Thank you for taking the time to review our paper, we really appreciate your comments. We completely agree that there are challenges with this approach and the method is reliant on several assumptions — in several cases these assumptions are due to a lack of available published data.

In response to the first paragraph, we must point out that theme (4) is incorrect. The method we use takes records of sedimentary P, expressed as loadings, to reconstruct both catchment P yield and mean lake water TP concentrations using a P model. These reconstructions are not based on runoff, elevation, and temperature scenarios but are based directly on lake sediment P records — and provide a continuous record of P yield and TP throughout the Holocene.

Runoff, elevation, and temperature come in where we attempt to look for associations between these environmental parameters and our reconstructed P yield and TP values for four time periods in the Holocene. These associations are explored in Figure 9 and Tables 1 and 2. To make this clearer we will cross refer to the tables in the figure caption (and vice versa) and address where this could be made clearer in the text.

The issue of changing focussing has been well studied but not addressed by us. Although theoretically a variable focussing factor could be used in the model, we feel that it could not be done with sufficient reliability to justify the added complexity to the model, and in many cases the data are not available. We can add a statement to this effect. We do not, however, see why omitting this should lead to the profiles being more alike.

Runoff is inferred from modern data using variable methods depending on data availability. This is described in the individual sites methodology section and in Table B1. In principle a variable value could be used, and we investigated using Hadley Centre model data to do this (also described in the model paper Moyle and Boyle 2021). In practice this method would be based on uncertain, low resolution data — and again we feel
that it could not be done with sufficient reliability to justify the added complexity to the model.

It is true that catchment P yield is influenced by runoff. However, our measure of this is not because it is based directly on lake sediment P loads, scaled for outflow loss using Rp, and normalised to catchment area. It is also true that runoff influences the estimated captured fraction. However, this does not vary strongly with typical variation in runoff through the Holocene.

We can make the methods section clearer to fully describe how runoff was calculated in each instance and can add a discussion of how variable runoff would affect the model outcome in the reliability and limitations section, including checking whether impounded tributaries have impacted on the lakes.

The question about statistical significance of the environmental drivers is an important one. Most are weak as you point out, but we do also find some statistically significant associations with both runoff and mean annual temperature, and feel that this is important to report. At the moment the dataset of lakes is small, and this section provides some initial reflections on these patterns. At the same time we are mindful that you find the section to be a distraction, at least due to its positioning in the document, and we would very much welcome further comment from peer reviewers on this.

Finally in reference to figure 7b, the line displayed on the graph is a 1:1 line for illustrative purposes – it was an oversight to leave this out of the caption and this will be corrected. We will also adjust the understandably misleading sentence on line 370 that refers to this figure (either adding a statistical test or changing “correlates” to “is related to”), again an oversight.

We don’t actually offer a statistical test of the relationship between TP and SI-TP, however this can be added if necessary. The distributions are skewed so regression can only be tested statistically on log transformed data. This yields an adjusted $r^2$ value 74.4%, with $F = 41.61$. A test in which each case is dropped in turn finds minimum values of $r^2 = 69.3\%$ and $F = 30.4$, so with only trivially reduced statistical significance. This supports our statement that SI-TP correlates with measured lake water TP.

Thanks again for your supportive and thoughtful review, this paper is part of Maddy’s PhD work and your comments have been useful and are very much appreciated!