Response to Reviewers

Reviewer 1:

- R1-1: "Why didn't the authors spin up a 2nd set of random forest models for cyanobacterial density, given that those data were available? It is not clear why they authors just point to the general relationship between chl-a and cyano density, rather than actually modeling it."

    - *Response:* Reviewer 1 makes an excellent point here. We are in the process of modelling cyanobacteria more directly; however, we feel that including that with this current manuscript is a bit beyond its scope. Given this we have removed the paragraph in the conclusion pertaining to cyanobacteria.

- R1-2: "The authors rightly acknowledge that the GIS only models do not have great predictive accuracy for any one lake, however the discussion does point to the relevance of this work for making lake-specific management decision. I think it would be very interesting if, in the discussion, the authors focussed on how this method could be applied to questions related to scaling-up process estimates (e.g., GPP contribution from lakes across the nation)."

    - *Response:* Another good point by Reviewer 1. The benefit of the GIS Only model is in its applicability to all lakes. We have added a paragraph to the Conclusions that mentions this (lines 336 -344).

- R1-3: "Issues related to uncertainty with the in situ and GIS only models might also be informative to such future research efforts."

    - *Response:* Related to R1-2. Uncertainity discussed in added paragaraph in Conclusions.

Reviewer 2:

- R2-1: "To broaden the interest of the paper to readers of ESA journals, the introduction could speak more broadly to the problem of ecological classification and its implication for resource management."

    - *Response:* We agree with Reviewer 2 that ecological classification and reserouce managment are a very important topic. We added some text and a citation to the introduction (lines 71-73), but did not include a broader discussion. We did not include this discussion for two reasons. First, we felt it was a bit beyond the scope of our paper. Our focus is the particular issue of modelling chlorophyll a and using a pre-defined classifcation of trophic state. We are not addressing the much larger issue of classifcation. Second, this is a large topic and distilling this to just a few lines in our introduction would have not provided enought information to be particularlly useful.

- R2-2: "Foremost, the approach of modeling trophic state as indicated by chlorophyll concentration in the all variables model strikes me as somewhat circular - essentially modeling one well known indicator of trophic state (chlorophyll) using other well-known indicators of trophic state (TP concentration, turbidity). To the point, Carlson's (1996) trophic state index incorporated both TP and chlorophyll concentration."

– *Response:* We disagree with Reviewer 2's concerns here. Using water quality parameters to predict chlorohpyll *a* has a long history. For instance, in Carlson's original paper on the Trophic State Index (1977), he incorporates information into the indexes using a linear regression between log TP as an independent variable and log chl *a* as the dependent. Many other authors also use nutrients as dependent variables with trophic state as an independent (e.g. Imboden and Gächter 1978, Salas and Martino 1991). Lastly, in Carlson (1996), a single trophic state is not calculated. Three separate TSI's for TP, Sechhi, and chl *a* are calculated and none of these use both chl *a* and TP. If we have misunderstood Reviewer 2's comments we apologize, but as we understand it we feel that the literature and mechanisms support using TP (and other water quality variables) as independent variables in a model predicting chlorophyll *a*.

- R2-3: "A further concern with the random forest model is that a common and important watershed-derived factor - color - was not included (Soranno et al. 2008). DOC might serve as a proxy, but only roughly so (Pace and Cole 2002)."

  – *Response:* We agree that color is an important factor and because of this did include it in our analysis. See Appendix 1 for a list of variables included in the initial model. We used a variable selection step to select variables to be included in the final model. For this particular data and model(s), adding color to the models did not increase the models predictive ability. As our primary goal was prediction of cholorphyll *a* and trophic state, we view this an acceptable approach. We do fear that our explantion of the variables and Appendix 1 was not clear and have changed the wording referring to the Appendix (line 165) of the paper to help alleviate this potential issue. Additionally, there are some significant differences between the paper mentioned by Reviewer 2 (Soranno et al. 2008) and our study. Namely, our study spans a much great spatial extent (the US vs Michigan) and thus we modeled a much wider range of color (median 12, range 0 - 150) and lake types. This alone could explain why color was not adding significantly to predictive ability.

- R2-4: "Another concern with the model is the presentation of its utility in application. The authors present a kappa coefficient of 0.6 as a benchmark for model quality, yet both models fell below that threshold (the GIS only model substantially so) - what is the implication? Even if the GIS only model is successful in specific instances (L322), the boundary conditions for those instances remain uncharacterized."

  – *Response:* We did a poor job a explaining kappa in this regard. Our reasons for including it was to compare across the two models as total accuracy is often considered a poorer metric for comparison. Additionally, given that we have estimates of uncertainty (prediction probability) this can be used to look at kappa, instead of total accuracy, as a fucntion of prediction probability. Indeed, this is likely the better approach. As such we have addressed the concerns of Reviewer 2 and have added some text to the section where we discuss kappa (lines 191-192) and have also changed Figure 10 and Tables 4 and 5 to show Kappa instead of Total Accuracy.

- R2-5: "A final concern is the choice of the 3 km buffer for the GIS analysis. Others have explored the scaling problem associated with landscape predictors of lake water quality in much greater detail (Cheruvelil et al. 2013 comes to mind, in particular), and the choice is not well-justified in the methods."

– *Response:* As Reviewer 2 notes, scaling is a very important topic with regard to landscape predictors. The models we describe here had the primary purpose of using random forests to 1) predict chl *a* and 2) test to see if the GIS derived variables could be used alone. A more careful consideration of various scales (e.g. the EDUs used in Cheruvelil et al. 2013) would be interesting; however, we do not feel that it would appreciably change the predictions. We feel this way for several reason. First, in some preliminary efforts we tried a variety of scales (300m, 1.5km, and 3km), and they had little impact on prediction accuracy. Second, much of the information provided by multiple scales is accounted for in our analysis with the use of local lake scale predictors, the 3 km buffers, and the ecoregions. Alternative scales may very well improve the overal prediction accuracy but our preliminary efforts indicated that those gains would be slight. We chose the 3km buffer as it made intuitive sense as representative of landuse impacts that would not be accounted for by the local lake-scale variables or the ecoregion. We do agree that this was not justified well in the methods and have added some text to be more clear (lines 113-120).

- R2-6: "First, the cyanobacteria analysis is presented only in the discussion, allowing limited scrutiny. If the analysis is of merit, it should be described in the methods and presented in the results."

    – *Response:* Another great point by Reviewer 2. We are in the process of modelling cyanobacteria and have come to the conclusion including the cyanobacteria figure and mention in the conlclusion is beyond the scope of this manuscript and have removed it.

- R2-7: "Also, regarding the partial dependence plots, how is TN concentration saturating with respect to chlorophyll concentration given the N:P - chlorophyll concentration relationship shown?"

    – *Response:* We are a bit unclear on Reviewer 2's concern. But as we understand the comment, this would require looking at the partical dependence of TN, TP, and N:P simultaneously. At this time we are unclear on how this would be done. Additionally, the relative roles of nitrogen, phosphorus and the N:P ratio is a broad area of research and fully exploring this is beyond the scope of this paper. Lastly, our sense is that partial dependence plots are increasingly difficult to interpret with more than one or two variables included. Partial dependence, as we have presented, is a common way around this and does provide a useful alternative to mutli-dimensional plots (Hastie et al. 2009).