

Article

# Analysis and Modelling of Taste and Odour Events in a Shallow Subtropical Reservoir

Edoardo Bertone <sup>1,\*</sup> and Kelvin O'Halloran <sup>2</sup>

<sup>1</sup> Griffith School of Engineering, Griffith University, Parklands Drive, Southport, Queensland 4222, Australia

<sup>2</sup> Scientific Services and Data Systems, Seqwater, 117 Brisbane St., Ipswich, Queensland 4305, Australia; Kelvin.O'Halloran@seqwater.com.au

\* Correspondence: edoardo.bertone@griffithuni.edu.au; Tel.: +61-7-5552-9044

Academic Editors: Luigi Berardi and Daniele Laucelli

Received: 10 June 2016; Accepted: 15 August 2016; Published: 19 August 2016

**Abstract:** Understanding and predicting Taste and Odour events is as difficult as critical for drinking water treatment plants. Following a number of events in recent years, a comprehensive statistical analysis of data from Lake Tingalpa (Queensland, Australia) was conducted. Historical manual sampling data, as well as data remotely collected by a vertical profiler, were collected; regression analysis and self-organising maps were used to determine correlations between Taste and Odour compounds and potential input variables. Results showed that the predominant Taste and Odour compound was geosmin. Although one of the main predictors was the occurrence of cyanobacteria blooms, it was noticed that the cyanobacteria species was also critical. Additionally, water temperature, reservoir volume and oxidised nitrogen availability, were key inputs determining the occurrence and magnitude of the geosmin peak events. Based on the results of the statistical analysis, a predictive regression model was developed to provide indications on the potential occurrence, and magnitude, of peaks in geosmin concentration. Additionally, it was found that the blue green algae probe of the lake's vertical profiler has the potential to be used as one of the inputs for an automated geosmin early warning system.

**Keywords:** geosmin; self-organising maps; taste and odour; vertical profiling system; water treatment

## 1. Introduction

The detection of Taste and Odour (T&O) compounds, such as geosmin or 2-methylisoborneol (MIB), can compromise the organoleptic quality of the water and divert consumers from its use despite not presenting any health-related risk. Although the human detection threshold can largely change from one individual to another [1], geosmin concentrations as low as 5 ng/L are detectable [2]. Understanding and modelling T&O compounds is a priority for water utilities, in order to produce treated water with high organoleptic quality and thus enhance the confidence and reliance of the consumers towards the drinking water supply system. As a result great economic and social benefits could be achieved if a model was developed that can predict in advance T&O events.

Modelling geosmin and other T&O compounds is extremely challenging, as the reasons for the appearance of these compounds are still largely unknown [3–5]. Variables and factors affecting T&O compound presence can be several; typically are different from location to location [1,6] and over time [1]. However, a number of models have been developed which, for a specific lake, can predict geosmin concentrations with acceptable accuracy based on water quality (e.g., [1,7–10]).

Although it has been pointed out (e.g., [11]) that there can be several possible sources of geosmin and MIB in a reservoir (e.g., vegetation and standing timber, actinomycetes), a typical event that can cause T&O complications is an algal bloom. During their growth and subsequent decay, algae, and specifically cyanobacteria [7], can produce metabolites including biotoxins and T&O

compounds [2,12]. However, the production of metabolites is related to the species and strain of the cyanobacteria blooming [13], and there is still large uncertainty related to which species can produce these compounds, since newer studies often prove older studies wrong (as explained in [14]). During a bloom, different species and strains are in competition and interacting with each other, through a number of nonlinear behaviours determined by factors, such as nutrient availability, presence of grazing zooplankton, or physical factors [5]. For example, in Hinze dam (South-East Queensland, Australia), Uwins et al. [11] reported steadily increasing geosmin concentration following an early spring bloom in *Anabaena* sp. Following the bloom, and decay of precipitating cells, geosmin was released. Actinobacteria on the other hand, despite potential for contributing to T&O compounds release, were inhibited in this production by high water temperature, high dissolved oxygen and low phosphorus levels. As a result of this complexity and high uncertainty, certain variables which are surrogate estimators of algal counts, such as chlorophyll-a, have been correlated with geosmin both positively, i.e., high chl-a levels linking to high geosmin concentrations (e.g., [15] although based on only few data points; [4,16]), and negatively, with higher geosmin levels measured where lower chl-a was detected (e.g., [1]).

In general, the presence of geosmin and MIB, which in many cases is linked to algal blooms, has been correlated to a large number of possible predictors. Aside the already mentioned chl-a, also the sum of green algae [4], regardless of the species, sometimes proved to be a good predictor. In that particular study however, single species and strains were not measured, thus it is unknown if geosmin was caused by the same specific type of algae, or if different ones result in similar geosmin concentrations. Other possible predictors include nutrients such as nitrogen and phosphorus [4,10,17–19], as well as metallic micronutrients, such as copper or manganese [4,17]. Other critical factors proved to be water temperature [17,19,20], light intensity [17,21], turbidity and water clarity [9,10], dissolved oxygen [7], rainfall [11] and oxidation-reduction potential [4].

The importance of light availability is related to the energy that light provides to enable photosynthetic fixation of dissolved inorganic carbon, which can be subsequently routed into the cellular synthesis of geosmin [1]. Additionally, although algal blooms, and thus T&O events, typically occurred in warm stratified seasons, some studies [22] also proved how high geosmin levels can be detected during lake circulation periods; interestingly, other studies [23] also found how low, instead of high, temperatures can stimulate the production of geosmin and their accumulation in cells due to lower chl-a demand, although high temperature or optimum light intensity would be necessary for more intracellular geosmin release.

This study aims to exploit historical sets of relevant data, and use cutting-edge data analytics to better understand, and model, the occurrence of T&O events in a relatively shallow, subtropical reservoir in Australia. As already mentioned, there is large uncertainty around T&O events, with the understanding and prediction of such events being site and season specific. Therefore, a full statistical analysis of the available historical data was performed to gain a specific understanding of the behaviour of the reservoir of interest. Given the relatively large amount of data, and recent advancements in the hydroinformatics field, it was possible to identify potential predictor variables of T&O events.

Based on the correlations found, a simple statistical model was also developed which enables a prediction of the magnitude of possible future high geosmin concentrations. Despite limited by the number of historical events available for analysis and the complexity of the system, the model can assist water treatment operators for an improved understanding and preparedness towards geosmin peak events; the results of this analysis also provide an example of potential geosmin production behaviour in similar reservoirs.

## 2. Materials and Methods

### 2.1. Study Location

Lake Tingalpa (Figure 1), bounded by Leslie Harrison dam, is located about 20 km south-east of Brisbane, in Queensland (Australia). The main purpose of the reservoir is to supply water to the Redland City, which has around 150,000 inhabitants. On average, it supplies 20% of the water demand in this area. The main inflow to the reservoir is Tingalpa Creek. The 535-m-long dam wall was completed in 1984, and the reservoir could supply 24,868 ML at full capacity; however, its capacity decreased to 13,206 ML on Friday, 1 August 2014, due to safety concerns with the integrity of the wall. The catchment area covers 87 km<sup>2</sup>. An intake tower allows the withdrawal of raw water, which is redirected to the Capalaba water treatment plant (WTP).

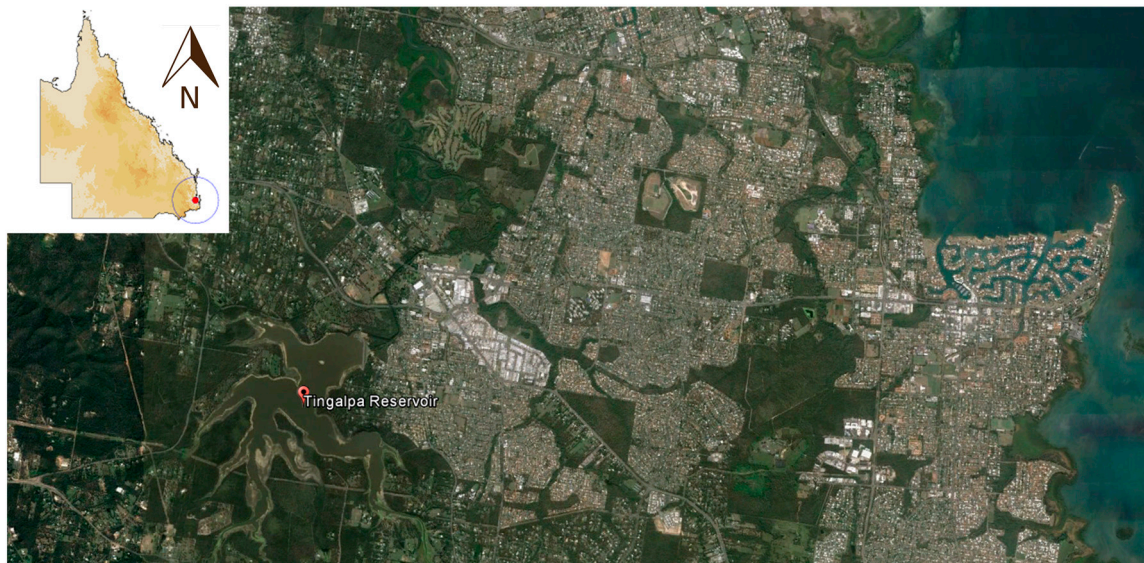


Figure 1. Lake Tingalpa location.

### 2.2. Data Collection

Through an effective collaboration with Seqwater, the main bulk water supplier in South-East Queensland, and custodian of the Capalaba WTP and Lake Tingalpa, historical data necessary for this research study was made available. Among a wide range of parameters, the most relevant ones for this study were essentially raw water values for algae biovolumes, metals (iron, aluminium and manganese, both soluble and total), nutrients, organics, physical properties (e.g., pH, alkalinity, water temperature, turbidity, true colour) and obviously T&O compounds (specifically geosmin and MIB). These data have a fortnightly frequency and cover the period from 2011 to 2016. Additionally, daily reservoir volume data were made available for the same period.

With regards to lake data, data collected from a Vertical Profiling System (VPS) located near the dam wall was also available. A VPS consists of a system of probes, connected to a floating buoy, which are winched up and down the reservoir and can collect information for the full water column, every hour with a 1-m vertical resolution, for a number of parameters, such as water temperature, pH, dissolved oxygen (DO), conductivity, but also, importantly for this study, chl-a and blue-green algae (BGA) counts.

### 2.3. Data Analysis

Firstly, data were analysed with a number of statistical techniques. Software such as Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and Matlab 2015a (Mathworks, Natick, MA, USA)

were used. The first step involved time series data analysis to identify critical historical T&O events; once these were identified, a number of potential input factors were also selected based on the literature. These included: total cyanophytes biovolume; *Microcystis*, *Merismopedia* and *Dolichospermum* biovolumes; nitrogen and total phosphorus; iron, aluminium and manganese; rainfall; reservoir volume; turbidity; pH; dissolved oxygen; and water temperature. A smaller dataset, containing data for T&O compounds and potential inputs around the events periods only, was created. Secondly, self-organising maps (SOMs) were used to investigate the actual presence of any significant correlation between those multiple variables. SOMs [24,25] are a type of unsupervised learning Artificial Neural Network (ANN) that produce topologically ordered maps based on their own classification of the presented data and the dimensionality reduction of the feature space. These ordered, colour contour maps enable the researcher to visually identify similar patterns between some of them; this would mean that similar levels of different variables typically matched over time. That is, the visual inspection of the component planes of the SOM provides a rapid and intuitive means of examining the covariance between the selected variables in order to obtain an increased understanding of the system [26], especially in complex problems with several potential variables involved (e.g., [27]). Based on the conclusions of these first steps, a more in-depth analysis could be performed, and a simple prediction model was developed.

### 3. Results and Discussion

#### 3.1. Time-Series Analysis

Figure 2 illustrates the temporal variation of sampled geosmin and MIB concentrations in the raw water of Lake Tingalpa since 2011. Figure 2a allows a better analysis of earlier, smaller events while Figure 2b includes the major, more recent events. Therefore, it can be seen how a number of geosmin events above the detection threshold (5 to 10 ng/L) were detected over time, however two extreme events occurred during late spring/summer 2015, leading to concentrations over 150 ng/L, and resulting in wide spread customer complaints. Table 1 displays how these events increased the calculated average geosmin concentrations. As a further consideration, MIB concentrations were consistently much lower than geosmin levels, thus the attention of this research shifted to geosmin only, after these preliminary outcomes.

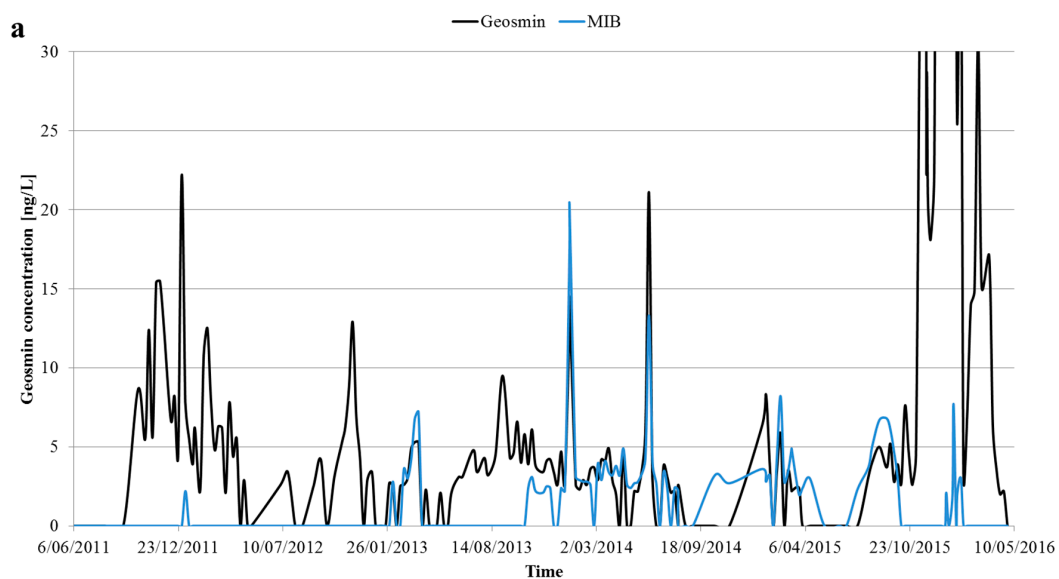
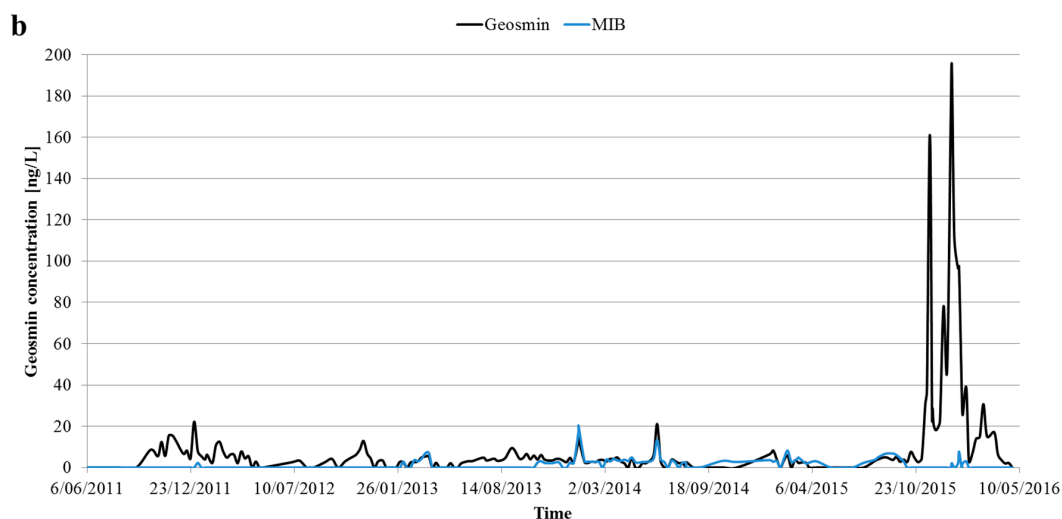


Figure 2. Cont.





**Figure 2.** Geosmin and MIB concentrations for raw water from Lake Tingalpa, 2011–2016, focusing on <30 ng/L events (a), and overall (b).

**Table 1.** Average Geosmin and MIB peak/overall values, including and excluding 2015 extreme Geosmin events.

Period	Average Geosmin	Average MIB	Average Peak Geosmin	Average Peak MIB
July 2011–May 2015	3.9 ng/L	1.6 ng/L	15 ng/L	17 ng/L
July 2011–May 2016	10.4 ng/L	1.5 ng/L	44 ng/L	17 ng/L

As a first step, data were filtered to include only the events periods. Then, based on the literature and available data, a number of potential predictors were included, such as: total cyanophytes, as well as particular species of them (such as *Microcystis aerug.* and spp., *Dolichospermum circ.* and *Merismopedia* spp.), water temperature, reservoir volume, iron, total oxidised nitrogen, turbidity, and past week change in reservoir volume (proportional to rain and evaporation among others). Other parameters that were expected to play a role (such as dissolved oxygen and phosphorus) showed instead only a very weak correlation with geosmin peak events. Additionally, since geosmin was occasionally already detected the week/fortnight before the peak, the independent variable considered was the increase rate of geosmin, assumed linear, in ng/L per day, calculated by dividing the overall weekly/fortnightly variation by the number of days in between the two sampling (i.e., 7/14). This gives a better representation of the dynamics leading to the peak, and accounts for initial concentrations that might have been detected.

### 3.2. Self-Organising Maps

The prepared smaller dataset was used to create self-organising maps (Figure 3). Table 2 also presents the numerical quantification of each of these variables during those critical events, with conditional formatting helping to visually inspect similar trends. The first thing that can be noticed from the SOMs, is how the map for geosmin peaks has very similar colour patterns (highlighted with green connections) to turbidity, total iron, and total oxidised nitrogen (NO<sub>3</sub> and NO<sub>2</sub>), as well as being inversely proportional to the reservoir volume (highlighted with a red connection). This means that any time there was a major geosmin peak, usually also those parameters recorded higher values; similarly, higher geosmin peak values were also recorded in times of low reservoir volume. From Figure 4, a confirmation can be found and a very clear correlation can be seen between the substantial volume reduction, which occurred in winter 2014, and the sharp increases in iron, nitrogen and turbidity, most likely due to less dilution and increased reservoir instability leading to higher mixing. As a consequence, the geosmin peak events also became much more drastic.

Increased nutrient availability is a well-documented factor cited in the literature that favours the production of geosmin. Turbidity was also found to positively correlate with geosmin in a number of studies [9,10]. Additionally, from the SOMs it can be also noticed how warmer water temperatures are typically linked to larger geosmin concentrations, as well as higher daily increase rates. This is again in agreement with the literature [17,19,20], as it implies, for instance, a higher intracellular geosmin release.

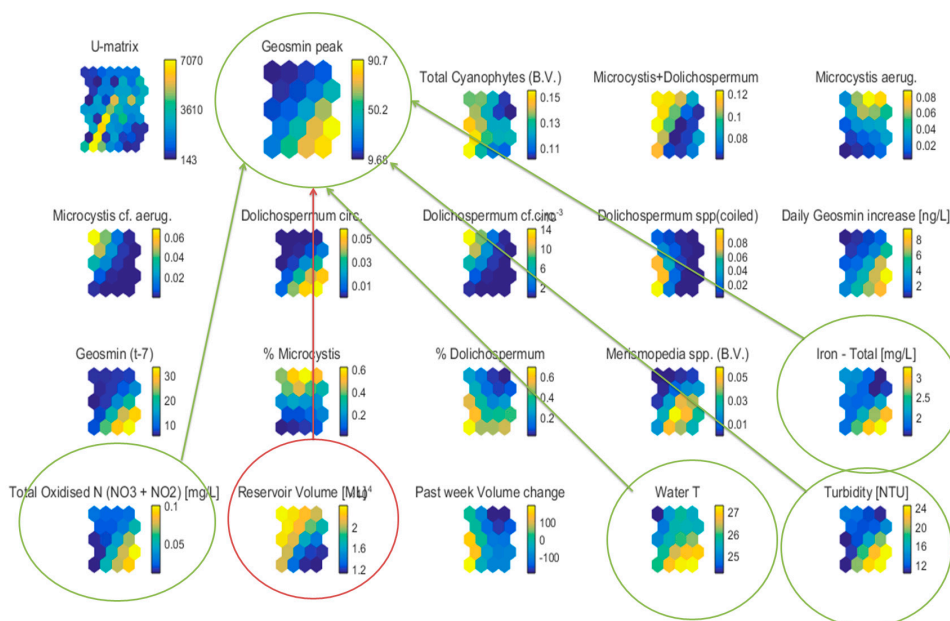
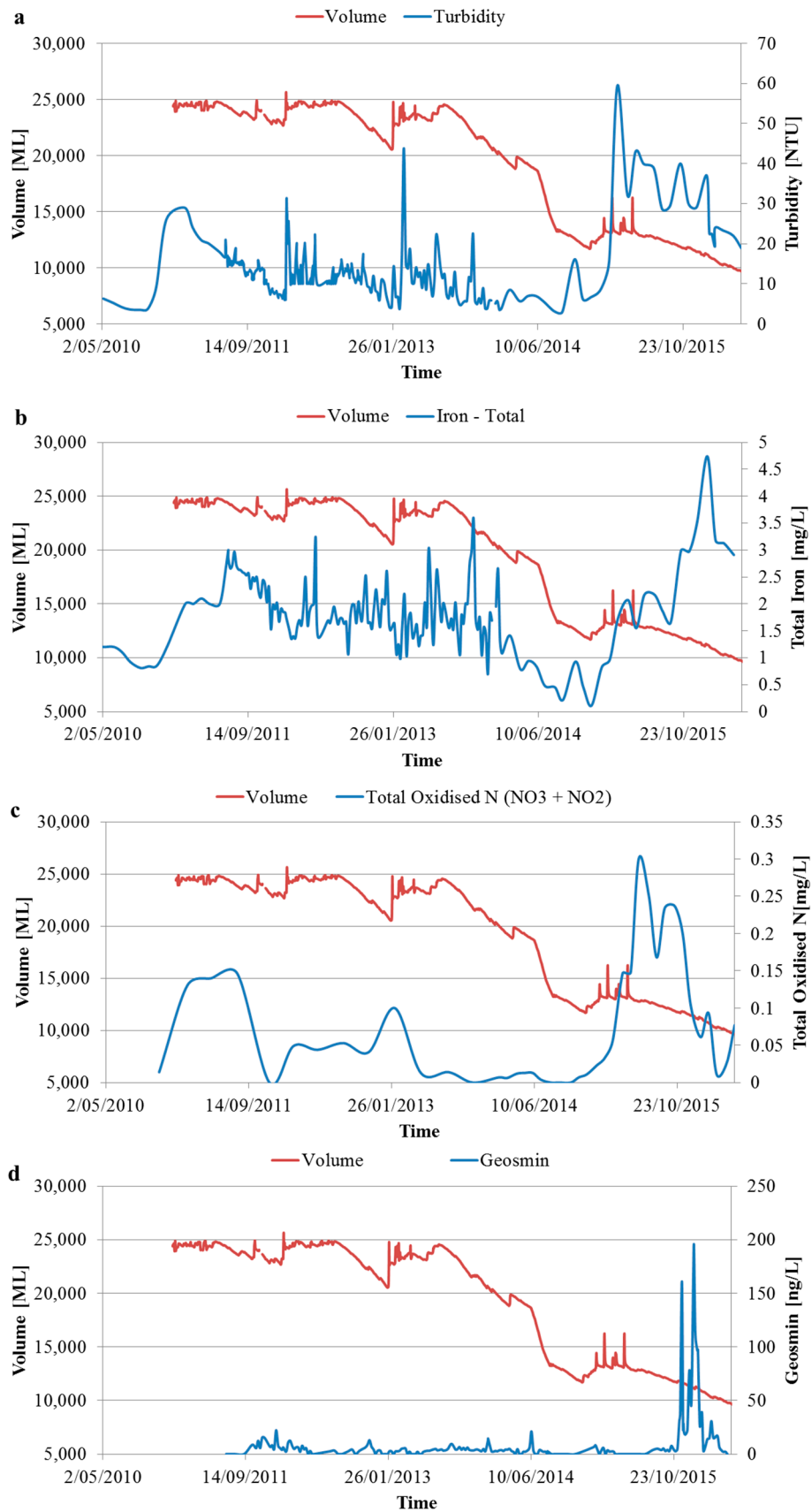


Figure 3. SOMs for geosmin and potential geosmin predictors. Raw water from Lake Tingalpa, 2011–2016.

Table 2. Geosmin peak events and main predictors’ values.

Day	Geosmin (ng/L)	Tot. Cyanophytes (B.V.)	Microcystis + Anabaena (mm <sup>3</sup> /L)	Geosmin (t–7) (ng/L)	Daily Geosmin Increase (ng/L)
10/11/2011	15.4	0.153	0.130	5.6	1.4
29/12/2011	22.2	0.163	0.133	4.5	2.5
16/02/2012	12.5	0.095	0.080	10.8	0.2
3/05/2012	0.0	0.277	0.271	2.9	–0.4
20/11/2012	12.9	0.156	0.156	8.8	0.6
3/09/2013	9.5	0.014	0.0	6.95	0.4
10/06/2014	21.1	0.031	0.0	5.8	2.2
14/10/2014	0.0	0.191	0.0	0.0	0.0
19/11/2015	161	0.183	0.161	35.3	20.6
31/12/2015	195	0.125	0.125	80.5	16.4
28/01/2016	38.7	0.091	0.091	25.4	1.9
1/3/2016	30.7	0.020	0.0	14.9	2.3
Day	Water Temperature (°C)	Total Iron (mg/L)	Reservoir Volume (ML)	Volume Variation (ML/week)	Total NO <sub>3</sub> + NO <sub>2</sub> (mg/L)
10/11/2011	29.5	2.38	23642	–406.19	0
29/12/2011	25.8	1.67	23199	291.62	0
16/02/2012	26.1	1.4	24414	187.56	49
3/05/2012	22.5	3.24	24537	230.42	44
20/11/2012	25.1	1.94	22415	–204.53	41
3/09/2013	20.4	1.53	23749	–164.70	14
10/06/2014	24.9	0.832	18578	–173.33	13
14/10/2014	25.9	0.926	12632	–120.35	0
19/11/2015	26.2	2.97	11728	–62.22	200
31/12/2015	26.7	3.53	11164	–37.27	104
28/01/2016	27.4	4.74	10862	–129.6	61
1/3/2016	29.3	3.18	10315	–120.14	93



**Figure 4.** Time series of reservoir volume and (a) turbidity; (b) total iron; (c) total oxidised nitrogen and (d) geosmin in the raw water, Lake Tingalpa 2010–2016.

### 3.3. Events Analysis

Table 2 allows for a better understanding of the role that cyanophytes blooms play in determining geosmin increment rates during peak events. The figure reports the biovolumes of total cyanophytes, as well as the sum of *Microcystis* species and *Dolichospermum circ.* (former *Anabaena*), typically associated to T&O compounds production [14,28,29]. This table also includes two entries (3 May 2012, and 14 October 2014) where, despite cyanophytes blooms, there was no detection of geosmin. In this way it was possible to gain a fuller picture of the relationships between cyanobacteria and geosmin events. In addition, for event extremely high algal counts, including cyanophytes, were recorded on 9 January 2014. Extremely high iron levels were also recorded. This event is not herein analysed or reported, since further investigation is required for this particular event, as the levels are incredibly high and sampling/analysis issues would need to be validated. As a confirmation, the sampling immediately preceding and following this critical one yielded values for all the parameters within absolutely normal ranges. Additionally, no geosmin was detected. Thus, given the uncertainty around it, this particular event was excluded from the next steps of the analysis.

Firstly, attention was focused on the two blooms not leading to geosmin peaks. On the 14 October 2014, a total cyanophytes biovolume of  $0.191 \text{ mm}^3/\text{L}$ , the second highest in the dataset, was recorded. However, no geosmin was detected in the nearest sampling. This can be explained by the fact that neither *Microcystis* nor *Dolichospermum* species were detected; the dominant species leading to that bloom was *Merismopedia* spp., which is not known to be able to produce geosmin. This would strongly suggest that the type of cyanobacteria causing the bloom provides a much more useful prediction of a potential T&O event than the total aggregate cyanobacteria count. Nevertheless, the second event not leading to a geosmin peak event occurred on the 3 May 2012 ( $0.271 \text{ mm}^3/\text{L}$ , the highest value in the dataset), and in this case it was largely (98%) due to *Microcystis* and *Dolichospermum circ.* What is evident here, besides a high volume, is the low temperature of the raw water ( $22.5 \text{ }^\circ\text{C}$ ). Although more research is needed, since contrasting outcomes emerged on the role of temperature in geosmin production, we can assume that below certain temperature thresholds, the production/release of geosmin by *Microcystis* species and *Dolichospermum circ.* is limited or inhibited. This also highlights how algal (cyanobacterial) blooms, despite often being a critical predictor for geosmin events, are only one of the factors of a far more complex system, where nutrients, dam level, temperature, and other parameters also play determinant roles.

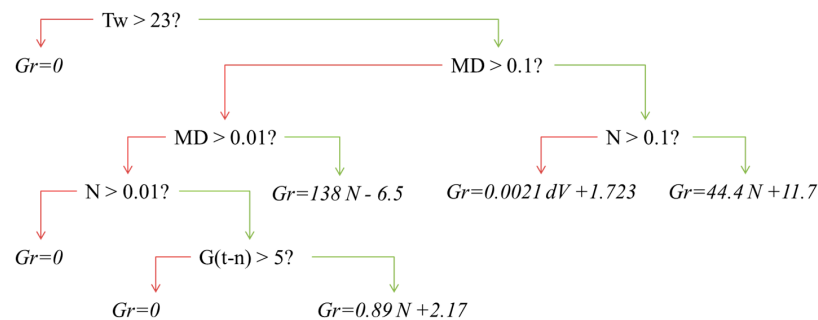
### 3.4. Model Development

Based on the available data, its analysis and on the considerations above, we developed a conceptual regression tree, illustrated in Figure 5. Since it is based on only few events, currently it cannot be deemed reliable to correctly estimate the exact amount of future peaks in geosmin. However, based on the analysis of historical events, it provides a structural hypothesis of relevant variables and processes, which determined the previous events or unexpected low values. Often, simpler regression models can outperform more complicated ones [30]. It can be seen how variables, such as water temperature and total amount of geosmin-producing cyanophytes, do not directly enter the regression equation (i.e., their value is not directly proportional to the geosmin daily increase rate), however they are determinant factors in setting up thresholds under/above which different processes are inhibited or supported. Additionally, certain variables, such as iron, turbidity and reservoir volume, were not included in the model due to multicollinearity (i.e., they exhibit a similar behaviour to nitrogen), although they are relevant predictors too.

The first cut is set by water temperature; regardless of other factors, there is no evidence of geosmin events when the raw water temperature was below  $23 \text{ }^\circ\text{C}$ . If the raw water is instead warmer than  $23 \text{ }^\circ\text{C}$ , a number of other options are possible. Under this initial constraint, the presence of geosmin-producing cyanophytes plays a critical role. If they are not detected, and also nitrogen is not present, then there would be no increase whatsoever in geosmin concentration. Even if oxidised nitrogen is detected, geosmin might slightly increase only if it was already present in some noticeable

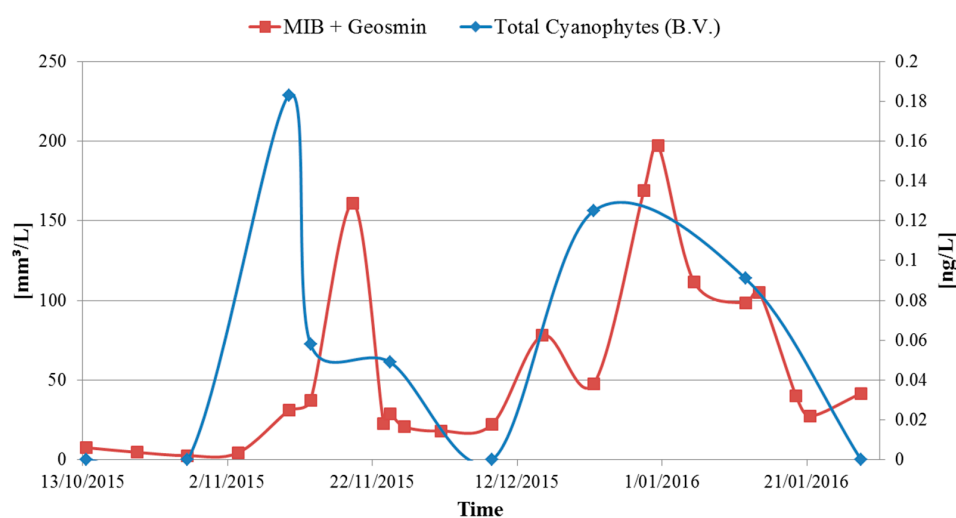


(>5 ng/L) amounts in the previous sample. This can be due to the occurrence of a long peak event, which can persist if enough nutrients are available. If instead the geosmin-producing cyanophytes are detected, and in large amounts (i.e., >0.1 mm<sup>3</sup>/L), then the rate of increase will depend on the available amount of nutrients (represented by oxidised nitrogen); if oxidised nitrogen is also present in large amounts (i.e., >0.1 mg/L), then the geosmin increase rate will be proportional to it; if instead, despite a bloom of geosmin-producing algae, there is no large amount of nutrients available, then the rate of increase would depend on the dam volume variation over the last week. In case of positive variation, rain and inflow were larger than evaporation and outflow, leading to higher increases in geosmin, which is in line with outcomes of previous studies [11]. The remaining scenario is given by medium-size geosmin-producing algal blooms (i.e., between 0.01 and 0.1 mm<sup>3</sup>/L); in this case, nitrogen plays again a critical role and its amount is proportional to the increase rate of geosmin.



**Figure 5.** Regression analysis tree for estimation of daily rate of increase of geosmin. Variables are: Gr = daily rate of increase of geosmin (ng/L); Tw = water temperature (°C); MD = total biovolume of *Microcystis* and *Dolichospermum* (mm<sup>3</sup>/L); N = total oxidised nitrogen (mg/L); G(t-n) = geosmin concentration previous sample. Green arrows represent the “yes” scenario while red arrows represent the “no” scenario.

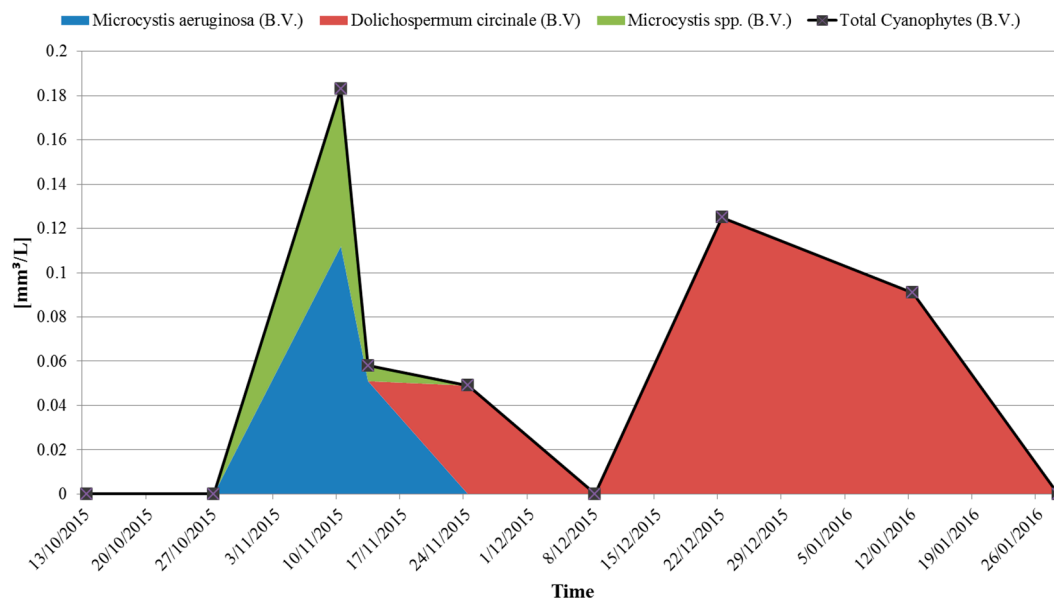
In order to validate the model, we focused on the two extreme events occurring at the end of 2015. Figure 6 shows how both geosmin events were preceded by earlier peaks in total cyanophytes.



**Figure 6.** Total Cyanophytes biovolume (mm<sup>3</sup>/L) and total count of geosmin and MIB (ng/L) in the raw water, Capalaba WTP, 13 October 2015, to 28 January 2016.

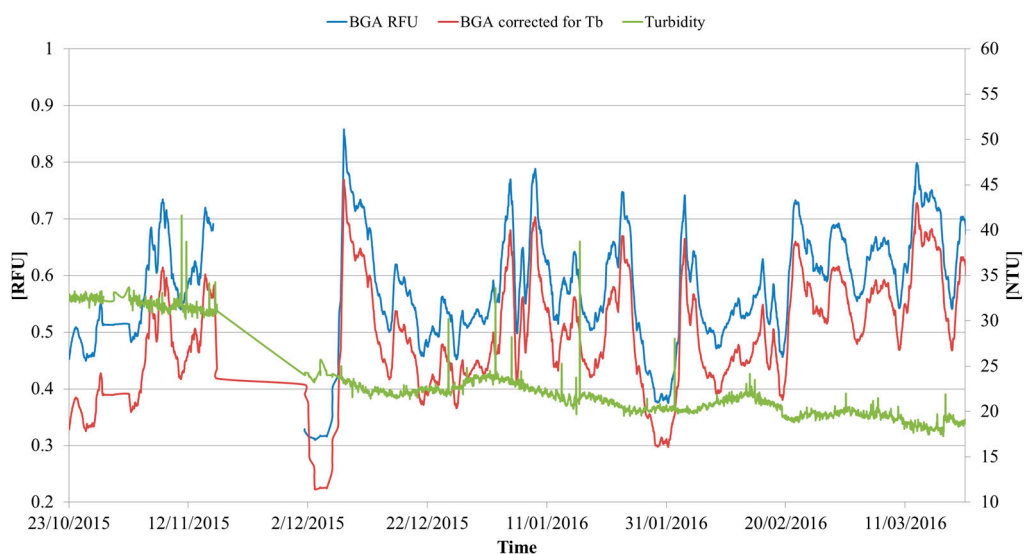
In addition, Figure 7 shows the speciation of the cyanophytes. It can be seen how both peaks were largely caused by species of *Microcystis* and *Dolichospermum*, although the two peaks were quite different to each other. Looking back at Figure 2, it seems that *Dolichospermum* is responsible for the

production of some MIB as well. However, as previously mentioned, the dominant compound is geosmin and this is produced by both *Microcystis* and *Dolichospermum* species.



**Figure 7.** Dominant classes and species of cyanophytes in the raw water, Capalaba WTP, 13 October 2015 to 28 January 2016.

Looking at the regression analysis tree developed in Figure 5, the temperature detected in the raw water during that period (which corresponds to late spring and summer months) was consistently above the threshold of 23 °C, thus allowing for geosmin event to occur. Additionally, as it can be seen from Figure 8, the total biovolume of *Microcystis* and *Dolichospermum* is above 0.1 mm<sup>3</sup>/L, thus leading towards the far right side of the tree. Finally, it was pointed out (Figure 4c) that the nutrient availability, in particular oxidised nitrogen, increased remarkably after the decrease in volume, reaching concentrations higher than 0.1 mg/L in the weeks preceding the geosmin event. Thus, these conditions led the developed model to predict sharp increases in geosmin levels; which actually occurred.



**Figure 8.** VPS BGA original and adjusted readings (RFU), and turbidity (NTU) at Lake Tingalpa, depth = 1 m, October 2015–March 2016.

Since the peaks in relevant cyanophytes typically occur a couple of weeks before the T&O events, it is possible to use these data, as well as the other information (water temperature, oxidised nitrogen, dam volume variation) required by the tree, to estimate the potential for the occurrence of geosmin peak events, and thus proactively adjust the treatment procedures accordingly.

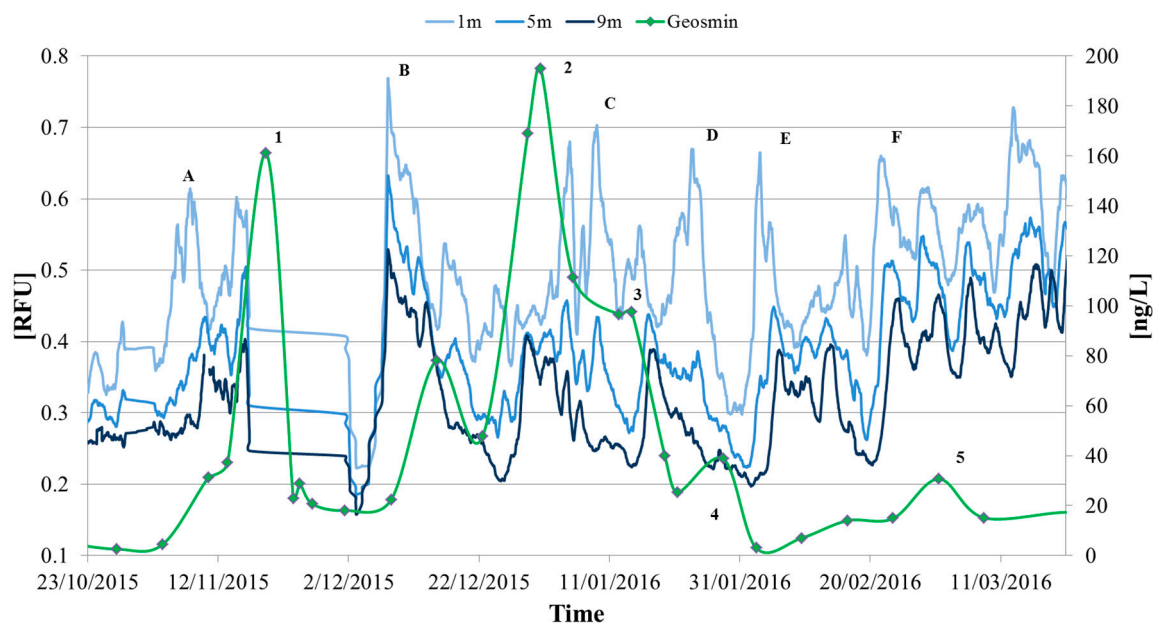
### 3.5. VPS Data Analysis and Predictive Potential

The historical data from the Vertical Profiling System (VPS) installed in Lake Tingalpa were analysed. The VPS was installed in 2013; hence it was possible to use its data for analysis of the more recent T&O events only. Additionally, during winter 2014, a number of sondes were changed from the previous YSI 6-Series Sensors to the newer EXO multiparameter water quality sonde product line, leading to newer units of measure, probe sensitivity to other water parameters and calibration procedures. Hence, for this study, we will focus on the two large 2015 events only. In particular, the aim was to analyse the phycocyanin-based blue-green algae (BGA-PC) sensor to see if it can be used to assist in predicting geosmin peak events.

The sensor measures blue-green algae in real-time through the *in vivo* fluorometry technique, which directly detects the fluorescence of a specific pigment in living algal cells and determines relative algal biomass. However, the sensor can be sensitive to variables such as turbidity, with newer EXO sondes less sensitive than the previous 6-series [31]. In Figure 8, the BGA readings (at depth = 1 m) adjusted for turbidity based on YSI indications [31] are reported. A proper calibration should be performed for this specific location, which evaluate the effect of not only turbidity, but also of other variables; however this was out of the scope of this work. Although hourly data was provided, the chart shows 24-h moving averages. Between 16 November 2015, and 2 December 2015, the VPS was not operational and therefore data is missing for those two weeks. It can be noticed how, before 16 November 2015, the average turbidity was slightly higher (around 30 NTU) than in the following months (around 20 NTU), thus, by adjusting the readings, the difference in the BGA peak values in November and in December is relatively smaller than initially measured. Nevertheless the difference is minimal, and such adjustment would be crucial only in case of extreme turbidity events.

Figure 9 illustrates the BGA values at 1, 5 and 9 m depths of Lake Tingalpa, as well as the geosmin values in the raw water redirected to the Capalaba WTP. Typically, the values closer to the surface are higher than at the bottom, consistent with the literature findings (e.g., more light, higher temperatures, etc.). Although not extremely clear, a pattern can be noticed, with BGA peaks often anticipating geosmin high values. For instance, peak A in BGA occurred 12 days before the peak 1 in geosmin. Peak 2 occurred 23 days after a very high peak (B) in BGA; despite the longer lag, a smaller peak in geosmin was already detected 7 days later. Peak C also occurred 5 days before peak 3, as well as peak D with peak 4. Peak 5, interestingly, did not yield any sharp, lagged geosmin peak, although there was a slow, constant increase after that. Finally, peak F anticipated peak 5 by 8 days.

It is clear from the previous paragraphs how there is high complexity and uncertainty involved in these T&O events, and other parameters (e.g., water temperature, nitrogen, dam level) are as important as cyanobacteria. Nevertheless, given the large amount of VPS data which are collected remotely (i.e., no need for samplers), and in real-time, collected in the reservoir, future work (e.g., accurate calibration with manual sampling data; analysis of new events) could focus on better exploring the potential of the VPS BGA probe to be used as an input for a prediction model which could provide early warnings of T&O events. Although it was evident that also the strain of cyanophytes is a critical predictor, more work is needed to understand if the fluorescence approach of the BGA probe allows for all the strains to be detected, or not. The correlations found between its readings and T&O events are interesting as they may imply that the BGA probe can mainly detect T&O-causing strains. Additionally, similar VPS-based models were already developed by the authors [30] for other parameters, leading to monetary and operational benefits for the water utility.



**Figure 9.** VPS BGA adjusted readings (RFU) at 1, 5 and 9 m depths in Lake Tingalpa, and geosmin levels (ng/L) in the raw water at the Capalaba WTP, October 2015–March 2016.

#### 4. Conclusions

A comprehensive analysis of data related to a number of T&O events was performed for Lake Tingalpa and the Capalaba WTP. Geosmin was found to be the dominant compound, and two extremely high peaks occurred in November/December 2015. One of the key-factors triggering geosmin events was the occurrence of cyanobacteria blooms; however, the species of cyanobacteria was also a critical factor, since some of them (e.g., *Merismopedia* spp.) did not produce geosmin. Importantly, blooms alone cannot fully explain the occurrence and magnitude of geosmin events; other factors such as water temperature, nitrogen and reservoir volume variations were found to be determinant input factors. In particular, it was noticed how higher geosmin peaks have been recorded since the reservoir volume was lowered in 2014, and in turn turbidity and nutrients increased. As a result, a simple regression analysis tree was developed and validated to provide predictive capabilities and better understanding of geosmin peak events. Although such model can be already used, with caution, by the plant operators for early prediction of T&O events, this requires manual data input based on the results of the latest lake sampling; nevertheless, analysis of VPS BGA probe data, in conjunction with other VPS data (e.g., water temperature, turbidity), showed potential to use only remotely collected data to provide early warnings for T&O events. The development of such tool will be the focus of future research.

**Acknowledgments:** This research work was conducted with the technical and financial support of Seqwater.

**Author Contributions:** Edoardo Bertone collected and analysed the data, developed the model and drafted the manuscript. Kelvin O'Halloran supervised the study, providing advice on methodology and results, and revised the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### Abbreviations

The following abbreviations are used in this manuscript:

T&O	Taste and Odour
MIB	2-methylisoborneol
Chl-a	Chlorophyll-a
Km	Kilometres

ML	Megalitres
WTP	Water Treatment Plant
VPS	Vertical Profiling System
BGA	Blue-Green Algae
Aerug.	Aeruginosa
Spp.	Species
Circ.	Circinalis

## References

1. Dzialowski, A.R.; Smith, V.H.; Huggins, D.G.; Denoyelles, F.; Lim, N.C.; Baker, D.S.; Beury, J.H. Development of predictive models for geosmin-related taste and odor in Kansas, USA, drinking water reservoirs. *Water Res.* **2009**, *43*, 2829–2840. [[CrossRef](#)] [[PubMed](#)]
2. Suffet, I.M.; Khiari, D.; Bruchet, A. The drinking water taste and odor wheel for the millennium: Beyond geosmin and 2-methylisoborneol. *Water Sci. Technol.* **1999**, *40*, 1–13.
3. Juttner, F.; Watson, S.B. Biochemical and ecological control of geosmin and 2-methylisoborneol in source waters. *Appl. Environ. Microbiol.* **2007**, *73*, 4395–4406. [[CrossRef](#)] [[PubMed](#)]
4. Parinet, J.; Rodriguez, M.J.; Serodes, J.B. Modelling geosmin concentrations in three sources of raw water in Quebec, Canada. *Environ. Monit. Assess.* **2013**, *185*, 95–111. [[CrossRef](#)] [[PubMed](#)]
5. Bruder, S.; Babbar-Sebens, M.; Tedesco, L.; Soyeux, E. Use of fuzzy logic models for prediction of taste and odor compounds in algal bloom-affected inland water bodies. *Environ. Monit. Assess.* **2014**, *186*, 1525–1545. [[CrossRef](#)] [[PubMed](#)]
6. Yu, J.; Zhao, Y.; Yang, M.; Lin, T.-F.; Guo, Z.; Gu, J.; Li, S.; Han, W. Occurrence of odour-causing compounds in different source waters of China. *J. Water Supply Res. Technol.* **2009**, *58*, 587–594. [[CrossRef](#)]
7. Sugiura, N.; Utsumi, M.; Wei, B.; Iwami, N.; Okano, K.; Kawauchi, Y.; Maekawa, T. Assessment for the complicated occurrence of nuisance odours from phytoplankton and environmental factors in a eutrophic lake. *Lakes Reserv. Res. Manag.* **2004**, *9*, 195–201. [[CrossRef](#)]
8. Smith, V.H.; Sieber-Denlinger, J.; deNoyelles, F., Jr.; Campbell, S.; Pan, S.; Randtke, S.J.; Blain, G.T.; Strasser, V.A. Managing taste and odor problems in a eutrophic drinking water reservoir. *Lake Reserv. Manag.* **2002**, *18*, 319–323. [[CrossRef](#)]
9. Christensen, V.G.; Graham, J.L.; Milligan, C.R.; Pope, L.M.; Ziegler, A.C. *Water Quality and Relation to Taste-and-Odor Compounds in North Fork Ninescaw River and Cheney Reservoir, South-Central Kansas, 1997–2003*; Scientific Investigations Report; U.S. Geological Survey: Pembroke, NH, USA, 2006.
10. Mau, D.P.; Ziegler, A.C.; Porter, S.D.; Pope, L.M. *Surface-Water-Quality Conditions and Relation to Taste-and-Odor Occurrences in the Lake Olathe Watershed, Northeast Kansas, 2000–02*; US Department of the Interior, US Geological Survey: Olathe, KS, USA, 2004.
11. Uwins, H.K.; Teasdale, P.; Stratton, H. A case study investigating the occurrence of geosmin and 2-methylisoborneol (MIB) in the surface waters of the Hinze Dam, Gold Coast, Australia. *Water Sci. Technol.* **2007**, *55*, 231–238. [[CrossRef](#)] [[PubMed](#)]
12. Smith, J.L.; Boyer, G.L.; Zimba, P.V. A review of cyanobacterial odorous and bioactive metabolites: Impacts and management alternatives in aquaculture. *Aquaculture* **2008**, *280*, 5–20. [[CrossRef](#)]
13. Taylor, W.D.; Losee, R.F.; Torobin, M.; Izaguirre, G.; Sass, D.; Khiari, D.; Atasi, K. *Early Warning and Management of Surface Water Taste and Odor Events*; American Water Works Association Research Foundation: Denver, CO, USA, 2006.
14. Qi, M.; Chen, J.; Sun, X.; Deng, X.; Niu, Y.; Xie, P. Development of models for predicting the predominant taste and odor compounds in Taihu Lake, China. *PLoS ONE* **2012**, *7*, e51976. [[CrossRef](#)] [[PubMed](#)]
15. Rosen, B.; MacLeod, B.; Simpson, M. Accumulation and release of geosmin during the growth phases of *Anabaena Circinalis* (Kutz.) rabenhorst. *Water Sci. Technol.* **1992**, *25*, 185–190.
16. Downing, J.A.; Watson, S.B.; McCauley, E. Predicting cyanobacteria dominance in lakes. *Can. J. Fish. Aquat. Sci.* **2001**, *58*, 1905–1908. [[CrossRef](#)]
17. Saadoun, I.M.K.; Schrader, K.K.; Blevins, W.T. Environmental and nutritional factors affecting geosmin synthesis by *Anabaena* sp. *Water Res.* **2001**, *35*, 1209–1218. [[CrossRef](#)]
18. Rashash, D.M. *Identification and Control of Odorous Algal Metabolites*; AWWA Research Foundation and American Water Works Association: Denver, CO, USA, 1996.



19. Srinivasan, R.; Sorial, G.A. Treatment of taste and odor causing compounds 2-methyl isoborneol and geosmin in drinking water: A critical review. *J. Environ. Sci.* **2011**, *23*, 1–13. [[CrossRef](#)]
20. Blevins, W.; Schrader, K.; Saadoun, I. Comparative physiology of geosmin production by *Streptomyces halstedii* and *Anabaena* sp. *Water Sci. Technol.* **1995**, *31*, 127–133. [[CrossRef](#)]
21. Tsuchiya, Y.; Matsumoto, A. Characterization of *Oscillatoria f. Granulata* producing 2-methylisoborneol and geosmin. *Water Sci. Technol.* **1999**, *40*, 245–250. [[CrossRef](#)]
22. Journey, C.A.; Arrington, J.M.; West, R.; Westcott, J.W.; Tuck, K. Geosmin occurrence in Lake William C. Bowen and municipal reservoir# 1, Spartanburg County, South Carolina, 2005 to 2006. In Proceedings of the South Carolina Water Resources Conference, Charleston, SC, USA, 14–15 October 2008.
23. Zhang, T.; Li, L.; Song, L.; Chen, W. Effects of temperature and light on the growth and geosmin production of *Lyngbya kuetzingii* (Cyanophyta). *J. Appl. Phycol.* **2009**, *21*, 279–285. [[CrossRef](#)]
24. Kohonen, T. The self-organizing map. *Proc. IEEE* **1990**, *78*, 1464–1480. [[CrossRef](#)]
25. Vesanto, J.; Himberg, J.; Alhoniemi, E.; Parhankangas, J. Self-organizing map in Matlab: The SOM toolbox. In Proceedings of the Matlab DSP Conference, Espoo, Finland, 16–17 November 1999; pp. 16–17.
26. Blokker, E.; Furnass, W.R.; Machell, J.; Mounce, S.R.; Schaap, P.G.; Boxall, J.B. Relating water quality and age in drinking water distribution systems using self-organising maps. *Environments* **2016**, *3*, 10. [[CrossRef](#)]
27. Bertone, E.; Stewart, R.A.; Zhang, H.; O'Halloran, K. Hybrid water treatment cost prediction model for raw water intake optimization. *Environ. Model. Softw.* **2016**, *75*, 230–242. [[CrossRef](#)]
28. Taş, S.; Okuş, E.; Aslan-Yılmaz, A. The blooms of a cyanobacterium, *Microcystis* cf. *Aeruginosa* in a severely polluted estuary, the Golden Horn, Turkey. *Estuar. Coast. Shelf Sci.* **2006**, *68*, 593–599. [[CrossRef](#)]
29. Bowmer, K.; Padovan, A.; Oliver, R.; Korth, W.; Ganf, G. Physiology of geosmin production by *Anabaena Circinalis* isolated from the Murrumbidgee River, Australia. *Water Sci. Technol.* **1992**, *25*, 259–267.
30. Bertone, E.; Stewart, R.A.; Zhang, H.; Bartkow, M.; Hacker, C. An autonomous decision support system for manganese forecasting in subtropical water reservoirs. *Environ. Model. Softw.* **2015**, *73*, 133–147. [[CrossRef](#)]
31. Exo Water Quality Field Sensors: Features, Specifications, and Comparability to YSI 6-Series Sensors. Available online: <http://www.exowater.com/media/pdfs/EXO-6Series-Sensor-Comparison.pdf> (accessed on 10 June 2016).



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).