

ESSD-2025-649

TPLake-MED: A Monthly Extent Dataset for Lakes on the Tibetan Plateau

Siyu Zhao, Xiang Zhao, Jiacheng Zhao, Xin Zhang, Xingyu Liu, Chengzhi Yao

In the following document, reviewer comments are presented in **blue**, and author responses are provided in **black**.

Reviewer #2

Comments:

This manuscript presents a timely dataset of monthly area changes for lakes larger than 10 km² on the Tibetan Plateau from 2000 to 2024, derived from the 500 m MODIS surface reflectance product. The authors extracted lake boundaries using a random forest classifier, combined with subsequent filtering and morphological post-processing. This publicly available dataset will provide a valuable foundation for understanding the dynamics of lake changes, their climatic and other environmental drivers, and their impacts on ecological systems and infrastructure security. However, I do have some comments and concerns regarding the current manuscript, and some of the issues need to be addressed before it can be accepted for publication.

Response:

We sincerely thank the reviewer for the comprehensive and insightful evaluation of our work, as well as for the constructive comments provided. Your recommendations have been extremely valuable in helping us significantly enhance the depth and scientific rigor of our manuscript. In response, we have carefully considered every comment and made extensive revisions to thoroughly address your concerns. Detailed, point-by-point responses to each of your comments are provided below.

My major comments are listed below:

Comments:

1. Over the past two decades, numerous studies have focused on lakes on the Tibetan Plateau, including research on lake boundary extraction, change patterns and their

driving factors, and the development of lake inventories. However, the current manuscript appears to cite only a few of these recent studies. I recommend that the authors conduct a more comprehensive literature review of lake-related research on the Tibetan Plateau from approximately the last 20 years. A selection of relevant studies that I have been involved in is provided below for the authors' consideration (but the authors should also review other related studies).

Liao, J., Shen, G., Li, Y., 2013. Lake variations in response to climate change in the Tibetan Plateau in the past 40 years. *International Journal of Digital Earth* 6, 534–549. <https://doi.org/10.1080/17538947.2012.656290>

Li, Y., Liao, J., Guo, H., Liu, Z., Shen, G., 2014. Patterns and Potential Drivers of Dramatic Changes in Tibetan Lakes, 1972–2010. <https://doi.org/10.1371/journal.pone.0111890>

Zhang, J., Hu, Q., Li, Y., Li, H., Li, J., 2021. Area, lake-level and volume variations of typical lakes on the Tibetan Plateau and their response to climate change, 1972–2019. *Geo-spatial Information Science* 24, 458–473. <https://doi.org/10.1080/10095020.2021.1940318>

Response:

Thank you very much for this constructive and insightful comment. We agree that our initial literature review lacked the historical context of lake evolution on the Tibetan Plateau. The papers you recommended (Liao et al., 2013; Li et al., 2014; Zhang et al., 2021) provide important long-term perspectives on lake expansion and climatic drivers. These long-term trends motivated us to produce higher frequency monthly data to capture short-term variations. We have added your recommended papers and other related studies to the manuscript.

The revised text in the **Introduction** section (**Lines 52–54**) is as follows:

“...Lakes are an integral component of the Qinghai–Tibet Plateau hydrological system, modulating regional water cycles and preserving climatic signals across multiple temporal scales (Li et al., 2018; Song et al., 2020). The Plateau hosts the largest and most numerous cluster of high elevation lakes globally (Zhao et al., 2022).

Over the past few decades, extensive research has documented dramatic lake

variations on the Tibetan Plateau, identifying rising temperatures, increased precipitation, and accelerated glacier retreat as primary drivers of this widespread lake expansion (Liao et al., 2013; Li et al., 2014; Song et al., 2014; Yang et al., 2017; Zhang et al., 2021). Variations in lake extent therefore act as critical indicators of regional hydrological dynamics and cryospheric change.”

Liao, J., Shen, G., and Li, Y.: Lake variations in response to climate change in the Tibetan Plateau in the past 40 years, *Int. J. Digit. Earth*, 6, 534–549, <https://doi.org/10.1080/17538947.2012.656290>, 2013.

Li, Y., Liao, J., Guo, H., Liu, Z., and Shen, G.: Patterns and potential drivers of dramatic changes in Tibetan lakes, 1972–2010, *PLoS ONE*, 9, e111890, <https://doi.org/10.1371/journal.pone.0111890>, 2014.

Song, C., Huang, B., Richards, K., Ke, L., and Phan, V. H.: Accelerated lake expansion on the Tibetan Plateau in the 2000s: Induced by glacial melting or other processes?, *Water Resour. Res.*, 50, 3170–3186, <https://doi.org/10.1002/2013WR014724>, 2014.

Yang, R., Zhu, L., Wang, J., Ju, J., Ma, Q., Turner, F., and Dong, H.: Spatiotemporal variations in volume of closed lakes on the Tibetan Plateau and their climatic responses from 1976 to 2013, *Climatic Change*, 140, 621–633, <https://doi.org/10.1007/s10584-016-1877-9>, 2017.

Zhang, J., Hu, Q., Li, Y., Li, H., and Li, J.: Area, lake-level and volume variations of typical lakes on the Tibetan Plateau and their response to climate change, 1972–2019, *Geo-spatial Inf. Sci.*, 24, 458–473, <https://doi.org/10.1080/10095020.2021.1940318>, 2021.

Comments:

2. The authors’ proposed lake extraction method is based on a random forest classifier applied to MODIS surface reflectance data, followed by filtering and morphological post-processing. While the method achieves good overall accuracy, it remains susceptible to issues related to cloud cover and topographic or cloud shadows. To mitigate the impact of cloud cover, the authors developed pre-processing steps that fill cloud masks using imagery from other time periods. However, it remains unclear how the method addresses topographic and cloud shadows. The authors note that previous studies have incorporated DEMs and derived terrain factors to help resolve this issue, yet they did not integrate such terrain information into their own approach. It would be interesting to know the rationale behind this decision—specifically, why terrain

features were not considered and whether the authors explored their potential to further improve classification accuracy in shadow-affected areas.

Response:

Thank you for your insightful question. We did explore the potential of integrating DEM during our algorithm design phase. However, we ultimately opted for an alternative strategy based on scale constraints, computational efficiency, and the dynamic physical nature of cloud shadows.

To systematically answer whether terrain features could further improve accuracy, we conducted a spatio-temporal ablation study using both a highly rugged canyon region (Yamdruk Lake, Region: 2) (Fig. 1a1, b1) and a flat control region (Selin Co, Region: 1) (Fig. 1c1, d1) across different time (high vs. low sun angles). We compared our spectral-only baseline with a model incorporating SRTM Elevation, Slope, and Aspect, using the synchronously computed water area from 10 m Sentinel-2 imagery as an independent reference (Fig. 1).

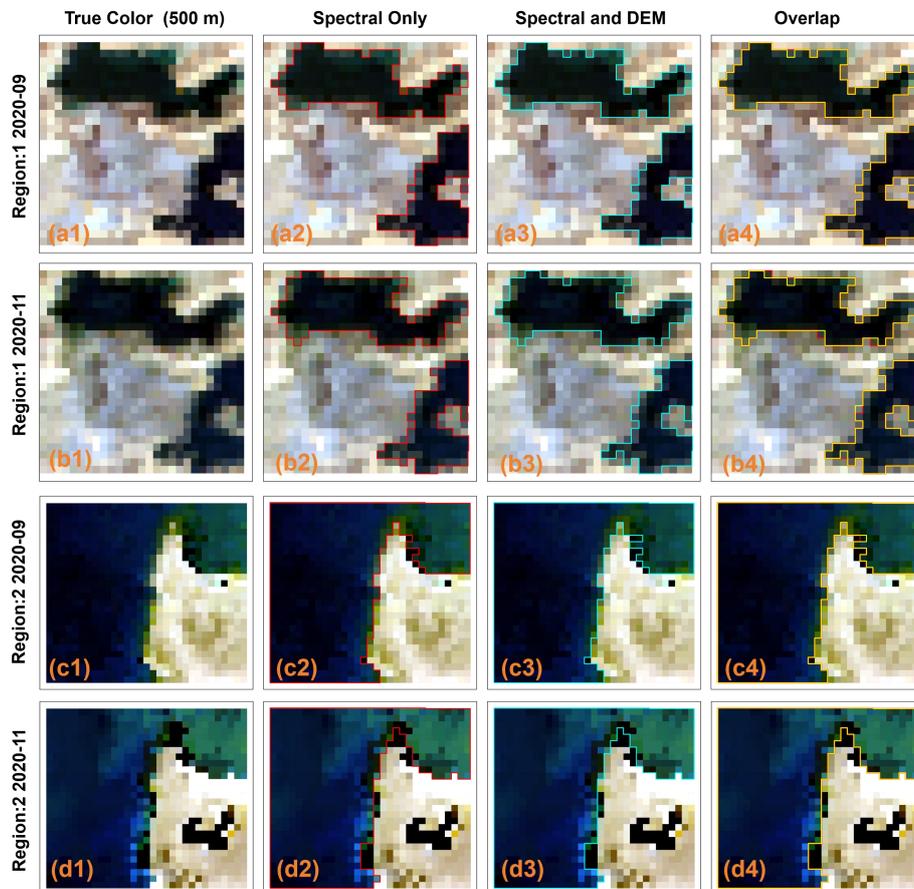


Figure 1: The examples of the visual ablation study for topographic feature integration

across rugged canyon (a, b) and flat (c, d) scenarios. (a1–d1) The true color composite of MODIS 500 m imagery. (a2–d2) The spectral-only baseline extraction result in red. (a3–d3) The DEM-enhanced extraction result in cyan. (a4–d4) The overlay result of both models. The background subfigures use MODIS/Terra MOD09A1 Collection 6.1 8-day (500 m) imagery; data ©NASA EOSDIS/LP DAAC.

Table 1: Ablation study of topographic features for water extraction.

Region	Time	Sun Angle & Shadows	Sentinel (km ²)	Spectral (km ²)	Error (%)	Spectral+DEM (km ²)	Error (%)
Region:1	2020-11	Low Sun Angle	220.06	218.99	-0.49	218.99	-0.49
	2020-09	Moderate Angle	227.70	226.75	-0.42	226.5	-0.53
Region:2	2020-11	Extreme Shadow	136.41	135.56	-0.62	134.08	-1.71
	2020-09	Moderate Angle	138.20	136.44	-1.27	133.45	-3.44

In flat regions (Region 1), the relative area errors against the Sentinel-2 baseline remain extremely low and nearly identical for both the spectral-only (-0.42% to -0.49%) and DEM-enhanced (-0.49% to -0.53%) models, indicating that terrain features offer negligible value where topography is absent. However, in rugged terrains (Region 2), the inclusion of DEM actively degrades accuracy by exacerbating omission errors. Even in the most extreme topographic shadow scenario (Yamdrok Lake in November), the relative error worsened from -0.62% (spectral-only) to -1.71% (Spectral+DEM). Under moderate sun angles in Region 2, this error further increased from -1.27% to -3.44% when terrain features were added. The visual overlay (Fig. 1a4–d4) confirms that this increased omission is primarily associated with severe shoreline erosion rather than the effective removal of shadow artifacts.

Furthermore, integrating high-resolution DEMs across a 25-year Tibetan Plateau time series severely inflates memory overhead and computation timeouts in Google Earth Engine. Crucially, static DEMs remain entirely ineffective against dynamic cloud shadows. This limitation makes the DEM approach functionally incomplete, as it still necessitates a separate pipeline for cloud shadow removal. To efficiently

mitigate both topographic and cloud shadows, we implemented a unified spatiotemporal post-processing pipeline. Our approach filters these artifacts through three constraints: (1) majority filtering and morphological operations to dissolve isolated pixel noise; (2) historical JRC water extents as a spatial baseline to exclude out-of-basin shadow artifacts; (3) a temporal consistency check to filter ephemeral noise. Because true lake expansion is continuous, anomalous water patches appearing in only a single month are effectively identified and removed as transient noise.

Comments:

3. All methods developed or compared in this manuscript (random forest, SVM, and CART) are pixel-based classification approaches. However, object-based image analysis (OBIA) may be more effective for lake extraction, particularly with the recent development of deep learning-based image segmentation models, such as UNet and DeepLabV3+. Unlike pixel-based methods, OBIA segments images into homogeneous objects rather than classifying individual pixels, which has the advantages of better preserving lake integrity, reducing salt-and-pepper noise, and incorporating spatial context and shape information. It would be interesting to know whether the authors considered OBIA or deep learning-based segmentation methods and their rationale for selecting a pixel-based approach for lake extraction.

Response:

Thank you for your valuable suggestion. While we completely agree that Object-Based Image Analysis (OBIA) and deep learning (DL) models excel in high-resolution imagery, our decision to employ a pixel-based Random Forest (RF) classifier was driven not only by computational pragmatism but by the intrinsic physical limitations of coarse-resolution imagery. We base our rationale on three main factors:

(1) OBIA relies on grouping pure pixels to extract shape and texture. However, at the 500 m MODIS scale, lake boundaries are heavily dominated by mixed pixels. Grouping these coarse pixels into larger objects inadvertently blurs the already fragile boundary gradients. While OBIA outperforms pixel-based methods on high-resolution

data, it actually yields lower classification accuracies than pixel-based classifiers when applied to coarse-resolution imagery (Gao et al., 2008).

(2) Generating a continuous 25-year monthly dataset across the entire Tibetan Plateau demands immense resources. Deploying complex DL inference on Google Earth Engine faces severe memory and timeout bottlenecks. Furthermore, constructing a pixel-perfect DL training dataset covering 25 years of seasonal ice and varying turbidity is practically unfeasible, whereas RF generalizes robustly across diverse spectral conditions with significantly fewer samples.

(3) We recognize the inherent salt-and-pepper noise of pixel-based methods. Instead of relying solely on the classifier, we engineered an automated spatiotemporal post-processing workflow. By integrating spatial majority filtering and morphological topology repair, we successfully simulated the contextual advantages of OBIA without suffering the accuracy degradation associated with coarse-scale segmentation.

Nonetheless, we acknowledge the transformative potential of deep learning. As a future research direction, we plan to conduct localized explorations using high-resolution imagery (e.g., 10 m Sentinel-2) and advanced DL models (such as UNet) to map lakes.

Gao, Y., Mas, J. F., Kerle, N., and Navarrete Pacheco, J. A.: A comparison of the performance of pixel-based and object-based classifications over images with various spatial resolutions, *OnLine J. Earth Sci.*, 2, 27–35, <https://doi.org/10.36478/ojesci.2008.27.35>, 2008.

Comments:

4. The authors used the 500 m MODIS dataset for lake classification. It is important to provide a quantitative estimate of the uncertainty introduced by this spatial resolution on lake boundary delineation and area calculations. For example, what is the impact of a one-pixel (or half-pixel) shift along the lake boundary on the estimated area for different lake sizes, particularly for smaller lakes near the 10 km² threshold? Such an uncertainty analysis would help readers better understand the magnitude of lake changes over time and assess whether observed variations are significant compared to the inherent limitations of the data due to the coarse resolution. I

recommend that the authors include this uncertainty assessment to strengthen the interpretation of lake-changing trends.

Response:

Thank you for your valuable suggestion. We fully agree that evaluating the uncertainty induced by the 500 m spatial resolution is paramount. To rigorously address this, we conducted an assessment of boundary shift uncertainties, alongside an evaluation of the signal-to-noise ratio (SNR). These analyses were stratified by lake size (10–50, 50–100, 100–500, and >500 km²) and utilized the entire 25-year dataset (2000–2024).

1. Quantitative Uncertainty of the 500 m Resolution

Following your specific recommendation, we simulated a “worst-case scenario” by assuming a continuous one-pixel (500 m) and half-pixel (250 m) shift along the extracted boundaries for all lakes (Fig. 2). To quantify this impact, the relative area error was estimated geometrically: we multiplied the extracted lake mean perimeter by the assumed boundary shift distance (either 250 m or 500 m) to calculate the maximum potential area of misclassification. This theoretical error value was then divided by the originally extracted lake area to determine the relative area error percentage. For large lakes (>500 km²), the median relative area error under a 500 m pixel shift is minimal (<10%), demonstrating the high robustness of our dataset for macroscopic analyses. However, for small lakes (10–50 km²), the median relative uncertainty fundamentally reaches approximately 40-50% due to the high perimeter to area ratio.

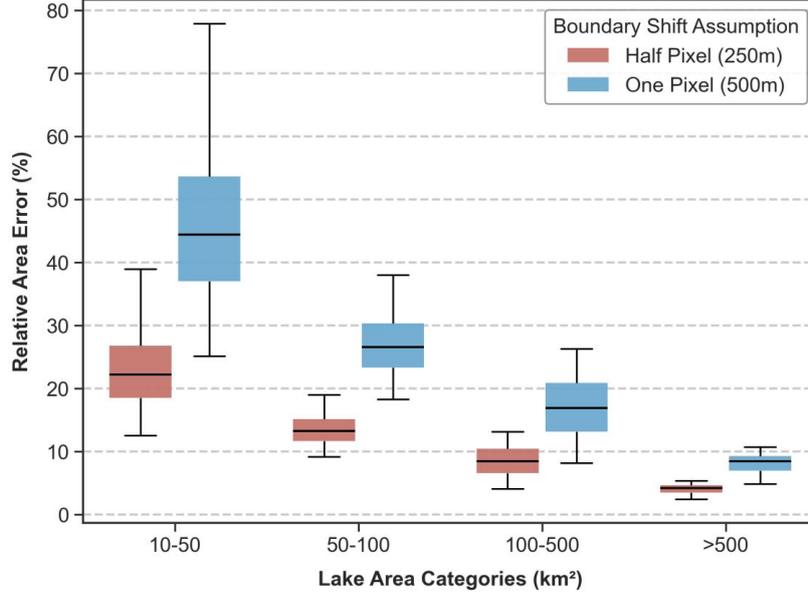


Figure 2: Quantitative assessment of lake area uncertainty induced by the 500 m spatial resolution. The boxplots illustrate the simulated relative area error (%) under hypothetical worst-case boundary shifts of a half-pixel (250 m, red) and one-pixel (500 m, blue), stratified by four distinct lake size categories. (Fig.5a in the revised manuscript)

2.Signal-to-Noise Ratio (SNR) of Real Lake Variations

To determine whether the observed long-term lake area variations are genuine physical signals or merely artifacts of spatial resolution noise, we calculated the SNR for each lake over the 25-year period (Fig. 3). SNR was defined as the ratio of historical absolute area fluctuation ($Area_{max} - Area_{min}$, Signal) to the one-pixel boundary uncertainty ($Perimeter_{mean} \times 0.5$, Noise), calculated as follows:

$$SNR = \frac{Area_{max} - Area_{min}}{Perimeter_{mean} \times 0.5}$$

where $Area_{max}$ is the maximum lake area, $Area_{min}$ is the minimum lake area, $Perimeter_{mean}$ is the mean perimeter of the lake across all valid observations, and 0.5 is the 500 m spatial resolution to unify the unit in kilometers. Specifically, the spatial uncertainty of the extracted water area is inherently proportional to its perimeter. If the classification algorithm misclassifies the boundary by exactly one full pixel outward or inward along the entire edge, the maximum potential false area

generated is the product of the lake's perimeter and the spatial resolution (Wang et al., 2012; Reza et al., 2023). Furthermore, to eliminate the potential exaggeration of area fluctuations caused by transient observational anomalies, the genuine physical variation was conservatively defined using the 95th and 5th percentiles of the historical area time series, rather than absolute extremes.

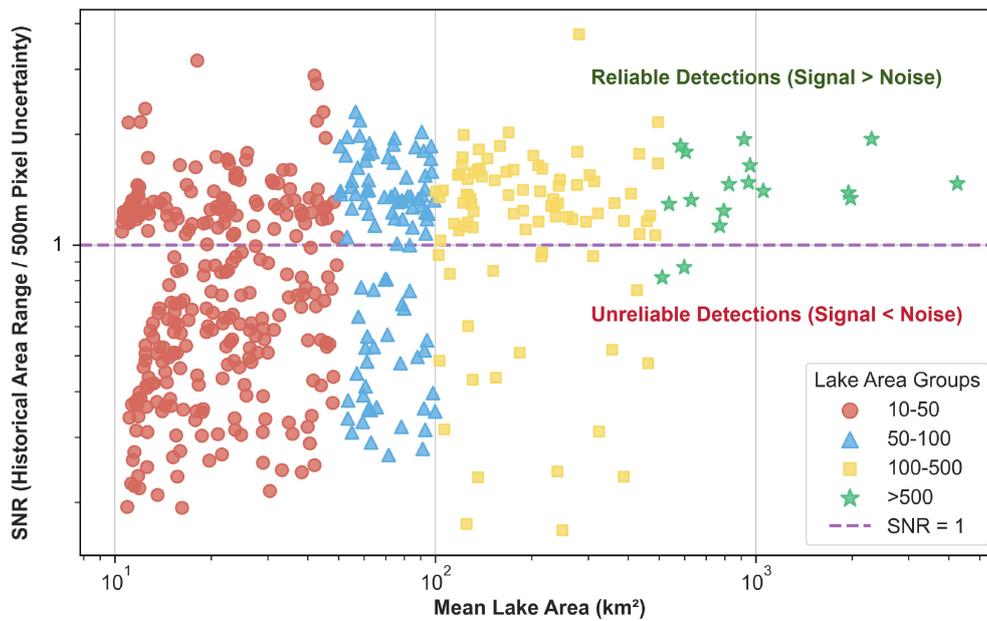


Figure 3: SNR evaluation of historical lake area variations over the 25-year period. Each marker represents an individual lake. The lakes are categorized into four size groups, distinguished by colors and marker shapes: 10–50 km² (red circles), 50–100 km² (blue triangles), 100–500 km² (yellow squares), and >500 km² (green stars). The horizontal purple dashed line indicates the detection threshold of SNR = 1. Regions above and below this line are annotated to denote reliable (Signal > Noise) and unreliable (Signal < Noise) area variation detections, respectively. (Fig.5b in the revised manuscript)

The evaluation of our dataset yields the following statistics:

>500 km² group: Median SNR = 1.41, with 88.2% of lakes exhibiting genuine variations exceeding the noise threshold (SNR > 1).

100–500 km² group: Median SNR = 1.29, with 76.4% valid variations.

50–100 km² group: Median SNR = 1.23, with 63.5% valid variations.

10–50 km² group : Median SNR = 0.73, with 40.4% valid variations.

The SNR analysis precisely defines the observational limits of our dataset. While it is true that for the majority of small lakes (median SNR < 1), the 500 m resolution noise may mask subtle hydrological dynamics, the compelling fact remains that over 40% of small lakes (115 lakes) successfully breached the strict noise threshold (SNR > 1). This evidence proves that even under extreme spatial resolution constraints, a significant portion of small lakes exhibits dramatic, undeniable area fluctuations. This strongly reinforces our initial scientific conclusion: small lakes demonstrate heightened sensitivity to environmental changes compared to larger, more stable water bodies.

We have integrated the uncertainty assessment into the newly reorganized **Results** section (**Section 4.1.2** Uncertainty assessment and Signal-to-Noise Ratio analysis).

Wang, X., Liu, S., Guo, W., Yao, X., Jiang, Z., and Han, Y.: Using remote sensing data to quantify changes in glacial lakes in the Chinese Himalaya, *Mountain Research and Development*, 32, 203–212, <https://doi.org/10.1659/MRD-JOURNAL-D-11-00044.1>, 2012.

Reza, M., Joshi, R. C., Nabiyal, C., and Negi, M.: Glacier retreat and proglacial lake dynamics of Darma Valley, Central Himalaya, India, *Bulletin of Geography. Physical Geography Series*, 25, <https://doi.org/10.12775/BGEO-2023-0009>, 2023.

Comments:

5. The authors used a selection of very large lakes as examples to illustrate lake boundaries and comparisons with other datasets. As expected, the impacts of spatial resolution, cloud cover, and topographic shadows on these large lakes are relatively minor, and the boundaries are consequently of high quality. However, this does not demonstrate the method's performance for smaller lakes, which are more sensitive to these sources of error. For these small lakes, even minor misclassifications along the boundaries can result in substantial relative errors in area estimates. Therefore, I recommend that the authors provide a more systematic accuracy assessment stratified by lake size groups (e.g., 10–50 km², 50–100 km², 100–500 km², etc.). This would allow readers to understand how classification performance varies with lake size and to assess the reliability of the dataset across its full range, particularly for the smallest

lakes where uncertainty is expected to be highest.

Response:

Thank you for your valuable suggestion. We agree that evaluating the algorithm solely on macroscopic lakes may obscure critical errors. To rigorously assess the method's performance across different spatial scales, particularly for small lakes that are highly sensitive to spatial resolution limitations and mixed pixels, we conducted a stratified accuracy assessment using independent high-resolution ground truth.

We selected a representative sample of 22 lakes stratified by the recommended size categories. For each lake, we generated reference masks using 10 m Sentinel-2 imagery from October 2020. Subsequently, we conducted dense stratified random sampling at a 30 m spatial resolution to evaluate the sub-pixel errors along the 500 m MODIS boundaries. To prevent statistical distortion from varying sample sizes across different lakes, we adopted a micro-average approach. Specifically, we aggregated the absolute counts of True Positives (TP), False Positives (FP), and False Negatives (FN) across all sample points within each specific size tier. Furthermore, to address the limitation of the original Figure 8, which exclusively featured macroscopic lakes, we selected one representative lake from each of the four size intervals to visually demonstrate the extraction performance across spring, summer, autumn, and winter.

Across all size categories, precision remains consistently high (>0.95), indicating that our algorithm and morphological post-processing effectively minimize false positives. For the smallest tier (10 to 50 km²), the F1-score decreases to 0.947. This reduction is primarily driven by a lower recall (0.927), which reflects the physical limitations of the 500 m sensor. The coarse spatial resolution intrinsically leads to false negatives along the highly convoluted and fragmented shorelines of small lakes.

The extracted red boundaries for Rena Co, Aweng Co, Hulu Lake, and Zhari Nam Co align closely with the water bodies depicted in the corresponding seasonal satellite imagery, accurately capturing shoreline inflections, peninsulas, and bay entrances regardless of lake size. Our extraction method demonstrates robust performance under varying and challenging surface conditions, including partial snow cover, cloud shadows, and low solar elevation. From a seasonal perspective, these

representative lakes exhibit the typical variation pattern of highland lakes, generally reaching their minimum area in spring, expanding to their maximum area in autumn, and gradually contracting in winter.

We have incorporated and revised the above validation analysis into the **Section 4.1.1** and updated the original Figure 8.

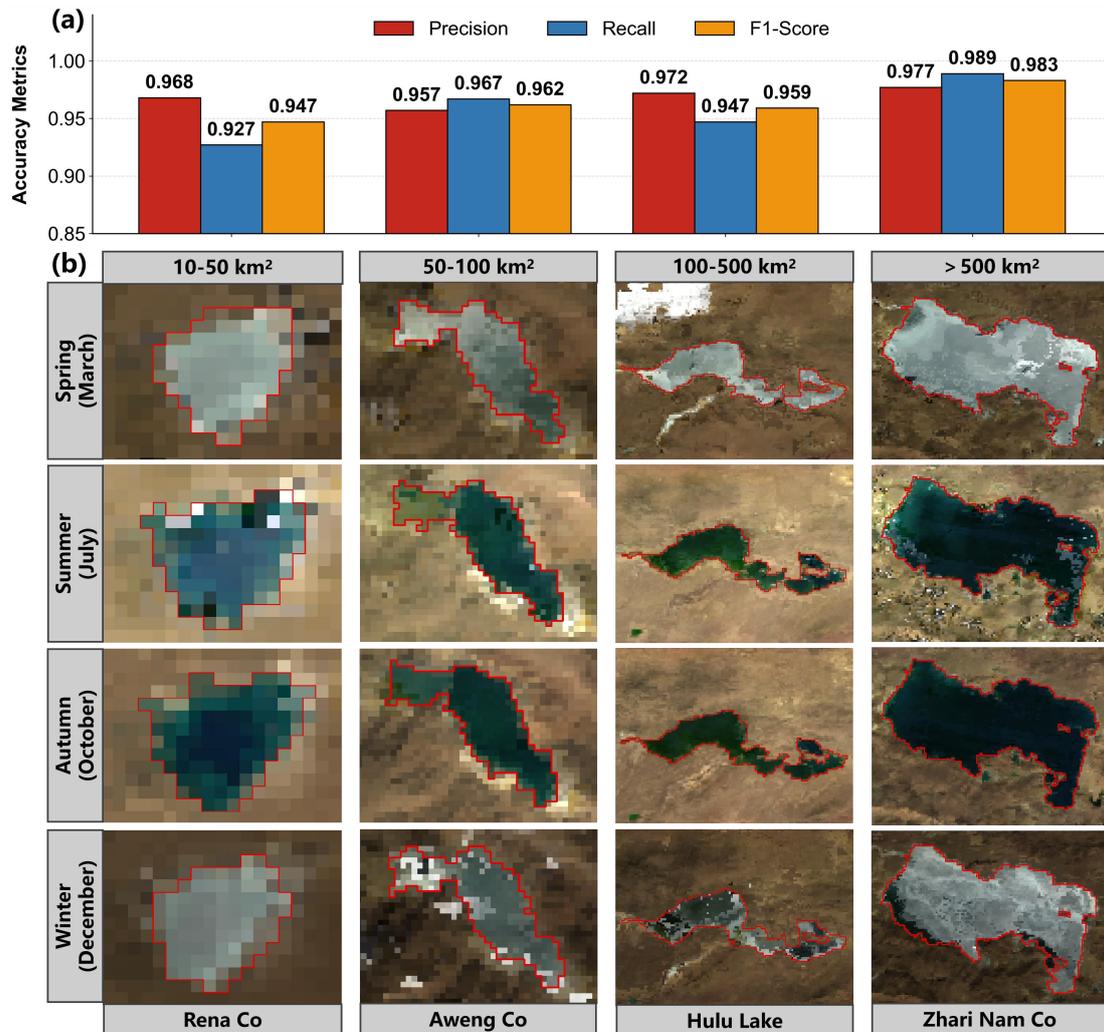


Figure 4: Comprehensive visual and quantitative validation of lake boundary extraction stratified by size categories. (a) Overall spatial accuracy metrics (Precision, Recall, and F1-score) evaluated across four distinct lake size tiers, derived from the comprehensive sub-pixel validation dataset. (b) Seasonal visual extraction results for four representative lakes serving as visual examples for each respective size tier: Rena Co (10–50 km²), Aweng Co (50–100 km²), Hulu Lake (100–500 km²), and Zhari Nam Co (>500 km²). Red outlines indicate the extracted continuous boundaries at the 500

m spatial resolution. The background utilizes MODIS/Terra MOD09A1 Collection 6.1 8-day (500 m) imagery; data © NASA. (**Fig.4** in the revised manuscript)

Some of my detailed comments are listed below:

Comments:

Line 21: The overall accuracy of 93.21% and F1 score of 0.927. I believe that the overall accuracy of 93.31% is based on the confusion matrix of lake and non-lake classification. It is meaningless for the accuracy of the lake boundaries if having imbalanced samples, for example, a large portion of the classification area is non-lake. I suggest that the authors just focus on the lake-specific metrics, such as precision, recall, and F-1 score in the abstract. I guess the F-1 score of 0.927 is for lakes only, which is more meaningful. It should also include other metrics, such as precision and recall.

Response:

Thank you for pointing this out. We agree that Overall Accuracy is a misleading metric in this highly imbalanced binary classification scenario, where the non-water background vastly outweighs the water targets. This metric inherently obscures the actual performance on the minority target class.

As the reviewer noted, the initially reported F1-score of 0.927 strictly represents the performance for the water class. To prevent any misinterpretation and align with rigorous statistical practices, we have removed the Overall Accuracy metric from the Abstract. Instead, we now exclusively report the water-specific metrics (Precision, Recall, and F1-score) to provide a transparent and objective assessment of the lake boundary extraction.

We have revised the corresponding sentence in the **Abstract (Lines 21–22)** as follows:

“A random forest classifier based on spectral indices was developed and validated with 533 water/non-water samples, **achieving a water-class Precision of 0.986, a Recall of 0.875, and an F1-score of 0.927.**”

Comments:

Line 42-52: See my general comments about the literature review

Response:

We have comprehensively revised the literature review in the Introduction to incorporate these foundational long-term studies. Please refer to our detailed response to Major Comment 1 for the specific textual updates.

Comments:

Line 54: what is water body index-based approach? Need to explain. Maybe the authors can explain it in the literature review part (different methods to map lake boundary).

Response:

Thank you for this suggestion. We have added a brief explanation of the water body index-based approach in the Introduction to clarify its fundamental mechanism before introducing specific indices.

The revised text in the **Introduction** section (**Lines 62–63**) is as follows:

“Traditional water body index-based approaches, **which rely on mathematical combinations of specific spectral bands to enhance water signals while suppressing background noise**, have found application in large scale dynamic monitoring...”

Comments:

Line 55: NDWI, MNDWI, AWEI: Need to define these terms when first using them.

Response:

We apologize for this omission. We have now explicitly defined these acronyms upon their first appearance in the Introduction to ensure clarity.

The revised text in the **Introduction** section (**Lines 65–66**) is as follows:

“Common indices such as **the Normalized Difference Water Index (NDWI; McFeeters, 1996)**, **the Modified Normalized Difference Water Index (MNDWI; Xu, 2006)**, and **the Automated Water Extraction Index**

(AWEI; Feyisa, 2014) are extensively utilised with multispectral data from MODIS, Landsat, and Sentinel-2 for water extraction.”

Comments:

Line 53-62: It remains unclear to me how machine learning methods can better handle cloud cover and topographic shadow when relying solely on satellite imagery. I suspect that the improved performance of these newer approaches stems from the incorporation of terrain features, rather than from the machine learning techniques themselves. While machine learning can enhance classification accuracy in general, it is unlikely to effectively mitigate issues related to cloud cover and terrain shadow without the integration of ancillary data such as DEM and other terrain derivatives. The authors should clarify the logic of this part.

Response:

Thank you for pointing out this ambiguity. We agree that machine learning models relying strictly on optical features cannot inherently resolve topographic or dynamic cloud shadows without auxiliary terrain data. The true advantage of machine learning over simple index-based methods lies in its ability to handle non-linear spectral complexity by effectively weighing various spectral signatures, rather than eliminating shadow artifacts. Accordingly, we have revised the Introduction to accurately reflect this mechanism. Furthermore, as detailed in our response to Major Comment 2, the mitigation of shadow misclassifications in our methodology is achieved exclusively through our spatiotemporal post-processing pipeline, not the Random Forest classifier itself.

The revised text in the **Introduction** section (**Lines 69–84**) is as follows:

“Conversely, water body classification methods integrating machine learning (e.g., random forest, support vector machine) **demonstrate favourable adaptability and generalisation capabilities by effectively weighing various spectral signatures to improve the accuracy of lake boundary identification in complex environments** (Li et al., 2021; Zhou et al., 2022). **To further mitigate shadow artifacts, subsequent studies have integrated these models with ancillary terrain**

data (Li et al., 2022; Liu et al., 2022; Yang et al., 2025). For instance, Li et al. (2021) integrated Landsat-8/9 multispectral data with DEM terrain features using a random forest algorithm to extract water bodies on the Tibetan Plateau, achieving an overall accuracy of 95.84%, significantly outperforming traditional methods like NDWI. Liu et al. (2022) integrated Landsat multispectral data and DEM texture features using a random forest algorithm to generate a water body distribution map for the Tibetan Plateau, achieving an overall accuracy of 92.9%, markedly superior to single index methods. Yang et al. (2025) utilised the Google Earth Engine (GEE) platform to classify wetland types using a random forest approach, integrating Landsat time series data from 2000 to 2024 with DEM texture features, which achieved an average overall accuracy of 88.45% and effectively distinguished lakes from marshy meadows. **However, while incorporating DEMs is effective for fine-scale imagery, relying on terrain features at a coarse spatial resolution (e.g., 500 m) introduces severe scale-dependent shoreline erosion and computational inefficiencies, and remains fundamentally ineffective against dynamic cloud shadows, thereby necessitating robust spatiotemporal post-processing to effectively eliminate these residual artifacts and classification noise.”**

Comments:

Line 62-75: There are numerous inconsistencies in the citation format throughout the manuscript. For example, some citations include the year (e.g., Wang et al., 2023), while others list only the author names without the year (e.g., Li et al. and Liu et al.). Please carefully review the entire manuscript to ensure all citations are formatted consistently according to the journal’s guidelines.

Response:

We apologize for this omission. We have conducted a thorough review of the entire manuscript and systematically corrected all in-text narrative citations to strictly adhere to the journal’s formatting guidelines, ensuring that the publication year is never omitted.

The revised text in Section **Introduction** is as follows:

- **Section 1 (Original Line 40):**

Original: “For instance, Li et al. integrated Landsat-8/9 multispectral data with DEM terrain features...”

Revised: “For instance, [Li et al. \(2022\)](#) integrated Landsat-8/9 multispectral data with DEM terrain features...”

- **Section 1 (Original Line 41):**

Original: “Liu et al. integrated Landsat multispectral data, DEM texture features...”

Revised: “[Liu et al. \(2022\)](#) integrated Landsat multispectral data, DEM texture features...”

- **Section 1 (Original Line 42):**

Original: “Yang et al. utilised the Google Earth Engine (GEE) platform to classify...”

Revised: “[Yang et al. \(2025\)](#) utilised the Google Earth Engine (GEE) platform to classify...”

- **Section 1 (Original Lines 45-46):**

Original: “For instance, Zhang et al. constructed triennial area time series... Pang et al. generated continuous area time series...”

Revised: “For instance, [Zhang et al. \(2018\)](#) constructed triennial area time series... [Pang et al. \(2021\)](#) generated continuous area time series...”

- **Section 1 (Original Lines 48-50):**

Original: “Zhang et al. acquired lake area data for 16 periods... Zhang et al. extracted monthly scale area... Ran et al. generated a dataset...”

Revised: “[Zhang et al. \(2022\)](#) acquired lake area data for 16 periods... [Zhang et al. \(2019\)](#) extracted monthly scale area... [Ran et al. \(2023\)](#) generated a dataset...”

We have ensured that this standard is uniformly applied across all sections of the revised manuscript.

Comments:

Line 82: The authors argue that “existing datasets emphasize large lakes, with

insufficient coverage of small and medium lakes”. However, their manuscript also only focuses on large lakes. It is better to revise the logic of this part.

Response:

Thank you for pointing this out. Characterizing our 10 km² threshold as covering “small and medium lakes” while criticizing other datasets for focusing on large lakes is contradictory and misrepresents our actual contribution.

Our intention was not to claim the extraction of micro-lakes, but rather to highlight a specific methodological bottleneck: existing high-frequency, long-term datasets struggle to maintain robust spatiotemporal consistency for lakes below the 50 km² threshold over multi-decadal scales. Pushing a 25-year continuous monthly dataset reliably down to the 10 km² threshold is the actual gap we aim to fill.

To accurately reflect this logic and eliminate the contradictory claim, we have revised the **Introduction (Lines 98–100)** as follows:

Original: “Furthermore, existing datasets emphasise larger lakes, with insufficient coverage of small and medium lakes. Spatial resolution and temporal continuity are often difficult to reconcile (Ma et al., 2022).”

Revised: “Furthermore, maintaining spatiotemporal consistency simultaneously is highly challenging; long-term high-frequency datasets struggle to robustly cover lakes below the 50 km² threshold over multi-decadal scales due to data gaps and methodological inconsistencies (Ma et al., 2022).”

Furthermore, to ensure our empirical validation fully aligns with this revised scope, we have now incorporated an accuracy assessment explicitly targeting smaller lakes (10–100 km²), as detailed in our response to Major Comment 5.

Comments:

Line 100: This sentence repeats the same sentence already presented in the introduction. Remove it to avoid redundancy.

Response:

Thank you for pointing this out. We agree that the repetitive use of descriptive terms such as “Roof of the World” and “Water Tower of Asia” was unnecessary and

distracting. We have removed these terms from the **Study Area** section to maintain a more objective and scientific tone focused strictly on geographical characteristics and lake distribution. The term is now mentioned only once in the **Introduction** to provide context for the region's significance.

Comments:

Line 118: define JRC and IoU. For the dataset, it is better to provide a reference or a website link.

Response:

We thank the reviewer for pointing out this omission. We have now explicitly defined both terms upon their first appearance in the manuscript and provided their foundational citations. “JRC” has been detailed as the Joint Research Centre Global Surface Water dataset, and we have added the appropriate citation (Pekel et al., 2016) along with its official website link. Additionally, we have explicitly defined “IoU” as the Intersection over Union and added its foundational reference (Jaccard, 1912).

The revised text in Section **Method (Lines 139–140)** is as follows:

Original: “Third, post-classification processing included noise reduction, missing pixel detection and filling, overlaying the JRC global surface water mask, morphological smoothing, boundary IoU constraints, and invalid geometry repair to optimize boundaries.”

Revised: “Third, post-classification processing included noise reduction, missing pixel detection and filling, overlaying the Joint Research Centre (JRC) global surface water mask (Pekel et al., 2016; <https://global-surface-water.appspot.com/>), morphological smoothing, boundary Intersection over Union (IoU; Jaccard, 1912) constraints, and invalid geometry repair to optimize boundaries.”

Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, *Nature*, 540, 418–422, <https://doi.org/10.1038/nature20584>, 2016.

Jaccard, P.: The distribution of the flora in the alpine zone, *New Phytol.*, 11, 37–50, <https://doi.org/10.1111/j.1469-8137.1912.tb05611.x>, 1912.

Comments:

Line 262: What is the uncertainty associated with lake boundaries extracted from the 500 m spatial resolution MODIS dataset? The authors should discuss how the relatively coarse resolution may affect the accuracy of lake area estimates, particularly for smaller lakes near the 10 km² threshold (about 40 pixels).

Response:

Thank you for highlighting this point. As this specific concern aligns directly with your broader question regarding spatial resolution limitations, we have addressed it comprehensively. We have conducted a detailed, size-stratified boundary shift uncertainty assessment and a Signal-to-Noise Ratio (SNR) analysis, specifically focusing on the vulnerability of smaller lakes near the 10 km² threshold.

Please refer to our detailed response to Major Comment 4.

Comments:

Line 274-275: The observation that lakes reach their maximum extent in September and October may also reflect a lag effect, as rainfall or meltwater from glaciers and permafrost may require time to flow into the lakes.

Thank you for this insightful hydrological perspective. We fully agree that the delayed peak in lake extent (September/October) compared to the peak precipitation and temperature season (July/August) is a classic manifestation of the hydrological lag effect on the Tibetan Plateau.

The routing of surface and subsurface runoff, compounded by the gradual thawing of the permafrost active layer, necessitates a buffer period. This physical phenomenon is strongly supported by recent studies on intra-annual lake dynamics, which quantitatively identified a one to three month lag period in the response of inland lake areas to monthly precipitation due to delayed hydrological feedback (Liang and Li, 2019; Wang et al., 2025). Furthermore, comprehensive studies on the Tibetan Plateau have explicitly attributed the autumn peak of lake levels to this pronounced routing lag of summer monsoon precipitation and meltwater (Chang et al.,

2022).

We have enriched the manuscript by explicitly incorporating this lag effect mechanism and citing these supporting studies to strengthen the physical interpretation of our monthly time-series data.

Chang, J., Wei, Z., Li, Y., Huang, C., and Zheng, D.: Seasonal trends and cycles of lake-level variations over the Tibetan Plateau using multi-sensor altimetry data, *J. Hydrol.*, 604, 127251, <https://doi.org/10.1016/j.jhydrol.2021.127251>, 2022.

Liang, K. and Li, Y.: Changes in lake area in response to climatic forcing in the endorheic Hongjian Lake basin, China, *Remote Sens.*, 11, 3046, <https://doi.org/10.3390/rs11243046>, 2019.

Wang, Y., Feng, M., Li, B., Sun, J., and Guo, Y.: Recent lake surface dynamics in the Hunshandake sandy land (2017–2022) and their response to climatic factors, *Ecol. Indic.*, 178, 113820, <https://doi.org/10.1016/j.ecolind.2025.113820>, 2025.

Comments:

Line 290-295: Rather than simply describing the relationship between lake changes and climate factors, I recommend that the authors conduct some more detailed statistical analysis to quantify their relationships to strength the lake change driving mechanisms.

Response:

Thank you for this recommendation. We completely agree that a quantitative statistical analysis is essential to substantiate the spatial heterogeneity of the lake driving mechanisms described in our manuscript. To comprehensively address this, we employed a two step statistical framework: 1) a Pearson correlation analysis to assess the temporal covariation, and 2) a standardized multiple linear regression model to quantify the exact relative contribution of each key climate driver (precipitation, temperature, and potential evaporation) to lake area variations. Based on the longitudinal gradient highlighted in our manuscript, we categorized all studied lakes into three distinct longitudinal zones: the Western Plateau, Central Plateau, and Eastern Plateau, utilizing ERA5-Land monthly meteorological time series data from 2000 to 2024. One month lag was applied to precipitation and temperature to account

for catchment concentration times, while concurrent evaporation was used.

In the western region, lake dynamics are jointly dominated by precipitation (39.8% contribution, $r=0.83$) and evaporation (34.1%). This statistically confirms our observation that western lakes are highly sensitive to significant seasonal precipitation variations and intense evaporation. Conversely, in the eastern region, temperature emerges as the primary driver (38.8% contribution, $r=0.74$), while the impact of evaporation drops significantly to 27.2%. Because temperature is the primary driver of cryospheric ablation, this quantifies our statement that eastern lakes are largely supplemented by glacial runoff and experience smoother changes under a more stable climate regime. The central region exhibits a remarkably balanced transition (Precipitation 35.1%, Temperature 32.7%, Evaporation 32.2%).

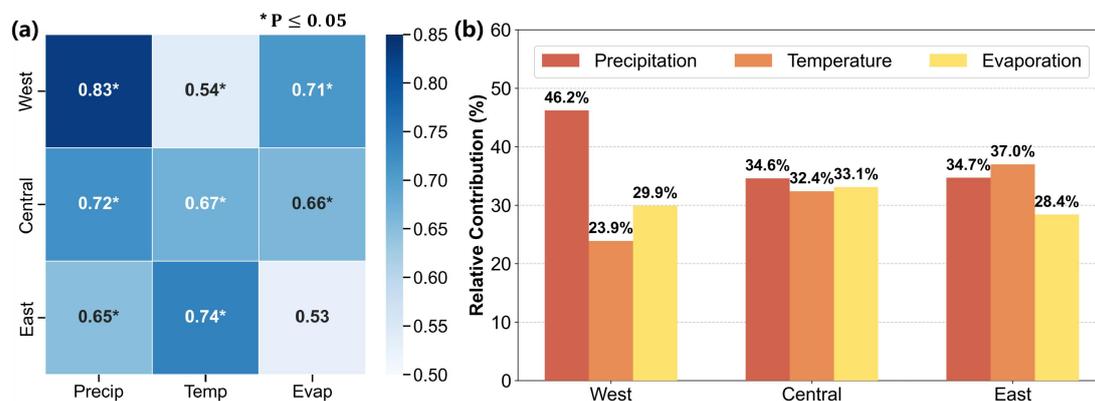


Figure 5: Driving mechanisms of lake area variations across the longitudinal gradient of the Tibetan Plateau. (a) Heatmap of Pearson correlation coefficients (r) between monthly lake area and key climate drivers. To account for catchment concentration times, a one month lag was applied to precipitation and temperature, while potential evaporation was considered concurrently. (b) Relative contributions (%) of the three climate drivers to lake area variations, derived from absolute standardized coefficients of multiple linear regression models.

We have added this composite figure in the Supplementary Materials (Fig. S1) and revised the corresponding paragraph in Section 4.3.2 as follows:

“This spatial variation was closely linked to regional climatic conditions. **To explicitly quantify these driving mechanisms across the longitudinal gradient, we conducted a Pearson correlation analysis coupled with a standardized multiple**

linear regression for all studied lakes within the western, central, and eastern zones (Fig. S1). The statistical results reveal that in the western region, lake dynamics are jointly dominated by seasonal precipitation pulses (39.8% contribution, $r=0.83$) and intense evaporation (34.1%), leading to pronounced lake water fluctuations. Conversely, moving eastward, temperature emerges as the primary driver (38.8% contribution, $r=0.74$), while the impact of evaporation drops significantly to 27.2%. Because temperature acts as the primary driver for cryospheric ablation, this confirms that the eastern lakes are largely supplemented by glacial runoff and experience relatively smoother changes under a more stable climate regime. The central region exhibits a remarkably balanced transition among precipitation (35.1%), temperature (32.7%), and evaporation (32.2%).”

Comments:

Line 315-317: What is the difference between “in the Plateau interior” and “across the interior”?

Response:

We apologize for the confusing and redundant phrasing. There is no strict geographical difference intended between the two phrases; this was a drafting oversight. Our intention was to express that lakes with extreme fluctuations (6.50–26.38 km²) are concentrated deep within the central plateau, whereas lakes with moderate variability (3.91–6.50 km²) are more generally scattered across the wider plateau area. To accurately describe this distribution without causing semantic confusion, we have changed “across the interior” to “across the broader plateau region.”

The revised text in **Section 4.3.2 (Lines 490–491)** is as follows:

Original: “Differences of 3.91–6.50 km² were widely distributed across the interior, indicating moderate variability;”

Revised: “Differences of 3.91–6.50 km² were widely distributed across the broader plateau region, indicating moderate variability;”

Comments:

Section 4.2: I recommend reorganizing the Results section to present the accuracy assessment and uncertainty analysis before the other results. This change will allow readers to evaluate the reliability of the dataset before understanding the patterns and trends of lake changes.

Response:

Thank you for pointing this out. We agree that presenting the spatiotemporal trends before rigorously establishing the validity and uncertainty boundaries of the dataset was a structural and logical flaw in our original draft.

To correct this and ensure readers can immediately evaluate the data's reliability, we have comprehensively reorganized the Results (Section 4). We have moved the validation content to the very beginning and integrated the newly conducted uncertainty assessment into this leading section.

The revised structural hierarchy in Section **Results** is as follows:

- 4.1 Validation and uncertainty analysis of lake extraction
 - 4.1.1 Classification accuracy and boundary validation
 - 4.1.2 Uncertainty assessment and Signal-to-Noise Ratio analysis
- 4.2 Comparison with existing products
- 4.3 Spatiotemporal patterns of lake area changes (2000–2024)

We have also updated all corresponding figure numbers and cross references throughout the manuscript to reflect this logical progression.

Comments:

Line 361-362: All lakes shown in Figure 8 are the largest lakes on the Tibetan Plateau, which are likely to yield better classification results and be less sensitive to issues such as cloud cover and topographic shadows. However, I believe the key test of the method's robustness lies in its performance on smaller lakes near the 10 km² threshold. These lakes comprise only 40 or more pixels and are inherently more vulnerable to misclassification caused by clouds, shadows, and mixed pixels.

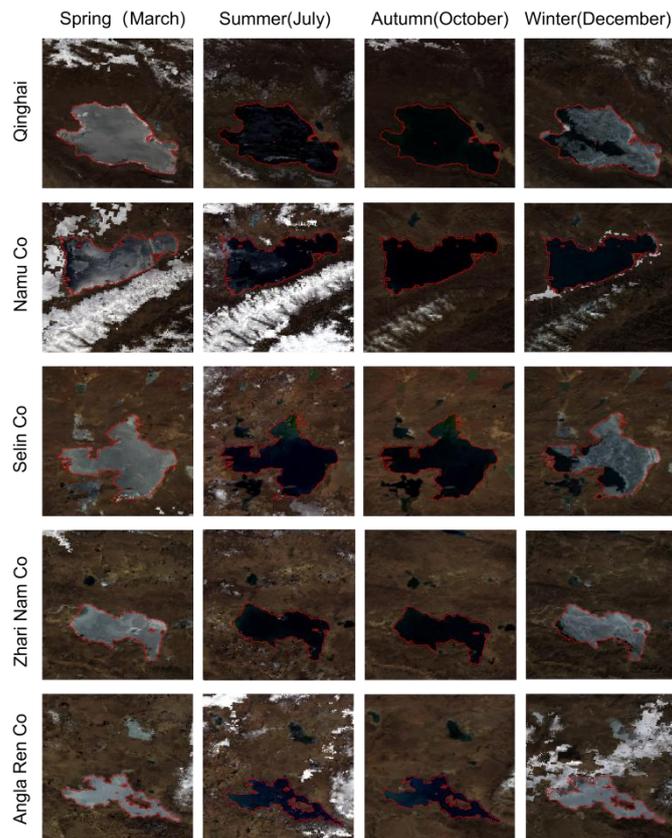
Response:

Thank you for pointing this out. We agree that displaying exclusively macroscopic lakes in the original Figure 8 inadvertently masked the classification challenges associated with smaller water bodies near the 10 km² threshold.

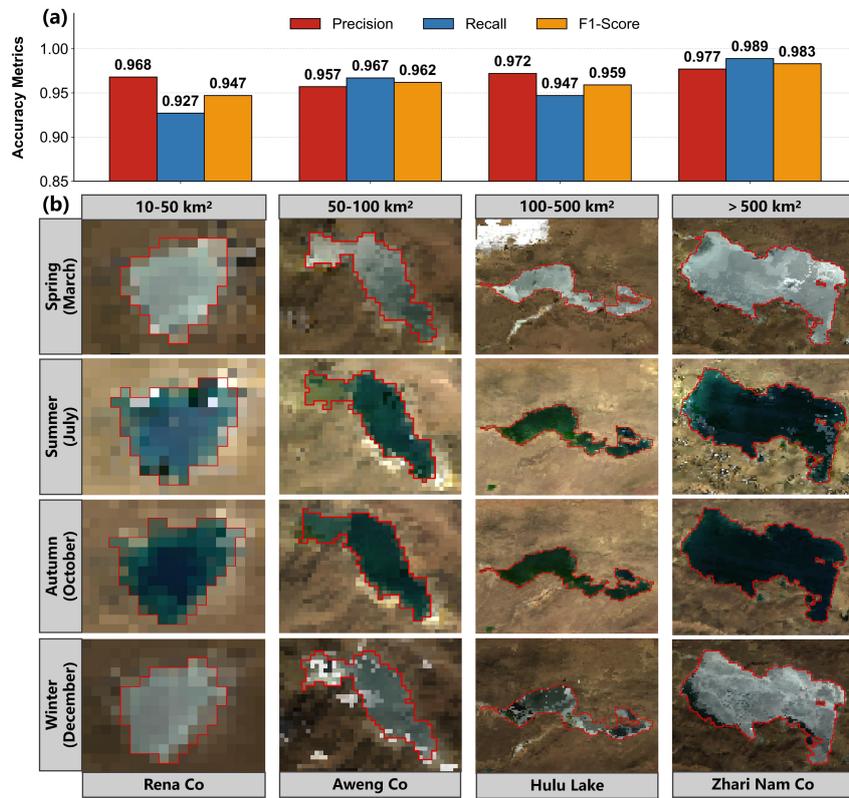
To directly address this, we have fundamentally redesigned Figure 8. We have discarded the approach of only showing the largest lakes. Instead, the revised figure now presents one representative lake from each of the four defined size intervals (including the 10–50 km² tier). This revision allows for a much more objective and transparent visual assessment of the method’s robustness across different spatial scales.

For the comprehensive quantitative validation methodology and the detailed spatial accuracy metrics stratified by these lake size groups, please refer to our detailed response to Major Comment 5. The updated figure and relevant descriptions have been integrated into **Section 4.1.1**. Both the original and the revised figures are presented below:

Original Figure 8:



Revised Figure 8:



Comments:

Line 365: “low solar elevation”? Would you mean “topographic shadows”?

Response:

Thank you for pointing this out. While a low solar elevation angle during winter is the underlying atmospheric cause, the direct physical challenge to water body classification is indeed the resulting “topographic shadows.” We have corrected this terminology to be physically precise and consistent with the remote sensing context.

The revised text in **Section 4.1.1 (Lines 330–331)** is as follows:

Original: “...under varying surface conditions, including snow cover, cloud shadow, and low solar elevation.”

Revised: “...under varying surface conditions, including snow cover, cloud shadow, and topographic shadows.”

Comments:

Line 371: What are seasonal climate changes? Maybe just seasonal changes. Also, Selin Co is larger than Nam Co now, so it should be a large lake, not a medium lake.

Response:

Thank you for pointing this out. The phrase “seasonal climate changes” incorrectly conflates long-term climate trends with short-term intra-annual variations. We have corrected this to “seasonal changes” for precise physical terminology. Describing Selin Co as a “medium lake” was a highly careless oversight. We are fully aware that due to its rapid and sustained expansion, Selin Co has long surpassed Nam Co. We have revised the text to accurately classify Selin Co as a major large lake, appropriately distinguishing it from actual medium lakes like Zhari Nam Co and Angla Ren Co.

Please note that during the revision process, as we updated the corresponding figures and restructured this section to provide a more comprehensive analysis, the specific paragraph containing these sentences was removed from the manuscript.

Comments:

Line 384: The comparison with other datasets using only 11 large lakes is not enough. As expected, the differences for these large lakes are relatively small, which does not provide a rigorous test of the method’s performance. It will be essential to compare some small lakes.

Response:

Thank you for this suggestion. We agree that that validating our results solely against large lakes is insufficient to assess performance limits at a 500 m spatial resolution, as this approach can mask potential boundary errors.

To address this, we expanded our validation dataset by randomly selecting 9 additional small lakes with areas ranging from 10 to 50 km². We directly compared the monthly area time series of these lakes from our dataset against the dataset provided by [Li et al. \(2025\)](#).

Figure 6 has been updated to include these small lakes, indicated by red triangles.

While the revised scatter plots across all four seasons show a slight expected increase in relative variance for smaller water bodies as a result of the amplified effect of mixed boundary pixels, the extracted area values for these small lakes maintain a remarkably high degree of quantitative agreement with the reference data. The overall statistical metrics across the expanded dataset remain excellent (e.g., $R^2 > 0.999$, $UMSE < 0.0034$), which conclusively demonstrates that our dataset preserves robust boundary delineation and sub-pixel accuracy even near the theoretical spatial resolution limits of the MODIS sensor.

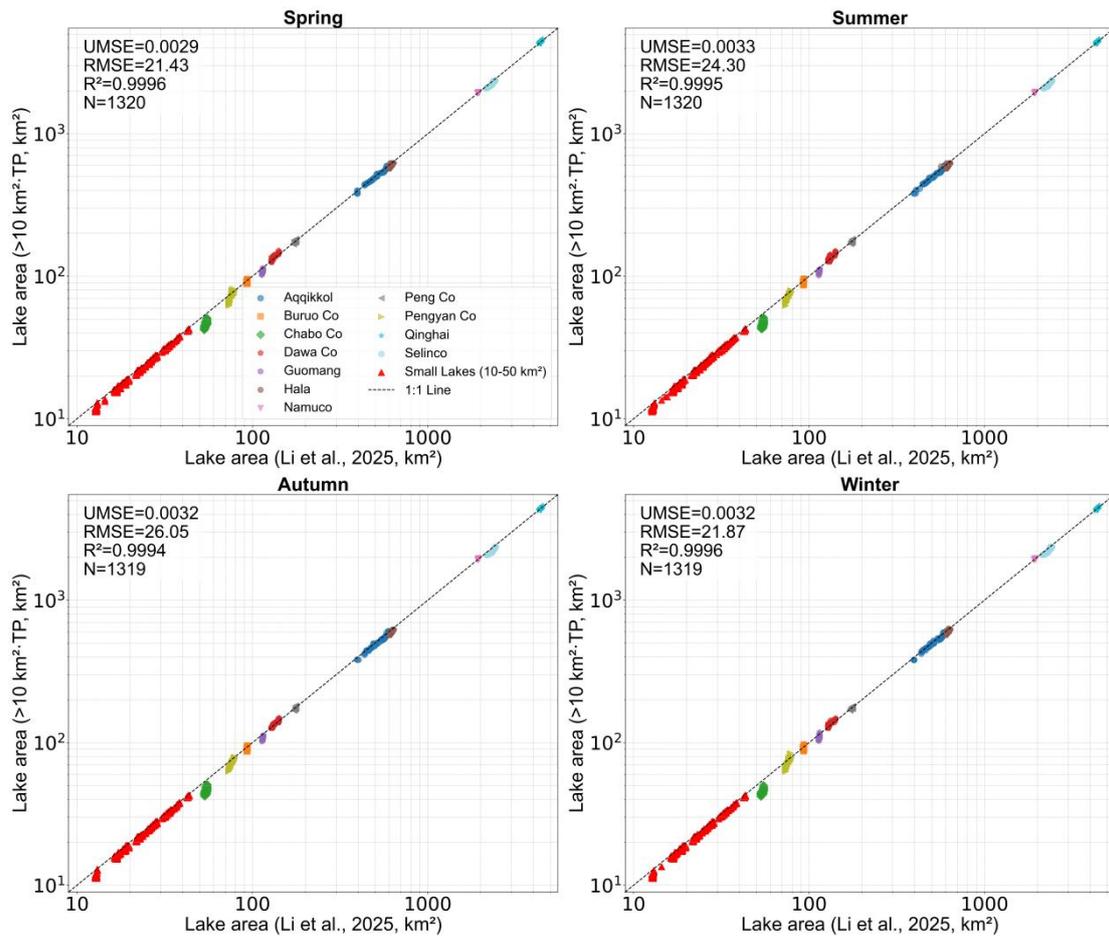


Figure 6: Scatter plot comparing different lake area extraction methods in spring, summer, autumn, and winter. (**Fig.6** in the revised manuscript)