Response to Reviewer comments
Manuscript Number: essd-2021-227-RC2
Manuscript Title: Water clarity annual dynamics (1984–2018) dataset across China derived from Landsat images in Google Earth Engine

Response to anonymous referee #2:
Anonymous referee #2:
The manuscript describes a water clarity product derived from Landsat images available through GEE. The manuscript is well organized and written, and the derived product would benefit the community that is interested in inland water management and the response of inland waters to climate change. Before this manuscript is published, this reviewer believes that addressing the following general comments would improve the quality of the manuscript.

1. In the manuscript, some statements have confliction, particularly about the used Landsat images. At some places, it is stated that TOA reflectance is used, but in figure captions, surface reflectance is stated. Please clarify.

Response: Thank you for your careful review, we have carefully checked this problem in the whole manuscript, and the detailed revisions can be seen below.

1) Lines 205-206: Figure 3: Model calibration and validation for SDD estimation with Landsat TOA reflectance product acquired by different Landsat sensors. (The word of "surface" has been revised into "TOA".)

2. Some sentences or statements are confusing, and need to be rephrased.

Response: Thank you for your careful review, the confused sentences or statements have been rephrased. Detailed revisions can be seen below.

1) Lines 75-76: "Yet, the wealth of ecological information contained in the archived Landsat images has not been fully explored." This sentence has been deleted.

2) Lines 118-119: In this study, we use band ratios (Green/NIR or Green/SWIR), Normalized Difference Water Index (NDWI), Tasseled Cap Transformation (TC), and a density slicing with multi-threshold approach to build a decision tree for delineating water body boundaries. (The word of "retrieving" has been revised into "delineating".)

3) Lines 154-155: Landsat imagery atmospheric correction is a key step for water quality inversion (Wang et al., 2009), particularly for monitoring of temporal variation at large scale. This sentence has been rephrased, that is "Landsat imagery atmospheric correction is a key step for water quality inversion (Wang et al., 2009), particularly for monitoring of temporal variation at large scale. The TOA products were produced using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) software within GEE (Schmidt et al., 2013)."

4) Lines 181-182: $Y_{observed,i}$ refers to the in situ SDD measurements, $\overline{Y_{observed,i}}$

refers to the average of observed Y, and $Y_{estimated,i}$ refers to the estimated SDD from the Landsat data. (The explanation of $\overline{Y_{observed,i}}$ has been added.)

5) Lines 311-313: "In the lakes of Type III, the ELR and the TQR had the largest proportions of lake areas when the lake SDDs were between 0.5-2 m and 2-3 m, respectively, accounting for 76.80% (SDD: 0.5-1m), 46.90% (SDD: 1-2 m) and 46.65% (SDD: 2-3 m) of the total lake area in each lake region, respectively (Fig.7b-d)." This sentence has been rephrased, that is "In the lakes of Type III, the ELR had the largest proportion of lake area when SDD was 0.5-2 m, and TQR the largest when SDD was 2-3 m. The percentages of lake area when SDD was 0.5-2 m in the ELR were 76.80% (SDD: 0.5-1m) and 46.90% (SDD: 1-2 m), while that in the TQR was 46.65% (SDD: 2-3 m)."

6) Lines 348-349: Because of the similarity of methods and images used in Zhang et al. (2021) and the present study, it provides a unique opportunity to compare in-situ measured SDD with SDD estimation obtained by Zhang et al. (2021) and in our study. This sentence has been rephrased, that is "Because of the similarity of methods and images used in Zhang et al. (2021) and the present study, it provides a unique opportunity to compare the lake SDD estimation model across China proposed by these two researches.

References:

Wang, M., Son, S.,Shi, W.: Evaluation of MODIS SWIR and NIR-SWIR atmospheric correction algorithms using SeaBASS data, *Remote Sens. Environ.*, *113*, 635-644, doi:10.1016/j.rse.2008.11.005, **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**.

3. There are some grammar errors, and the suggestions and comments from this reviewer can be found in the annotated pdf document.

Response: Thank you for your careful review, the grammar errors have been revised. Detailed revisions can be seen below.

1) Lines 45-46: Across the country, the number of stations dedicated to the monitoring of water quality in lakes (59) and reservoirs (52) is very limited in comparison to the national inventory of lakes and reservoirs. (The word of "are" has been revised into "is".)

2) Line 48: Commonly expressed as Secchi disk depth (SDD) (Carlson, 1977), water clarity provides both a practical and comprehensive measure of the trophic state of aquatic ecosystems (Olmanson et al., 2008; Richardson et al., 2010). (The word of "a" between "and" and "comprehensive" has been deleted.)

3) Lines 57-58: Remote sensing has been widely used for monitoring the spatiotemporal dynamics of SDD at regional and national scales. (The word of

"scale" has been revised into "scales".)

4) Lines 112-113: The percentage distribution of lakes, based on the number of lakes and lakes surface area in the five lake regions is shown in the pie charts. (The word of "are" has been revised into "is".)

5) Lines 113-114: The left one (green box) shows about all lakes extracted from Landsat images (b), while the lower left corner one (red box) displays about lakes with SDD records more than 10 years (c). (The word of "by" has been revised into "from".)

6) Lines 137-139: For the first two datasets, SDD data derived from field surveys (2004-2018) were matched with the top of atmosphere reflectance (TOA) data collected by Landsat satellites overpassing a lake/reservoir within 7 days of field site visit. (The word of "air" has been revised into "atmosphere".)

7) Lines 142-143: For the third dataset, the cloud-free TOA images whose dates were closest to time recorded on the lake survey reports were selected to match the measured SDD, which were between May and October during the period of field survey. (The word of "date" has been revised into "dates".)

8) Line 170: At last, 10,814 lakes (size > 0.01 km^2) were examined for the interannual dynamics of SDD. (The phrase of "used to examine" has been revised into "were examined for".)

9) Lines 183-184: Once the annual mean SDD maps were generated, the average of SDD for each pixel within a lake was calculated for the observation period (1984-2018). (The word of "were" has been revised into "was".)

10) Lines 202-203: Therefore, the estimation of SDD using images acquired by Landsat series of sensors provides a reliable method to examine historical trend in SDD through time series analysis. (The word of "the" has been added before "estimation".)

11) Lines 215-216: Although the number of lakes with SDD < 2 m was more numerous (80.9% of lakes), the total area of lakes with SDD between 0-0.5 m and > 4 m was the largest, accounting for 24% and 24.3% of the total area in each category, respectively (Fig. 4c). (The word of "were" has been revised into "was".)

12) Lines 224-225: The lakes in the NLR were located in the northwest and southwest of the region. In the YGR, the lakes were clustered in the southern and northeast of the region (i.e., mid-east of Sichuan province and most of Yunnan and Guangxi province). (The word of "are" has been revised into "were".)

13) Lines 233-234: (e) the proportion of lake number at different SDD levels in the five lake regions. (The word of "with" has been revised into "at".)

14) Lines 239-240: During 1984-2018, the lakes in the NLR exhibited the lowest SDD (mean: 0.60±0.09 m), followed by the ELR (mean: 1.23±0.17 m). (The word of "the" has been added before "lakes".)

15) Lines 253-255: Among the three types of lakes — lakes with SDD showing significant increasing (Type I), decreasing (Type II) and nonsignificant (Type III) trends from 1984 to 2018, the lake SDDs in the Type I were mainly concentrated in 0.5-3 m, in the Type II were dominated by 0-2 m, and in the Type III widely distributed in 0-3 m. (The word of "were" has been added before "mainly".)

16) Line 260: The titles of the horizontal axis in Figures 5d, 5e, and 5f have been revised from "lake number" to "lake number percentage".
    Lines 264-266: The proportions of lake numbers with different SDD values (0-0.5 m, 0.5-1 m, 1-2 m, 2-3 m, 3-4 m, and >4 m): (d) lakes with SDD showing significant increasing, (e) lakes with SDD showing significant decreasing, (f) lakes with SDD showing no significant trend. (The word of "showing" has been added after "SDD".)

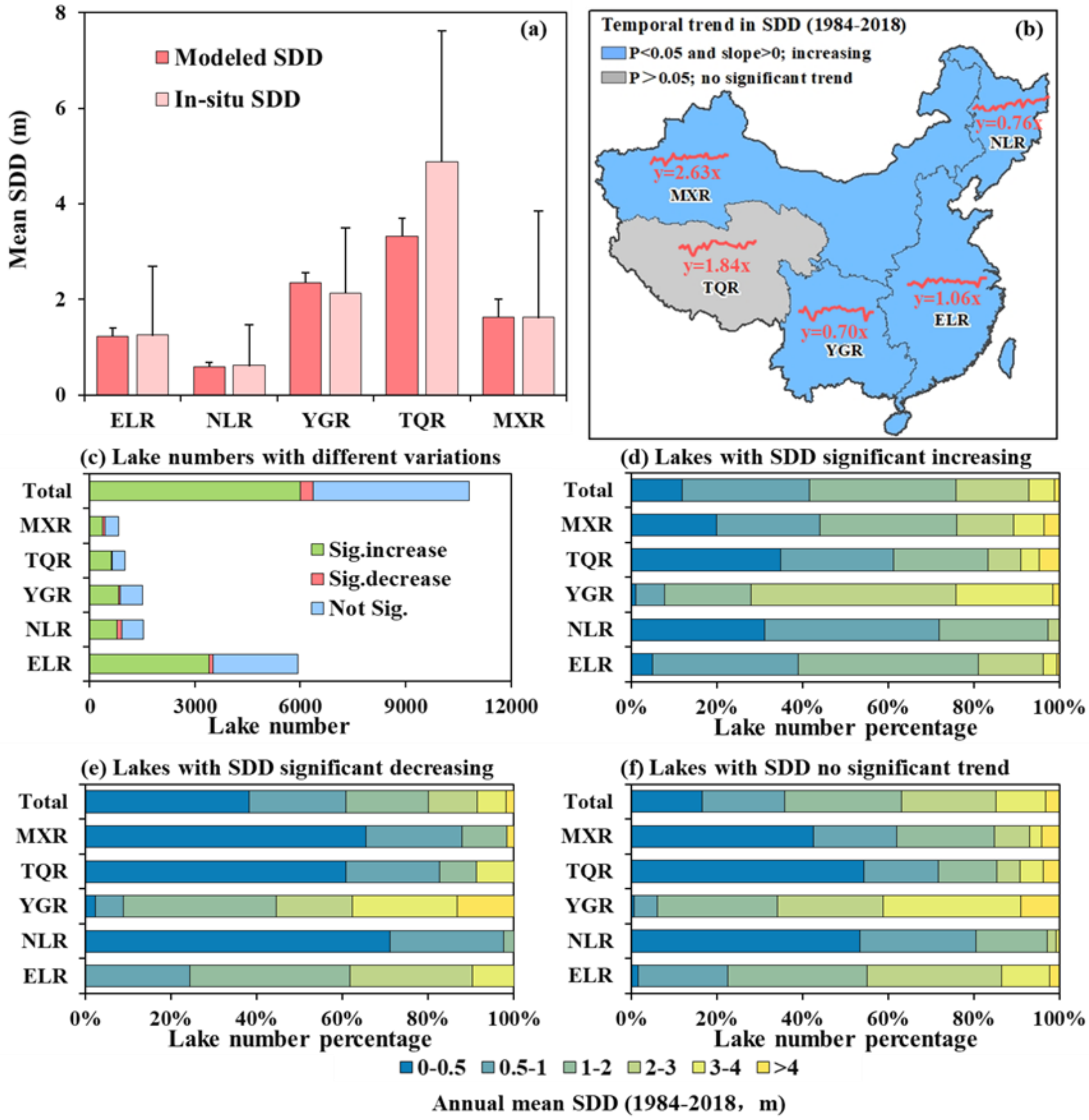

Figure 5: The interannual dynamics of lake SDDs in China from 1984-2018. (a) the multi-year average SDD values of the modelled and in-situ SDDs in the five lake regions. (b) the interannual trends of mean lake SDDs in five lake regions based on the 5% significant level and slope that is the coefficient of simple linear regression. (c) the number of lakes with SDD showing statistically significant ($p < 0.05$) increasing (Type I), decreasing (Type II) and nonsignificant (Type III) trends. The proportions of lake numbers with different SDD values (0-0.5 m, 0.5-1 m, 1-2 m, 2-3 m, 3-4 m, and >4 m): (d) lakes with SDD showing significant increasing, (e) lakes with SDD showing significant decreasing, (f) lakes with SDD showing no significant trend.

17) Line 267: 6.2 Lake SDDs versus different lake sizes in China. (The word of "size" has been revised into "sizes".)

18) Lines 268-269: The annual mean SDD and lake area were both separated into six levels, and the proportions of lakes with different areas in each SDD category were demonstrated in Fig. 6. (The word of "the" between "in" and "Fig.6" has been deleted.)

19) Lines 275-276: Among the three types of lakes in each SDD category, there exists the similarity in the distribution of lakes with different sizes between the Type I and Type III, while that of Type II differentiated from these two types of lakes (Fig. 6). (The words of "in the" between "of" and "Type II" have been deleted.)

20) Lines 280-281: In the MXR, the number of lakes covering an area of 1-10 km^2 in the three types of lakes was much larger than that of other sizes among the lakes with SDDs in 0-3 m range (Fig. 6a-d). (The words of "were" and "more" have been revised into "was" and "larger", respectively. The word of "in" between "of" and "other" has been deleted.)

21) Lines 281-284: When the lake SDDs were > 3 m in this lake region, most of three types of lakes were dominated by the lakes covering an area of 0.01-1 km^2, apart from the lakes of Type III with SDD values > 4 m that the proportion of lakes with an area of 1-10 km^2 was slightly higher than that with an area of 0.01-1 km^2 (Fig. 6e-f). (The phrase of "in the Type III" between "4m" and "that" has been deleted.)

22) Lines 292-293: when SDDs were in the 0.5-1 m category, the number of lakes with an area between 1-10 km^2 and 10-50 km^2 was the largest, the percentages of which both were 40.00% (Fig. 6b). (The words of "were" and "most" have been revised into "was" and "largest", respectively.)

23) Lines 305-307: Spatially, the lakes in the Type I and Type III were mainly distributed in the central of the ELR, the western of the NLR, the mid-west of the TQR and the mid-east of the MXR, while those in the Type II were concentrated on the western of the NLR and eastern of the MXR. (The word of "that" has been revised into "those".)

24) Lines 313-314: In the lakes of Type II, the region that had the largest proportions of lake numbers and areas was inconsistent in each SDD category (0.5-3 m). (The word of "were" has been revised into "was".)

25) Lines 317-318: when the SDDs were in 2-3 m range, the lake number in the NLR was the largest and the total lake area in the ELR was the maximum (Fig.7d). (The word of "numbers" has been revised into "number". The phrases of "were the most" and "were the largest" have been revised into "was the largest" and "was the maximum", respectively)

26) Lines 322-323: In the Type II of lakes with SDD falling in the range 0.5-3 m, their distributions were scattered over part of the central and southeast coastal of the ELR, and southwest of the YGR (Fig.7b-d). (The phrase of "were between" has been revised into "falling in the range".)

27) Lines 327-328: Spatially, the lakes of Type I and Type III were concentrated at the junction of the ELR, YGR and MXR, the southeast coastal of the ELR, the southern of the YGR, and the western of the TQR. (The word of "the" between "of" and "Type I" has been deleted.)

28) Lines 328-329: The lakes of Type III were mainly distributed in the part of the southeast coastal of the ELR and the southern of the YGR. (The word of "were" has been added before "mainly".)

29) Lines 331-332: In the lakes of Type II, a few lakes existed in the MXR and YGR. (The phrase of "there were" has been deleted before "a few".)

30) Lines 339-340: The proportions of total lake area and lake number in each lake region were shown in the pie charts and histogram, respectively. (The word of "showed" has been revised into "shown".)

31) Line 348: Because of the similarity of methods and images used in Zhang et al. (2021) and the present study. (The words of "method" and "Zang" have been revised into "methods" and "Zhang", respectively.)

32) Lines 355-356: The dataset of water clarity of lakes developed in this study consists of one .shp file document containing the annual mean values of water clarity in each lake (size > 1 ha) during 1990-2018, with a temporal resolution of 5-year. (The word of "time" has been revised into "temporal".)

33) Lines 377-379: In-situ water clarity data collected in lakes across China during 2004-2018 was used to calibrate and validate SDD models that incorporate top of atmosphere reflectance product and Google Earth Engine to map the spatiotemporal dynamics of SDD over a 35-year time span (1984-2018). (The word of "air" has been revised into "atmosphere".)

References:

Carlson, R. E.: A Trophic State Index for Lakes, *Limnol. Oceanogr., 22*, 361-369, doi:10.2307/2834910, **1977**.

Olmanson, L. G., Bauer, M. E.,Brezonik, P. L.: A 20-year Landsat water clarity census of Minnesota's 10,000 lakes, *Remote Sens. Environ., 112*, 4086-4097, doi:10.1016/j.rse.2007.12.013, **2008**.

Richardson, T. L., Lawrenz, E., Pinckney, J. L., Guajardo, R. C., Walker, E. A., Paerl, H. W.,MacIntyre, H. L.: Spectral fluorometric characterization of phytoplankton community composition using the Algae Online Analyser (R), *Water Res., 44*, 2461-2472, doi:10.1016/j.watres.2010.01.012, **2010**.