Manuscript Number: essd-2021-227
Manuscript Title: Water clarity annual dynamics (1984–2018) dataset across China derived from Landsat images in Google Earth Engine

Response to editor comments:
Publish subject to minor revisions
The reviewer had no further comments and recommended publication of the manuscript. However, I believe the manuscript still requires minor revision to address issues mainly in its writing. Below you may find few specific issues highlighted by the editor. I would encourage the authors to carefully check the manuscript and maybe have a professional language service to help improve the writing.
Response: Thank you for your professional suggestions, we have carefully checked the whole manuscript and found a professional language service institution to improve the writing. The detailed revision could be seen in the version of manuscript marked-up with change.


EDITSPRINGS

EDITORIAL CERTIFICATE

This document certifies that the manuscript listed below was edited for proper English language, grammar, punctuation, spelling, and overall style by one or more of the highly qualified native English speaking editors at EditSprings.

Manuscript title:

An inland water clarity dataset of China made using Landsat observation from 1984 to 2018

Authors:

Hui Tao, Kaishan Song, Ge Liu, Qiang Wang, Zhidan Wen, Pierre-Andre Jacinthe, Xiaofeng Xu, Jia Du, Yingxin Shang, Sijia Li, Zongming Wang, Lili Lyu, Junbin Hou, Xiang Wang, Dong Liu, Kun Shi, Baohua Zhang, Hongtao Duan

Date Issued:

Nov 26 2021

Certificate Number:

ES-202111201311137996

This certificate can be verified on www.editsprings.com/query.asp. This document certifies that the manuscript listedabove was edited for proper English language, grammar,punctuation, spelling, and overall style by one or more of thehighly qualified native English speaking editors at EditSprings. Neither the research content nor the authors' intentionswere altered in any way during the editing process. Documents receiving this certification should be English-ready forpublication; however, the author has the ability to accept or reject our suggestions and changes.

EditSprings provides a range of editing, translation and manuscript services for researchers and publishersaround the world. Our top-quality PhD editors are all native English speakers from famous institutions crossthe U.S., Britain, Canada and so on.Our editors come from nearly every research field and possess thehighest qualifications to edit research manuscripts written by non-native English speakers.

1. The title seems containing too much information. I would suggest the authors to simplify it to it to highlight the key information of the dataset. On the other hand, the dataset is about lakes, but title did not point it out.
Response: Thank you for your valuable suggestion. We have changed the original title into a new one, i.e., An inland water clarity dataset of China made using Landsat observation from 1984 to 2018.

2. Abstract, line 22-23. “In 2018,……”the sentence is confusing. Please rephrase it.
   Response: Thank you for this careful review, we have rephrased this sentence in lines 25-27, i.e., In 2018, we found the number of lakes with SDD < 2 m accounted for the largest proportion (80.93%) of the total lakes, but the total area of lakes with SDD of <0.5 m and > 4 m were the largest, both accounting for about 24.00% of the total lakes, respectively.

3. I would suggest remove the word lake when mentioning the regions to avoid confusion. For example, changing “lakes in the Tibetan-Qinghai Plateau lake region” to “lakes in the Tibetan-Qinghai Plateau region”.
   Response: Thanks for your patient review. We have removed the word of “lake” when mentioning the regions in the paper, and the detailed revision could be seen in the following:
   Lines 27-29: During 1984-2018, lakes in the Tibetan-Qinghai Plateau region (TQR) had the clearest water with an average value of 3.32±0.38 m, while that in the Northeastern region (NLR) exhibited the lowest SDD (mean: 0.60±0.09 m).

   Lines 31-35: At the five lake regions, except for the Inner Mongolia-Xinjiang region (MXR), more than half of the total lakes in every other region exhibited significant increasing trends. In the Eastern region (ELR), NLR and Yungui Plateau region (YGR), almost more than 50% of the lakes that displayed increase or decrease in SDD were mainly distributed in the area range of 0.01-1 km^2, whereas that in the TQR and MXR were primarily concentrated in large lakes (> 10 km^2).

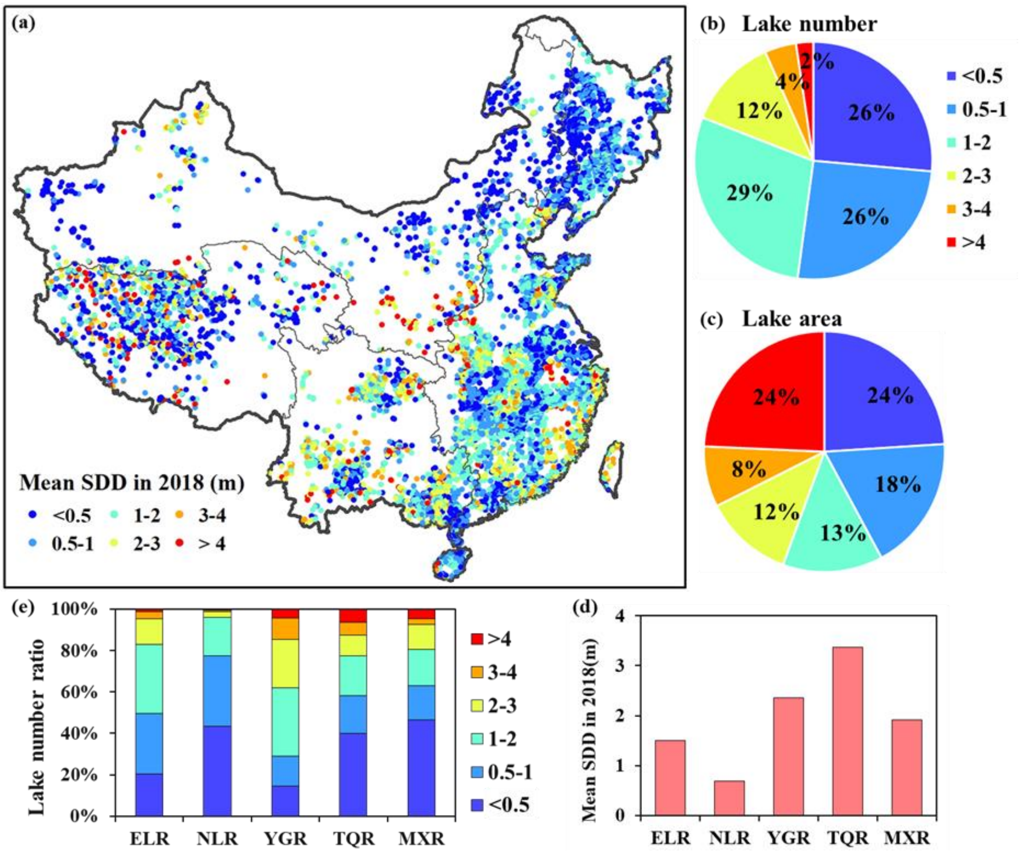
4. I would suggest the authors replace “0-0.5 m” with “< 0.5 m” to be consistent with “> 4 m”.
   Response: Thank you for this suggestion, we have replaced “0-0.5 m” with “<0.5m” in the whole manuscript (including the Figures), and the detailed revision could be seen in the following:

   Lines 25-27: In 2018, we found the number of lakes with SDD < 2 m accounted for the largest proportion (80.93%) of the total lakes, but the total area of lakes with SDD of <0.5 m and > 4 m were the largest, both accounting for about 24.00% of the total lakes, respectively.

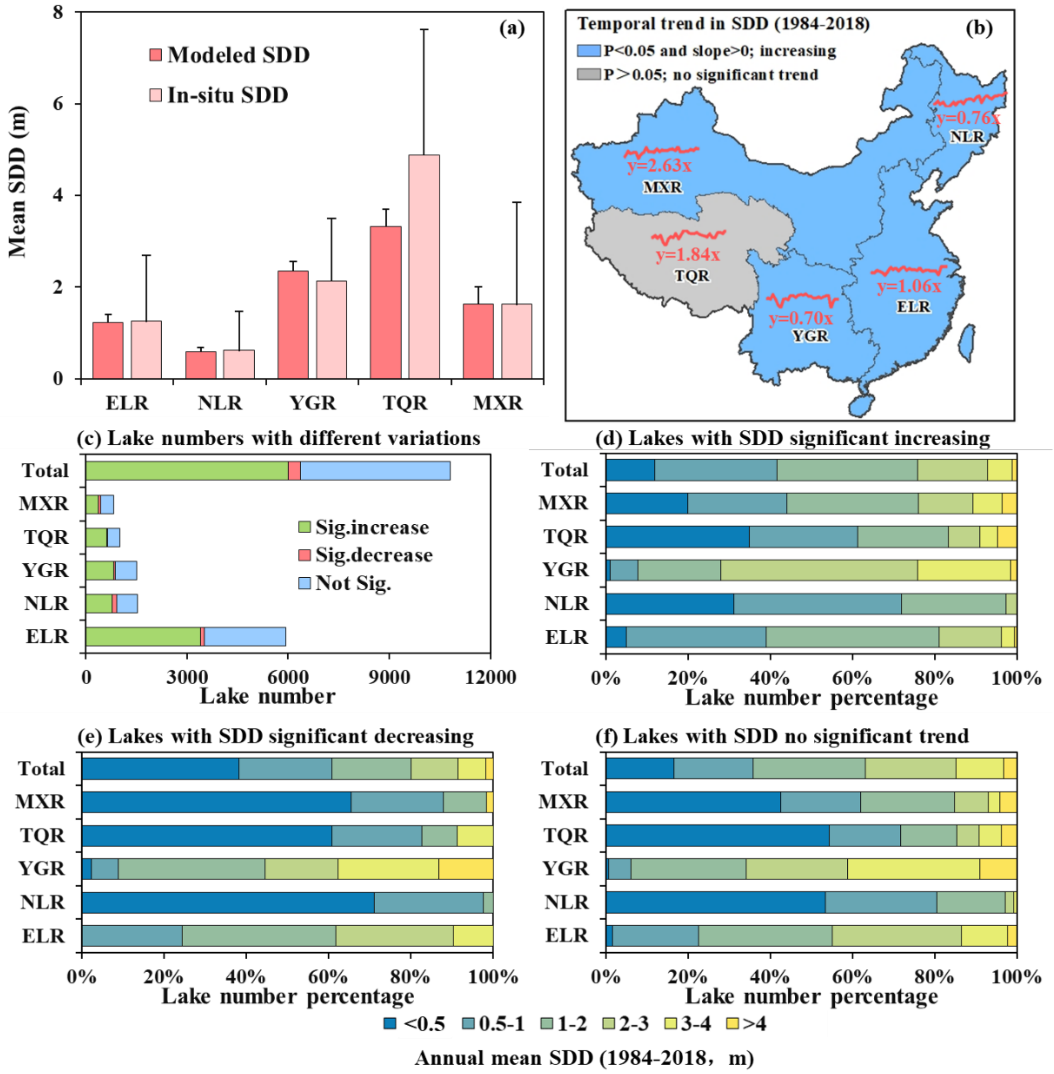
   Lines 249-250: Based on their mean SDD, all lakes across China in 2018 were divided into six levels, i.e., <0.5 m, 0.5-1 m, 1-2 m, 2-3 m, 3-4 m, and >4 m, with 26.4%, 25.7%, 28.8%, 12.5%, 4.3%, and 2.3% of lakes in each SDD level, respectively (Fig. 4b).

   Lines 251-252: Although the number of lakes with SDD < 2 m was more numerous (80.9% of lakes), the total area of lakes with SDD of <0.5 m or > 4 m was the largest, accounting for 24% and 24.3% of the total area in each category, respectively (Fig. 4c).

Lines 268-270: Annual mean SDD of lakes (>0.01 km²) across China in 2018. (a) Spatial distribution of lakes with SDD values. (b) Proportion of lake number with SDD values for six levels (i.e., <0.5 m, 0.5-1 m, 1-2 m, 2-3 m, 3-4 m, and >4 m).



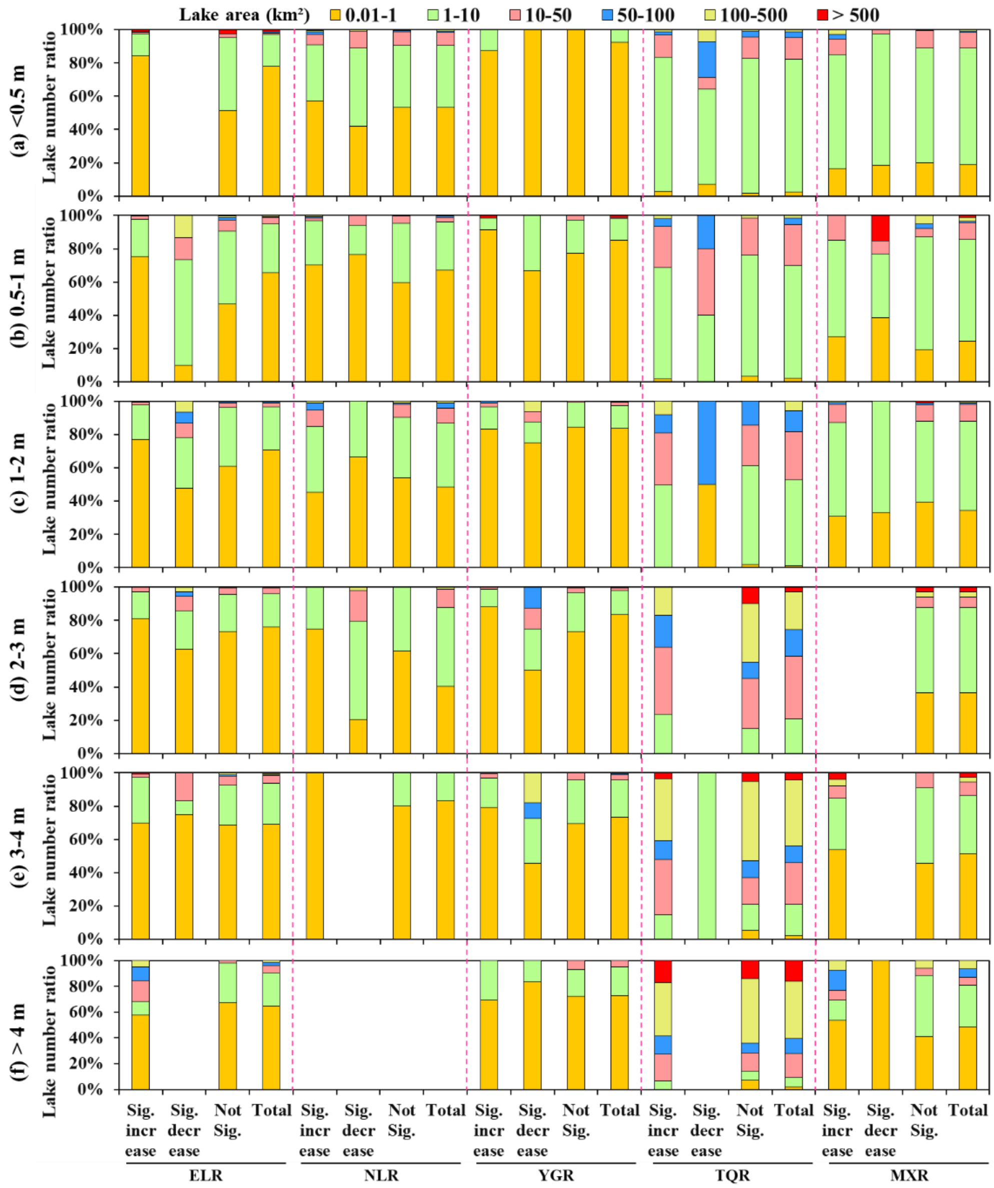
Lines 289-291: Figure 5: The interannual dynamics of lake SDDs in China during 1984-2018. (a) Multi-year average SDD values of the modelled and in-situ SDDs in the five lake regions. (b) Interannual trends of mean lake SDDs in five lake regions based on the 5% significant level and slope representing the coefficient of simple linear regression. (c) Number of lakes with SDD showing statistically significant ($p < 0.05$) increasing (Type I), decreasing (Type II) and nonsignificant (Type III) trends. Proportions of lake numbers with different SDD values (<0.5 m, 0.5-1 m, 1-2 m, 2-3 m, 3-4 m, and >4 m) for: (d) lakes with SDD showing significant increasing trend; (e) lakes with SDD showing significant decreasing trend; and (f) lakes with SDD showing no significant trend.



Lines 320-322: The lakes of Type II, located in the three lake regions, with SDD values of 0.5-1 m in the ELR, and of <0.5 m and 2-3 m in the NLR were dominated by the area size of 1-10 km^2, while the remaining lakes were mostly with the area range of 0.01-1 km^2 (Fig. 6a-f).

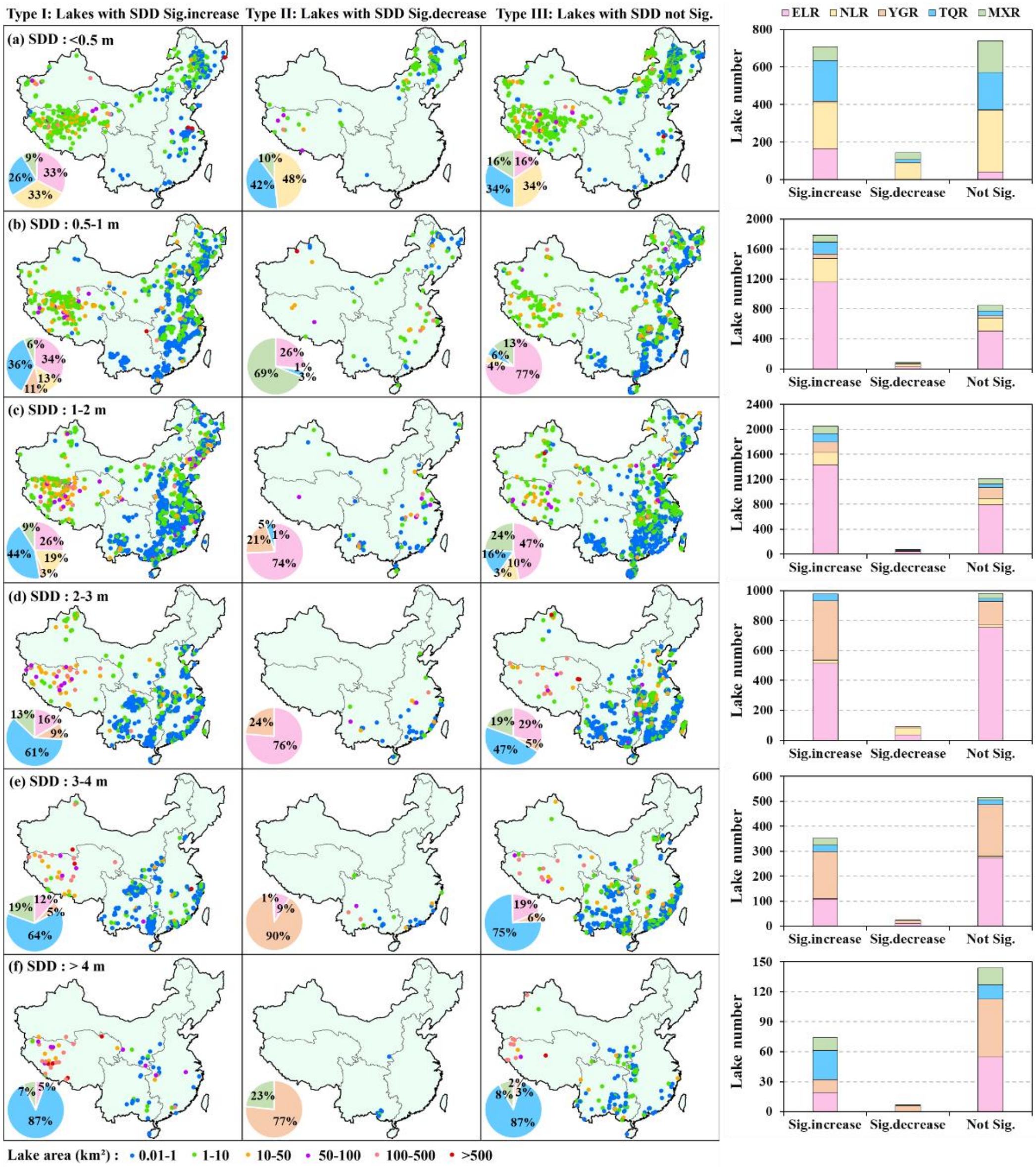
Lines 335-336: For the lakes of Type II in the TQR, the lakes with SDDs in the <0.5 m category were distributed in the area range of 10-50 km^2, followed by that of 50-100 km^2 (Fig. 6a).

Lines 343-345: Figure 6: Proportions of lake numbers in different areas in the six SDD categories. The six SDD categories are: (a) <0.5 m; (b) 0.5-1 m; (c) 1-2 m; (d) 2-3 m; (e) 3-4 m; (f) >4 m. The SDD values are the average of estimated results in each lake during 1984-2018. In the five lake regions, the lakes are further divided into three types — lakes with SDD showing significant increasing (Type I), decreasing (Type II) and nonsignificant (Type III) trends during 1984-2018.



Lines 349-351: In the SDD of <u>≤0.5 m</u> category (Fig.7a), the NLR had the largest lake numbers and areas in the three types of lakes, accounting for 34.51% and 33.20% in Type I, 63.19% and 48.17% in Type II and 44.46% and 34.38% in Type III of the number of lakes and areas in the lake region, respectively.

Lines 386-390: Figure 7: Spatial distribution of lakes with multi-year average SDD values during 1984-2018. The SDD values were divided into six levels: (a) <u>≤0.5 m</u>; (b) 0.5-1 m; (c) 1-2 m; (d) 2-3 m; (e) 3-4 m; (f) >4 m. The lakes were separated into three types of lakes—lakes with SDD showing significant increasing (Type I), decreasing (Type II) and nonsignificant (Type III) trends during 1984-2018.Proportions of total lake area and lake number in each lake region are shown in the pie charts and histogram, respectively.



5. Line 131. "delineating water body boundaries…" please confirm if it was to delineate the water body boundaries or simply classify water and non-water at the step?
Response: Thank you for this question, the word of "delineating" is something of a misnomer, and we have replaced it with "extract" in lines 133 and 139, i.e., We extracted the boundaries of these changing lakes using Landsat images during 1984-2018.; First, we used the MNDWI, combined with Tasseled Cap Transformation (TC) and a density slicing with multi-threshold approach, to build a decision tree for extracting water body boundaries using the ENVI software package (Rokni et al., 2014; Xu et al., 2006).

References:

Rokni, K., Ahmad, A., Selamat, A.,Hazini, S.: Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery, Remote Sens., 6, 4173-4189, doi:10.3390/rs6054173, **2014**.

Xu, H.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, Int. J. Remote Sens., 27, 3025-3033, doi:10.1080/01431160600589179, **2006**.

6. Line 140, "we mainly viewed the Landsat (5/7/8) and Google Earth images to confirm the changing region" the sentence is confusing.
   Response: Thank you for this professional question, we have described this sentence clearer than before in lines 128-130, i.e., As for the reservoirs, we mainly viewed and compared the Landsat natural color images on the website of Earthdata Search (https://search.earthdata.nasa.gov/) and historical images embedded in Google Earth to confirm the changing region, respectively.

7. Line 141, "With respect to the small lakes with an area < 1km^2, we assumed that their boundaries didn't change during the study period" How the small lakes boundaries was determined?
   Response: Thank you for your patient review. We may not express the meaning of this sentence clearly to make reader confused, and we have revised this sentence in lines 131-132, i.e., For the small lakes with area < 1 km^2 obtained from the study of Song et al. (2020) we assumed their boundaries to remain unchanged during the study period. In the beginning of this paragraph, we gave a brief introduction about lake boundaries: Following Song et al. (2020), the lake boundaries (lakes and reservoirs) with area > 0.01 km^2 across China were derived from Landsat 8 OLI images mainly acquired in 2016, and detailed description on boundary extraction is available in that study.

References:

Song, K., Liu, G., Wang, Q., Wen, Z., Lyu, L., Du, Y., Sha, L.,Fang, C.: Quantification of lake clarity in China using Landsat OLI imagery data, *Remote Sens. Environ.*, *243*, 111800, doi:10.1016/j.rse.2020.111800, **2020**.

8. Line 153, "We divided water bodies into lakes, reservoirs, and rivers according to their shoreline features, and also through referencing to the Global Reservoirs and Dams database (Lehner et al., 2011), Chinese Reservoirs and Dams database, and high-resolution images from Google Earth to tell rivers and reservoirs from water bodies." Maybe the authors distinguished the lakes from reservoirs, and rivers mainly by visual interpretation?
   Response: Thank you for this question, we indeed distinguished the lakes from reservoirs, and rivers mainly by visual interpretation, which took a lot of manpower and time. We have added these key words in lines 143-146, i.e., According to the shoreline features, we divided water bodies into lakes, reservoirs and rivers. By referring to the Global Reservoirs and Dams database (Lehner et al., 2011), Chinese Reservoirs and Dams database and high-resolution images from Google Earth, we distinguished rivers and reservoirs from water bodies mainly by visual interpretation.

References:

Lehner, B., Liermann, C. R., Revenga, C., Voeroesmarty, C., Fekete, B., Crouzet, P., Doell, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Roedel, R., Sindorf, N.,Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, *Front. Ecol. Environ.*, *9*, 494-502, doi:10.1890/100125, **2011**.

9. Line 160, "one pixel buffer inward of water boundary was removed for lakes with an area ≤ 1 km², and two pixels for lakes with an area > 1 km² in order to avoid the influence of adjacent land on water bodies." the numbers of pixels buffered in this study are not well enough explained.

   Response: Thank you for your valuable comment. We have made a detailed explanation in lines 152-156, i.e., In our study, one pixel (two pixels) buffer inward of water boundary was removed for lakes with area ≤ 1 km² (> 1 km²) in order to avoid the influence of adjacent land on water bodies that can result in mixed land-water pixels. The determination of the number of pixel buffered was referenced to the method proposed in the study of Wang et al. (2018) who made a comparison of water-leaving reflectance in the transects selected from the land-water boundaries to identify a suitable buffer distance.

References:

Wang, S. L., Li, J. S., Zhang, B., Spyrakos, E., Tyler, A. N., Shen, Q., Zhang, F. F., Kutser, T., Lehmann, M. K., Wu, Y. H.,Peng, D. L.: Trophic state assessment of global inland waters using a MODIS-derived Forel-Ule index, *Remote Sens. Environ.*, *217*, 444-460, doi:10.1016/j.rse.2018.08.026, **2018**.

10. Line 173-174. Please confirm if all of the Landsat data were processed with LEDAPS, especially the Landsat 8 OLI data. LEDAPS usually only applies to TM and ETM+.

    Response: Thank you for your professional comment. We are sorry that the improper description of TOA products was given here, and we have corrected the mistake in lines 185-188 and lines 410-411. As for the LEDAPS, it is the Landsat surface reflectance products that were atmospherically corrected from raw digital values in LEDAPS and Landsat Surface Reflectance Code (LaSRC) software (Schmidt et al., 2013; Zhang et al., 2021).

    Lines 185-188: The TOA products within GEE were produced using the equations developed by Chander et al. (2009), and the function of these equations was to convert calibrated digital numbers to absolute units of TOA reflectance. The description of Landsat TOA products could be seen on the GEE platform (https://developers.google.com/earth-engine/datasets/catalog/landsat).

    Lines 410-411: The calibrated TOA reflectance products within the GEE were produced using the equations developed by Chander et al. (2009).

References:

Chander, G., Markham, B. L.,Helder, D. L.: Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors, *Remote Sens. Environ., 113*, 893-903, doi:10.1016/j.rse.2009.01.007, **2009**.

Schmidt, G. L., Jenkerson, C. B., Masek, J., Vermote, E.,Gao, F.: Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description, *U.s.geological Survey*, **2013**.

Zhang, Y., Zhang, Y., Shi, K., Zhou, Y.,Li, N.: Remote sensing estimation of water clarity for various lakes in China, *Water Res., 192*, 116844-116844, doi:10.1016/j.watres.2021.116844, **2021**.

11. Line 178, missed space between 1,301 and pairs.
    Response: Thank you for pointing out this mistake, and we have corrected it in line 193, i.e., ……1,301 pairs of in-situ SDD and TOA……

12. Section 6.1 "Average and temporal trend in lakes SDD". Correct "lakes" to "lake". Also, please clarify what the average was applied to, spatially or temporally?
    Response: Thank you for this question, we have revised the title of section 6.1in line 273, i.e., 6.1 Temporal average and trend in lakes SDD.

13. Section 6.2 "Lake SDDs versus different lake sizes in China". I would suggest the authors to rephase the title. Or at least remove the word "different".
    Response: Thank you for this question, we have revised the title of section 6.2 in line 309, i.e., 6.2 Lake SDDs versus lake sizes in China.

14. Line 324, "total lake numbers" could be misleading, because the number is calculated from a group of subjects. Total number usually means another level of aggregation. For example, the total numbers of lakes in the regions is calculated as the number of lakes in all of the regions combined, instead of the number of lakes in a particular region. Please confirm and correct the expresses through the manuscript if it is needed.
    Response: Thank you for pointing out the problem of conceptual confusion, we have corrected these mistakes in the whole manuscript, and the detailed revision could be seen in the following:

    lines 348-352: The spatial distributions of lakes and their <u>number of lakes</u> and areas of the three types of lakes in five lake regions are presented in the Fig.7. In the SDD of <0.5 m category (Fig.7a), the NLR had the largest lake numbers and areas in the three types of lakes, accounting for 34.51% and 33.20% in Type I, 63.19% and 48.17% in Type II and 44.46% and 34.38% in Type III of <u>the number of lakes</u> and areas in the lake region, respectively.

    Lines 363-364: When the SDDs ranged from 1-2 m, <u>the number of lakes</u> and area in the ELR were the largest (Fig.7c).

    Lines 379-380, Regarding the SDD of >4 m category (Fig.7f), the TQR had the largest lake number and area in the lakes of Type I, accounting for 39.19% of the number of lakes and 87.34% of the total lake area, respectively.

Lines 381-383: For the lakes of Type III, the YGR had the most lakes and the TQR had the largest total lake area, accounting for 40.28% of the number of lakes and 87.00% of the total lake area, respectively.

15. Line 379, "… existed some errors". Please rephase it to avoid ill expression. Also, change "validation model" to either "validation" or "model validation".
Response: Thank you for your helpful comment, we have corrected these mistakes in lines 406-309, i.e., On the one hand, the SDD estimation model proposed in this study contained some errors, where the model validation yielded these results: R2=0.80, RMSE = 92.7 cm, RMSE% = 57.6%, MAE= 54.9 cm.

16. Data avaliablity section. replace ".shp file" with "shapefile file".
Response: Thank you for pointing out this mistake, we have corrected it in line 416, i.e., The dataset of water clarity of lakes developed in this study consists of one shapefile file document containing the annual mean values of water clarity in each lake (size > 0.01 km^2) during 1990-2018, with a temporal resolution of 5-year.