Dear Editor,

Thank you for your letter and the reviewer comments concerning our manuscript entitled 'Remote Quantification of the Trophic Status of Chinese Lakes' (*hess-2022-91*).

We have studied comments carefully and have made correction which we hope meet with approval. Revised portion are marked in blue in the revised manuscript. The main corrections and the responses to the reviewer comments are described below.

Thanks again for your time and help.

Yours sincerely

Sijia Li and Zhidan Wen on behalf of the authors.

Responses to Anonymou Referee#1' Comments

Dear reviewer,

We would like to express our sincere appreciation for your careful reading and helpful comments. These comments are all valuable and helpful for revising and improving our paper, as well as the important guiding significance to our researches. We have studied comments carefully and addressed the points noted below.

Revised portion are marked in blue in the paper. The main corrections in the paper and the responds to the reviewer's comments and remarks are as flowing.

"In this manuscript, the authors present a methodological framework, using stepwise multiple regression analysis to find some band ratios and establish the XGBoost of machine learning approaches to estimate lakes TSI across China. Transferability and applications of XGBoost were tested in three different water classification scenarios, which provides a new idea for complex water color remote sensing modeling of class II water. This manuscript is well written and organized and this work is very meaningful, the method used is reasonable, and the model performance is good. I only have a few minor comments.

1. Line 37, 475, one or two references here can be helpful.

   Response: Thanks for your patient review. According to your suggestion, we add a reference (Guo et al., 2020; *Environmental Pollution,* Photo-induced phosphate release during sediment resuspension in shallow lakes: A potential positive feedback mechanism of eutrophication) in Line 37; a reference (Wen et al., 2016; *HESS,* Influence of environmental factors on spectral characteristic of chromophoric dissolved organic matter (CDOM) in Inner Mongolia Plateau, China). We hope that these revisions and the improved text will be satisfactory.

2. Line 339, 445, linear regression was used to identify significantly sensitive spectral variables related to TSI, the authors state that blue/red, green/red band ratios showed a good regression coefficient (R2 >0.59) with TSI. Have the results (R2 >0.59) been compared with other band combinations? It is best? What about other band ratios? The selection process of the optimum band should be described in detail. The tables or figures with comparative results should be given.

   Response: Thanks for your patient review. Linear regression was used to identify significantly sensitive band combinations could increase robustness of algorithms. According to you suggestions, we add some linear regression coefficients (2-tiailed) between band combinations and TSI in TableS4 (Supplementary Material). We hope that these revisions and the improved text will be satisfactory.

TableS4 The linear regression coefficients (2-tiailed) between band combinations and TSI

Band combinations	R^2	Band combinations	R^2
Band 4/ Band 5	0.71[**]	Band 2+Band 3	0.54[**]
Band 3/ Band 5	0.68[**]	Band 2+Band 4	0.58[**]
Band 2/ Band 6	0.60[**]	Band 2+Band 6	0.50[**]
Band 1/ Band 6	0.59[**]	Band 3-Band 4	0.40[**]
Band 1/ Band4	0.47[**]	Band 3-Band 5	0.40[**]
Band 3/ Band 2	0.53[**]	Band 4-Band 5	0.50[**]
Band 6/ Band 5	0.58[**]	Band 4-Band 6	0.49[**]
Band 1-Band 3	0.48[**]	Band 5-Band 6	0.49[**]

[**] presents the significance level <0.01

3. Line 382, What is the principle of selecting 555 representative lakes for mapping? Are the mapping images of 555 lakes consistent in date? As you know, the TSI derived from images of different seasons, cloud be completely different.

Response: Thanks for your patient review. Wang and Dou et al., (1998) conducted a national early investigation of lake water qualities, and grouped them into five limnetic regions. The results demonstrated that 51.2% lakes of eutrophication in China (Wang & Dou, 1998). Hence, in our study, some historical records of Chl-a, SDD and TP from in comparison to earlier national investigation by Wang and Dou (1998) can be collected and compared. Although some lakes disappeared referring to the numbers recently according to statistics from Ma et al. (2011), these typical Chinese lakes have deteriorated considerably in terms of water quality (calculated TSI) at an alarming rate (Table S6).These can be found in revised manuscript Line 553-562.

Thanks for your suggestion. We selected the MSI images and XGBoost algorithm to map our TSI considering cloud-free images in autumn (Sep. and Oct) with 7-day time window. Then some typical lakes were lengthen the time-window. This is because China stretches over 135°-73°E and 53°-3° N, by way of example in DaGuangBa (Table 1, 18°58'N,109°00'E), there are only 75 scenes level 1C products per year due to less orbits, in comparison to 175 scenes per year in Võrtsjärv (Estonia, 58°15'N, 26°1'E). For higher latitudes countries, the revisit time of MSI is almost every second day due to the overlapping orbits (Toming et al., 2016). There are few clouds in autumn, and some lakes in Tibet are covered by ice in June and July. Hence, most of good images were acquired in autumn to maintain synchronization (Olmanson et al., 2008; Song et al., 2020). The revised text can be found in Line 379 and 381. We hope that these revisions and the improved text will be satisfactory.

4. Line 484-493, the results showed the support vector machine performed worse than XGBoost and random forest. Why? I suggest that specific reasons need to

be explained clearly, from the mechanism of the algorithm or drawbacks or advantages.

Response: Thanks for your patient review. This is because RF and XGBoost are integrated algorithms based on multiple single trees and decision trees. Then the trees are unpruned and diverse, leading to Unpruned and diverse trees lead to a high resolution in the feature space. For continuous features, it means a smoother decision boundary. According to your suggestion, we added these in revised manuscript (Line 436-440). We hope that these revisions and the improved text will be satisfactory.

5. Line 490, the references should be cited here.

Response: Thanks for your patient review. According to your comment, we cited references in revised manuscript (Chen et al., 2016). We hope that these revisions and the improved text will be satisfactory.

Line 404, Fig.8 did not been cited in text.

Response: We are sorry for our careless mistakes. In revised manuscript, we corrected these (Line 391). We hope that these revisions and the improved text will be satisfactory.