Dear Dr. Ghadouani

Thank you for uploading the revised manuscript and the clear structure of your response. I think the manuscript has been improved and many of the critical points mentioned during the review have been properly addressed. Nevertheless, a careful reading still shows some issues that need improvement.

Below I listed my comments regarding your responses according to your enumeration. In addition, there are some further comments listed at the end. As you may see there is a recurring theme centered around the statistical model that you have used. Based on the information provided I have doubts that the model is adequate. I suggest that you apply a mixed model approach instead (in SPSS MIXED instead of GLM (http://www.spss.ch/upload/1126184451_Linear%20Mixed%20Effects%20Modeling%20in%20SPSS.pdf) ; a nice intro in R is available by Bates (http://lme4.r-forge.r-project.org/lMMwR/lrgprt.pdf)).

If my arguments are not valid, please provide an answer - also in the manuscript – that clarifies the situation also for the readers.

Comments to Reviewer 1

#1: ok (although map might yield more insight of it was combined with some land use data, for example)

2, 3: ok

4: On this issue, I don't agree. First, I cannot find the answers you provide in your response in the text (neither on L. 409 – 417 nor between 437 and 445). Second, I think the RDA you have carried out should be included. Whether you include it into the main text or provide it as Supporting Information is up to you to decide. Third, the issue of absolute cyanobacteria biomass and relative fractions seems more important to me than what is appears from the current paper. I ask myself the simple question whether it is worse to have a lower biomass of cyanobacteria (and related microcystin) that contributes a larger fraction to total biomass as compared to a lower fraction but higher absolute biomass. Perhaps, I missed that point. If yes, please indicate where to find it. Otherwise, provide a discussion of this issue in the manuscript (or explain why the question does not make sense).

5-7: ok

Comments to Reviewer 2:

1-6: ok

7: Here, it is not clear how the GLM was actually set up. On L. 260 – 261, it is stated that the three replicates per site and sampling date have been used for calculating an average. Fig. 5 and Fig. 6 and Tab. 2 however, indicate that all single measurements have been used (N = 48, which does not correspond to 3 lakes times 6 sampling dates). Because there were 3 distinct sampling points in each lake (L. 152 – 153), I assume you have also used the location as a random factor (see also comment on top and at the end). Please specify your model precisely by providing the equation. This may go into the Supporting Information.

On the other hand, in Fig. 7 only 5(?) data points are depicted for Lake Bibra. Why that?

8: Again, this is not sufficiently clear: did you analyse autocorrelations for data of each location? The data in Fig. 5 to 7 strongly suggest that there were strong temporal changes in the chemical status of the lake. If this was true, how could distinguish autocorrelation from a trend with such a short time series? Please explain and show some actual data (as supporting information) for illustration.

9-13: ok

Comments to the Editor:

1-4: ok

5: see comment above.

6: ok

7: Thanks for adding Tab. 3 and Fig. 5 – 7. This is very useful!

There is a question of what the regression for all lakes actually means: does this regression correspond to the fixed effect if you consider the three lakes and the three sampling locations in each lake as random effects (in a mixed model)? To be honest: the regression line depicted for all lakes in Fig. 5 – 7 often looks like it was calculated based on the assumption that all 48 data points were independent (which they aren't). As a consequence the slope seems often to be controlled by the large range observed in Lake Bibra. This can be illustrated by Fig. 6A or 6B. The slopes for all three lakes individually are steeper than for the entire data set. How can that be? Please specify the model used for the regression for the individual lakes and for the entire data set (see also comment on top and at the end).

8-10: ok

In addition to comments to your answers, I came across some additional points that should be addressed. I list them below:

1: Abstract: The text is not very elegant. For example, the first sentence ends with "… management strategies." The second one starts with "In the management …".

For that section I have tried to come up with a reworded version like "Toxic cyanobacterial blooms in urban lakes present serious health hazards to humans and animals and require effective management strategies. Managing such blooms requires a sufficient understanding of the controlling environmental factors. A range of them has been proposed in the literature as potential triggers for cyanobacterial biomass development and microcystins formation. …".

Feell free to make use it or leave it – it is just a suggestion. Not more.

2: L. 36: Please insert "… in three urban lakes in Western Australia."

3: L. 41 – 42: Please explain a bit more in detail.

4: L. 54: Do all urban lakes suffer from this problem? Be more careful with the wording.

5: L. 65: spatial basis: do you mean within a lake or within a geographical region? Be more precise.

6: L. 68: Do you mean populations or planktonic communities here?

7: L. 77, 80: phosphorus/iron concentrations

8: L. 97 – 99: Be more specific why you expect site-specific relationships. Is it because of additional factors (not explicitly accounted for) that distinguish these lakes or is it because of the interplay of the factors that are explicitly considered? Perhaps you can consider this aspect also for the discussion part.

9: L. 145: For which period have these mean values been reported?

10: L. 177: Be consistent with the spelling: is it Fluoroprobe or FluoroProbe?

11: L. 246: Knowing the temperature range in the water implies actual measurements. Why haven't you used them directly?

12: L. 244 – 248: This is not really data processing or statistical analysis. Please move to 2.2.

13: L. 248 – 252: This part is not sufficiently clear. Please provide the actual model equation in the SI and show some of the corresponding data (see also comment above (Reviewer 2) and below regarding Sec. 3.1).

14: L. 252 – 257: Please move to 2.2.

15: According to Fig. 5 – 8 and Table 3, you have used the individual data points and not the average values. Please clarify in the text.

16: Section 3.1: I think this part contains important information for understanding the situation in the three lakes. Tab. 1 demonstrates that many variables had a large range. I assume that this implied severe changes over time (see also comment by Reviewer 2). Please describe here the most prominent temporal patterns in the data (including covariance of water chemistry parameters).

In addition, provide explanations why the nutrient levels change so dramatically. This may be important for understanding the context of algae dynamics as well.

17: L. 326 - 328: You make a general statement regarding nutrient levels. However, you did not include N species in this list. A positive correlation with the TN:TP ratio can be due to relatively low TP concentrations or due to relatively high TN concentrations. Please specify.

18: L. 341: Consistency can take different forms: there could be a consistent pattern in that the correlations for all lakes are either positive or negative for a given independent variable. Having the same slope and/or intercept is a much stronger criterion for consistency. Perhaps you may differentiate here because for some correlations even the sign of the slope is different between lakes (see Fig. 5B or C).

19: L. 352 – 353: There is a negative correlation between TDP and cyanobacterial fraction in Lake Bibra.

20: L. 430, 451: Skip “variability of”: the cyanobacterial fraction is correlated directly with the mentioned water quality parameters. If you actually refer to the variability then you need to show actual data.

21: Table 1: Please explain the ANOVA that you list in the last column: What was actually compared?

22: Table 2: The caption seems to be at odds with the description on L. 266 – 271. Please specify the model used for this analysis. Make sure that you properly take into account that not all data are independent: that’s what is your manuscript about if you consider the three lakes! Additionally, you may also have to account for the three sampling locations in each lake. Accordingly, a mixed model would be an appropriate approach. Such a model would directly yield the input you need for Table 3.

23: Table 3: Units in Column 1 can be skipped.

24: Fig. 5 – 7: It seems that the regression for all lakes is calculated with an ordinary linear regression using all data points (see comment # 7) without taking the lake-specific dependence into account. Based on Table 3 I assume that this holds true also for the regression lines of the single lakes. There are two comments regarding these regressions: First, describe more precisely what how the regression you show have actually been derived. Second, use a regression approach that correctly accounts for the dependencies of the data. Because your data demonstrate that the data are influenced by the lakes, pooling all data for an ordinary regression is not appropriate. A mixed model would be a valid option.

Sincerely

Dr. Christian Stamm
Editor HESS
Environmental Chemistry
Eawag
Ueberlandstr. 133
CH-8600 Dübendorf, Switzerland