Response to Reviewer #2

Many thanks to Reviewer #2 for a very useful review. Below, we provide comments (in red text) in response to each point raised and suggestions for changes we would make during revisions.

General comments:

This study presents an application of a Continuous Bayesian Network (CBN) to seasonal (6-month average) algal forecasting in a northern lake. This is likely the first use of CBN for this purpose. In general, the model performs similarly to a traditional (discretized) BN and a naïve model (using the mean from the previous year). It could be a good fit for this special issue, but I do have several concerns, as outlined below.

I'm not really sure that there is a strong contribution, as the CBN does not perform particularly well.
Response: Studies using Bayesian Networks (BN) are dominated by discrete networks (as we lay out in the introduction). However, I think one of the main points of the paper is that we found that a very simple continuous BN, a Gaussian BN (GBN), had many advantages compared to a discrete BN: it was much quicker and less subjective to develop, and the fitted conditional probability densities (CPDs) at each node were more robust. I hope that just demonstrating this simple, easy-to-use alternative to discrete BNs for water quality modelling is an important contribution to the field, given the booming popularity of BNs in environmental modelling. If you agree, then we can try to highlight this point more throughout the text (abstract, introduction, discussion, conclusions). You are of course right that a challenge in arguing the case is that the GBN performance wasn't very impressive in this case study. But it performed as well as/better than the discrete BN, and I do think that other groups looking to develop BNs may be interested in our experiences. Particularly if we tidy up our scripts (which we intend to do), and make them available via e.g. Zenodo.

Also, the model appears to be based on existing software (an R package), so there isn't new methods development. If the objective of the study is to provide a thorough demonstration of CBNs for algal bloom modeling, that could potentially be an important contribution. In this case, I'd like to see more demonstrations of how the CBN approach (e.g., Figure 7) can be advantageous for studying a system or supporting management. In my opinion, the current discussion is too focused on skill assessment (e.g., R2), which probably doesn't do justice to the CBN approach. Also, probabilistic predictions using various linear covariates can also be obtained through multiple linear regression (frequentist or Bayesian), so why use a CBN? I think there are potentially good reasons for using a CBN, but they aren't compellingly demonstrated in the current manuscript.
Response: We think that the focus on seasonal forecasting of lake water quality is also an important objective. Seasonal forecasting of water quality is in its infancy, despite its potential management relevance. We could only find 3 existing papers which looked at seasonal water quality forecasting – two were focused on river water quality and the third was part of the project which lead to this submission (and thanks for pointing out the Lake Erie seasonal forecasting, we hadn't been aware of that; do you know if there is an associated paper?).

And then you're right, that another objective is to demonstrate the use of a GBN for seasonal forecasting. I do think a thorough skill assessment is important when you are trying to develop a forecasting tool. However, you are right that the BN approach has many other benefits, and we could easily add something to demonstrate some of these. The GBN we developed was used in a prototype seasonal forecasting tool for this lake, and perhaps adding a new section (e.g. at the end of Section 3) which focused on the use of the GBN for management would help do more justice to the value of the BN approach? E.g. we could add: (1) a figure with an

example forecast for a historic year for the lake, to show the kinds of probabilistic information included (which would be harder to obtain from standard multiple linear regression methods), as well as the text interpretations we developed with local managers; (2) a brief discussion of how this kind of forecast could support management. Otherwise, it's worth noting that the CBN approach shares the same (well-documented) benefits as the discrete BN approach, which are already mentioned in the introduction (line 76 - 83).

Also, I'd like to see more discussion of how this effort compares to other CBN (or BN) applications for water quality or environmental sciences, more broadly.
Response: We can certainly add to the literature review in the introduction/discussion sections.

Major comments:

The paper includes a tangential analysis on making predictions at smaller time scales (e.g., Lines 208-215). I recommend removing this material, as it doesn't seem relevant to the main focus of this paper (no CBN was used). Furthermore, this additional analysis doesn't provide new insights (that aren't available through existing phytoplankton literature). It seems a bit "tacked on". If you do keep this analysis, the data should be presented (as in Figure 2 for the six-month model).
Response: We suggest moving these additional analyses (and section 3.1.2) to an appendix, as well as adding figures to present the data. I think it is nice information to include in the paper, so would rather not remove it altogether: the "pre-peak" temporal aggregation backs up the choice of variables used in the (coarse) 6 month temporal aggregation, whilst the monthly analysis highlights that it is wrong to assume that those variables which are responsible for within-year variation can be used to predict between-year variation. But you are right that it is a bit of a tangent, and so moving it to an appendix should improve the paper.

The variable selection process seems ad hoc (Section 3.1.1), making it somewhat hard to follow and likely difficult to reproduce
Response: I think our variable selection process was more robust than is usual in BNs. Many BNs are developed using a top-down approach, where researchers decide in advance which variables and relationships they think are important. We think our approach achieved a good balance of combining robust statistical methods for selecting variables (feature importance analysis) with process-based knowledge, to filter out relationships which are likely spurious, or on one occasion to add in variables which we think are of particular interest. However, I think we can simplify and cut down the text in this results Section, and improve Table 4 (results of the feature importance analysis).

Some of the explanations seem questionable. For example, the article cites previous literature showing that "windier summers" are relevant, but the CBN uses winds from the previous 6 months (prior to summer), right?
Response: There has been some misunderstanding about which periods the different variables relate to (e.g. the "wind_speed" feature is mean wind speed from the May-Oct growing season, i.e. the same time period as lake water quality is summarised over). We can amend the text clarify this, and the different time periods that were used, in Tables 1-2.

I have two general suggestions. First use clear and consistent terminology that clarifies which time periods you are talking about (also use consistent notation across the different figures and tables). Second, drop wind from the 6-month analysis altogether. Much of the text is a somewhat tortuous explanation (at least for this reader) of reasons to include/exclude wind speed, while in reality, the authors readily acknowledge that wind speed is only relevant at smaller time scales (e.g., Lines 443-445: "wind would likely only have an immediate and relatively short-lived effect…"), not ~6 months in advance.

Response: It is important that the time periods referred to are clear, so we will certainly take a careful look through the text, tables and figures and amend things, as well as improving the consistency in our terminology.

Your second suggestion is perhaps not so relevant, now that we have clarified that "wind_speed" relates to the same time period as lake chemistry? I certainly would like to keep wind speed in the analysis. We know that wind is potentially important in these systems for a variety of reasons, and I think it is feasible that a generally windier May-Oct period may result in a generally less stratified lake, and therefore e.g. lower maximum cyanobacteria during that same period. However, we will certainly smooth out the text and simplify things so that it reads more easily.

Detail-oriented comments:

Line 11: Clarify in the abstract that you are predicting a May-October average (rather than daily predictions).
Will do

Line 20: The term "purely parameterized" is used multiple times throughout this manuscript, but I don't understand what it means or how it is justified. As noted above, the parameterization process seems somewhat ad hoc to me.
It means that the conditional probability densities in the GBN and the conditional probability tables in the discrete BN were fitted just using the data, rather than using expert knowledge. We will clarify in the text.

Line 23: Suggest clarifying what is meant by a "naïve forecast" here.
Will do

Line 44: Models for Lake Erie cyanobacteria blooms (including Bayesian models) predict the maximum bloom size months in advance.
Thanks, we'll look into. We were relying on Rousso et al. (2020) here.

Line 56: Could you explain why "colour" is particularly relevant to water treatment or provide a reference?
Will do

Figure 1: Suggest including arrows to show dominant flow directions.
Good suggestion

Table 2: Clarify what averaging periods were used.
Good suggestion, and something Reviewer #1 struggled with too, and we are planning to re-do Tables 1 and 2.

Line 273: Clarify what normality test was used.
Will do

Figure 3, 4, 5: Clarify why only certain features are shown in each figure.
OK

Table 4: The "Feature subset" column is confusing. Use consistent terminology and explain in the caption.
Agreed, will do

Line 370-371: Revise for clarity.
OK

Line 422: Suggest “wind-related” instead of “related” for clarity.
OK

Line 458: The term “credible” usually refers to the uncertainty in a parameter. It could be good to present actual parameter uncertainties (e.g., credible intervals). Also, I don’t think relationships matching the simple bivariate correlations necessarily makes them “credible” in any sense. For example, see literature on Simpson’s Paradox.
Replace “credible” with “plausible”. I don’t know of a good way to generate credible intervals of fitted GBN parameters, but am open to suggestions. More generally, we will re-write this part in a more nuanced way, whilst still maintaining the main message that the fitted coefficients look plausible.

Line 470: Again, I’m not sure using simple bivariate correlations to evaluate a more sophisticated model makes sense.
Response: Again, we can make this more nuanced. But we think that the main message (that the fitted CPTs in the discrete BN were less plausible than those in the continuous BN) is backed up by our exploratory data analysis, and are just a result of the small sample size used to fit the conditional probability tables.

Table 6: To me, making some numbers bold isn’t effective for highlighting unexpected results. It really depends on which particular pair of numbers is being compared. Also, I wouldn’t describe some of these relationships as a “physical” response.
Response: Agreed, we will re-do this table, or consider replacing it with a figure showing CPTs for the whole network (the numbers will change anyway, as we plan on exploring different ways of specifying the priors in the discrete BN).

Line 569: This statement seems too strong and/or requires clarification.
Response: The clarification came in the preceding paragraph, but do you disagree? But we can certainly qualify the statement, e.g. adding to the end of the line: “…a number of studies will have over-estimated its importance, by assuming that the within-year relationship between temperature and algal dynamics can be used to infer future algal responses to increases in summer temperature under climate change”, or similar.

Line 644: This is clearly true (based on the general nature of a GBN), but it wasn’t really explored in this study. I’m not sure why it is a conclusion.
Response: One of the main points of the paper is that continuous BNs, which have benefits over discrete BNs that go beyond pure performance, should be considered more widely when people are developing BNs. Of course you’re right that we didn’t focus on this aspect, so we can re-write to make it clear that it was a general motivation for the study.

Line 659: This seems like a bit of a stretch. I’m not sure that any “expert” can predict an extreme event ~6 months in advance. Maybe the authors mean something else, but I can’t imagine what.
Response: Where data is lacking, BNs are often developed using “expert knowledge”, where researchers decide, for example, that extra nodes should be included, and on the CPT or CPDs at certain nodes (rather than using data counts or fitted distributions). This means that we could, for example, have added in a “winter discharge” – “growing season lake TP” relationship, and decided on the coefficients that define that relationship using a best guess. So the expert knowledge in this case is not about predicting an event 6 months in advance (which would indeed be a stretch), but adding in a best guess of expected (but largely unobserved) behaviour, in an attempt to make the model better able to predict outside the bounds of the training data. We can modify the text on lines 599-600 and 659 to clarify this.