siderophores and refractory DOM on cyanobacterial biomass in oligotrophic lakes. *Freshw. Biol.* 59, 1423–1436. <https://doi.org/10.1111/fwb.12355>
- Sotero-Martins, A., Carvajal, E., Albuquerque dos Santos, J.A., Moura, P.G., Handam, N.B., Kotowski Filho, N.P., Jardim, R., Ferrão Filho, A. da S., 2021. Events linked to Geosmin and 2-methylisoborneol (2-MIB) in a Water Supply in the State of Rio de Janeiro, Brazil: a case study. *Int. J. Hydrol.* 5, 214–220.  
<https://doi.org/10.15406/ijh.2021.05.00283>
- Souffreau, C., Verbruggen, H., Wolfe, A.P., Vanormelingen, P., Siver, P.A., Cox, E.J., Mann, D.G., Vijver, B. Van de, Sabbe, K., Vyverman, W., 2011. A time-calibrated multi-gene phylogeny of the diatom genus *Pinnularia*. *Mol. Phylogenet. Evol.* 61, 866–879. <https://doi.org/10.1016/j.ympev.2011.08.031>
- Soule, T., Garcia-Pichel, F., 2019. Cyanobacteria. *Encycl. Microbiol.* 4, 799–917.  
<https://doi.org/10.1016/B978-0-12-809633-8.20886-8>
- Srinivasan, R., Sorial, G.A., 2011. Treatment of taste and odor causing compounds 2-methyl isoborneol and geosmin in drinking water : A critical review. *J. Environ. Sci.* 23, 1–13. [https://doi.org/10.1016/S1001-0742\(10\)60367-1](https://doi.org/10.1016/S1001-0742(10)60367-1)
- Stancheva, R., 2019. *Planothidium sheathii*, a new monoraphid diatom species from rivers in California, USA. *Phytotaxa* 393, 131–140.  
<https://doi.org/10.11646/phytotaxa.393.2.4>
- Steffen, M.M., Li, Z., Effler, T.C., Hauser, L.J., Boyer, G.L., Wilhelm, S.W., 2012. Comparative Metagenomics of Toxic Freshwater Cyanobacteria Bloom Communities on Two Continents. *PLoS One* 7, 1–9.  
<https://doi.org/10.1371/journal.pone.0044002>
- Stoermer, E.F., Julius, M.L., 2003. Centric Diatoms, in: Wehr, J.D., Sheath, R.G.

(Eds.), *Freshwater Algae of North America - Ecology and Classification*. Elsevier Inc., pp. 559–594. <https://doi.org/https://doi.org/10.1016/B978-0-12-741550-5.X5000-4>

Strait, J.M., 2015. Investigation of the performance of biologically-active GAC filters for Taste and Odor removal. University of Minnesota.

Štrojsová, A., Vrba, J., Nedoma, J., Komárková, J., Znachor, P., 2003. Seasonal study of extracellular phosphatase expression in the phytoplankton of a eutrophic reservoir. *Eur. J. Phycol.* 38, 295–306. <https://doi.org/10.1080/09670260310001612628>

Stumpf, R.P., Wynne, T.T., Baker, D.B., Fahnenstiel, G.L., 2012. Interannual Variability of Cyanobacterial Blooms in Lake Erie. *PLoS One* 7, 1–11. <https://doi.org/10.1371/journal.pone.0042444>

Su, M., Andersen, T., Burch, M., Jia, Z., An, W., Yu, J., Yang, M., 2019. Succession and interaction of surface and subsurface cyanobacterial blooms in oligotrophic/mesotrophic reservoirs: A case study in Miyun Reservoir. *Sci. Total Environ.* 649, 1553–1562. <https://doi.org/10.1016/j.scitotenv.2018.08.307>

Su, M., Gaget, V., Giglio, S., Burch, M., An, W., Yang, M., 2013. Establishment of quantitative PCR methods for the quantification of geosmin-producing potential and *Anabaena* sp. in freshwater systems. *Water Res.* 47, 3444–3454. <https://doi.org/10.1016/j.watres.2013.03.043>

Subhash, Y., Sasikala, C., Ramana, C. V., 2014. *Sphingopyxis contaminans* sp. nov., isolated from a contaminated Petri dish. *Int. J. Syst. Evol. Microbiol.* 64, 2238–2243. <https://doi.org/10.1099/ijs.0.061937-0>

Sugiura, N., Iwami, N., Inamori, Y., Nishimura, O., Sudo, R., 1998. Significance of attached cyanobacteria relevant to the occurrence of musty odor in Lake Kasumigaura. *Water Res.* 32, 3549–3554. [https://doi.org/10.1016/S0043-1354\(98\)00153-5](https://doi.org/10.1016/S0043-1354(98)00153-5)

Sugiura, N., Utsumi, M., Wei, B., Iwami, N., Okano, K., Kawauchi, Y., Maekawa, T., 2004. Assessment for the complicated occurrence of nuisance odours from phytoplankton and environmental factors in a eutrophic lake. *Lakes Reserv. Res. Manag.* 9, 195–201. <https://doi.org/10.1111/j.1440-1770.2004.00246.x>

- Sumitomo, H., 1988. Odor Decomposition by the Yeast *Candida*. *Water Sci. Technol.* 20, 1157–162.
- Suurnäkki, S., Gomez-Saez, G. V., Rantala-Ylinen, A., Jokela, J., Fewer, D.P., Sivonen, K., 2015. Identification of geosmin and 2-methylisoborneol in cyanobacteria and molecular detection methods for the producers of these compounds. *Water Res.* 68, 56–66. <https://doi.org/10.1016/j.watres.2014.09.037>
- Takamura, N., Iwakuma, T., Yasuno, M., 1985. Photosynthesis and primary production of *Microcystis aeruginosa* Kütz. in Lake Kasumigaura. *J. Plankton Res.* 7, 303–312. <https://doi.org/10.1093/plankt/7.3.303>
- Tamura, T., 2014. The Family Sporichthyaceae, in: Rosenberg, E., DeLong, E.F., Stackebrandt, E., Thompson, F. (Eds.), *The Prokaryotes*. Springer, Berlin, Heidelberg, pp. 883–888. [https://doi.org/10.1007/978-3-642-30138-4\\_182](https://doi.org/10.1007/978-3-642-30138-4_182)
- Tanaka, A., Oritani, F., Uehara, A., Saito, A., Kishita, H., Niizeki, Y., Yokota, H., Fuchigami, K., 1996. Biodegradation of a musty odour component, 2-methylisoborneol. *Water Res.* 30, 759–761.
- Tanaka, D., Takahashi, T., Yamashiro, Y., Tanaka, H., Kimochi, Y., Nishio, M., Sakatoku, A., Nakamura, S., 2017. Seasonal variations in bacterioplankton community structures in two small rivers in the Himi region of central Japan and their relationships with environmental factors. *World J. Microbiol. Biotechnol.* 33, 1–16. <https://doi.org/10.1007/s11274-017-2377-4>
- Taranu, Z.E., Gregory-Eaves, I., Leavitt, P.R., Bunting, L., Buchaca, T., Catalan, J., Domaizon, I., Guilizzoni, P., Lami, A., MCGowan, S., Moorhouse, H., Morabito, G., Pick, F.R., Stevenson, M.A., Thompson, P.L., Vinebrooke, R.D., 2015. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecol. Lett.* 18, 375–384. <https://doi.org/10.1111/ele.12420>
- Taylor, W.D., Losee, R.F., Torobin, M., Izaguirre, G., Sass, D., Khiari, D., 2006. *Early Warning and Management of Surface Water Taste-and-Odor Events*. Denver, US.
- Thacker, M., Karthick, B., 2022. Response of Diatoms to the Changing Water Quality in the Myristica Swamps of the Western Ghats, India. *Diversity* 14, 1–21.

<https://doi.org/10.3390/d14030202>

- Thompson, A.W., Foster, R.A., Krupke, A., Carter, B.J., Musat, N., Vaultot, D., Kuypers, M.M., Zehr, J.P., 2012. Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science* (80-. ). 6101, 1546–1560. <https://doi.org/10.1126/science.1222700>. PMID: 22997339
- Thomsen, P.F., Willerslev, E., 2015. Environmental DNA e an emerging tool in conservation for monitoring past and present biodiversity. *Biol. Conserv.* 183, 4–18.
- Tikhonov, G., Opedal, O.H., Abrego, N., Lehtikoinen, A., de Jonge, M.M.J., Oksanen, J., Ovaskainen, O., 2019. Joint species distribution modelling with the r-package Hmsc. *Methods Ecol. Evol.* 11, 442–447.
- Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., Schindler, D., Schlesinger, D.S., Swackhamer, D., 2001. Forecasting agriculturally driven global environmental change. *Science* (80-. ). 292, 281–284. <https://doi.org/10.1126/science.1057544>
- Tilman, D., Kilham, S.S., Kilham, P., 1982. COMMUNITY ECOLOGY : The Role of Limiting Nutrients. *Annu. Rev. Ecol. Syst.* 13, 349– 372.
- Touloupakis, E., Tartari, G., Zittelli, G.C., Torzillo, G., 2020. Growth and photosynthetic performance of *Chlamydomodium fusiforme* cells cultivated in BG11 and Bristol media. *J. Appl. Phycol.* 32, 145–152. <https://doi.org/10.1007/s10811-019-01900-y>
- Tringe, S.G., Hugenholtz, P., 2008. A renaissance for the pioneering 16S rRNA gene. *Curr. Opin. Microbiol.* 11, 442–446.
- Tromas, N., Taranu, Z.E., Martin, B.D., Willis, A., Fortin, N., Greer, C.W., Shapiro, B.J., 2018. Niche separation increases with genetic distance among bloom-forming cyanobacteria. *Front. Microbiol.* 9, 1–12. <https://doi.org/10.3389/fmicb.2018.00438>
- Tsao, H.W., Michinaka, A., Yen, H.K., Giglio, S., Hobson, P., Monis, P., Lin, T.F., 2014. Monitoring of geosmin producing *Anabaena circinalis* using quantitative PCR. *Water Res.* 49, 416–425. <https://doi.org/10.1016/j.watres.2013.10.028>
- Tucker, C.S., 2000. Off-flavor problems in aquaculture. *Rev. Fish. Sci.* 8, 45–88.



<https://doi.org/10.1080/10641260091129170>

- Tung, S.C., 2004. The effect of oxidants on 2-MIB concentration with the presence of cyanobacteria. *Water Sci. Technol.* 49, 281–288.
- Tuo, S.H., Lee Chen, Y.L., Chen, H.Y., Chen, T.Y., 2017. Free-living heterocystous cyanobacteria in the tropical marginal seas of the western North Pacific. *J. Plankton Res.* 39, 404–422. <https://doi.org/10.1093/plankt/fbx023>
- Tuomainen, J.M., Hietanen, S., Kuparinen, J., Martikainen, P.J., Servomaa, K., 2003. Baltic Sea cyanobacterial bloom contains denitrification and nitrification genes, but has negligible denitrification activity. *FEMS Microbiol. Ecol.* 45, 83–96. [https://doi.org/10.1016/S0168-6496\(03\)00131-4](https://doi.org/10.1016/S0168-6496(03)00131-4)
- Van de Vijver, B., Crawford, R.M., 2020. *Melosira jeanbertrandiana*, a new *Melosira* species (Bacillariophyceae) from the sub-Antarctic region. *Bot. Lett.* 167, 50–56. <https://doi.org/10.1080/23818107.2019.1688677>
- Van de Waal, D.B., Smith, V.H., Declerck, S.A.J., Stam, E.C.M., Elser, J.J., 2014. Stoichiometric regulation of phytoplankton toxins. *Ecol. Lett.* 17, 736–742. <https://doi.org/10.1111/ele.12280>
- Van Hannen, E.J., Mooij, W., Van Agterveld, M.P., Gons, H.J., Laanbroek, H.J., 1999. Detritus-dependent development of the microbial community in an experimental system: Qualitative analysis by denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* 65, 2478–2484. <https://doi.org/10.1128/aem.65.6.2478-2484.1999>
- van Rensburg, S.J., Barnard, S., Krüger, M., 2016. Challenges in the potable water industry due to changes in source water quality: Case study of midvaal water company, South Africa. *Water SA* 42, 633–640. <https://doi.org/10.4314/wsa.v42i4.14>
- Vandermaesen, J., Lievens, B., Springael, D., 2017. Isolation and identification of culturable bacteria, capable of heterotrophic growth, from rapid sand filters of drinking water treatment plants. *Res. Microbiol.* 168, 594–607. <https://doi.org/10.1016/j.resmic.2017.03.008>
- Vedler, E., Heinaru, E., Jutkina, J., Viggor, S., Koressaar, T., Remm, M., Heinaru, A., 2013. *Limnobacter* spp. as newly detected phenol-degraders among Baltic Sea

surface water bacteria characterised by comparative analysis of catabolic genes. *Syst. Appl. Microbiol.* 36, 525–532.

<https://doi.org/10.1016/j.syapm.2013.07.004>

Veech, J.A., 2013. A probabilistic model for analysing species co-occurrence. *Glob. Ecol. Biogeogr.* 22, 252–260. <https://doi.org/10.1111/j.1466-8238.2012.00789.x>

Velichko, N., Chernyaeva, E., Averina, S., Gavrilova, O., Lapidus, A., Pinevich, A., 2015. Consortium of the “bichlorophyllous” cyanobacterium *Prochlorothrix hollandica* and chemoheterotrophic partner bacteria: Culture and metagenome-based description. *Environ. Microbiol. Rep.* 7, 623–633.

<https://doi.org/10.1111/1758-2229.12298>

Veraart, A.J., Romani, A.M., Tornés, E., Sabater, S., 2008. Algal response to nutrient enrichment in forested oligotrophic stream. *J. Phycol.* 44, 564–572.

<https://doi.org/10.1111/j.1529-8817.2008.00503.x>

Villafane, V.E., Sundback, K., Figueroa, F.L., Helbling, E.W., 2003. Photosynthesis in the aquatic environment as affected by UVR, in: Helbling, E.W., Zagarese, H., Webb, A.R., Neale, P.J., Kieber, D.J., Wetzel, R.G., Blumthaler, M., Hargreaves, B.R., Zepp, R.G., Osburn, C.L., Morris, D.P., Diamond, S.A., Peake, B.M., Scully, N.M., Buma, A.G.J., Boelen, P., Jeffrey, W.H., Villafane, V.E., Sundback, K., Figueroa, F.L., Little, E., Fabacher, D., Banaszak, A., Leech, D.M., Johnsen, S., Hessen, D.O., Sommaruga, R., Leavitt, P.R., Hodgson, D., Pienitz, R., Williamson, C.E., Helbling, E.W., Zagarese, H., Hader, D.-P., Jori, G., *Photobiology, T.E.S. of (Eds.), UV Effects in Aquatic Organisms and Ecosystems. The Royal Society of Chemistry*, pp. 359–381.

<https://doi.org/10.1039/9781847552266-00357>

Visser, P.M., Ibelings, B.W., Bormans, M., Huisman, J., 2016. Artificial mixing to control cyanobacterial blooms: a review. *Aquat. Ecol.* 50, 423–441.

<https://doi.org/10.1007/s10452-015-9537-0>

Vitousek, P.M., Mooney, H.A., Lubchenco, J., Melillo, J.M., 1997. Human domination of Earth's ecosystems. *Science (80- )*. 277, 494–499.

Vrede, T., Ballantyne, A., Mille-Lindblom, C., Algesten, G., Gudas, C., Lindahl, S., Brunberg, A.K., 2009. Effects of N : P loading ratios on phytoplankton community composition, primary production and N fixation in a eutrophic lake.

- Freshw. Biol. 54, 331–344. <https://doi.org/10.1111/j.1365-2427.2008.02118.x>
- Walsby, A.E., 1994. Gas vesicles. *Microbiol. Rev.* 58, 94–144.
- Walters, K.J., Whitaker, M.J., Singh, S.P., Montgomery, B.L., 2013. Light intensity and reactive oxygen species are centrally involved in photoregulatory responses during complementary chromatic adaptation in *Fremyella diplosiphon*. *Commun. Integr. Biol.* 6, 1–5. <https://doi.org/10.4161/cib.25005>
- Wang, C., Wang, Z., Qiao, X., Li, Z., Li, F., Chen, M., Wang, Y., Huang, Y., Cui, H., 2013. Antifungal activity of volatile organic compounds from *Streptomyces alboflavus* TD-1. *FEMS Microbiol. Lett.* 341, 45–51. <https://doi.org/10.1111/1574-6968.12088>
- Wang, C.M., Cane, D.E., 2008. Biochemistry and molecular genetics of the biosynthesis of the earthy odorant methylisoborneol in *Streptomyces coelicolor*. *J. Am. Chem. Soc.* 130, 1–6. <https://doi.org/10.1021/ja803639g>
- Wang, R., Huajun, C., ChaoXi, J., BingWu, Y., 2014. Occurrence and degradation of 2-methylisoborneol and geosmin in water supply reservoir in winter. *China Environ. Sci.* 34, 896–903.
- Wang, S.H., Dzialowski, A.R., Meyer, J.O., Denoyelles, F., Lim, N.C., Spotts, W.W., Huggins, D.G., 2005. Relationships between cyanobacterial production and the physical and chemical properties of a Midwestern Reservoir, USA. *Hydrobiologia.* <https://doi.org/10.1007/s10750-004-4665-x>
- Wang, Zhicong, Huang, S., Li, D., 2019. Decomposition of cyanobacterial bloom contributes to the formation and distribution of iron-bound phosphorus (Fe-P): Insight for cycling mechanism of internal phosphorus loading. *Sci. Total Environ.* 652, 696–708. <https://doi.org/10.1016/j.scitotenv.2018.10.260>
- Wang, Z., Li, R., 2015. Effects of light and temperature on the odor production of 2-methylisoborneol-producing *Pseudanabaena* sp. and geosmin-producing *Anabaena ucrainica* (cyanobacteria). *Biochem. Syst. Ecol.* <https://doi.org/10.1016/j.bse.2014.12.013>
- Wang, Zhongjie, Song, G., Li, Y., Yu, G., Hou, X., Gan, Z., Li, R., 2019. The diversity, origin, and evolutionary analysis of geosmin synthase gene in cyanobacteria. *Sci. Total Environ.* 689, 789–796.

<https://doi.org/10.1016/j.scitotenv.2019.06.468>

Wang, Z., Song, G., Shao, J., Tan, W., Li, Y., Li, R., 2016. Establishment and field applications of real-time PCR methods for the quantification of potential MIB-producing cyanobacteria in aquatic systems. *J. Appl. Phycol.*

<https://doi.org/10.1007/s10811-015-0529-1>

Wang, Z., Xu, Y., Shao, J., Wang, J., Li, R., 2011. Genes associated with 2-methylisoborneol biosynthesis in cyanobacteria: Isolation, characterization, and expression in response to light. *PLoS One* 6, 1–9.

<https://doi.org/10.1371/journal.pone.0018665>

Wasmund, N., Nausch, G., Feistel, R., 2013. Silicate consumption: An indicator for long-term trends in spring diatom development in the Baltic Sea. *J. Plankton Res.* 35, 393–406. <https://doi.org/10.1093/plankt/fbs101>

Watson, S., Juttner, F., 2019. Taste and Odour in Source and Drinking Water: Causes, Controls, and Consequences, *Taste and Odour in Source and Drinking Water: Causes, Controls, and Consequences*. IWA Publishing.

<https://doi.org/10.2166/9781780406664>

Watson, S.B., 2010. Algal Taste and Odour, in: *Algae: Source to Treatment*. American Water Works Association, Denver, CO, pp. 329–376.

Watson, S.B., 2004. Aquatic taste and odor: A primary signal of drinking-water integrity. *J. Toxicol. Environ. Heal. - Part A* 67, 1779–1795.

<https://doi.org/10.1080/15287390490492377>

Watson, S.B., 2003. Cyanobacterial and eukaryotic algal odour compounds: Signals or by-products? A review of their biological activity. *Phycologia* 42, 332–350.

<https://doi.org/10.2216/i0031-8884-42-4-332.1>

Watson, S.B., Charlton, M., Rao, Y.R., Howell, T., 2007. Off flavors in large water bodies: physics, chemistry and biology in synchrony. *Water Sci Technol* 55, 1–8.

<https://doi.org/https://doi.org/10.2166/wst.2007.155>

Watson, S.B., Monis, P., Baker, P., Giglio, S., 2016. Biochemistry and genetics of taste- and odor-producing cyanobacteria. *Harmful Algae* 54, 112–127.

<https://doi.org/10.1016/j.hal.2015.11.008>

Watson, S.B., Ridal, J., 2004. Periphyton: a primary source of widespread and

- severe taste and odour. *Water Sci. Technol.* 49, 33–39.
- Watson, S.B., Ridal, J., Boyer, G.L., 2008. Taste and odour and cyanobacterial toxins: Impairment, prediction, and management in the Great Lakes. *Can. J. Fish. Aquat. Sci.* 65, 1779–1796. <https://doi.org/10.1139/F08-084>
- Watson, S.B., Satchwill, T., Dixon, E., Mccauley, E., 2001. Under-ice blooms and source-water odour in a nutrient-poor reservoir: Biological, ecological and applied perspectives. *Freshw. Biol.* 46, 1553–1567. <https://doi.org/10.1046/j.1365-2427.2001.00769.x>
- Webber, M.A., Atherton, P., Newcombe, G., 2015. Taste and odour and public perceptions: what do our customers really think about their drinking water? *J. Water Supply Res. Technol. - AQUA* 64, 802–811. <https://doi.org/https://doi.org/10.2166/aqua.2015.067>
- Webster, G., O’Sullivan, L.A., Meng, Y., Williams, A.S., Sass, A.M., Watkins, A.J., Parkes, R.J., Weightman, A.J., 2015. Archaeal community diversity and abundance changes along a natural salinity gradient in estuarine sediments. *FEMS Microbiol. Ecol.* 91, 1–18. <https://doi.org/10.1093/femsec/fiu025>
- Wei, X., Zhang, X., Jiang, C., 2022. Charting the landscape of the environmental exposome. *iMeta* 1, 1–44. <https://doi.org/10.1002/imt2.50>
- Westerhoff, P., Rodriguez-Hernandez, M., Baker, L., Sommerfeld, M., 2005. Seasonal occurrence and degradation of 2-methylisoborneol in water supply reservoirs. *Water Res.* 39, 4899–4912. <https://doi.org/10.1016/j.watres.2005.06.038>
- WHO, 2016. Management of cyanobacteria in drinking-water supplies: Information for regulators and water suppliers, *Who/Fwc/Wsh/15.03*.
- Wickham, H., 2016. *ggplot2. Use R!*, in: *Data Analysis*. Springer, Cham, pp. 1–13.
- Wienhausen, G., Noriega-Ortega, B.E., Niggemann, J., Dittmar, T., Simon, M., 2017. The exometabolome of two model strains of the *Roseobacter* group: A marketplace of microbial metabolites. *Front. Microbiol.* 8, 1–15. <https://doi.org/10.3389/fmicb.2017.01985>
- Wijewardene, L., Wu, N., Qu, Y., Guo, K., Messyasz, B., Lorenz, S., Riis, T., Ulrich, U., Fohrer, N., 2021. Influences of pesticides, nutrients, and local environmental

- variables on phytoplankton communities in lentic small water bodies in a German lowland agricultural area. *Sci. Total Environ.* 780, 1–14.  
<https://doi.org/10.1016/j.scitotenv.2021.146481>
- Wilhelm, S.W., Maxwell, D.P., G, T.C., 1996. Growth, iron requirements, and siderophore production in iron-limited *Synechococcus* PCC 7002. *Limnol. Oceanogr.* 41, 89–97.
- Wilhelm, S.W., Trick, C.G., 1994. Iron-limited growth of cyanobacteria: Multiple siderophore production is a common response. *Limnol. Oceanogr.* 39, 1979–1984. <https://doi.org/10.4319/lo.1994.39.8.1979>
- Winder, M., Hunter, D.A., 2008. Temporal organization of phytoplankton communities linked to physical forcing. *Oecologia* 156, 179–192.  
<https://doi.org/10.1007/s00442-008-0964-7>
- Winter, J.G., Desellas, A.M., Fletcher, R., Heintsch, L., Morley, A., Nakamoto, L., Utsumi, K., 2011. Algal blooms in Ontario, Canada: Increases in reports since 1994. *Lake Reserv. Manag.* 27, 107–114.  
<https://doi.org/10.1080/07438141.2011.557765>
- Wood, S., Williams, S.T., White, W.R., 2001. Microbes as a source of earthy flavours in potable water - A review. *Int. Biodeterior. Biodegrad.*  
[https://doi.org/10.1016/S0964-8305\(01\)00064-6](https://doi.org/10.1016/S0964-8305(01)00064-6)
- Woodhouse, J.N., Kinsela, A.S., Collins, R.N., Bowling, L.C., Honeyman, G.L., Holliday, J.K., Neilan, B.A., 2016. Microbial communities reflect temporal changes in cyanobacterial composition in a shallow ephemeral freshwater lake. *ISME J.* 10, 1337–1351. <https://doi.org/10.1038/ismej.2015.218>
- Woodhouse, J.N., Ongley, S.E., Brown, M. V, Neilan, B.A., 2013. Microbial diversity and diazotrophy associated with the freshwater non-heterocyst forming cyanobacterium *Lyngbya robusta*. *J. Appl. Phycol.* 25, 1039–1045.  
<https://doi.org/10.1007/s10811-012-9909-y>
- Woodhouse, J.N., Ziegler, J., Grossart, H.P., Neilan, B.A., 2018. Cyanobacterial community composition and bacteria-bacteria interactions promote the stable occurrence of particle-associated bacteria. *Front. Microbiol.* 9.  
<https://doi.org/10.3389/fmicb.2018.00777>

- Worm, J., Søndergaard, M., 1998. Dynamics of heterotrophic bacteria attached to *Microcystis* spp. (Cyanobacteria). *Aquat. Microb. Ecol.* 14, 19–28.  
<https://doi.org/10.3354/ame014019>
- Woyke, T., Chertkov, O., Lapidus, A., Nolan, M., Lucas, S., del Rio, T.G., Tice, H., Cheng, J.F., Tapia, R., Han, C., Goodwin, L., Pitluck, S., Liolios, K., Pagani, I., Ivanova, N., Huntemann, M., Mavromatis, K., Mikhailova, N., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Brambilla, E.M., Rohde, M., Mwirichia, R., Sikorski, J., Tindall, B.J., Göker, M., Bristow, J., Eisen, J.A., Markowitz, V., Hugenholtz, P., Klenk, H.P., Kyrpides, N.C., 2011. Complete genome sequence of the gliding freshwater bacterium *Fluviicola taffensis* type strain (RW262T). *Stand. Genomic Sci.* 5, 21–29. <https://doi.org/10.4056/sigs.2124912>
- Wu, A., Wang, Y., Friese, K., Zhang, L., Han, C., Kang, D., Shen, Q., 2021. Spatial and Seasonal Distribution of 2-Methylisoborneol in a Large Eutrophic Shallow Lake, China. *Water. Air. Soil Pollut.* 232. <https://doi.org/10.1007/s11270-021-05340-8>
- Wu, H., Li, A., Wang, J., Li, X., Cui, M., Yang, N., Liu, Y., Zhang, L., Wang, X., Zhan, G., 2022. A novel electrochemical sensor based on autotrophic and heterotrophic nitrifying biofilm for trichloroacetic acid toxicity monitoring. *Environ. Res.* 210, 112985. <https://doi.org/10.1016/j.envres.2022.112985>
- Wu, J.T., Juttner, F., 1988. Effect of environmental factors on geosmin production by *Fischerella muscicola*. *Water Sci Technol* 20, 143–148.  
<https://doi.org/https://doi.org/10.2166/wst.1988.0235>
- Xiao, Z., Zhang, S., Yan, P., Huo, J., Aurangzeib, M., 2022. Microbial Community and Their Potential Functions after Natural Vegetation Restoration in Gullies of Farmland in Mollisols of Northeast China. *Land* 11, 1–18.  
<https://doi.org/https://doi.org/10.3390/land11122231>
- Xie, M., Ren, M., Yang, C., Yi, H., Li, Z., Li, T., Zhao, J., 2016. Metagenomic analysis reveals symbiotic relationship among bacteria in microcystis-dominated community. *Front. Microbiol.* 7, 1–10. <https://doi.org/10.3389/fmicb.2016.00056>
- Xiong, Z., Lai, L., Ding, Y., Yang, L., Geng, Y., Pavlostathis, S.G., Shao, P., Zhang, Y., Luo, X., 2022. Corn cob biocarriers with available carbon release for *Chlamydomodium* sp. microalgae towards enhanced nitrogen removal from low

- C/N rare earth element tailings (REEs) wastewater. *Chemosphere* 307, 135673. <https://doi.org/10.1016/j.chemosphere.2022.135673>
- Xu, Z., Te, S.H., He, Y., Gin, K.Y.H., 2018. The characteristics and dynamics of cyanobacteria-heterotrophic bacteria between two estuarine reservoirs - Tropical versus subtropical regions. *Front. Microbiol.* 9, 1–13. <https://doi.org/10.3389/fmicb.2018.02531>
- Xue, Q., Utsumi, M., Shimizu, K., Chen, H., Li, M., Xu, C., Zhang, Z., Zhang, Q., Sugiura, N., 2011. Biodegradation of geosmin by biofilm from water treatment plant in winter season. *Int. J. Integr. Biol.* 11, 30–34.
- Xuwei, D., Min, Q., ren, R., Jiarui, L., Xiaoxue, S., Ping, X., Jun, C., 2019. The relationships between odors and environmental factors at bloom and non-bloom area in Lake Taihu, China. *Chemosphere* 218, 569–576. <https://doi.org/10.1016/j.chemosphere.2018.11.121>
- Yagi, M., 2005. 35 years' history of off-flavor problems in the southern basin of Lake Biwa, in: *The 7th IWA Symposium on Off-Flavours in the Aquatic Environment*. Cornwall, Canada.
- Yagi, M., Nakashima, S., Muramoto, S., 1988. Biological degradation of musty odor compounds, 2-methylisoborneol and geosmin in a bio-activated carbon filter. *Water Sci. Technol.* 20, 255–260.
- Yamada, Y., Kuzuyama, T., Komatsu, M., Shin-ya, K., Omura, S., Cane, D.E., Ikeda, H., 2015. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci.* 112, 857–862. <https://doi.org/10.1073/pnas.1422108112>
- Yamamoto, Y., Nizuma, S., Kuroda, N., Sakamoto, M., 1993. Occurrence of heterotrophic bacteria causing lysis of cyanobacteria in a eutrophic lake. *J. Phycol.* 41, 215–220.
- Yamamoto, Y., Tsukada, H., 2009. Measurement of in situ specific growth rates of microcystis (cyanobacteria) from the frequency of dividing cells. *J. Phycol.* 45, 1003–1009. <https://doi.org/10.1111/j.1529-8817.2009.00723.x>
- Yang, J.S., Yuan, D.X., Weng, T.P., 2010. Pilot study of drinking water treatment with GAC, O3/BAC and membrane processes in Kinmen Island, Taiwan. *Desalination* 263, 271–278. <https://doi.org/10.1016/j.desal.2010.06.069>



- Yang, S.Z., Xiong, X., Feng, G. Da, Li, H.P., Zhu, H.H., 2017. Reclassification of *Sphingopyxis contaminans* as *Sphingorhabdus contaminans* comb. Nov. and emended description of the genus *Sphingorhabdus*. *Int. J. Syst. Evol. Microbiol.* 67, 4328–4331. <https://doi.org/10.1099/ijsem.0.002201>
- Yang, T., Lee, C.S., Cho, J.-Y., Bae, M.-J., Kim, E.-J., 2022. Comparison of Bacterial Assemblages Associated with Harmful Cyanobacteria under Different Light Conditions. *Microorganisms* 10, 2150. <https://doi.org/10.3390/microorganisms10112150>
- Yang, Y., Li, S., Gao, Y., Chen, Y., Zhan, A., 2019. Environment-driven geographical distribution of bacterial communities and identification of indicator taxa in Songhua River. *Ecol. Indic.* 101, 62–70. <https://doi.org/10.1016/j.ecolind.2018.12.047>
- Youn, Seok Jea, Kim, H.N., Yu, S.J., Byeon, M.S., 2020. Cyanobacterial occurrence and geosmin dynamics in Paldang Lake watershed, South Korea. *Water Environ. J.* 34, 634–643. <https://doi.org/10.1111/wej.12547>
- Youn, S J, Yu, S.J., Byeon, M.S., 2020. Occurrence characteristics of stephanodiscus and synedra in relation to water temperature and concentrations of nutrients during spring diatom bloom in lake Paldang, Korea. *Appl. Ecol. Environ. Res.* 18, 5135–5147. [https://doi.org/10.15666/aeer/1804\\_51355147](https://doi.org/10.15666/aeer/1804_51355147)
- Young, W.F., Horth, H., Crane, R., Ogden, T., Arnott, M., 1996. Taste and Odour threshold concentrations of potential potable water contaminants. *Water Res.* 30, 331–340.
- Yuan, Q., Wang, P., Wang, X., Hu, B., Tao, L., 2022. Phytoremediation of cadmium-contaminated sediment using *Hydrilla verticillata* and *Elodea canadensis* harbor two same keystone rhizobacteria *Pedospaeraceae* and *Parasegetibacter*. *Chemosphere* 286, 1–11. <https://doi.org/10.1016/j.chemosphere.2021.131648>
- Yuan, R., Zhou, B., Shi, C., Yu, L., Zhang, C., Gu, J., 2012. Biodegradation of 2-methylisoborneol by bacteria enriched from biological activated carbon. *Front. Environ. Sci. Eng. China* 6, 701–710. <https://doi.org/10.1007/s11783-011-0367-6>
- Zamorano, D., Peredo-Parada, M., Lillo, D.J., Parodi, J., Díaz, C.A., 2019. *Mat*

- thickness associated with *Didymosphenia geminata* and *Cymbella* spp. in the southern rivers of Chile. *PeerJ* 2019, 1–20. <https://doi.org/10.7717/peerj.6481>
- Zamyadi, A., Henderson, R.K., Stuetz, R., Newcombe, G., Newtown, K., Gladman, B., 2016. Cyanobacterial management in full-scale water treatment and recycling processes: reactive dosing following intensive monitoring. *Environ. Sci. Water Res. Technol.* 2, 362–375. <https://doi.org/10.1039/c5ew00269a>
- Zemb, O., Achard, C.S., Hamelin, J., De Almeida, M.L., Gabinaud, B., Cauquil, L., Verschuren, L.M.G., Godon, J.J., 2020. Absolute quantitation of microbes using 16S rRNA gene metabarcoding: A rapid normalization of relative abundances by quantitative PCR targeting a 16S rRNA gene spike-in standard. *Microbiologyopen* 9, 1–19. <https://doi.org/10.1002/mbo3.977>
- Zepernick, B.N., Gann, E.R., Martin, R.M., Pound, H.L., Krausfeldt, L.E., Chaffin, J.D., Wilhelm, S.W., 2021. Elevated pH Conditions Associated With *Microcystis* spp. Blooms Decrease Viability of the Cultured Diatom *Fragilaria crotonensis* and Natural Diatoms in Lake Erie. *Front. Microbiol.* 12, 1–12. <https://doi.org/10.3389/fmicb.2021.598736>
- Zhang, H., Sekiguchi, Y., Hanada, S., Hugenholtz, P., Kim, H., Kamagata, Y., Nakamura, K., 2003. *Gemmatimonas aurantiaca* gen. nov., sp. nov., a Gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Gemmatimonadetes phyl. nov. *Int. J. Syst. Evol. Microbiol.* 53, 1155–1163. <https://doi.org/10.1099/ijs.0.02520-0>
- Zhang, J., Li, L., Qiu, L., Wang, X., Meng, X., You, Y., Yu, J., Ma, W., 2017. Effects of Climate Change on 2-Methylisoborneol Production in Two Cyanobacterial Species. *Water* 9, 859. <https://doi.org/10.3390/w9110859>
- Zhang, S., 2017. Biologically active filters: An advanced treatment process for pharmaceuticals and personal care products. New Jersey Institute of Technology.
- Zhang, T., Li, L., Song, L., Chen, W., 2009. Effects of temperature and light on the growth and geosmin production of *Lyngbya kuetzingii* (Cyanophyta). *J. Appl. Phycol.* 21, 279–285. <https://doi.org/10.1007/s10811-008-9363-z>

- Zhang, T., Li, L., Zheng, L., Song, L., 2017. Effects of nutritional factors on the geosmin production of *Lyngbya kuetzingii* UTEX 1547 (Oscillatoriales, Cyanobacteria). *Phycologia* 56, 221–229. <https://doi.org/10.2216/16-98.1>
- Zhang, Z., Nair, S., Tang, L., Zhao, H., Hu, Z., Chen, M., Zhang, Yao, Kao, S.J., Jiao, N., Zhang, Yongyu, 2021. Long-term survival of *Synechococcus* and heterotrophic bacteria without external nutrient supply after changes in their relationship from antagonism to mutualism. *MBio* 12, 1–15. <https://doi.org/10.1128/mBio.01614-21>
- Zheng, X., Wang, Y., Yang, T., He, Z., Yan, Q., 2020. Size-fractioned aggregates within phycosphere define functional bacterial communities related to *Microcystis aeruginosa* and *Euglena sanguinea* blooms. *Aquat. Ecol.* 54, 609–623. <https://doi.org/10.1007/s10452-020-09762-0>
- Zhi, Y., Wu, Q., Du, H., Xu, Y., 2016. Biocontrol of geosmin-producing *Streptomyces* spp. by two *Bacillus* strains from Chinese liquor. *Int. J. Food Microbiol.* 231, 1–9. <https://doi.org/10.1016/j.ijfoodmicro.2016.04.021>
- Zhou, A., Chen, Y.I., Zane, G.M., He, Z., Hemme, C.L., Joachimiak, M.P., Baumohl, J.K., He, Q., Fields, M.W., Arkin, A.P., Wall, J.D., Hazen, T.C., Zhou, J., 2012. Functional characterization of Crp/Fnr-type global transcriptional regulators in *Desulfovibrio vulgaris hildenborough*. *Appl. Environ. Microbiol.* 78, 1168–1177. <https://doi.org/10.1128/AEM.05666-11>
- Zhou, B., Yuan, R., Shi, C., Yu, L., Gu, J., Zhang, C., 2011. Biodegradation of geosmin in drinking water by novel bacteria isolated from biologically active carbon. *J. Environ. Sci.* 23, 816–823. [https://doi.org/10.1016/S1001-0742\(10\)60458-5](https://doi.org/10.1016/S1001-0742(10)60458-5)
- Zhou, X., Zhang, K., Zhang, T., Li, C., Mao, X., 2017. An ignored and potential source of taste and odor (T&O) issues—biofilms in drinking water distribution system (DWDS). *Appl. Microbiol. Biotechnol.* 101, 3537–3550. <https://doi.org/10.1007/s00253-017-8223-7>
- Zimba, P. V., Dionigi, C.P., Millie, D.F., 1999. Evaluating the relationship between photopigment synthesis and 2-methylisoborneol accumulation in cyanobacteria. *J. Phycol.* 35, 1422–1429. <https://doi.org/10.1046/j.1529-8817.1999.3561422.x>

Zimmerman, W.J., Soliman, B.H., Rosen, B.H., 1995. Growth and 2-methylisoborneol production by the cyanobacterium *Phormidium* LM689. *Water Sci. Technol.* 31, 181–186.

Zuo, Z., 2019. Why algae release volatile organic compounds - The emission and roles. *Front. Microbiol.* 10, 1–7. <https://doi.org/10.3389/fmicb.2019.00491>

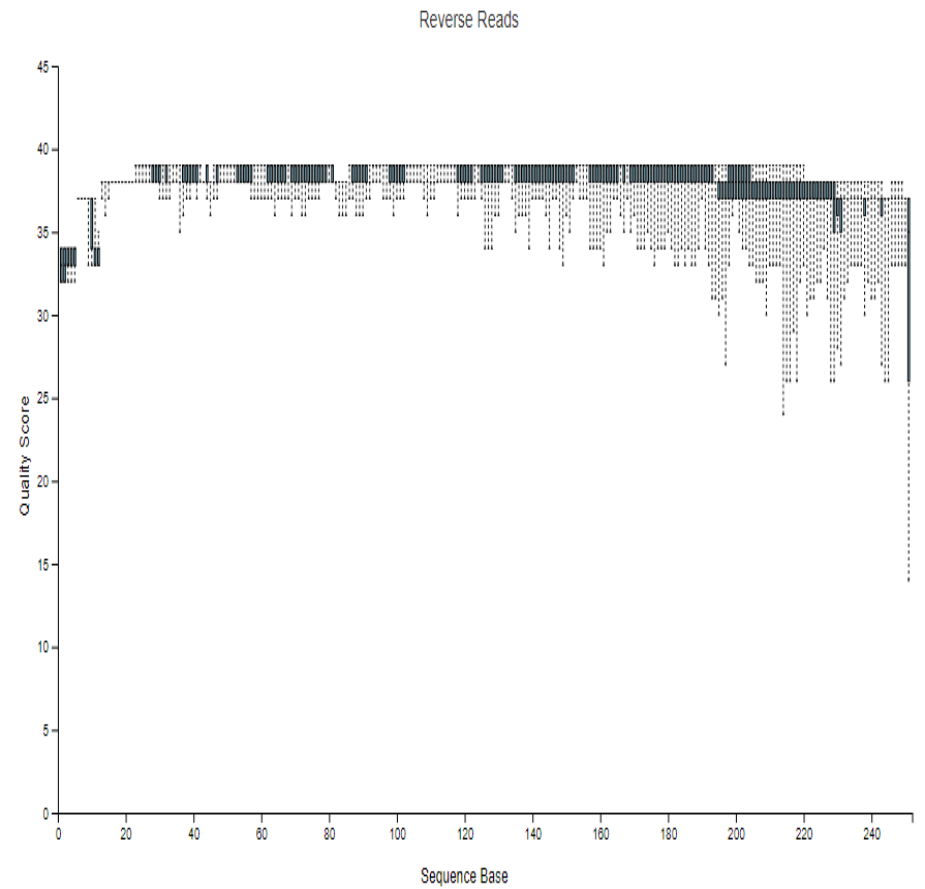
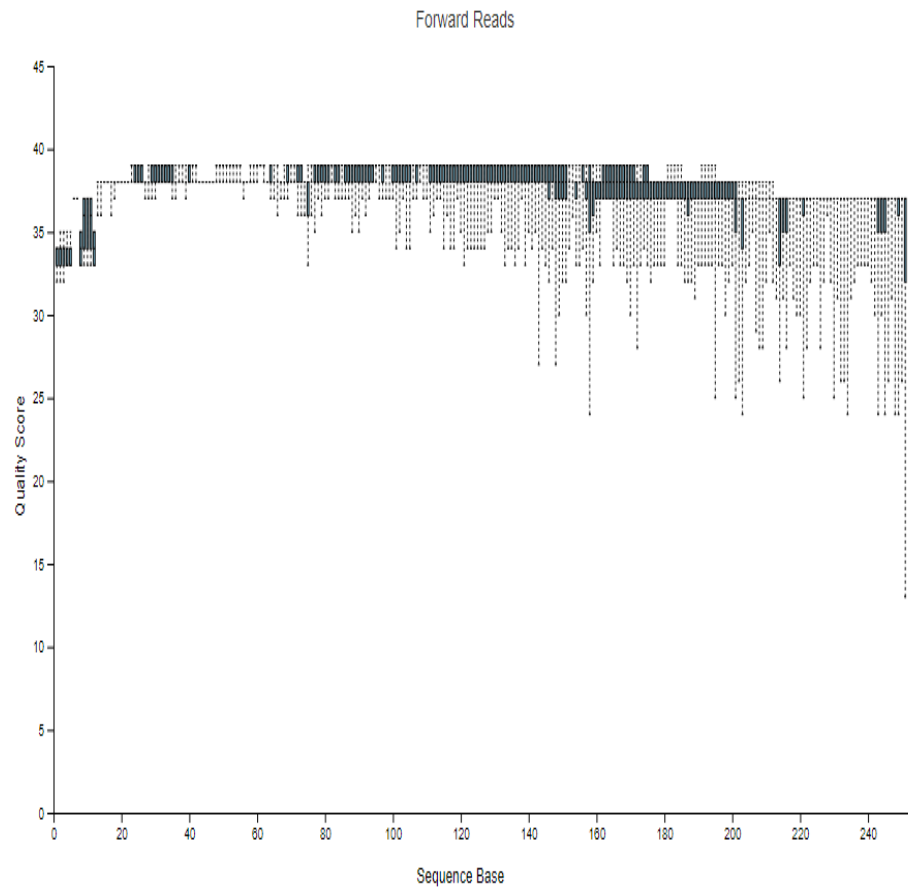
## Appendix

Appendix 1.1: List of genera capable of the biodegradation of geosmin and 2-MIB, with full taxonomy and corresponding reference.

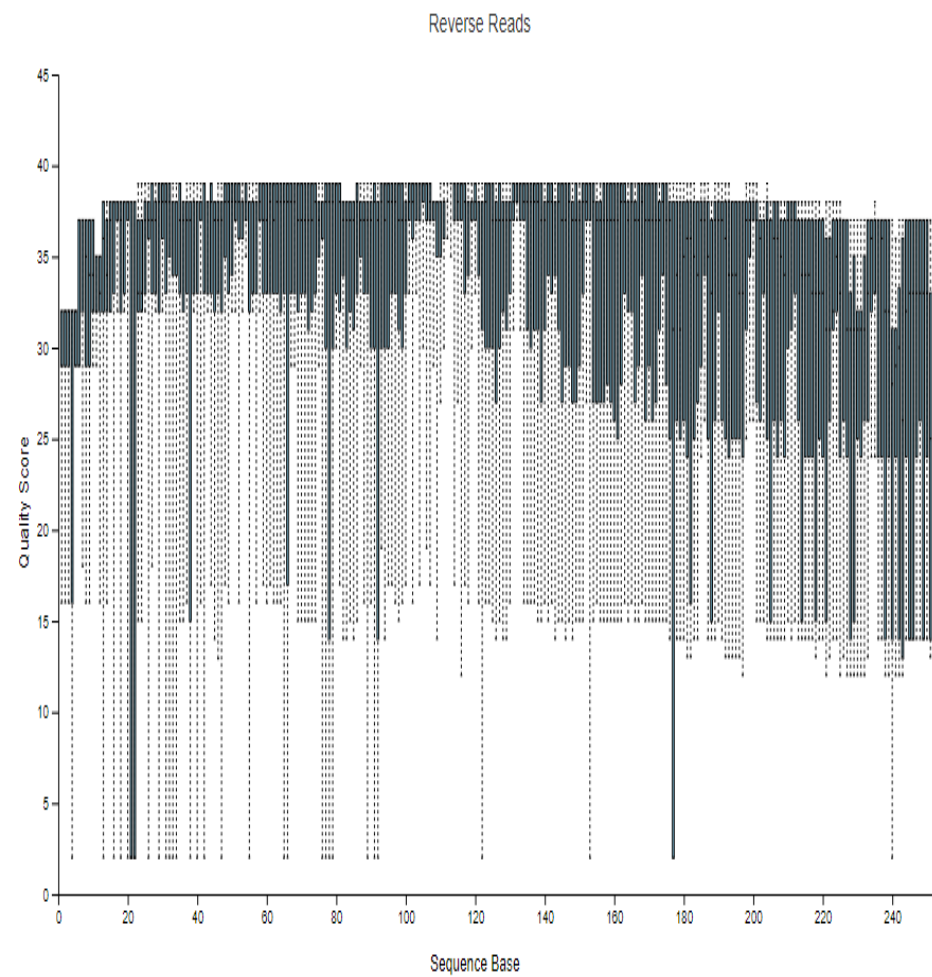
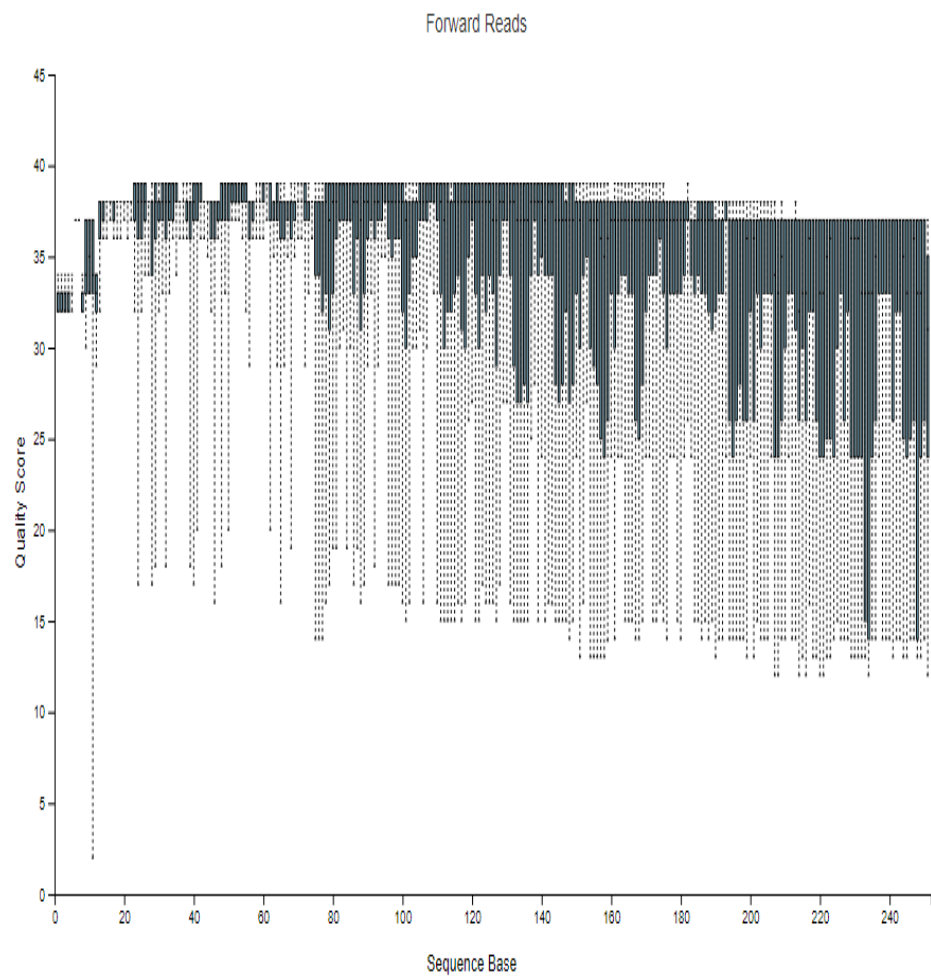
| T&O Degrading Genera     | Degrading compound | Phylum                | Class                      | Order                    | Family                    | Reference                    |
|--------------------------|--------------------|-----------------------|----------------------------|--------------------------|---------------------------|------------------------------|
| <i>Acidovorax</i>        | M                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>   | <i>Comamonadaceae</i>     | (Clercic, 2014)              |
| <i>Afipia</i>            | G                  | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Hyphomicrobiales</i>  | <i>Nitrobacteraceae</i>   | (Baek et al., 2015)          |
| <i>Arthrobacter</i>      | G                  | <i>Actinomycetota</i> | <i>Actinomycetia</i>       | <i>Micrococcales</i>     | <i>Micrococcaceae</i>     | (Clercic et al., 2022)       |
| <i>Bacillus</i>          | G + M              | <i>Bacillota</i>      | <i>Bacilli</i>             | <i>Bacillales</i>        | <i>Bacillaceae</i>        | (Clercic et al., 2021)       |
| <i>Bradyrhizobium</i>    | G                  | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Hyphomicrobiales</i>  | <i>Nitrobacteraceae</i>   | (Baek et al., 2015)          |
| <i>Brevibacterium</i>    | M                  | <i>Actinomycetota</i> | <i>Actinomycetia</i>       | <i>Micrococcales</i>     | <i>Brevibacteriaceae</i>  | (Yuan et al., 2011)          |
| <i>Candida</i>           | M                  | <i>Ascomycota</i>     | <i>Saccharomycetes</i>     | <i>Saccharomycetales</i> | <i>Saccharomycetaceae</i> | (Sumitomo, 1988)             |
| <i>Chryseobacterium</i>  | G                  | <i>Bacteroidota</i>   | <i>Flavobacteriia</i>      | <i>Flavobacteriales</i>  | <i>Weeksellaceae</i>      | (Zhou et al., 2011)          |
| <i>Comamonas</i>         | M                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>   | <i>Comamonadaceae</i>     | (Guttman and van Rijn, 2012) |
| <i>Curvibacter</i>       | M                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>   | <i>Comamonadaceae</i>     | (Clercic, 2014)              |
| <i>Enterobacter</i>      | M                  | <i>Pseudomonadota</i> | <i>Gammaproteobacteria</i> | <i>Enterobacterales</i>  | <i>Enterobacteriaceae</i> | (Clercic et al., 2021)       |
| <i>Flavobacterium</i>    | M                  | <i>Bacteroidota</i>   | <i>Flavobacteriia</i>      | <i>Flavobacteriales</i>  | <i>Flavobacteriaceae</i>  | (Yuan et al., 2011)          |
| <i>Hydrogenophaga</i>    | G                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>   | <i>Comamonadaceae</i>     | (Strait, 2015)               |
| <i>Leptothrix</i>        | G                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>   | <i>Comamonadaceae</i>     | (Clercic, 2014)              |
| <i>Janthinobacterium</i> | M                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>   | <i>Oxalobacteraceae</i>   | (Clercic, 2014)              |
| <i>Myroides</i>          | M                  | <i>Bacteroidota</i>   | <i>Flavobacteriia</i>      | <i>Flavobacteriales</i>  | <i>Flavobacteriaceae</i>  | (Clercic, 2014)              |
| <i>Novosphingobium</i>   | G + M              | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Sphingomonadales</i>  | <i>Erythrobacteraceae</i> | (Hoefel et al., 2006)        |
| <i>Pseudomonas</i>       | G + M              | <i>Pseudomonadota</i> | <i>Gammaproteobacteria</i> | <i>Pseudomonadales</i>   | <i>Pseudomonadaceae</i>   | (Clercic et al., 2021)       |
| <i>Psychroflexus</i>     | M                  | <i>Bacteroidota</i>   | <i>Flavobacteriia</i>      | <i>Flavobacteriales</i>  | <i>Flavobacteriaceae</i>  | (Clercic, 2014)              |
| <i>Rhodococcus</i>       | G + M              | <i>Actinomycetota</i> | <i>Actinomycetia</i>       | <i>Mycobacteriales</i>   | <i>Nocardiaceae</i>       | (Guttman and van Rijn, 2012) |
| <i>Rhodoferax</i>        | M                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>   | <i>Comamonadaceae</i>     | (Clercic, 2014)              |

Appendix 1.1 cont'd: List of genera capable of the biodegradation of geosmin and 2-MIB, with full taxonomy and corresponding reference.

| T&O Degrading Genera    | Degrading compound | Phylum                | Class                      | Order                   | Family                   | Reference                    |
|-------------------------|--------------------|-----------------------|----------------------------|-------------------------|--------------------------|------------------------------|
| <i>Shinella</i>         | M                  | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Hyphomicrobiales</i> | <i>Rhizobiaceae</i>      | (Du et al., 2017)            |
| <i>Sinorhizobium</i>    | G                  | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Hyphomicrobiales</i> | <i>Rhizobiaceae</i>      | (Zhou et al., 2011)          |
| <i>Sphingobium</i>      | M                  | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Sphingomonadales</i> | <i>Sphingomonadaceae</i> | (Ho et al., 2007)            |
| <i>Sphingomonas</i>     | G + M              | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Sphingomonadales</i> | <i>Sphingomonadaceae</i> | (Ho et al., 2007)            |
| <i>Sphingopyxis</i>     | G                  | <i>Pseudomonadota</i> | <i>Alphaproteobacteria</i> | <i>Sphingomonadales</i> | <i>Sphingomonadaceae</i> | (Hoefel et al., 2006)        |
| <i>Stenotrophomonas</i> | G                  | <i>Pseudomonadota</i> | <i>Gammaproteobacteria</i> | <i>Xanthomonadales</i>  | <i>Xanthomonadaceae</i>  | (Zhou et al., 2011)          |
| <i>Tenacibaculum</i>    | M                  | <i>Bacteroidota</i>   | <i>Flavobacteriia</i>      | <i>Flavobacteriales</i> | <i>Flavobacteriaceae</i> | (Clercic, 2014)              |
| <i>Thiomonas</i>        | M                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>  | <i>Comamonadaceae</i>    | (Clercic, 2014)              |
| <i>Variovorax</i>       | M                  | <i>Pseudomonadota</i> | <i>Betaproteobacteria</i>  | <i>Burkholderiales</i>  | <i>Comamonadaceae</i>    | (Guttman and van Rijn, 2012) |

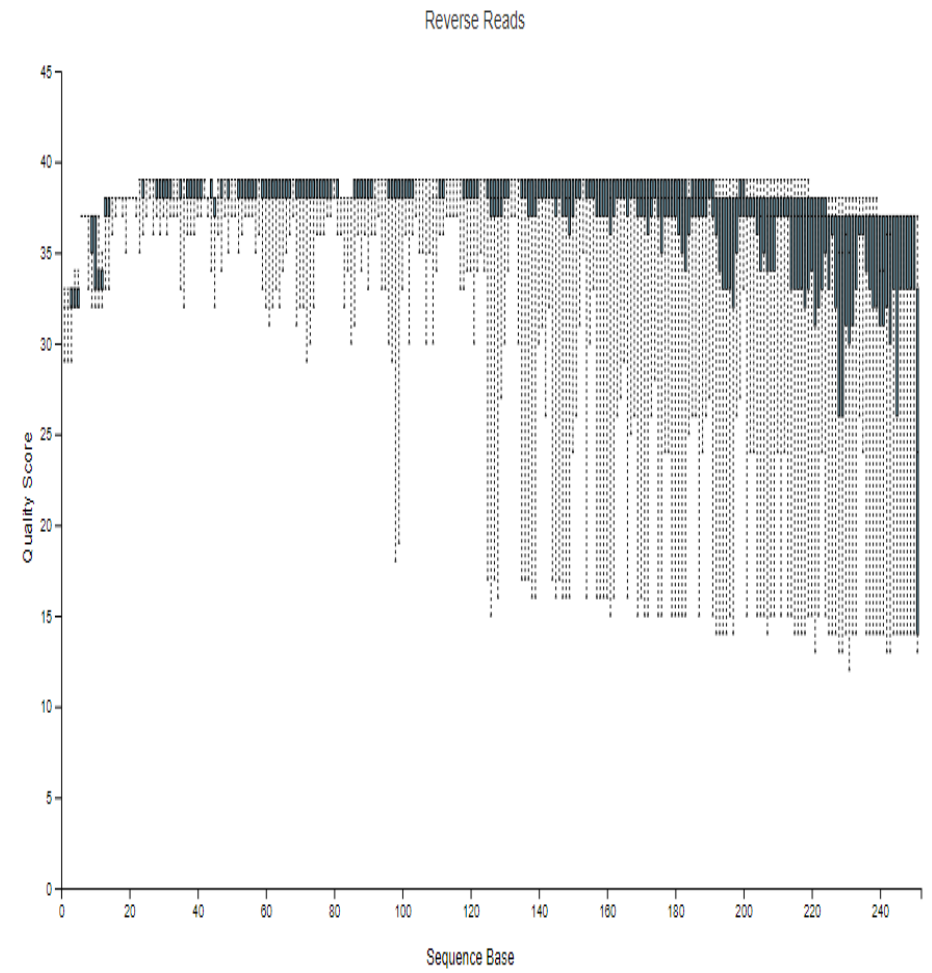
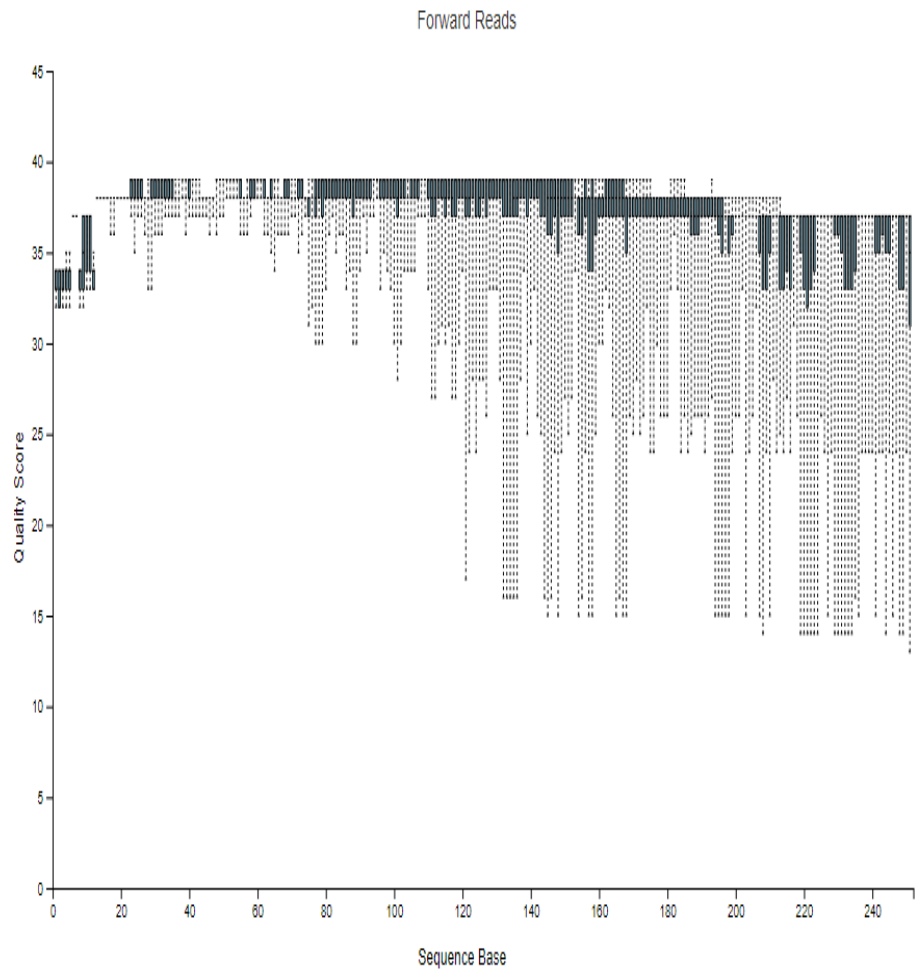


Appendix 2.1: Quality scores for the forwards and reverse reads for the MiSeq pilot run for 16S rRNA gene.

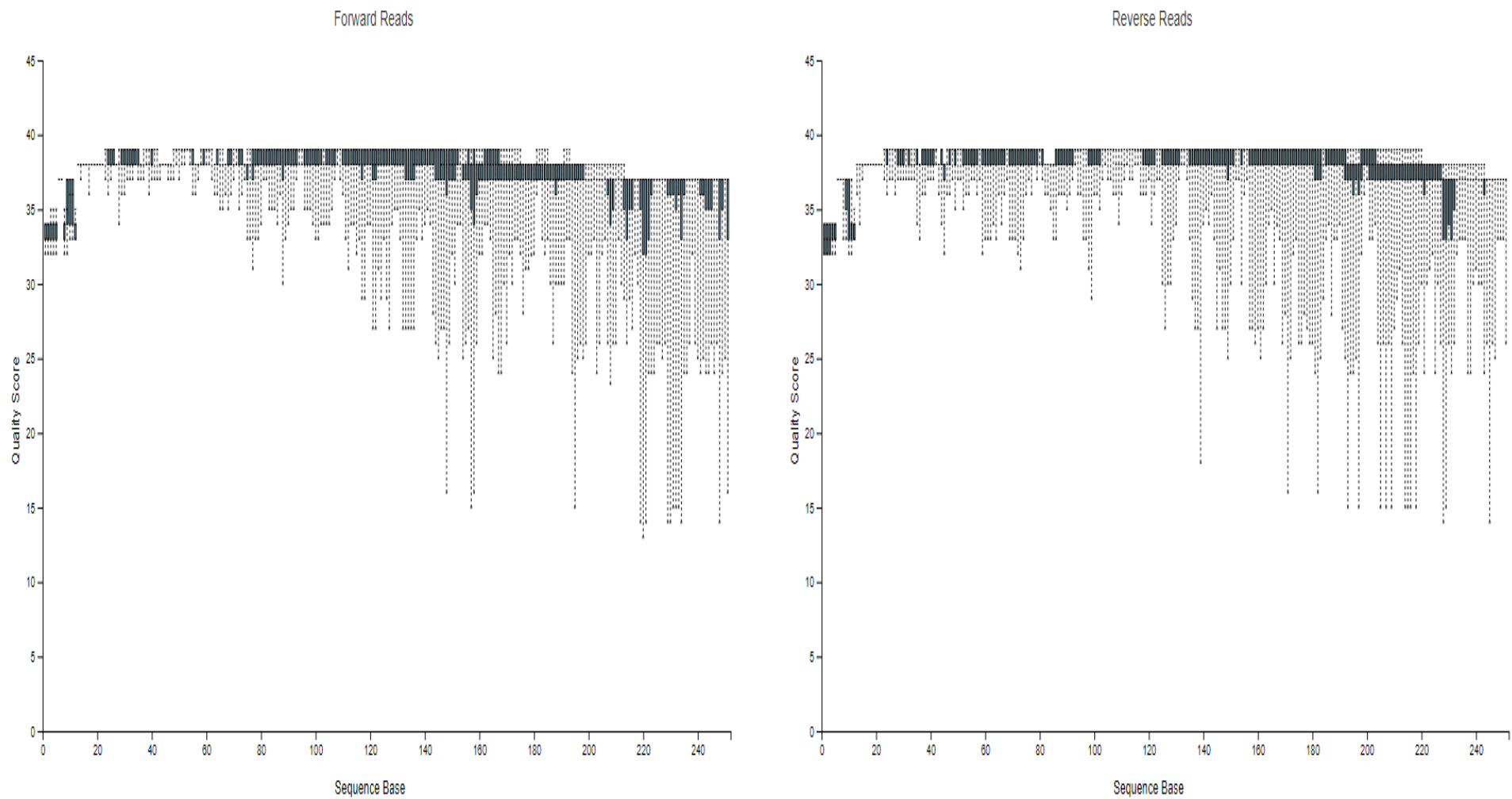


Appendix 2.2: Quality scores for the forwards and reverse reads for the MiSeq Plate 1-2 run for 16S rRNA gene.

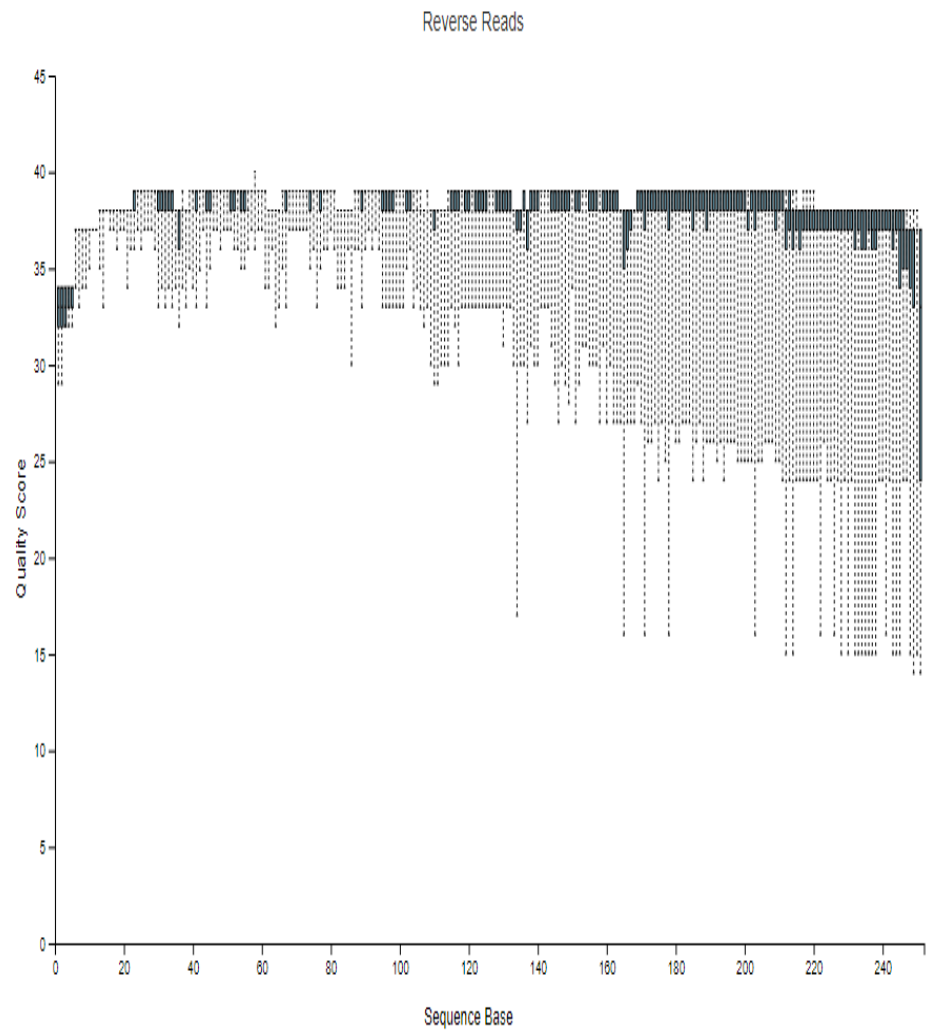
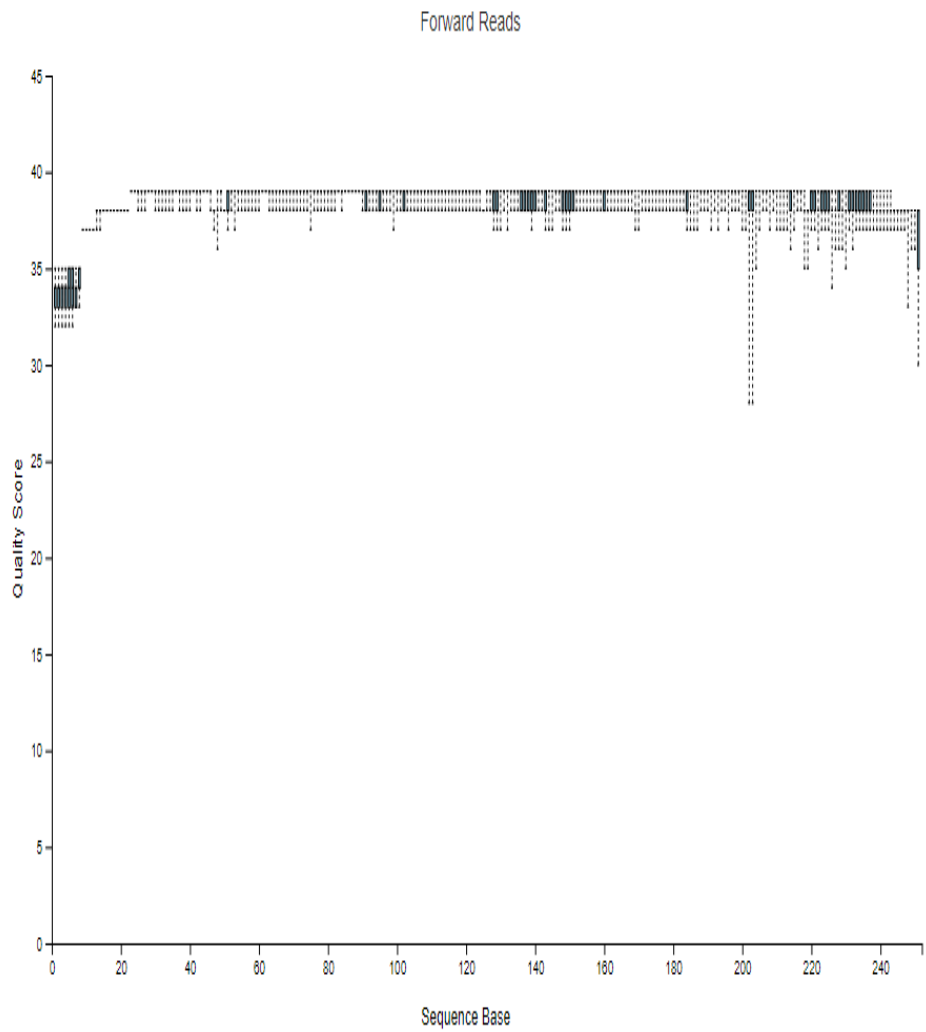




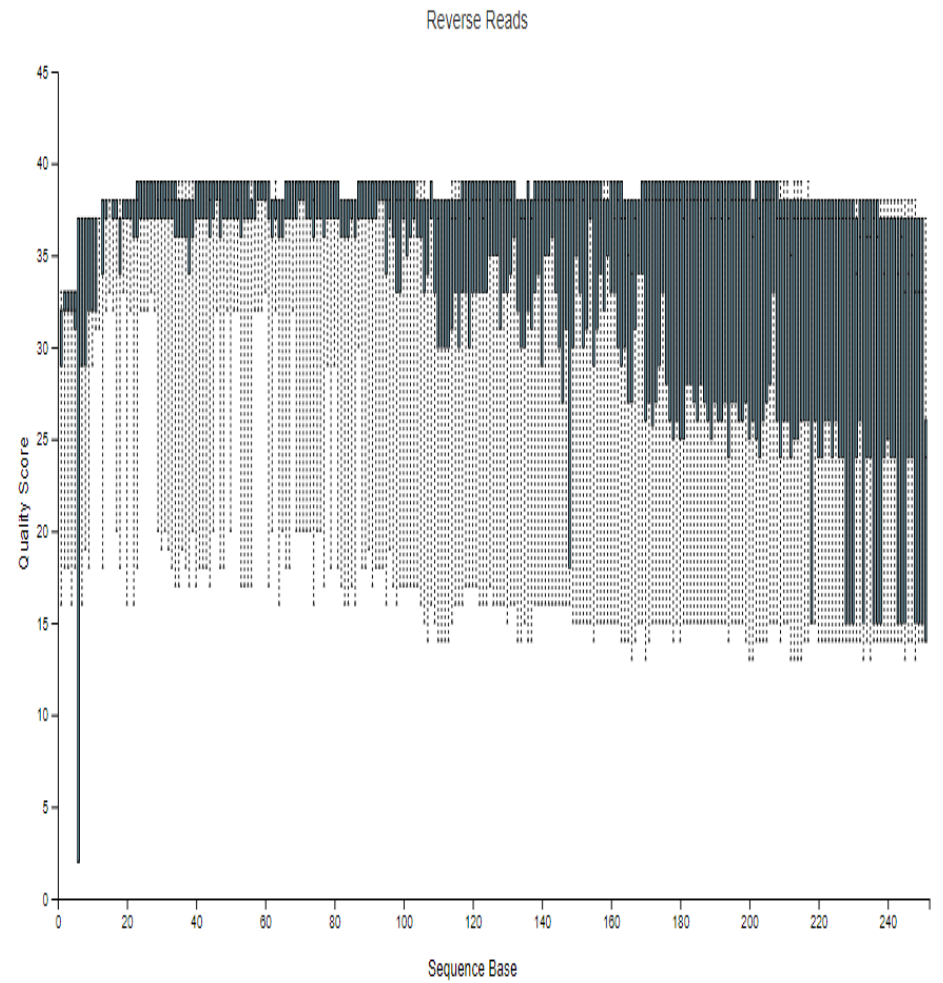
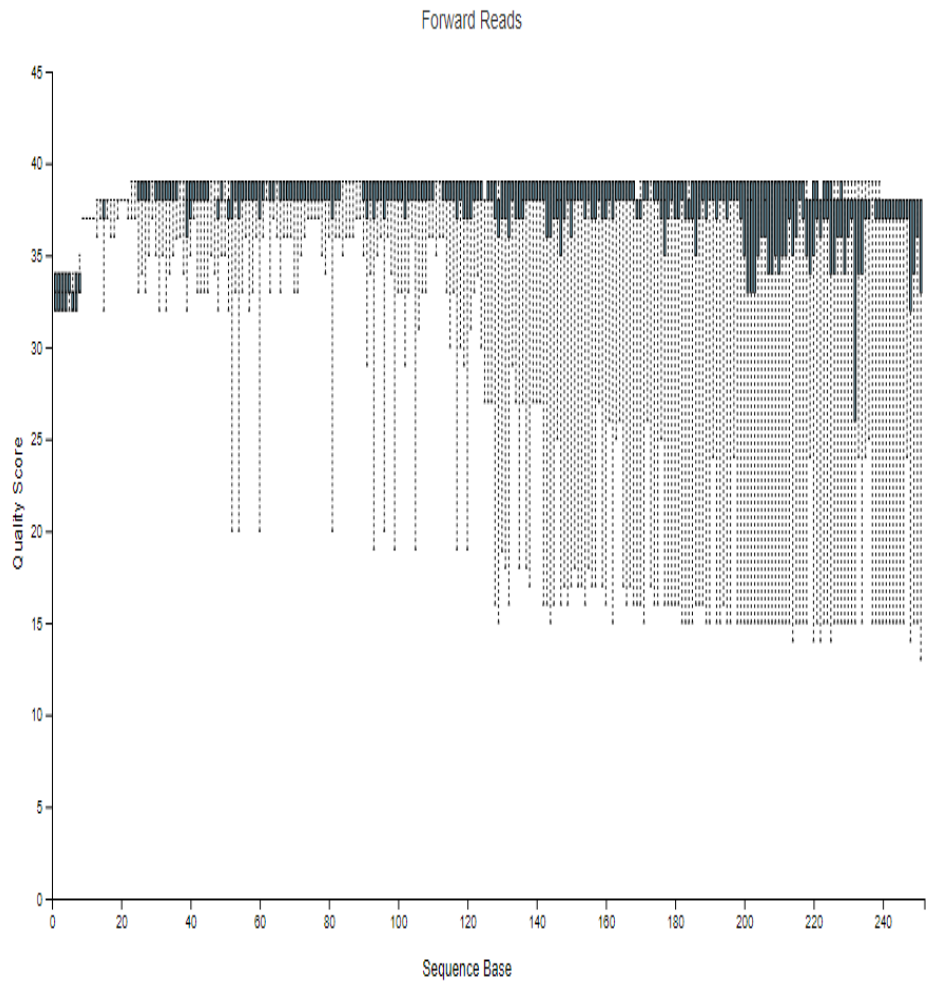
Appendix 2.3: Quality scores for the forwards and reverse reads for the MiSeq Plate 4-6 run for 16S rRNA gene.



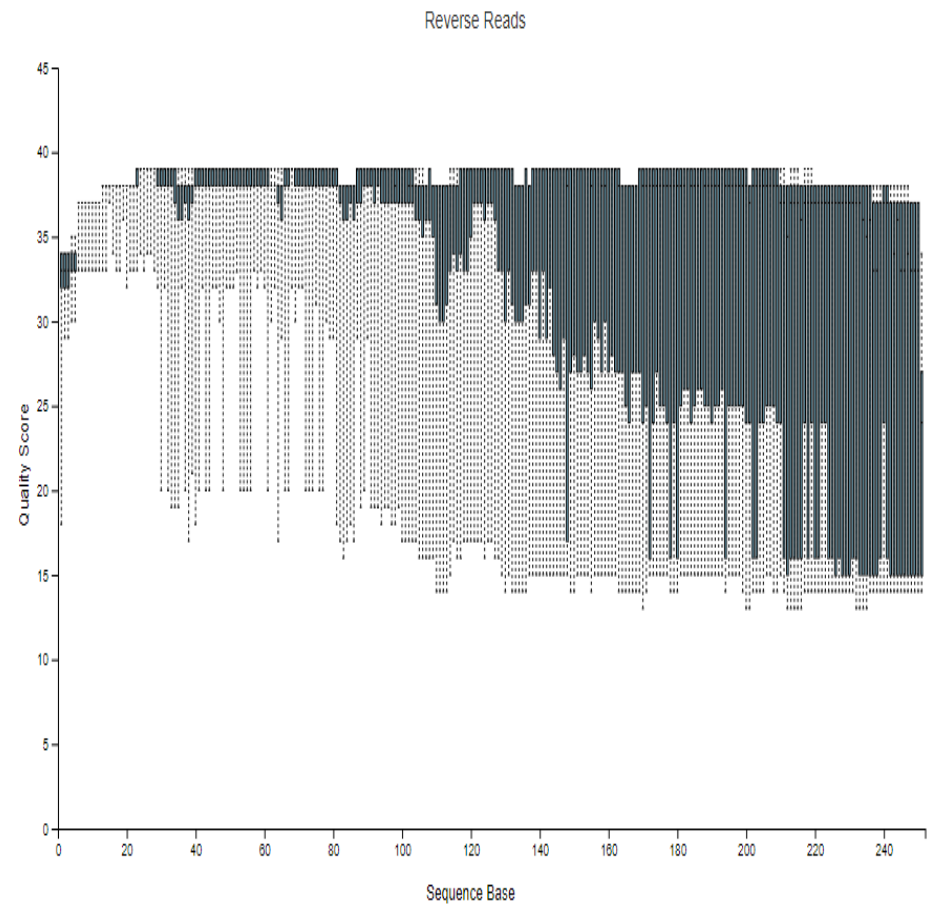
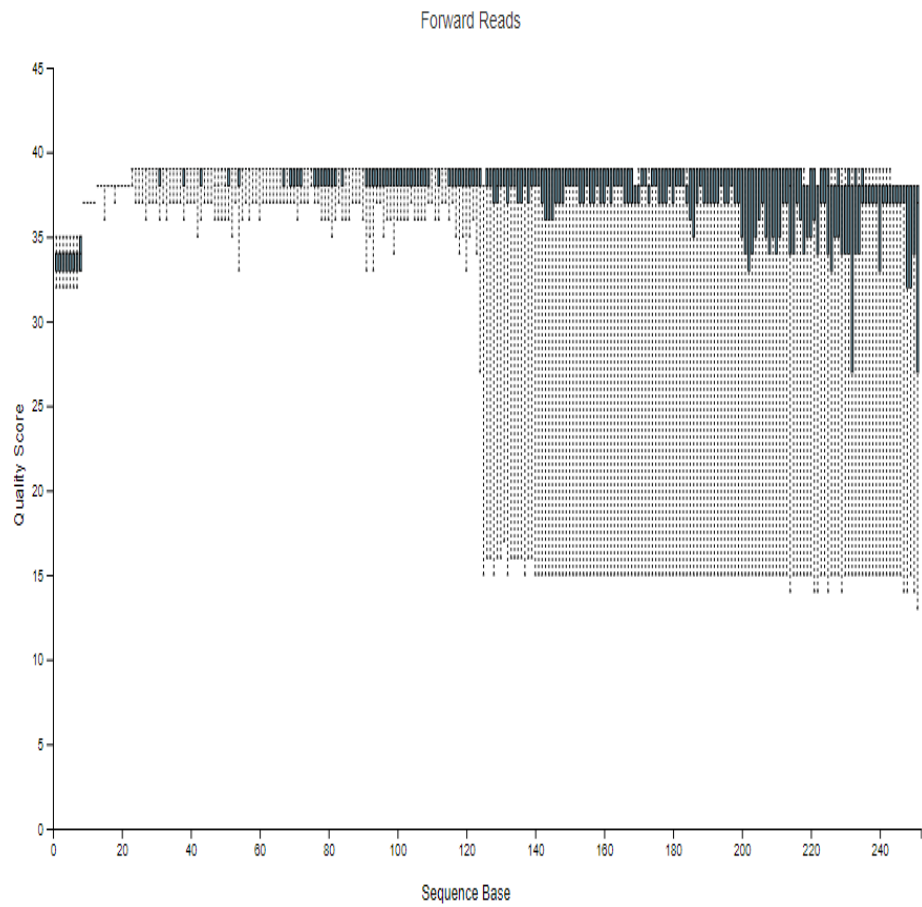
Appendix 2.4: Quality scores for the forwards and reverse reads for the MiSeq Plate 3 + 7 run for 16S rRNA gene.



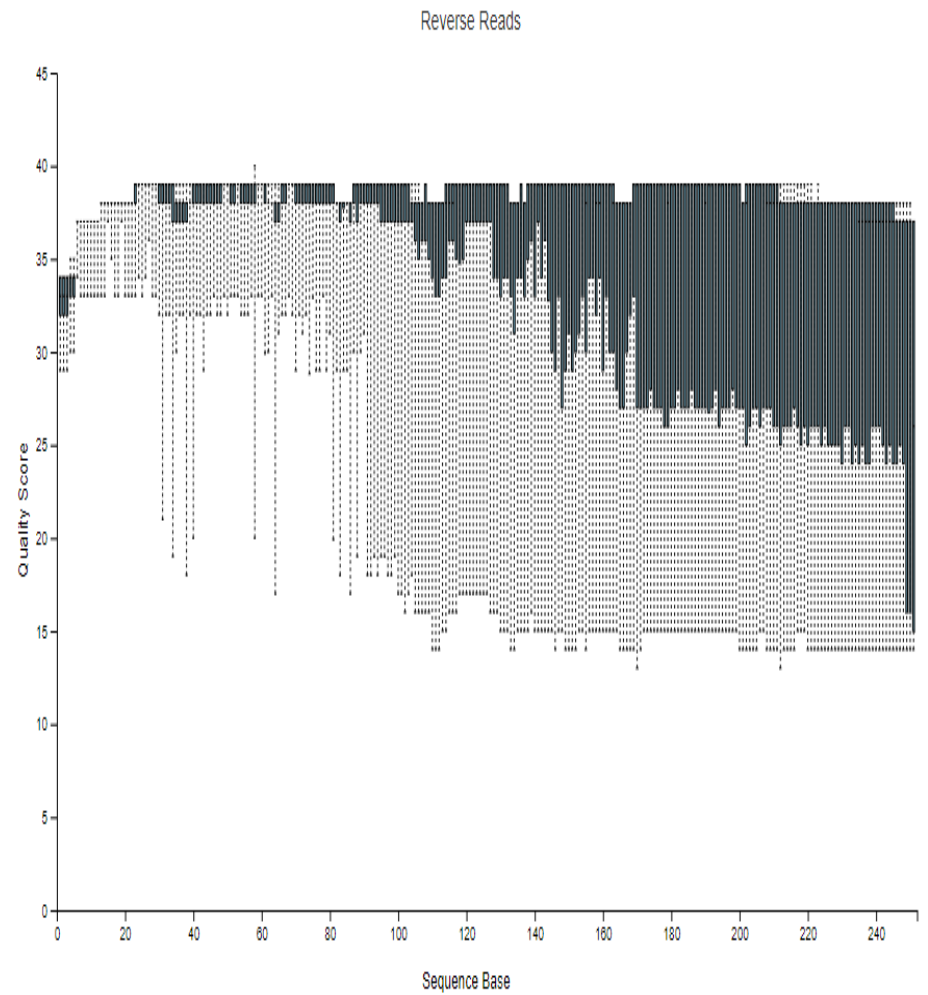
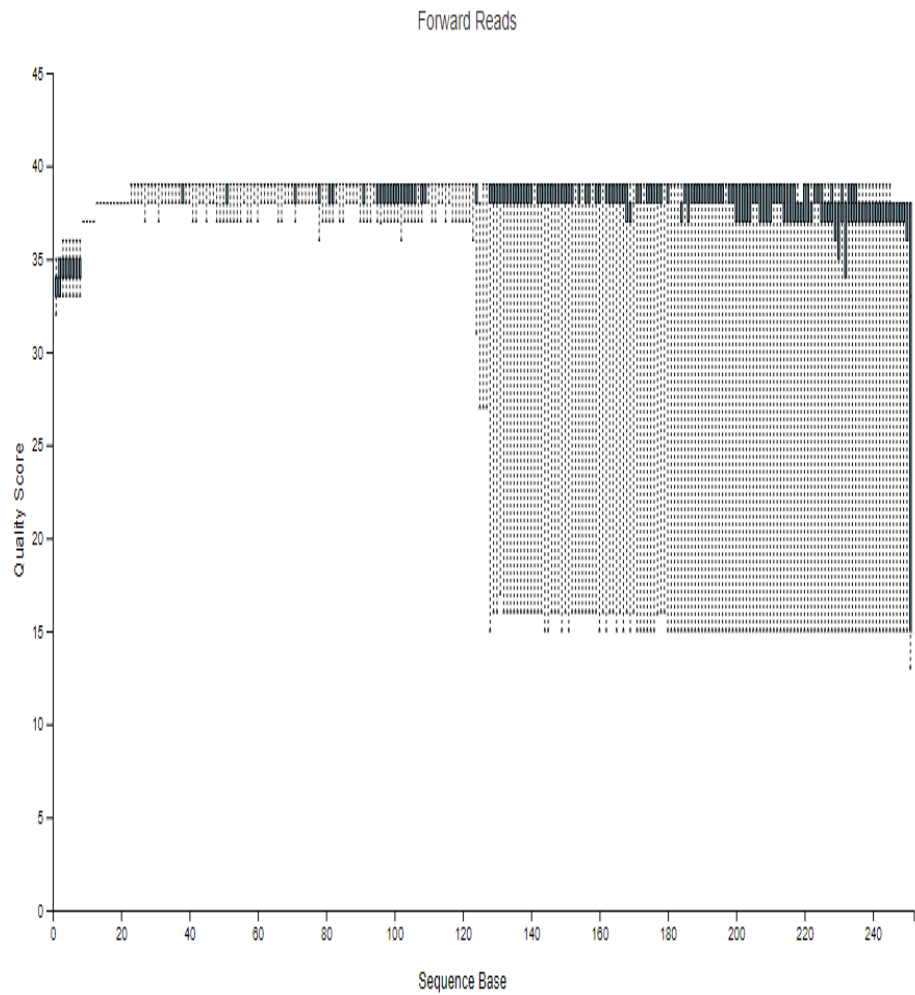
Appendix 2.5: Quality scores for the forwards and reverse reads for the MiSeq Pilot run for *rbcL* gene.



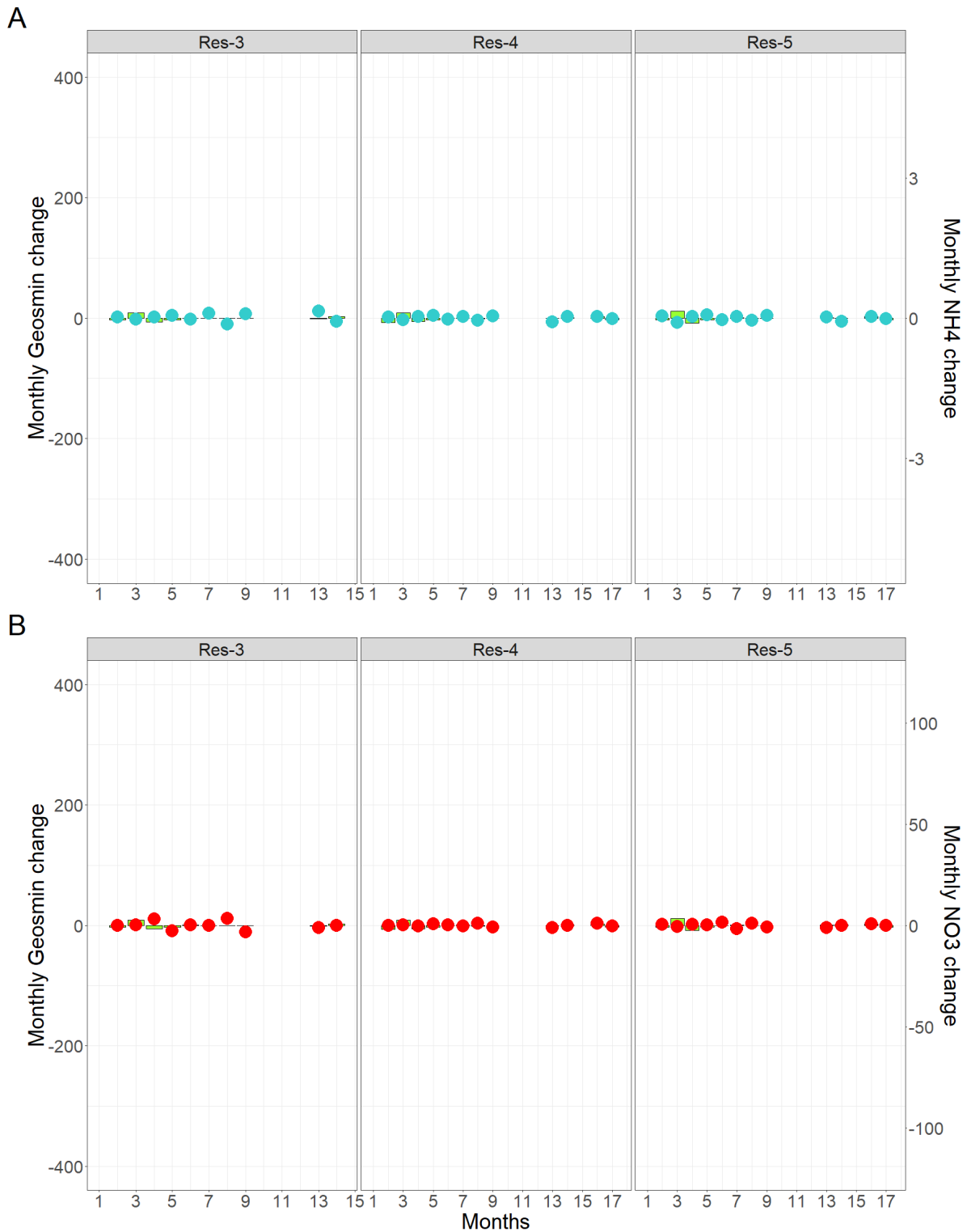
Appendix 2.6: Quality scores for the forwards and reverse reads for the MiSeq Plate 1-2 run for *rbcL* gene.



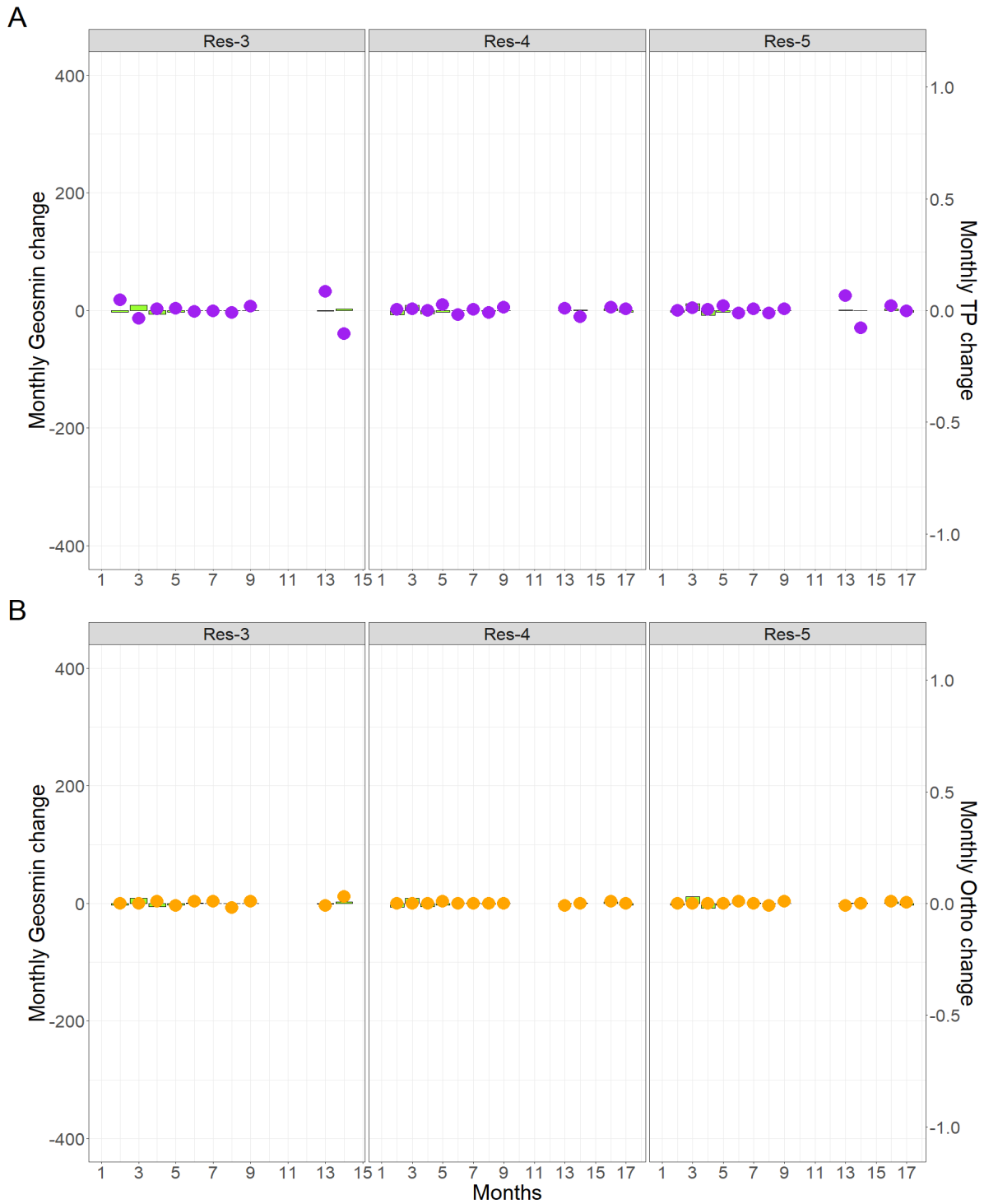
Appendix 2.7: Quality scores for the forwards and reverse reads for the MiSeq Plate 4-6 run for *rbcL* gene.



Appendix 2.8: Quality scores for the forwards and reverse reads for the MiSeq Plate 3 + 7 run for *rbcl* gene.

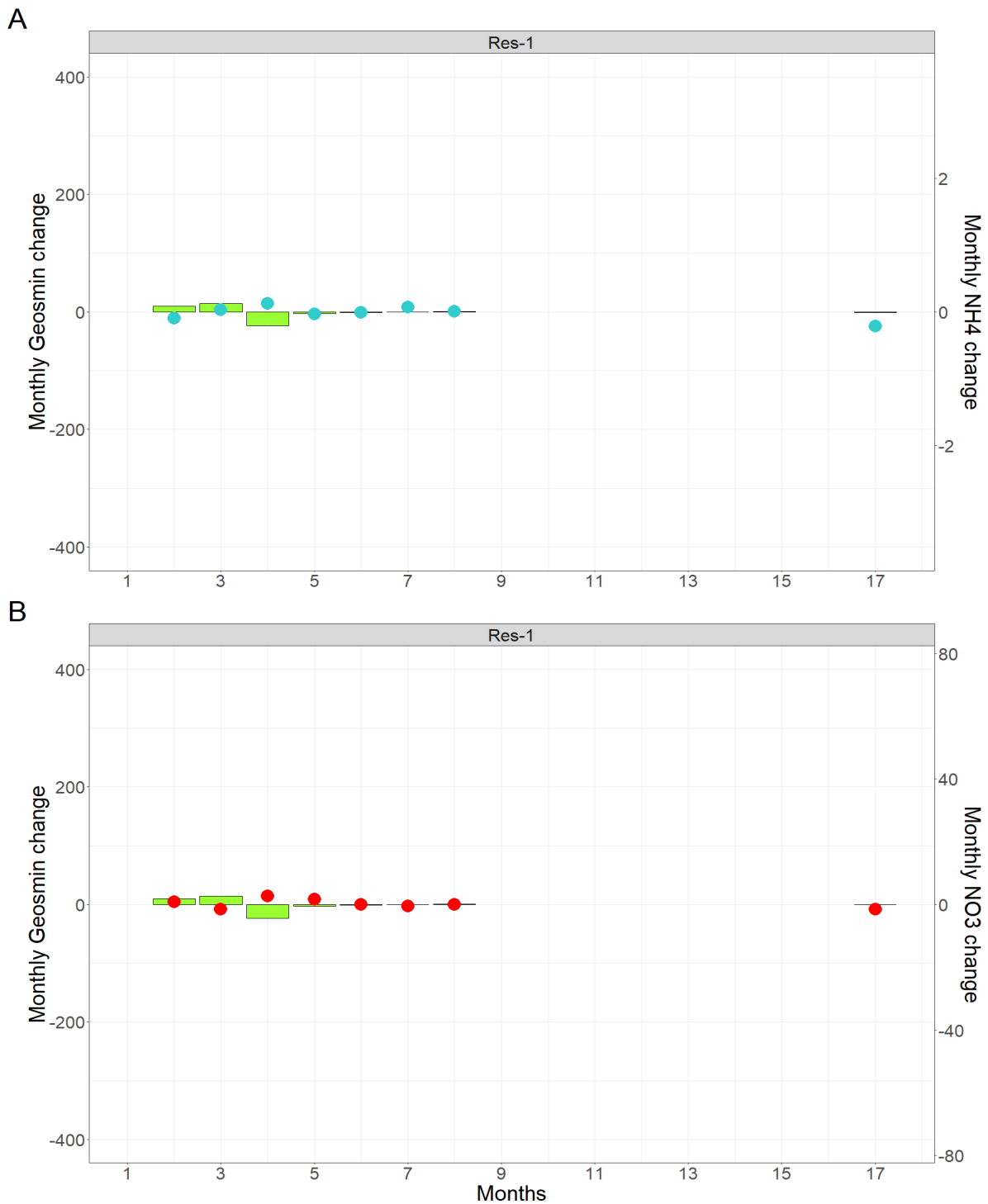


Appendix 3.1: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Cefni reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir’s monthly geosmin change.

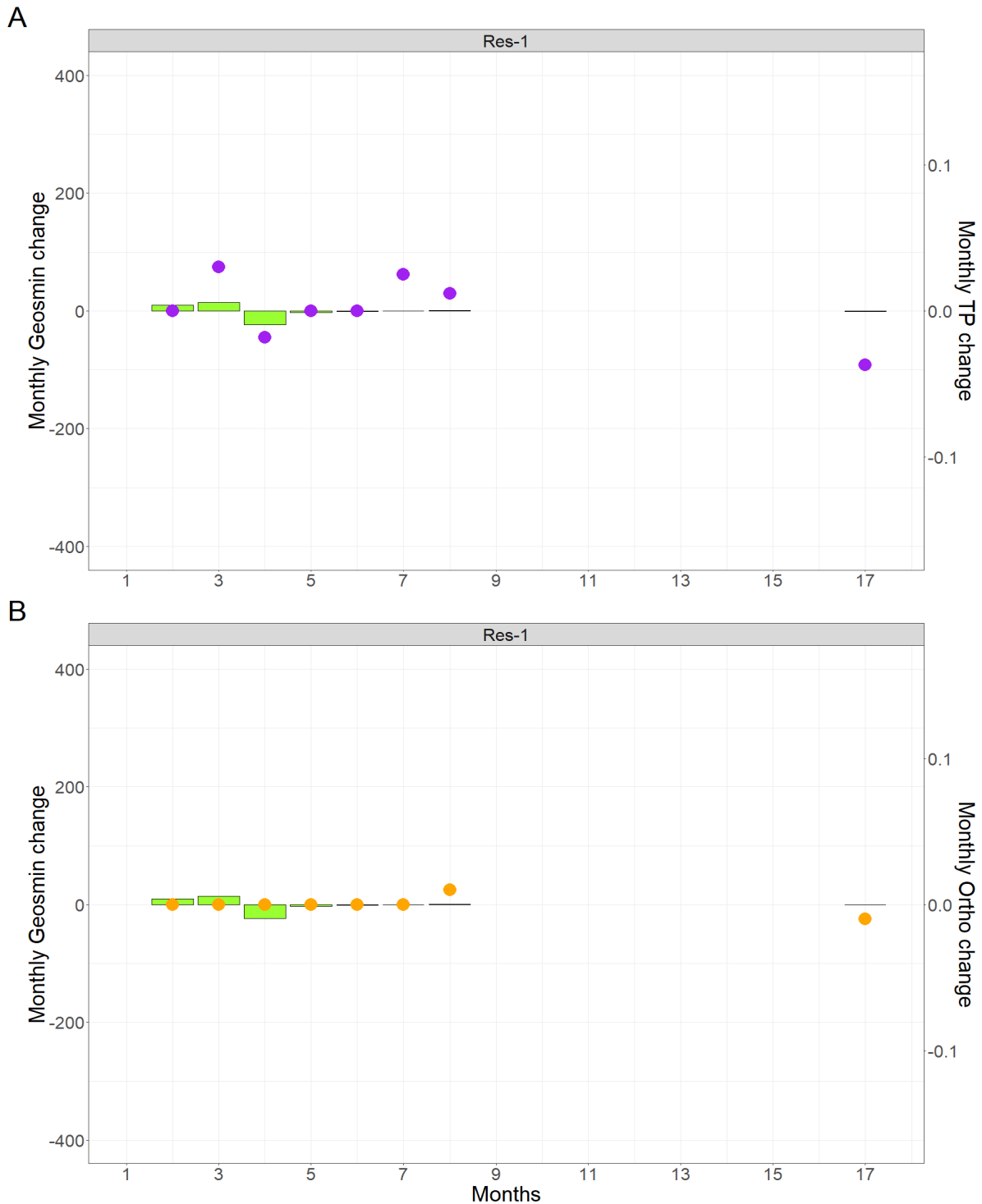


Appendix 3.2: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) and B – monthly changes in orthophosphate concentrations ( $\text{mg L}^{-1}$ ) (orange dots) given from each consecutive sampled month of the study period in Cefni reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir's monthly geosmin change.

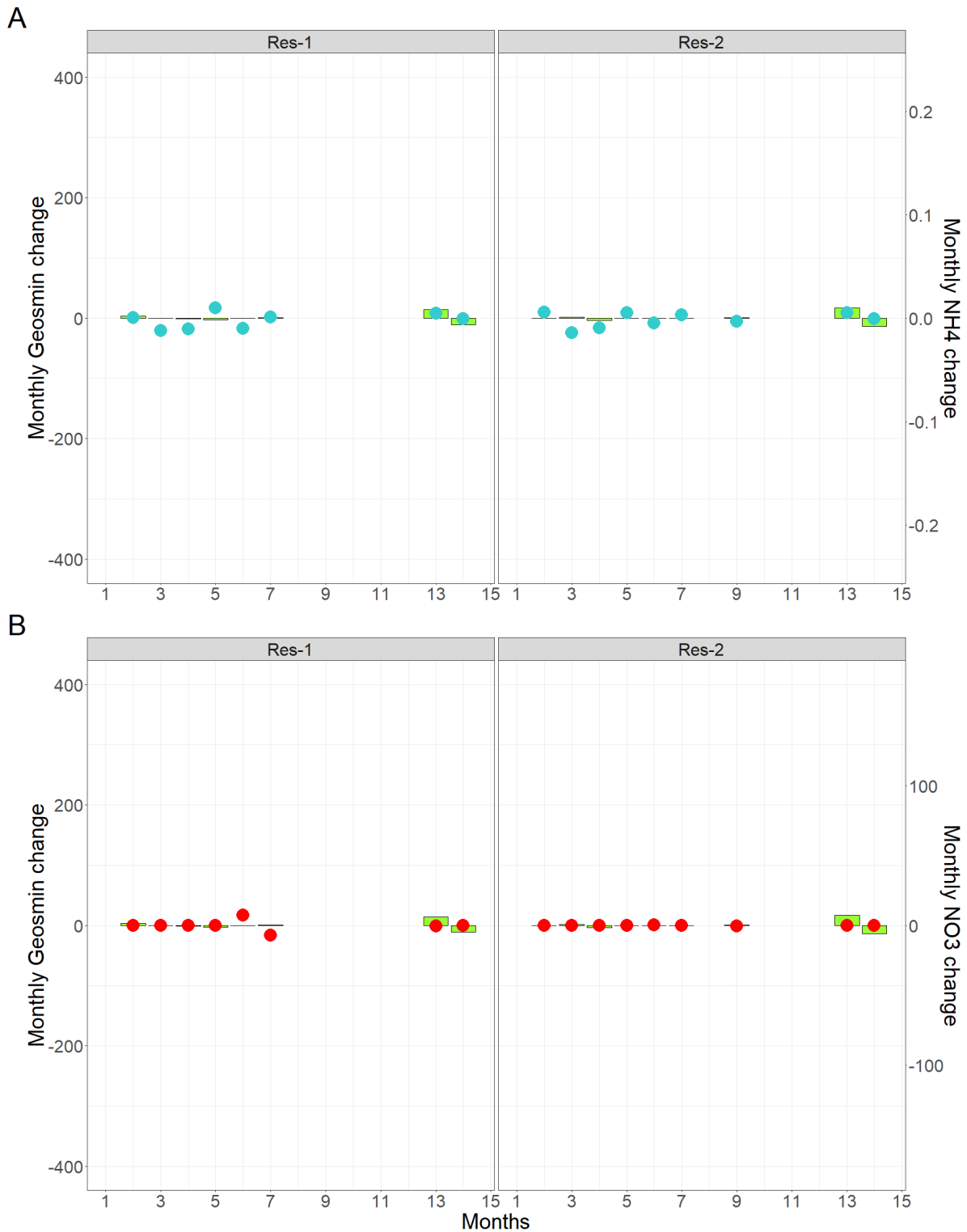




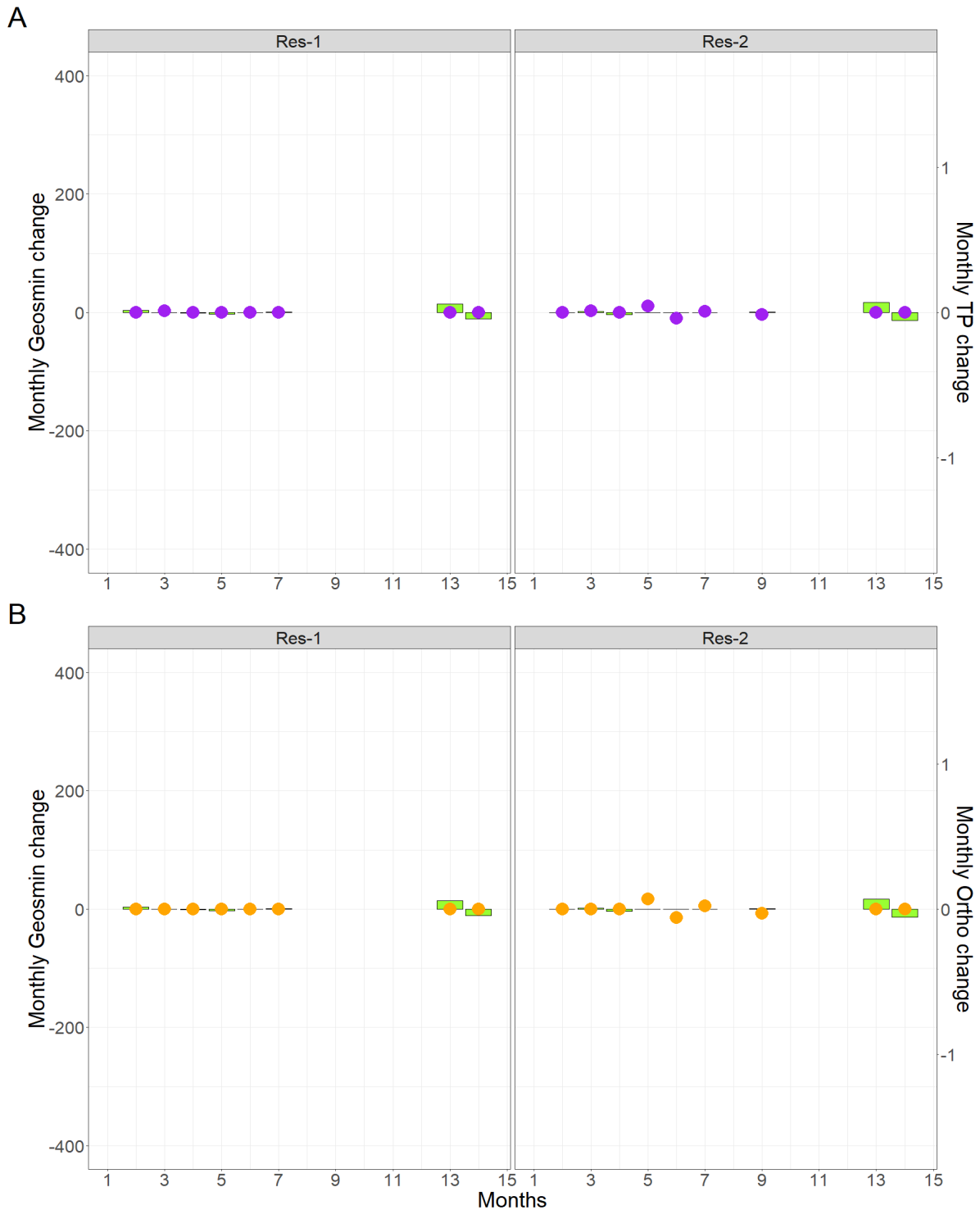
Appendix 3.3: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in NH<sub>4</sub><sup>+</sup> concentrations (mg L<sup>-1</sup>) (blue dots) and B – monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (mg L<sup>-1</sup>) (red dots) given from each consecutive sampled month of the study period in Dolwen reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir's monthly geosmin change.



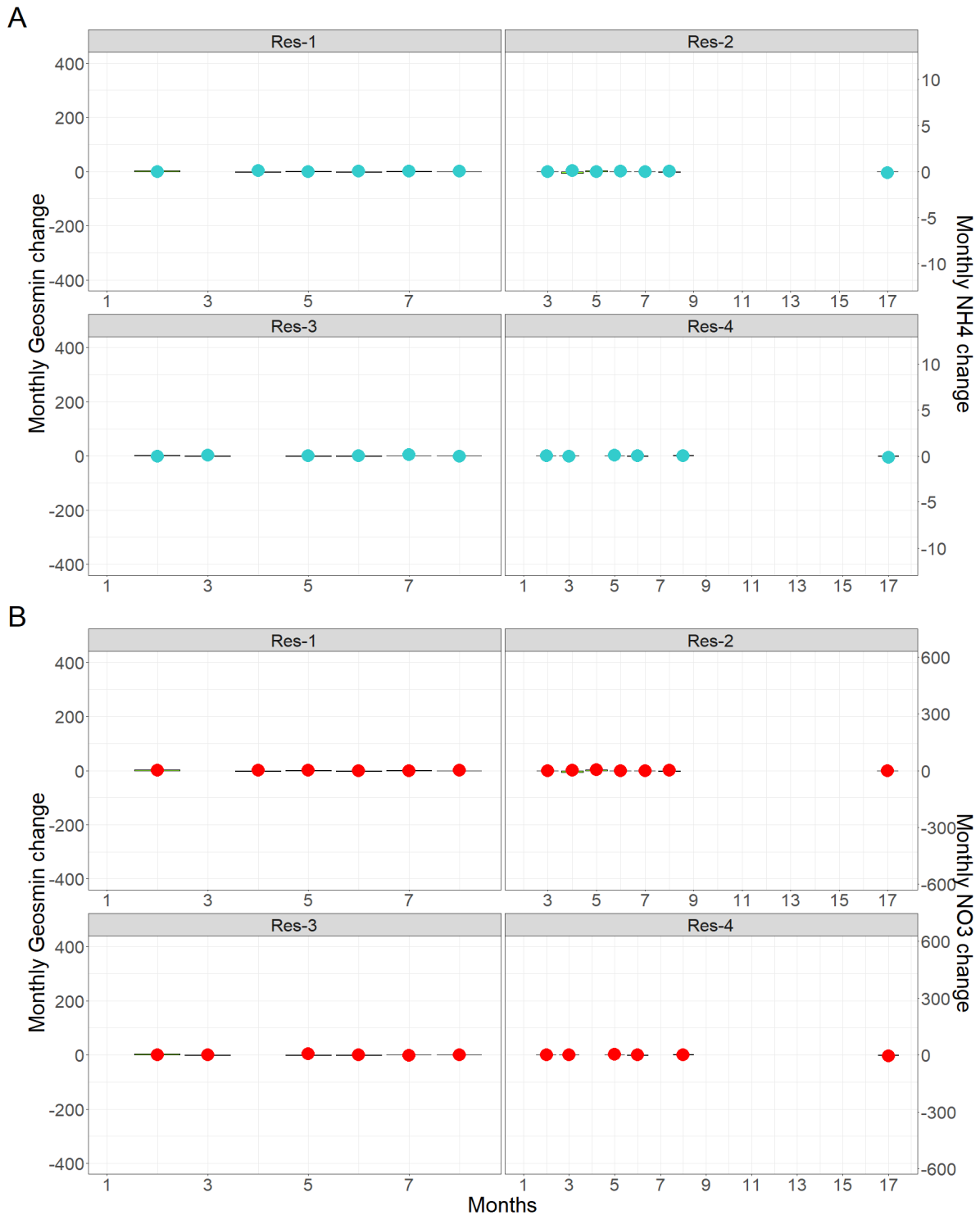
Appendix 3.4: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) and B – monthly changes in orthophosphate concentrations (mg L<sup>-1</sup>) (orange dots) given from each consecutive sampled month of the study period in Dolwen reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir’s monthly geosmin change.



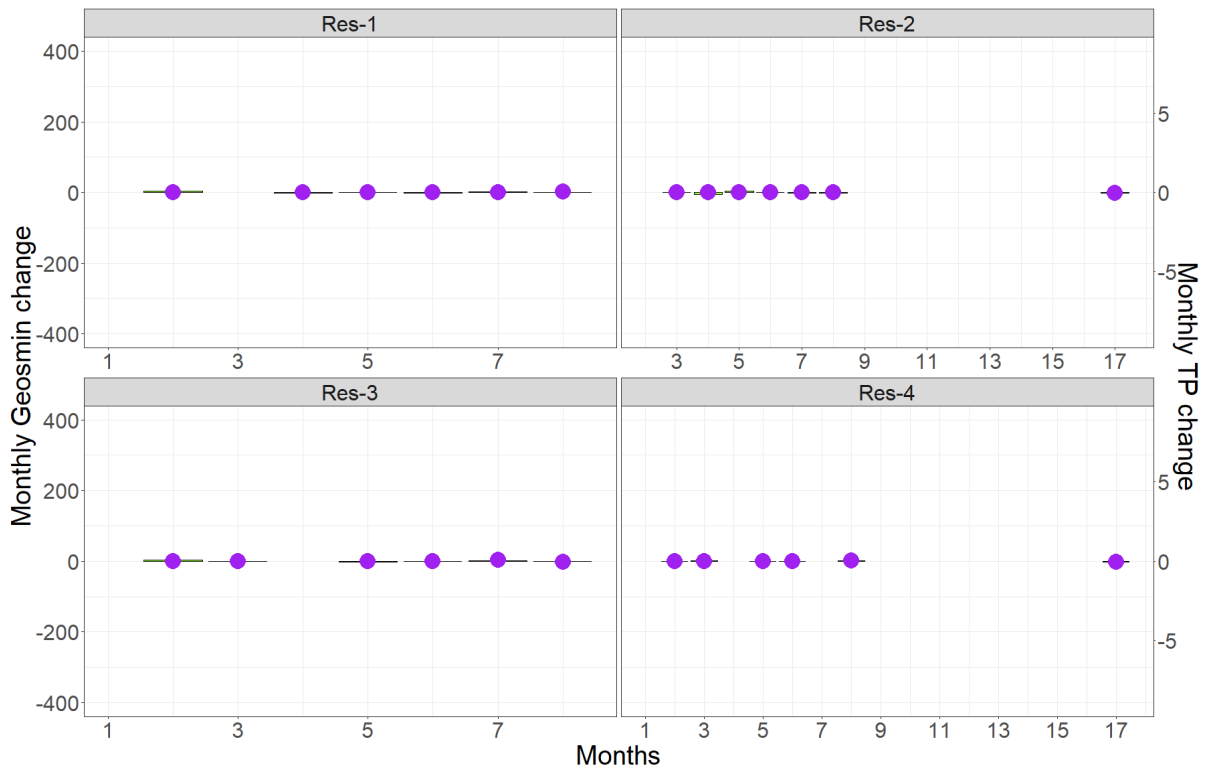
Appendix 3.5: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in  $\text{NH}_4^+$  concentrations ( $\text{mg L}^{-1}$ ) (blue dots) and B – monthly changes in  $\text{NO}_3^-$  concentrations ( $\text{mg L}^{-1}$ ) (red dots) given from each consecutive sampled month of the study period in Llwyn On reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir’s monthly geosmin change.



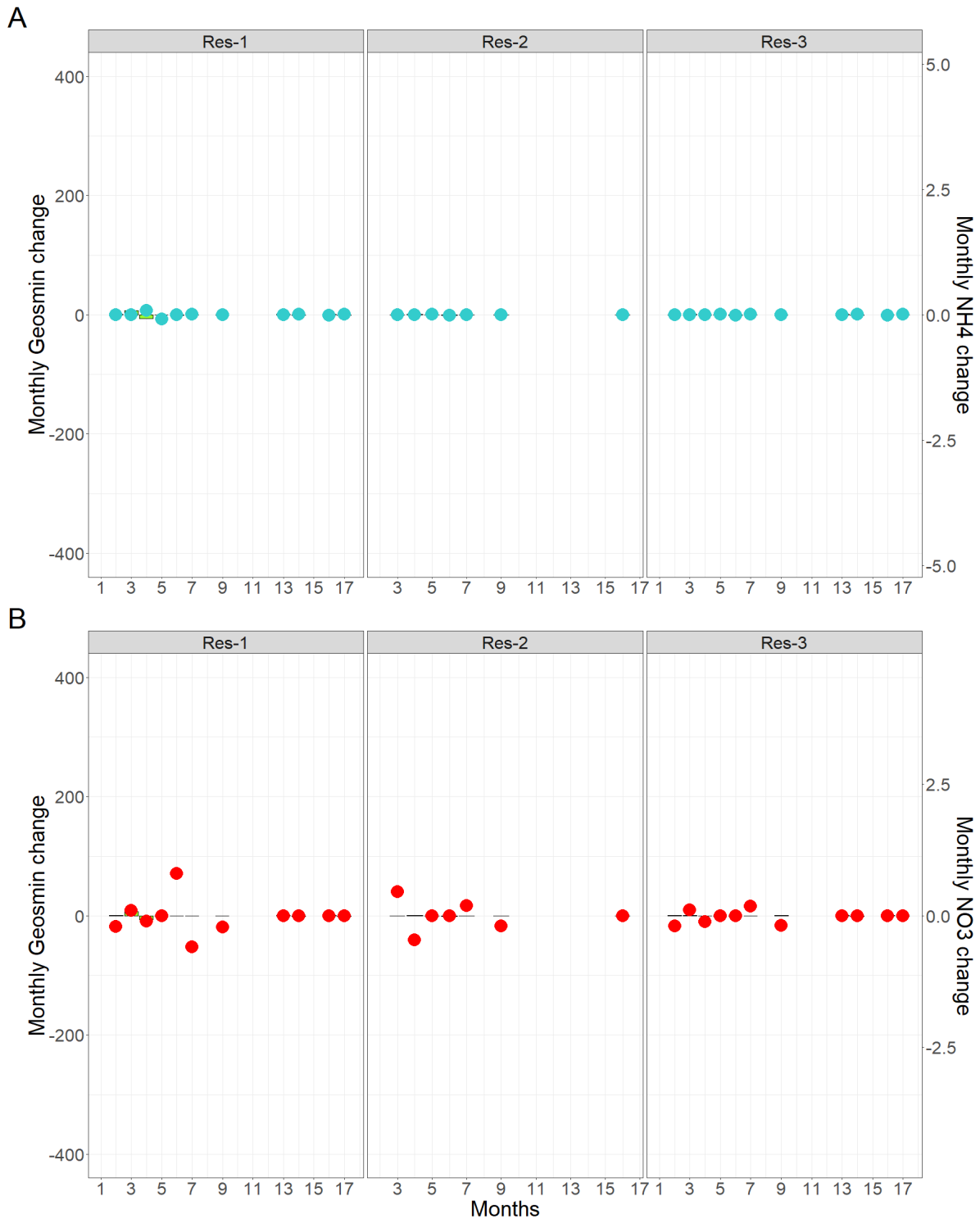
Appendix 3.6: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) and B – monthly changes in orthophosphate concentrations ( $\text{mg L}^{-1}$ ) (orange dots) given from each consecutive sampled month of the study period in Llwyn On reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir's monthly geosmin change.



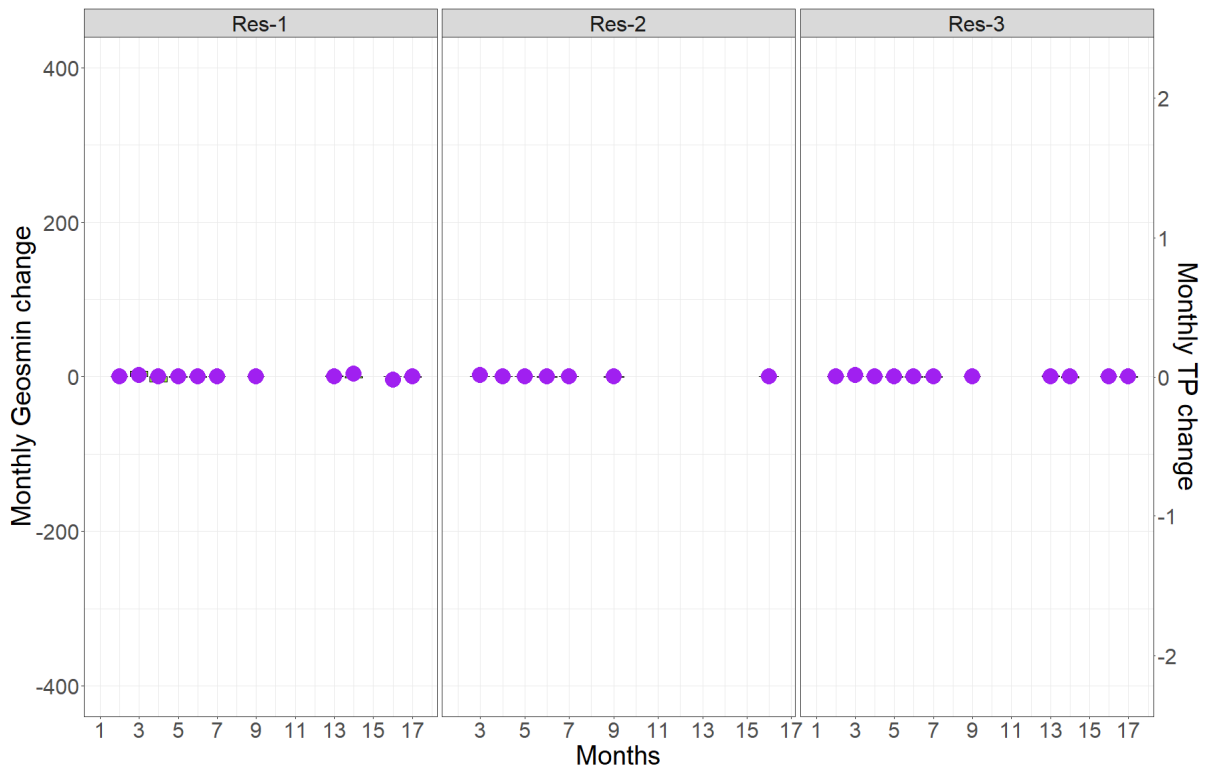
Appendix 3.7: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in NH<sub>4</sub><sup>+</sup> concentrations (mg L<sup>-1</sup>) (blue dots) and B – monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (mg L<sup>-1</sup>) (red dots) given from each consecutive sampled month of the study period in Plas Uchaf reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir’s monthly geosmin change.



Appendix 3.8: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in total phosphorous (TP) concentrations (mg L<sup>-1</sup>) (purple dots) given from each consecutive sampled month of the study period in Plas Uchaf reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir’s monthly geosmin change.

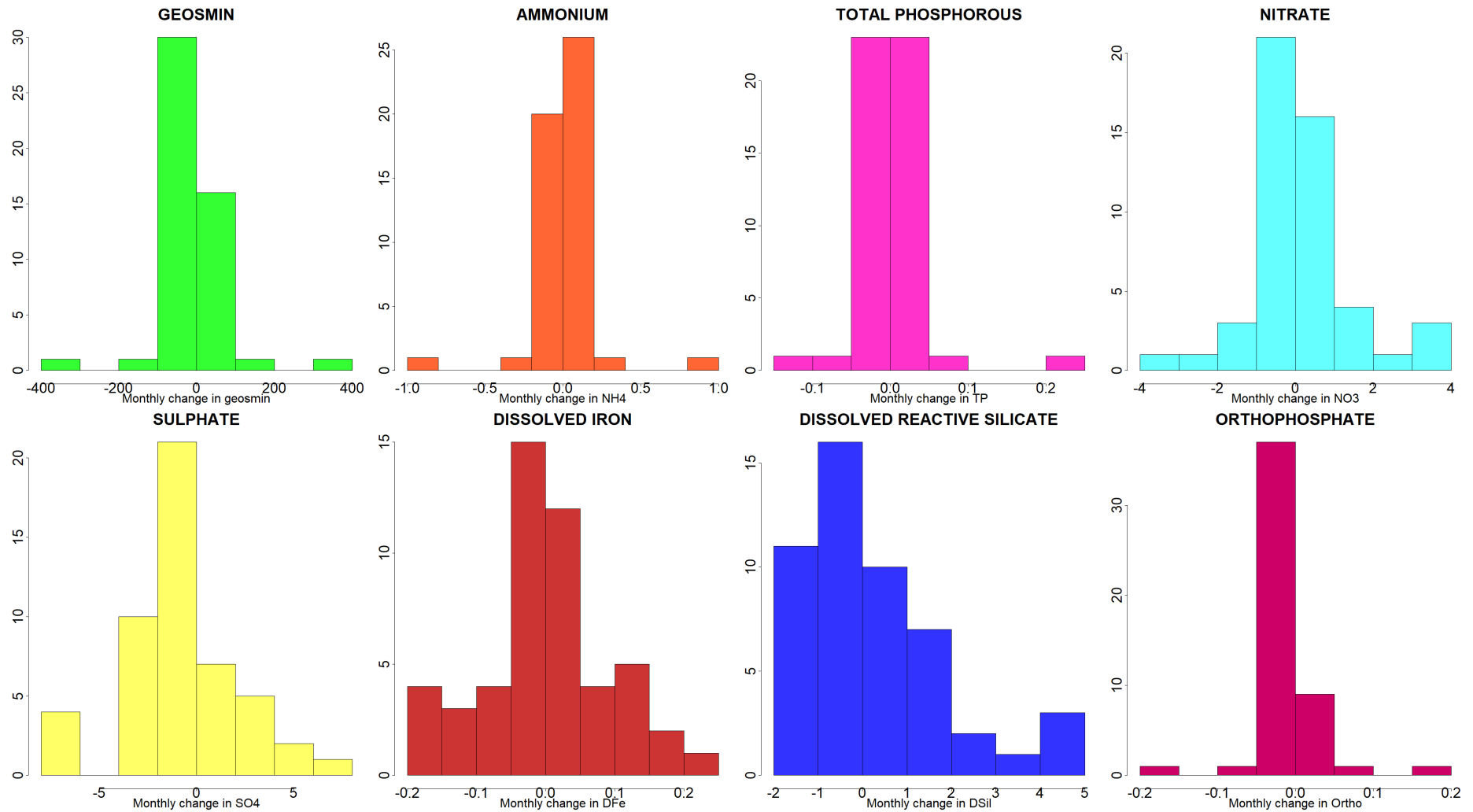


Appendix 3.9: Monthly changes in geosmin concentrations (ng L<sup>-1</sup>) (green bars) and A – monthly changes in NH<sub>4</sub><sup>+</sup> concentrations (mg L<sup>-1</sup>) (blue dots) and B – monthly changes in NO<sub>3</sub><sup>-</sup> concentrations (mg L<sup>-1</sup>) (red dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir’s monthly geosmin change.

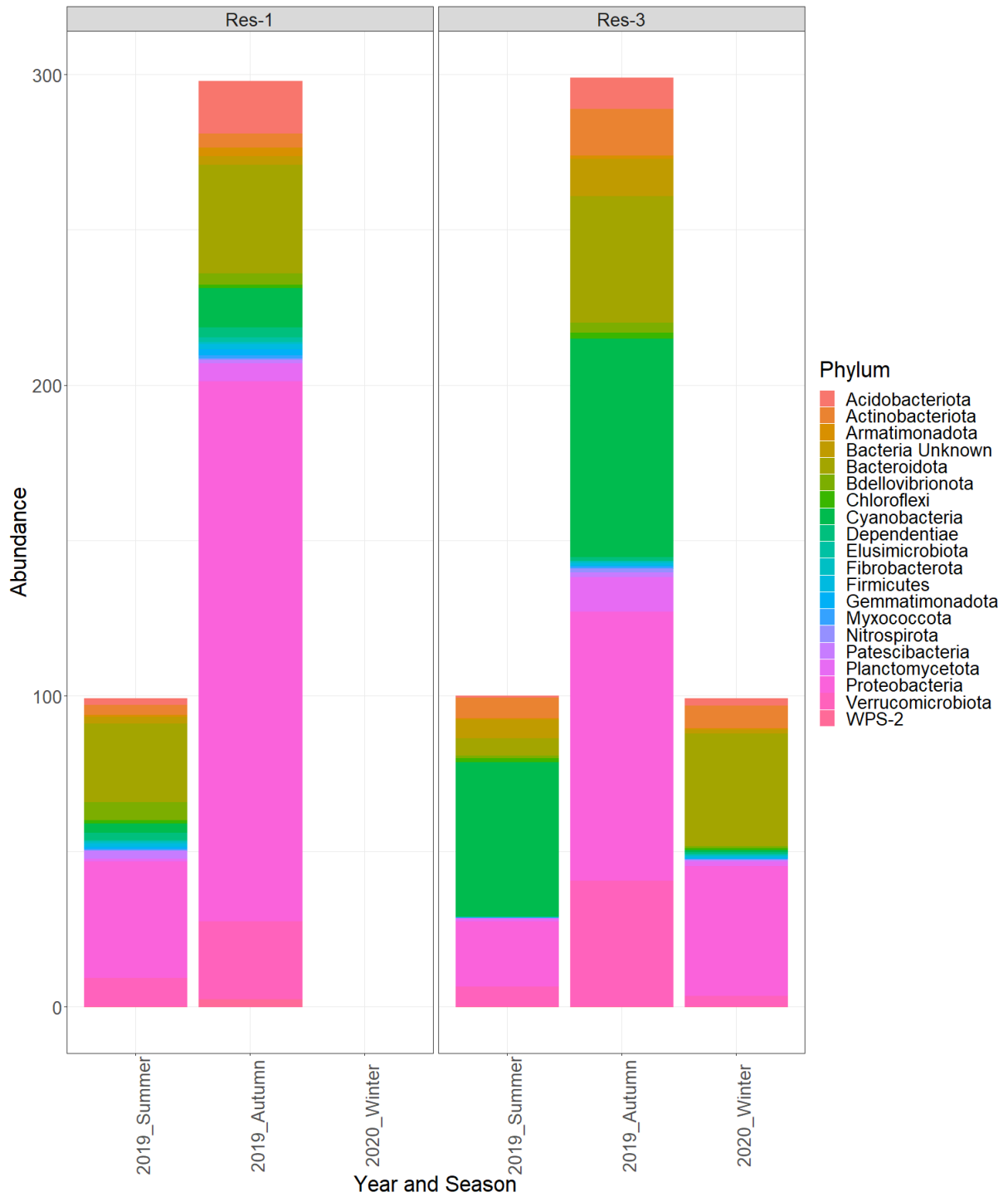


Appendix 3.10: Monthly changes in geosmin concentrations ( $\text{ng L}^{-1}$ ) (green bars) and A – monthly changes in total phosphorous (TP) concentrations ( $\text{mg L}^{-1}$ ) (purple dots) given from each consecutive sampled month of the study period in Pontsticill reservoir locations. Monthly geosmin change on the left hand y-axis scaled to Alaw reservoir’s monthly geosmin change.

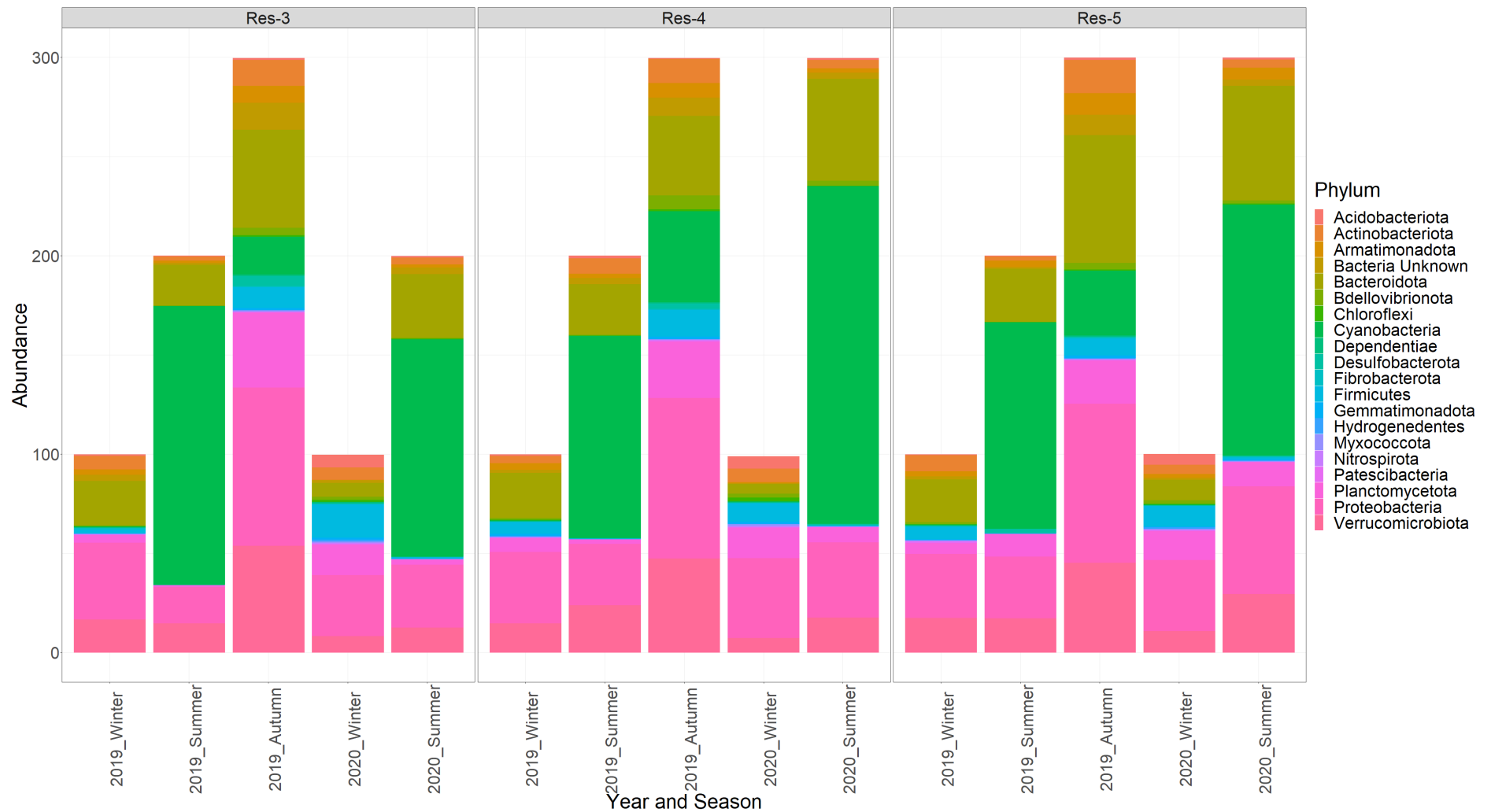




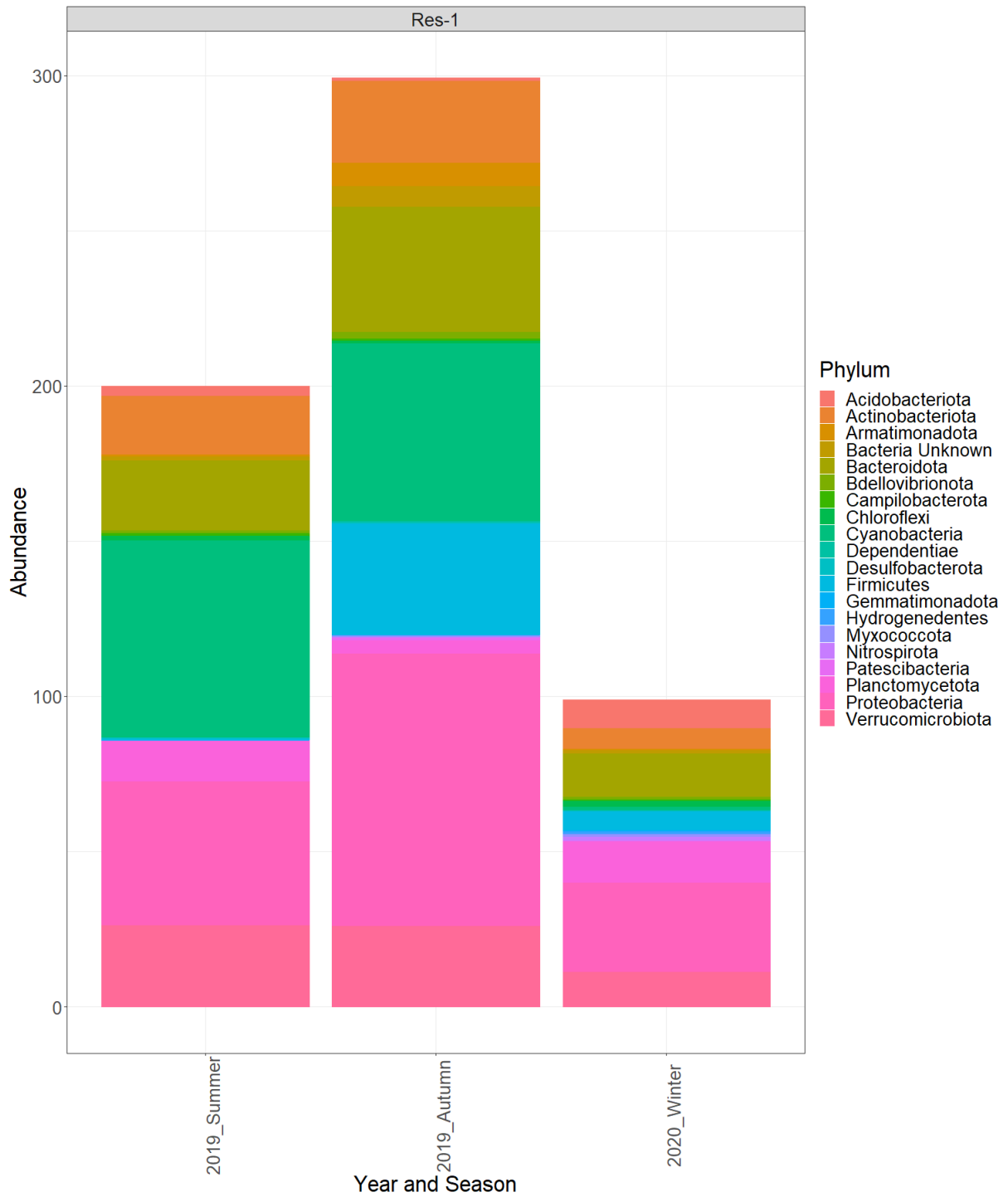
Appendix 3.11: Histogram of monthly changes in nutrients and geosmin concentrations from all reservoirs, showing distribution of data. The y-axis depicts the frequency.



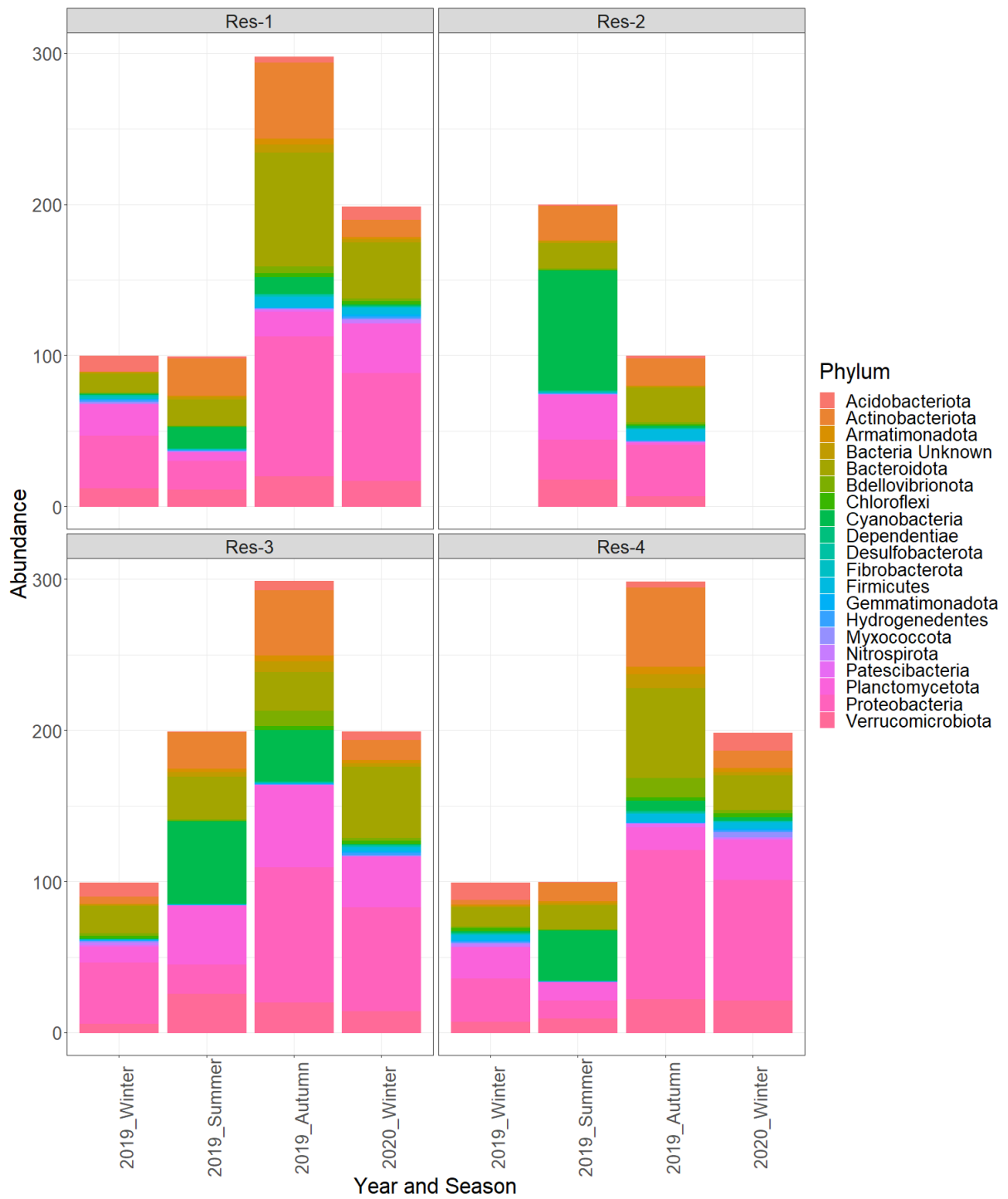
Appendix 4.1: Alwen top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.



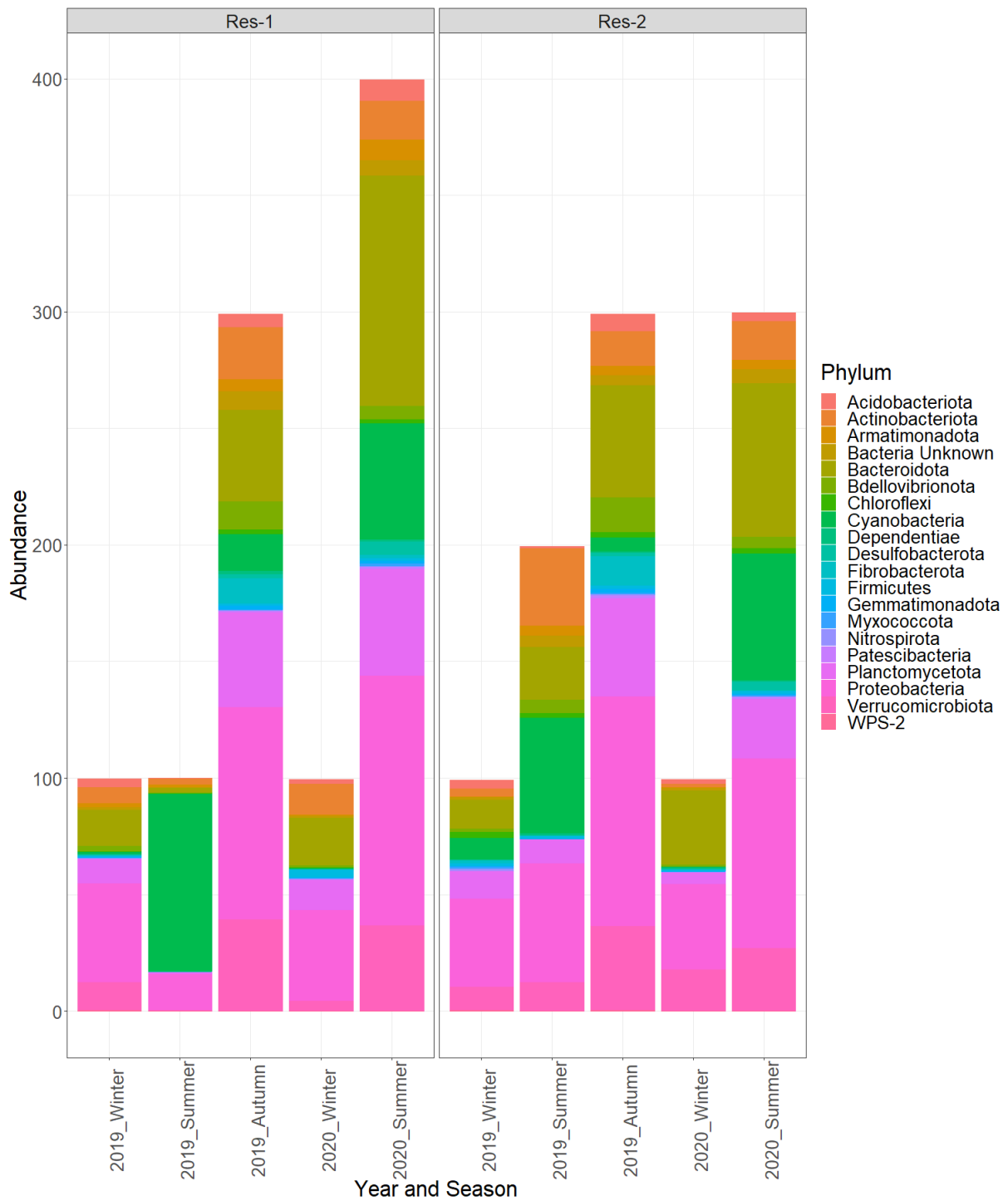
Appendix 4.2: Cefni top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.



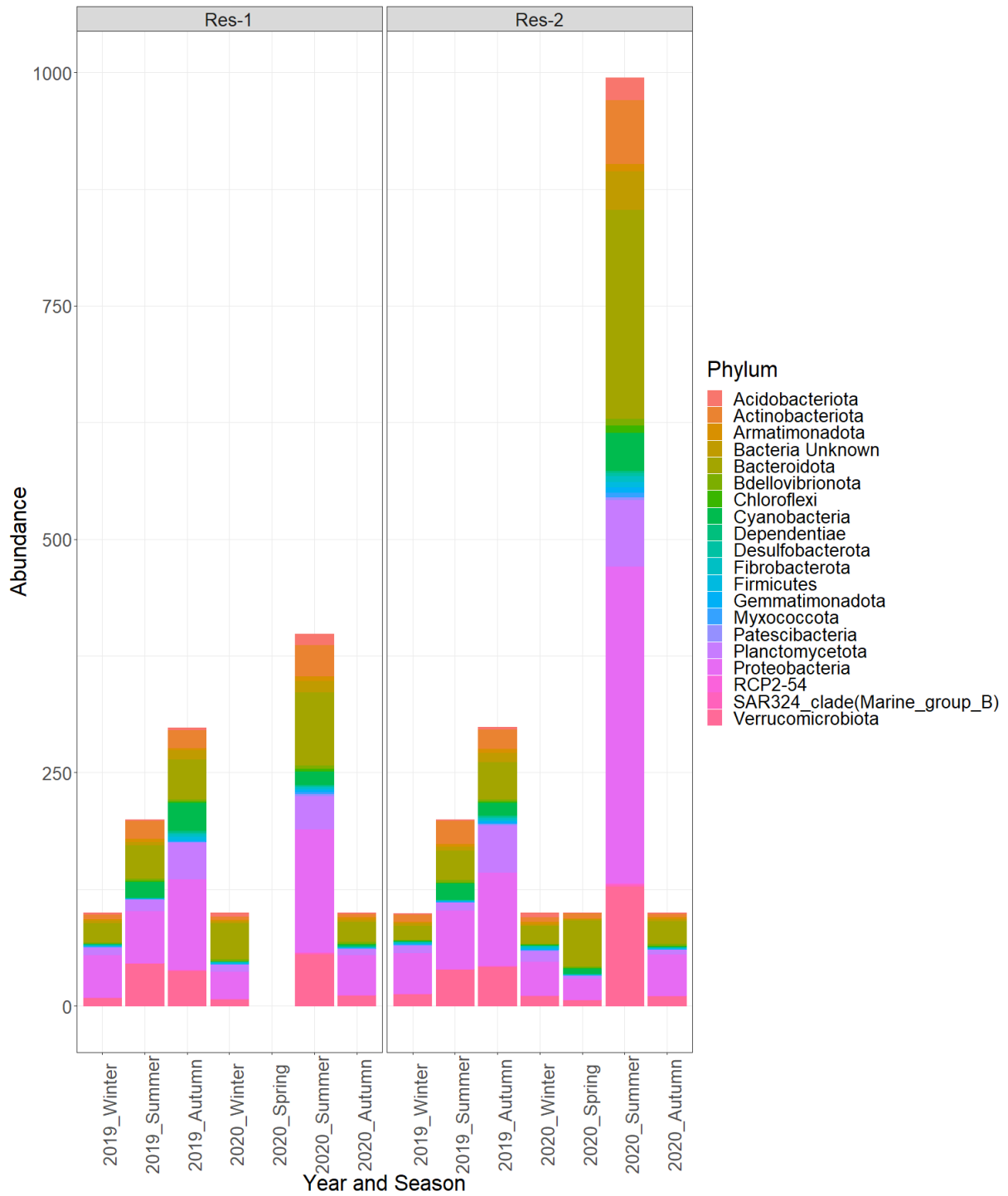
Appendix 4.3: Dolwen top 20 16S rRNA phyla abundance by year and season.



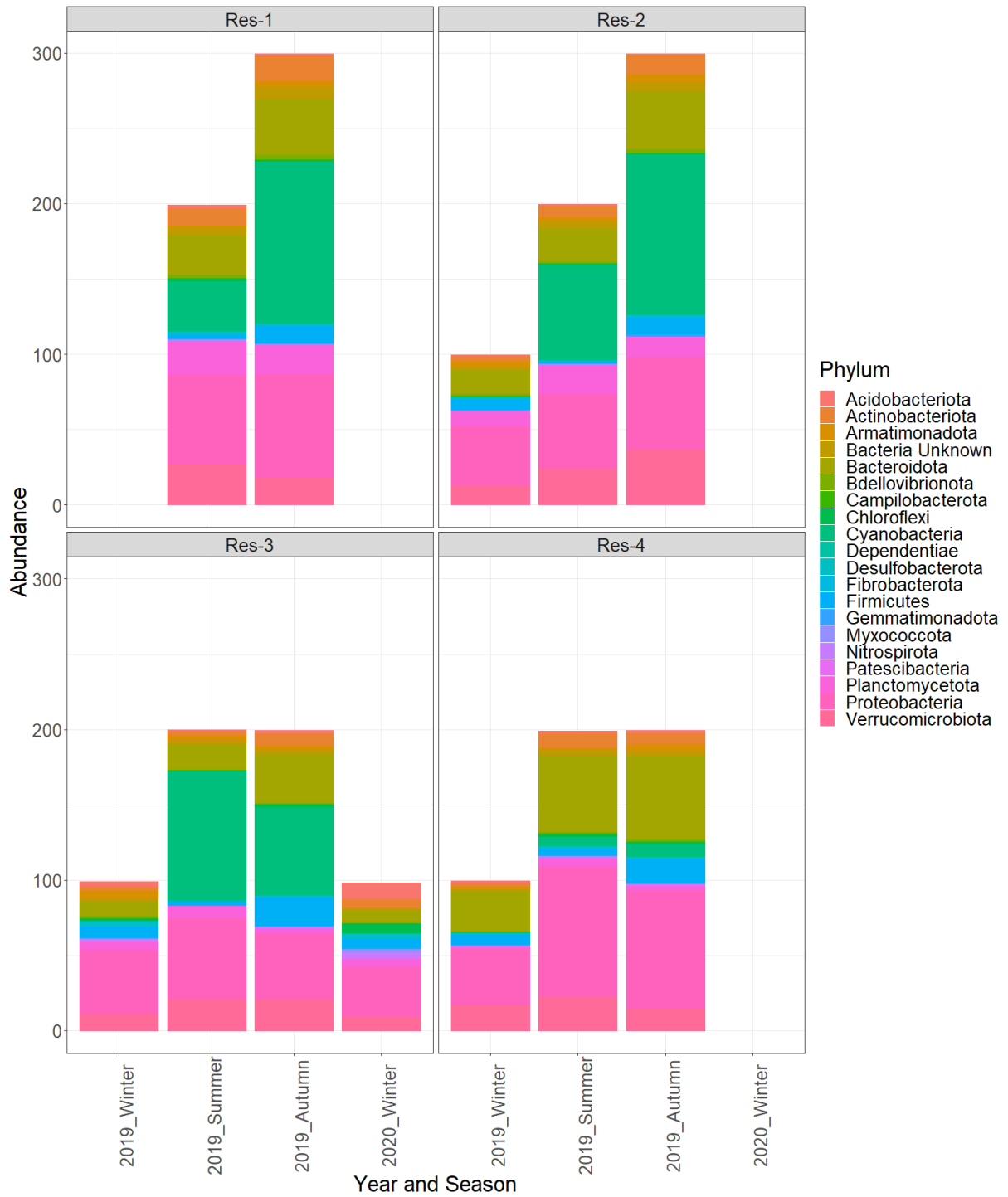
Appendix 4.4: Llandgefedd top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.



Appendix 4.5: Llwyn On top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.

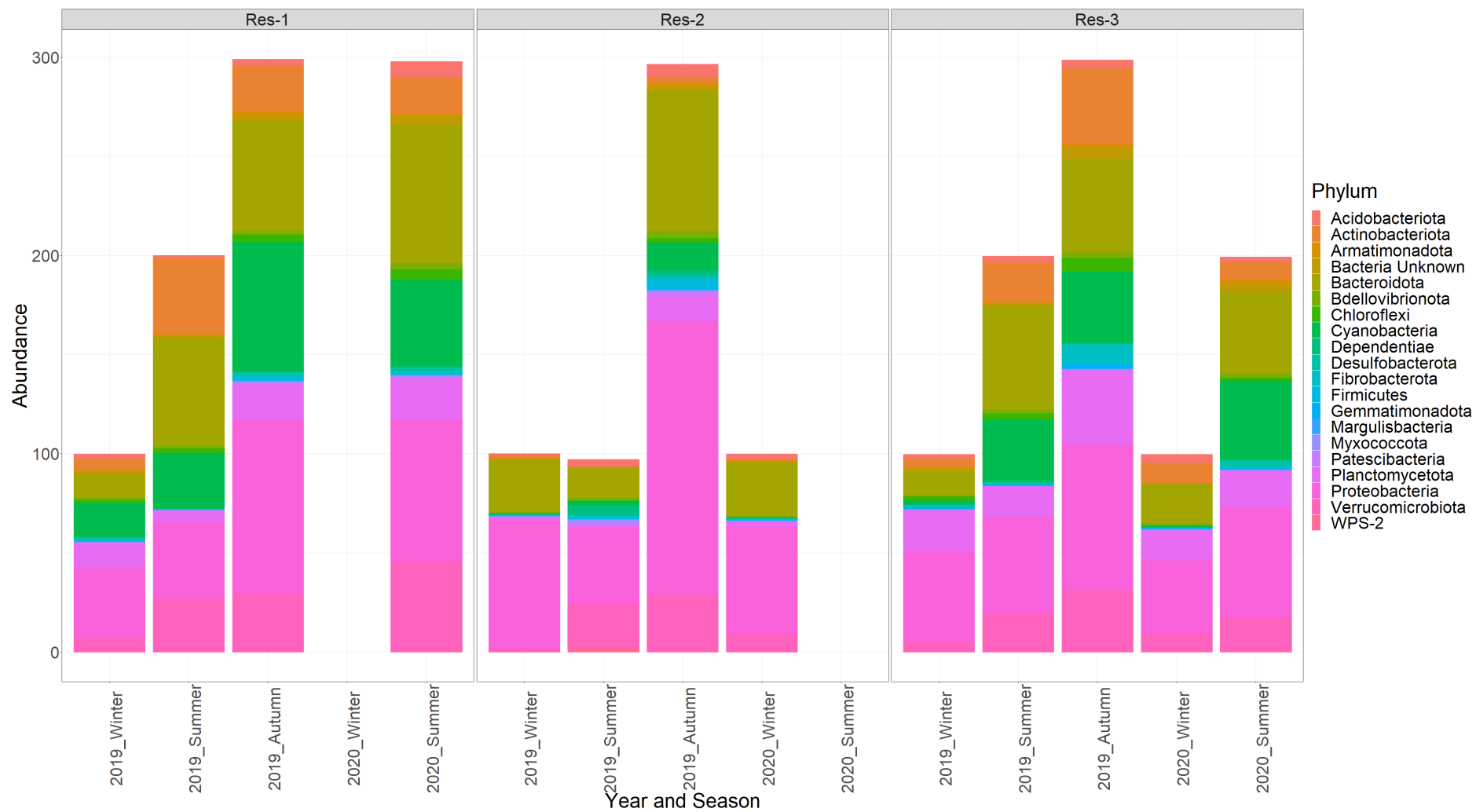


Appendix 4.6: Penwryn top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.



Appendix 4.7: Plas Uchaf top 20 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.

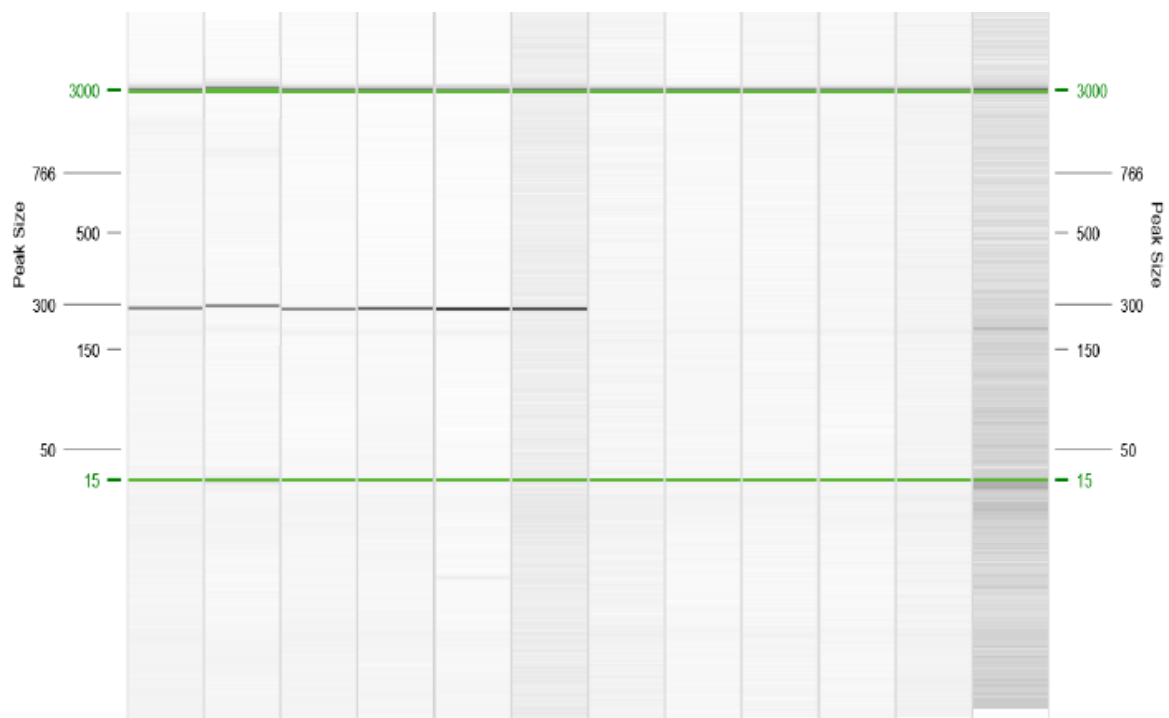




Appendix 4.8: Pontsticill 16S rRNA phyla abundance by year and season, facet wrapped by sample site location.



Appendix 5.1: QIAxcel screenshot of six replicates of PCR 16S rRNA product at approximately 1490 bp (lanes 1-6) and six replicates of PCR *geoA* product at approximately 905 bp (lanes 7-12) with 15 bp – 3000 bp alignment markers.



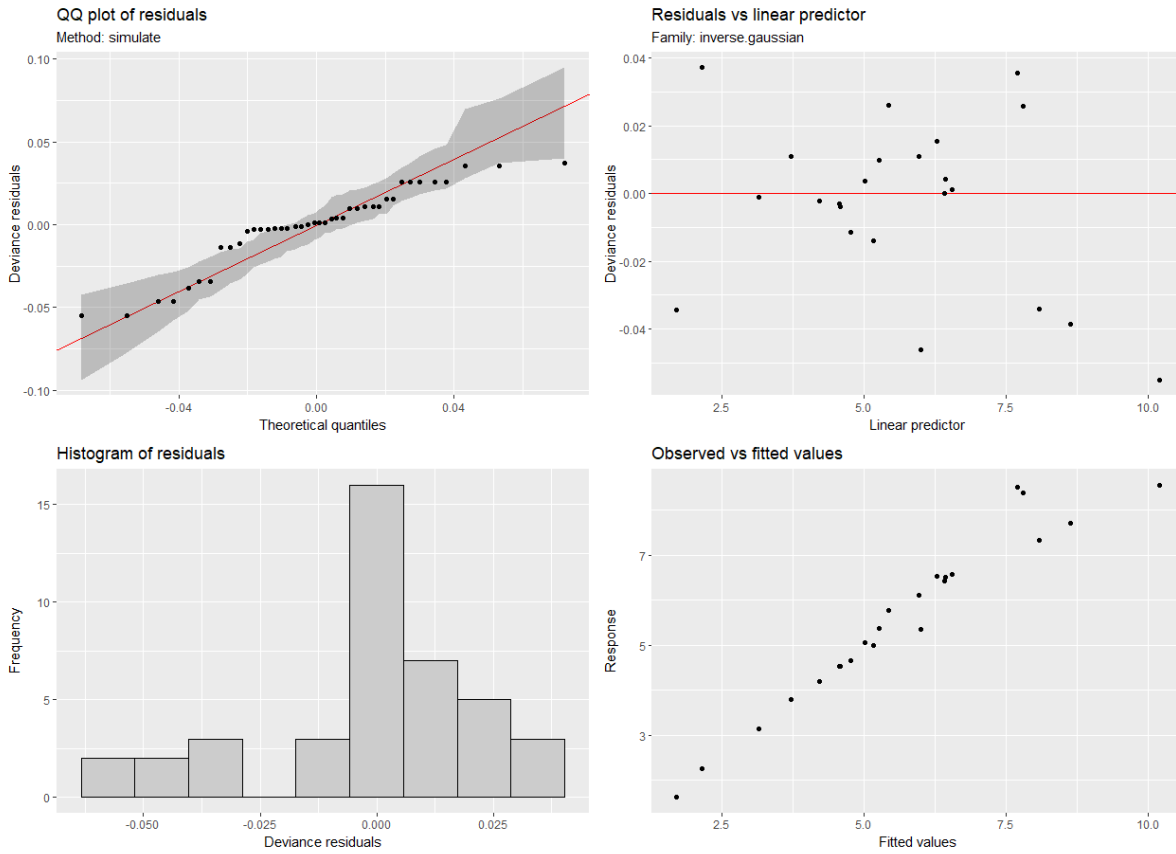
SI-5.2: QIAxcel screenshot of six replicates of PCR *mic* product at approximately 307 bp (lanes 1-6) with 15 bp – 3000 bp alignment markers and 6 blanks (lanes 7-12) .

Appendix 5.3: A list of all 30 species that can be detected by the primers MIBS02F and MIBAHR from the statistical software R using the package primerTree on 12/02/2021.

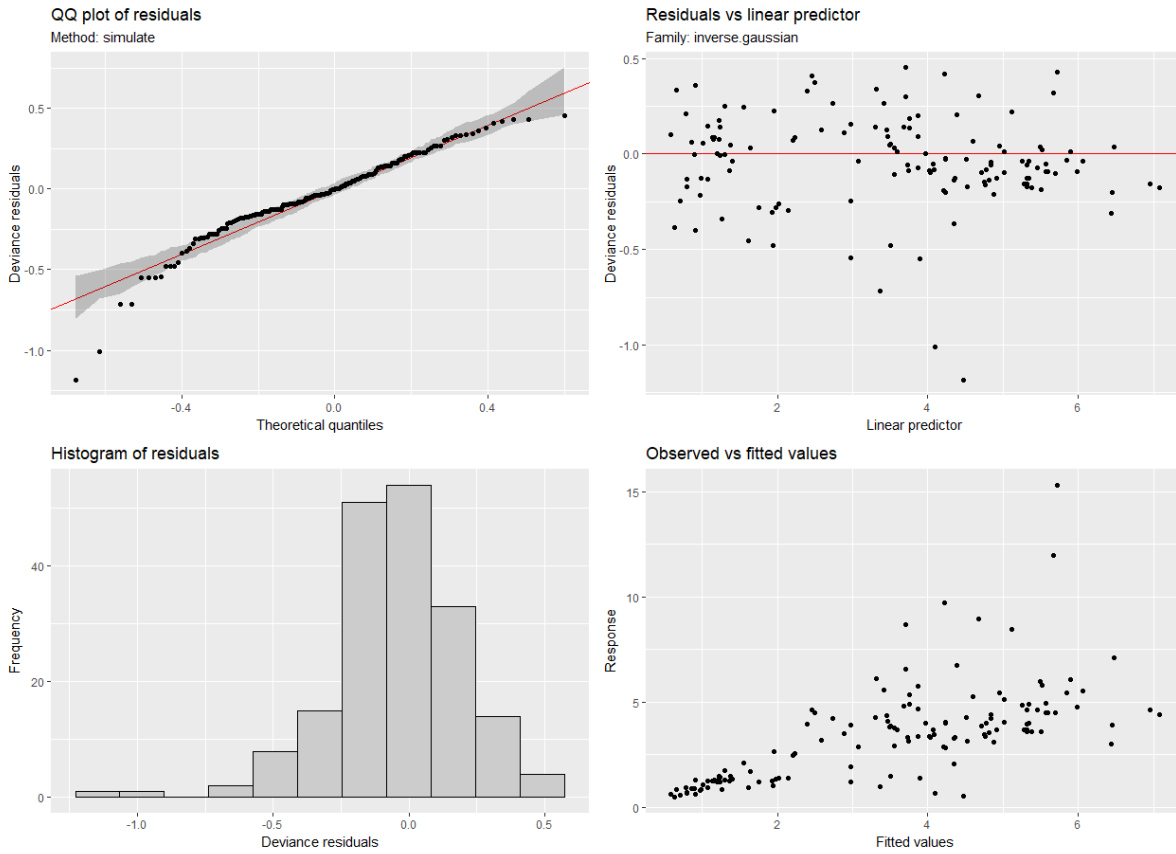
| Species                                       |
|---|
| <i>Planktothricoides raciborskii</i> CHAB3331 |
| <i>Pseudanabaena</i> sp. dqh15                |
| <i>Pseudanabaena galeata</i>                  |
| <i>Planktothricoides raciborskii</i>          |
| <i>Planktothricoides raciborskii</i>          |
| <i>Planktothricoides raciborskii</i>          |
| <i>Planktothricoides raciborskii</i>          |
| <i>Planktothricoides raciborskii</i>          |
| <i>Planktothricoides raciborskii</i>          |
| <i>Planktothricoides raciborskii</i>          |
| <i>Planktothricoides raciborskii</i>          |
| <i>Leptolyngbya</i> sp. A2                    |
| <i>Leptolyngbya</i> sp. A2                    |
| <i>Planktothricoides</i> sp. TH3              |
| <i>Planktothricoides raciborskii</i> CHAB3332 |
| <i>Planktothrix</i> sp. 328                   |
| <i>Pseudanabaena cinerea</i> NIES-4062        |
| <i>Pseudanabaena cinerea</i> NIES-4063        |
| <i>Pseudanabaena yagii</i> NIES-4237          |
| <i>Pseudanabaena yagii</i> NIES-4238          |
| <i>Oscillatoria prolifera</i> OpAlghanmi-77   |
| <i>Microcoleus pseudautumnalis</i> Ak1609     |
| <i>Pseudanabaena galeata</i> NRERC-312        |
| <i>Pseudanabaena</i> sp. AIFI-4               |
| <i>Pseudanabaena yagii</i> NIVA-CYA 111       |
| <i>Pseudanabaena</i> sp. PD8                  |
| <i>Planktothrix</i> sp. PD12                  |
| <i>Oscillatoria</i> sp. 327/2                 |
| <i>Pseudanabaena limnetica</i>                |
| <i>Oscillatoria limosa</i> LBD 305b           |

Appendix 5.4: A list of all species identified in the statistical software programme R using the package primerTree when comparing different published primers on 12/02/2021.

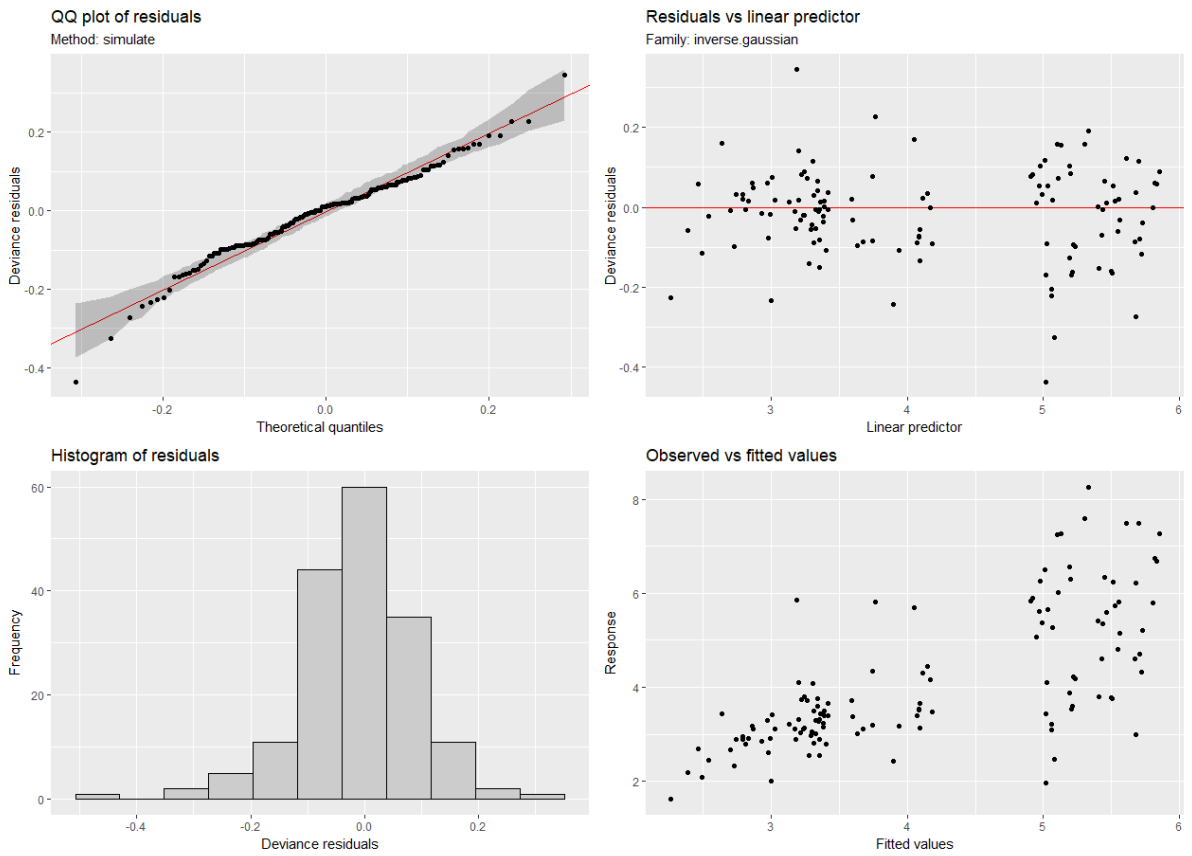
| Species identified                               |  |   |   |
|--|--|---|---|
| MIBS02F and MIBS02R<br>(Chiu et al., 2016)       | Mtcf and Mtrc<br>(Wang et al., 2011)             | MIB3324F and MIB4050R<br>(Suurnäkki et al., 2015) | MIB3313F and MIB4226R<br>(Suurnäkki et al., 2015) |
| <i>Planktothricoides raciborskii</i><br>CHAB3331 | <i>Planktothricoides raciborskii</i><br>CHAB3331 | <i>Planktothricoides raciborskii</i><br>CHAB3331  | <i>Planktothricoides raciborskii</i><br>CHAB3331  |
| <i>Pseudanabaena</i> sp. dqh15                   | <i>Pseudanabaena</i> sp. dqh15                   | <i>Pseudanabaena</i> sp. dqh15                    | <i>Pseudanabaena</i> sp. dqh15                    |
| <i>Pseudanabaena galeata</i>                     | <i>Pseudanabaena galeata</i>                     | <i>Pseudanabaena galeata</i>                      | <i>Pseudanabaena galeata</i>                      |
| <i>Planktothricoides raciborskii</i>             | <i>Pseudanabaena cinerea</i> NIES-4062           | <i>Planktothricoides raciborskii</i>              | <i>Microcoleus pseudautumnalis</i> Ak1609         |
| <i>Planktothricoides raciborskii</i>             | <i>Pseudanabaena cinerea</i> NIES-4063           | <i>Planktothricoides raciborskii</i>              | <i>Pseudanabaena yagii</i> NIVA-CYA 111           |
| <i>Planktothricoides raciborskii</i>             | <i>Pseudanabaena yagii</i> NIES-4237             | <i>Planktothricoides raciborskii</i>              | <i>Pseudanabaena limnetica</i>                    |
| <i>Planktothricoides raciborskii</i>             | <i>Pseudanabaena yagii</i> NIES-4238             | <i>Planktothricoides raciborskii</i>              | <i>Oscillatoria limosa</i> LBD 305b               |
| <i>Planktothricoides raciborskii</i>             | <i>Microcoleus pseudautumnalis</i> Ak1609        | <i>Planktothricoides raciborskii</i>              |   |
| <i>Planktothricoides raciborskii</i>             | <i>Pseudanabaena</i> sp. AIFI-4                  | <i>Planktothricoides raciborskii</i>              |   |
| <i>Planktothricoides raciborskii</i>             | <i>Pseudanabaena yagii</i> NIVA-CYA 111          | <i>Planktothricoides raciborskii</i>              |   |
| <i>Planktothricoides raciborskii</i>             | <i>Pseudanabaena limnetica</i>                   | <i>Leptolyngbya</i> sp. A2                        |   |
| <i>Leptolyngbya</i> sp. A2                       | <i>Oscillatoria limosa</i> LBD 305b              | <i>Planktothrix</i> sp. 328                       |   |
| <i>Leptolyngbya</i> sp. A2                       | <i>Cyaniris semiargus</i>                        | <i>Microcoleus pseudautumnalis</i> Ak1609         |   |
| <i>Planktothricoides</i> sp. TH3                 |  | <i>Pseudanabaena galeata</i> NRERC-312            |   |
| <i>Planktothricoides raciborskii</i><br>CHAB3332 |  | <i>Pseudanabaena yagii</i> NIVA-CYA 111           |   |
| <i>Planktothrix</i> sp. 328                      |  | <i>Pseudanabaena limnetica</i>                    |   |
| <i>Oscillatoria prolifera</i> OpAlghanmi-77      |  | <i>Oscillatoria limosa</i> LBD 305b               |   |
| <i>Microcoleus pseudautumnalis</i> Ak1609        |  |   |   |
| <i>Pseudanabaena galeata</i> NRERC-312           |  |   |   |
| <i>Pseudanabaena yagii</i> NIVA-CYA 111          |  |   |   |
| <i>Oscillatoria</i> sp. 327/2                    |  |   |   |
| <i>Pseudanabaena limnetica</i>                   |  |   |   |
| <i>Oscillatoria limosa</i> LBD 305b              |  |   |   |



Appendix 5.5: Alaw *geoA*:16S copy numbers mL<sup>-1</sup> GAM diagnostic plot. Top left includes the normal Q-Q plot, top right is the residuals vs. fitted values, bottom left is the histogram of residuals and bottom right is the response vs. fitted values.



Appendix 5.6: All reservoirs except Alaw *geoA*:16S copy numbers mL<sup>-1</sup> GAM diagnostic plot. Top left includes the normal Q-Q plot, top right is the residuals vs. fitted values, bottom left is the histogram of residuals and bottom right is the response vs. fitted values.



Appendix 5.7: All reservoirs *mic:16S* copy numbers mL<sup>-1</sup> GAM diagnostic plot. Top left includes the normal Q-Q plot, top right is the residuals vs. fitted values, bottom left is the histogram of residuals and bottom right is the response vs. fitted values.