Response to Comments by Anonymous Referee #2:

General comments:

Thanks for inviting me to review this manuscript. The authors developed a daily lake ice cover and lake ice phenology dataset for 32,800 lakes in global cold regions from 2002 to 2024 using MODIS data. The main contribution is the use of a cloud-gap-filling method to improve the temporal continuity and spatial coverage of MODIS lake ice observations. This study provides a new remote-sensing lake ice dataset covering a large number of lakes. However, as a data description paper, the current manuscript lacks sufficient validation and description of data accuracy, uncertainty, quality-control information, and applicability. I therefore recommend major revision.

Response:

We thank the reviewer for this detailed and constructive assessment. We agree that the original manuscript did not provide sufficiently comprehensive information on the evaluation, uncertainty, limitations, and applicability of LI-CCR. We have therefore substantially revised the manuscript to strengthen the supporting evidence and improve the transparency of the dataset.

First, we added a global analysis of the remaining unavailable observations after each processing step, stratified by continent, climate group, lake area, and freeze–thaw stage. Reconstruction consistency was then evaluated using a MODIS-internal Terra–Aqua paired experiment and an external cross-product comparison with the Sentinel-3-derived LIE-NH dataset, with the results further stratified by season, continent, climate group, lake area, and freeze–thaw stage.

Second, we expanded the LIP evaluation by adding an independent observational comparison with GLRIPD and a large-sample, same-source consistency assessment against the MODIS-derived GLIP dataset. These analyses include stratification by lake area, latitude, and climate group, together with a year-by-year consistency assessment for 2002–2019.

Third, we propagated the component-based uncertainty in daily LIC into the derived phenological dates and evaluated the sensitivity of LIP estimates to alternative LIC thresholds and persistence criteria. We also clarified the provenance and quality-control flags provided for the LIC and LIP records and revised the relevant terminology

Finally, we clarified that the initial ice–water classification was obtained by binary recoding of the MODIS Collection 6.1 NDSI_Snow_Cover layer rather than by recalculating NDSI or developing a new spectral classifier. Detailed responses and the corresponding manuscript revisions are provided below.

Major comment 1:

The statement in Lines 75–76 is inaccurate. Previous studies have not focused only on large lakes, and several recent remote-sensing studies have included large numbers of small and medium-sized lakes. The authors should include the relevant recent literature, revise this statement, and more clearly explain the additional contribution of this study compared with existing large-scale lake ice cover and lake ice phenology products.

Response:

We thank the reviewer for this important correction. We agree that the original wording was inaccurate and did not adequately acknowledge recent remote-sensing studies covering large numbers of small and medium-sized lakes. We have therefore expanded the discussion and citations of recent large-scale lake-ice studies in the Introduction. The revised text now identifies long revisit intervals, cloud contamination, and low illumination as the general constraints on spatially extensive daily optical monitoring of lake ice.

We have also clarified the specific contribution of LI-CCR relative to the existing large-scale lake ice products. LI-CCR applies a MODIS observability assessment to reduce shoreline mixed-pixel contamination and uses s staged spatiotemporal procedure to estimate observation gaps caused by cloud cover and polar night. It provides gap-filled daily lake ice coverage (LIC) for 32,800 lakes across both hemispheres from 2002 to 2024, together with annual ice-cover status, four ice phenology dates, ice cover duration (ICD), complete ice cover duration (CICD), probability of complete ice-cover occurrence (PCIO), component-based uncertainty estimates, and quality-control information. Thus, the principal contribution of LI-CCR lies in the combined provision of long-term daily records, broad lake coverage, and a comprehensive set of lake-ice indicators within a single, internally consistent dataset.

References cited in this response:

Tong, J., Lin, Y., Fan, C., Liu, K., Chen, T., Zeng, F., Zhan, P., Ke, L., Gao, Y., and Song, C.: Fine-scale monitoring of lake ice phenology by synthesizing remote sensed and climatologic features based on high-resolution satellite constellation and modeling, Science of the Total Environment, 912, 169002, https://doi.org/10.1016/j.scitotenv.2023.169002, 2024.

Liu, C., Huang, H., Hui, F., Zhang, Z., and Cheng, X.: Fine-Resolution Mapping of Pan-Arctic Lake Ice-Off Phenology Based on Dense Sentinel-2 Time Series Data, Remote Sensing, 13, 2742, https://doi.org/10.3390/rs13142742, 2021.

Major comment 2:

The validation of LIP mainly relies on a passive microwave product covering 56 large lakes and a model-based product covering 132 lakes on the Tibetan Plateau. This is not sufficient to validate a dataset covering 32,800 lakes. The authors should further compare their results with GLRIPD observations and other large-scale Landsat and MODIS lake ice products, with particular attention to the small and medium-sized lakes emphasized in this study. For example, Wang et al. (2021) (High-Resolution Mapping of Ice Cover Changes in Over 33,000 Lakes Across the North Temperate Zone) derived lake ice cover for more than 33,000 lakes using Landsat data, while Wang et al. (2022b) (Continuous loss of global lake ice across two centuries revealed by satellite observations and numerical modeling) investigated lake ice phenology for more than 30,000 lakes using MODIS data and numerical simulations. Other global or regional datasets covering small and medium-sized lakes may also be available and should be considered. Comparisons with other large-scale lake ice model products would also be more representative than relying only on a regional model product for the Tibetan Plateau. The validation should not be limited to overall statistical metrics. Spatial patterns, temporal variations, and differences among lake sizes, latitudes, climate regions, and lake types should also be presented.

Response:

We thank the reviewer for this important and constructive comment. We agree that the original LIP validation did not adequately represent the spatial extent and lake-size distribution of LI-CCR. We therefore retained the existing comparisons with the passive-microwave and regional model-based datasets, and added an independent comparison with in-situ-based GLRIPD dataset and a large-sample consistency assessment against the MODIS-derived GLIP dataset. Because only freeze-up end (FUE) and break-up end (BUE) were provided among these datasets, the new analyses focused on these two phenology dates. Sections 2.6, 3.4.3, and 4.1.3 have been revised accordingly.

We also considered the Landsat-based product of Wang et al. (2021) but it reports ice-cover occurrence rather than lake-wide phenology dates. A direct event-by-event comparison with LI-CCR was therefore not possible. We have nevertheless included this study in the revised Introduction to provide a more complete account of existing large-scale lake-ice products.

First, we added an independent comparison with GLRIPD observations, including 30 matched lakes for FUE and 42 for BUE. Consistency was stronger for BUE than for FUE (Sect.4.1.3, Figure 6). The weaker FUE consistency may partly reflect heterogeneous definitions of ice-on among in situ records, differences between local visual observations and the lake-wide LIC threshold used by LI-CCR, and greater spatial and short-term variability during freeze-up (Sharma et al., 2020).

Second, we compared LI-CCR with the MODIS-derived GLIP dataset (Wang et al., 2022b), covering more than 25,000 matched lakes. Because both products rely primarily on MODIS observations, we describe this comparison as a same-source, large-sample consistency assessment rather than an independent validation. The results showed broad consistency between the two methodologically distinct MODIS retrieval frameworks, particularly for BUE (Sect.4.1.3, Figure 6).

We further stratified the GLIP comparison by lake area, latitude, and Köppen climate group (Sect.4.1.3, Table 4). Small and medium-sized lakes did not show systematically weaker consistency than large lakes. Differences were more strongly associated with latitude and climate, with higher consistency at 45–75° N and in cold-climate regions. The year-by-year comparison for 2002–2019 showed stable consistency without a progressive increase in bias or deterioration in performance (Figure S1). We did not conduct a separate lake-type stratification because a globally consistent lake-level classification encompassing relevant attributes was not available for all matched lakes. We therefore used lake area, latitude and climate group to characterize broad spatial and environmental differences.

Together, these additions substantially improve the spatial, temporal, and lake-size representativeness of the LIP assessment. The corresponding changes are presented in Sects. 2.6, 3.4.3, and 4.1.3 and in Supplementary Fig. S1.

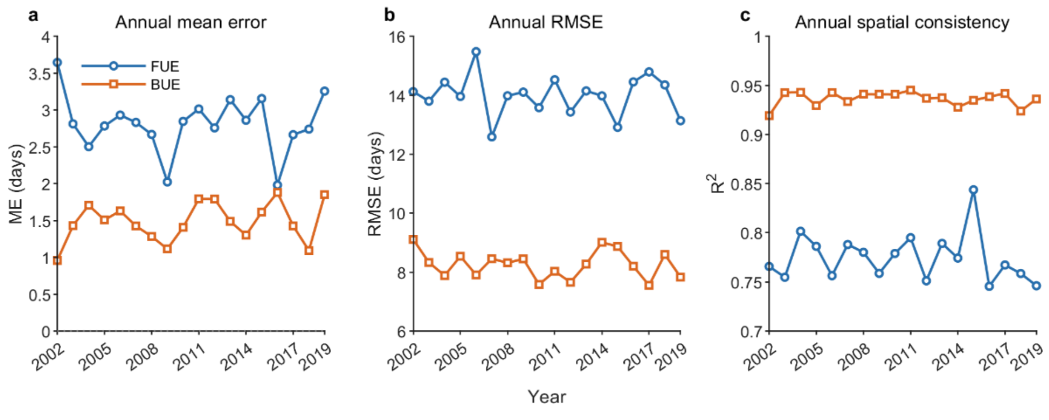

Figure S1: Temporal consistency between LI-CCR and GLIP lake ice phenology during 2002–2019.

References cited in this response:

Wang, X., Feng, L., Qi, W., Cai, X., Zheng, Y., Gibson, L., Tang, J., Song, X.-p., Liu, J., Zheng, C., and Bryan, B. A.: Continuous Loss of Global Lake Ice Across Two Centuries Revealed by Satellite Observations and Numerical Modeling, Geophysical Research Letters, 49, 10.1029/2022gl099022, 2022b.

Sharma, S., Meyer, M. F., Culpepper, J., Yang, X., Hampton, S., Berger, S. A., Brousil, M. R., Fradkin, S. C., Higgins, S. N., Jankowski, K. J., Kirillin, G., Smits, A. P., Whitaker, E. C., Yousef, F., and Zhang, S.: Integrating Perspectives to Understand Lake Ice Dynamics in a Changing World, Journal of Geophysical Research-Biogeosciences, 125, 10.1029/2020jg005799, 2020.

Wang, X., Feng, L., Qi, W., Cai, X., Zheng, Y., Gibson, L., Tang, J., Song, X.-p., Liu, J., Zheng, C., and Bryan, B. A.: Continuous Loss of Global Lake Ice Across Two Centuries Revealed by Satellite Observations and Numerical Modeling, Geophysical Research Letters, 49, 10.1029/2022gl099022, 2022b.

Major comment 3:

The NDSI threshold and the visible and near-infrared reflectance constraints described in Section 3.2.1 are highly similar to the standard snow and ice classification procedure used in the MOD10A1/MYD10A1 products. The authors must clarify how their classification differs from the existing MOD10A1/MYD10A1 classification and what additional improvement is achieved by reclassifying the MODIS observations. In addition, the Data section only introduces MOD10A1/MYD10A1 and does not identify the source of the original spectral bands required to recalculate NDSI and apply these thresholds. It is therefore unclear whether the procedure described in Section 3.2.1 was actually implemented in this study or merely restates the product algorithm.

Response:

We thank the reviewer for identifying this ambiguity. We agree that the original description incorrectly explained that we recalculated NDSI from the original MODIS reflectance bands and developed a new spectral classification procedure.

In fact, we used the Collection 6.1 NDSI_Snow_Cover layer from MOD10A1/MYD10A1, for which the NDSI calculation and associated spectral and quality screening had already been completed during product generation. We did not use the original band reflectance to recalculate NDSI. Within the HydroLAKES masks, we recoded NDSI_Snow_Cover values of 40-100 as lake ice and values of 0-39 as open water. This operation generated the binary ice–water input for the subsequent LI-CCR processing and did not constitute a new spectral classifier.

We have therefore revised Sects. 2.1 and 3.2.1 to describe the actual input data and binary recoding procedure. The methodological contribution of LI-CCR lies in the subsequent processing, including the MODIS lake-observability assessment, Terra–Aqua integration, staged spatiotemporal cloud-gap filling, treatment of low-illumination periods, noise correction, and generation of continuous daily lake-level LIC and phenology products.

Major comment 4:

The cloud-gap-filling assessment in Table 2 includes only several lakes in Asia and may not represent its performance across global regions and lake types. The authors should evaluate the filling proportion and reconstruction performance across representative regions, climate zones, lake sizes, and freeze–thaw stages. The contribution of each processing step should be readily available from the existing results, and such a stratified assessment is necessary to support the global applicability of the dataset.

Response:

We thank the reviewer for this important comment. We agree that the original Table 2, which included only several Asian lakes, was insufficient to characterize the performance across the global study domain. We therefore replaced the original assessment with a global analysis using daily Terra and Aqua observations throughout 2008 and substantially revised Sect. 4.1.1 and Table 2.

First, we assessed the improvement in data availability. For the availability assessment, we calculated the pixel-weighted proportion of unavailable observations in the original Terra and Aqua data and after each processing step. The results were stratified by continent, Köppen climate group, lake area, and freeze–thaw stage. Table 2 now reports the results after each step, thereby showing the contribution of the individual processing stages. Across the global study domain, the remaining unavailable-observation proportion decreased from approximately 60.0 % in the original observations to 20.8 % after completion of the workflow.

The improvement in data availability was not spatially or seasonally uniform. Larger lakes generally retained more valid spatial information for gap filling, whereas high-latitude lakes had more remaining data gaps during winter because of polar night and low illumination. Among freeze–thaw stages, the cumulative relative reduction ranged from 40.6 % during freeze-up to 71.1 % during stable ice-free conditions. After completion of the workflow, the remaining unavailable-observation proportion was 45.6 % during freeze-up and 31.2 % during stable ice-cover conditions, compared with less than 20 % during stable ice-free and break-up periods.

We also conducted an external cross-product comparison with the Sentinel-3-derived LIE-NH dataset, stratified by continent, climate group, lake area, and season, as described in our response to Major Comment 5 and in Sects. 3.4.2 and 4.1.2.

We did not conduct a separate lake-type stratification because a globally consistent lake-level classification encompassing attributes was not available for all lakes. We therefore used lake area, climate group, continent, and freeze–thaw stage to characterize major spatial and environmental differences.

Together, these complementary analyses distinguish improvements in data availability from reconstruction consistency and provide a substantially more representative assessment of the spatial, climatic, lake-size, and seasonal variability in the performance of the workflow.

Major comment 5:

Terra and Aqua use similar sensors, spectral bands, and classification methods and may therefore share common errors. Their comparison can be used to assess cross-sensor consistency and the internal performance of the algorithm, but it cannot independently validate the accuracy of cloud-gap filling. The authors should use different types of sensors or independent reference data for further validation. The sample information and spatiotemporal distribution of the current cloud-gap-filling validation are also insufficiently described.

Response:

We thank the reviewer for this important comment. We agree that the Terra–Aqua comparison cannot provide an independent validation of cloud-gap reconstruction. We have revised the manuscript to describe this experiment as a MODIS-internal paired evaluation.

We have also clarified the experimental design in Sect. 3.4.1. Rather than comparing Terra and Aqua classifications when both observations were clear-sky, we identified co-located lake pixels where Terra provided a valid clear-sky ice–water classification while the same-day Aqua observation was cloud-covered. The cloud-gap-filling workflow was then applied to the Aqua observation, and the reconstructed classification at each originally cloudy pixel was compared with the corresponding clear-sky Terra classification, which served as a pseudo-reference. This design evaluates reconstruction at actual cloud gaps and is analogous to pseudo-reference experiments used in previous MODIS snow-cover cloud-removal studies (Gafurov and Bárdossy, 2009; Li et al., 2019). The evaluation included all eligible lake pixels across the global study domain throughout 2008

To provide an external assessment based on a different sensor and processing chain, we compared the MODIS-derived classifications with the Sentinel-3-derived LIE-NH product. The comparison covered the Northern Hemisphere from July 2021 through June 2022, including more than 220 million valid same-day, spatially matched lake-pixel pairs under clear-sky conditions and approximately 31 million matches where the original MODIS observation was cloudy but LIE-NH provided a valid cloud-free classification. Overall consistency was 98.6 % for the clear-sky MODIS classifications and 98.1 % for the reconstructed classifications at originally cloudy MODIS pixels. The results were further stratified by season, continent, Köppen climate group, and lake area (in Sect. 4.1.2).

Sections 2.5, 3.4.1–3.4.2, and 4.1.1–4.1.2 have been revised accordingly to distinguish the MODIS-internal paired evaluation from the external Sentinel-3 comparison and to clarify their respective experimental designs, sample coverage, evidence levels, and limitations.

References cited in this response:

Li, X., Jing, Y., Shen, H., and Zhang, L.: The recent developments in cloud removal approaches of MODIS snow cover product, Hydrol. Earth Syst. Sci., 23, 2401–2416, https://doi.org/10.5194/hess-23-2401-2019, 2019.

Gafurov, A. and Bardossy, A.: Cloud removal methodology from MODIS snow cover product, Hydrology and Earth System Sciences, 13, 1361-1373, 10.5194/hess-13-1361-2009, 2009.

Major comment 6:

The discussion of algorithm performance and errors relies mainly on the Qinghai Lake case study. Qinghai Lake is a high-altitude plateau lake with distinctive environmental conditions and may not adequately represent lakes in other climate regions or lake types. A reduction in the number of cloudy pixels also does not necessarily indicate that the reconstructed values are accurate. The current analysis therefore does not constitute a complete error assessment.

Response:

We thank the reviewer for identifying limitations in the original assessment. We agree that the Qinghai Lake case study was not representative of the global study domain and that reductions in unavailable observations do not demonstrate the accuracy of the reconstructed classifications. We therefore removed the Qinghai Lake-based assessment and restructured the evaluation to algorithm performance of cloud removal and reconstruction consistency under cloudy conditions.

To evaluate the algorithm performance, we replaced the original illustrative cases with a global analysis using daily Terra and Aqua observations throughout 2008, as detailed in our response to Major Comment 4. The pixel-weighted proportion of remaining unavailable observations in the original data and after each processing step was quantified by continent, Köppen climate group, lake area, and freeze–thaw stage (Sect. 4.1.1 and Table 2). This analysis characterizes the contribution of each processing step and identifies spatial and seasonal differences in the capacity of the workflow to fill observational gaps.

Reconstruction consistency was evaluated separately using an external cross-product comparison with the Sentinel-3-derived LIE-NH dataset, as detailed in our response to Major Comment 5. Overall consistency with LIE-NH was 98.6 % under clear-sky conditions and 98.1 % for classifications reconstructed at originally cloudy MODIS pixels. Consistency remained high across all four seasons and was further stratified by lake area, continent, and climate group (Sect. 4.1.2). No systematic deterioration was observed for small lakes, while cloudy-condition consistency ranged from 96.2 % to 98.1 % among continents and from 94.4 % to 98.0 % among climate groups. We did not conduct a separate lake-type stratification because a globally consistent classification was not available for all evaluated lakes. We therefore used lake area, continent, and climate group to characterize broad environmental differences.

The revised algorithm performance and reconstruction consistency results are now presented systematically in Sects. 4.1.1 and 4.1.2 and Tables 2 and 3. These analyses provide a substantially broader assessment than the original Qinghai Lake case, showing that the performance and consistency varies among environmental settings and is generally lower during rapidly changing freeze-up and break-up periods.

Major comment 7:

The authors should evaluate how errors in daily LIC propagate into LIP dates and conduct sensitivity analyses to justify the use of the 10% and 90% LIC thresholds and the three-consecutive-day rule. The current accuracy assessment focuses mainly on cloud cover and cloud-gap filling, while other potential error sources, including thin ice, wet snow, turbid water, sun glint, terrain shadows, and shoreline mixed pixels, are insufficiently discussed.

Response:

We thank the reviewer for this important comment. We agree that the original manuscript did not adequately quantify the propagation of daily LIC uncertainty into the derived LIP dates, assess the sensitivity of LIP to the selected LIC thresholds and persistence criterion, or discuss errors inherited from the MODIS-based ice–water classification. We addressed these three issues separately.

First, we added a dedicated LIC-to-LIP uncertainty-propagation analysis in Sect. 3.5.2 and report the results in Sect. 4.2.2. For each lake, FUS, FUE, BUS, and BUE were independently re-extracted from the lower and upper bounds of daily LIC range. The earliest and latest resulting dates defined a time range for each phenological event. FUS and BUE were the least sensitive, with mean range widths of 0.76 and 0.32 d, respectively, and more than 95 % of their ranges were within 1 d. FUE and BUS were more sensitive, with mean widths of 2.00 and 1.63 d; nevertheless, 77.9 % and 75.6 % of their ranges were within 1 d. These range quantify the propagation of the component-based LIC uncertainty of our cloud removal method.

Second, we clarified the rationale for using the 10 % and 90 % LIC thresholds and the three-day persistence criterion. The two LIC thresholds are commonly used in satellite remote sensing to represent the onset and offset of lake ice, and using them makes our results comparable with others. Anyway, the thresholds reduce the sensitivity to small amounts of residual shoreline ice or open water and to shoreline mixed pixels (Reed et al., 2009; Xu et al., 2024). The persistence criterion reduces spurious threshold crossings caused by retrieval noise and short-term fluctuations in ice cover (Hallerbäck et al., 2022).

Following Kropáček et al. (2013), we further evaluated four threshold–persistence combinations: 10 %/90 % and 5 %/95 % thresholds, each combined with one- or three-day persistence. Because GLIP only provides FUE and BUE estimates, the comparative sensitivity assessment was limited to these two events. For FUE, the 10 %/90 %–3 d configuration produced the highest R^2 and lowest RMSE. For BUE, it shared the highest R^2 with the 5 %/95 %–3 d configuration and produced a similar RMSE, together with a smaller absolute ME. We therefore selected the 10 %/90 % thresholds with three-day persistence as the best overall balance across FUE and BUE and also to follow the common approach, rather than implying that this configuration was optimal for every individual metric. The analysis is summarized in Sect. 4.1.3, and the complete results are provided in Supplementary Table S1.

Table S1: Sensitivity of FUE and BUE estimates to LIC thersholds and persistence criteria, evaluated against the GLIP dataset.

Events	LIC thersholds	Persistence days	R^2	ME (d)	RMSE (d)
	10 %/90 %	1 d	0.77	0.4	13.7
FUE	**10 %/90 %**	**3 d**	**0.79**	**2.6**	**13.3**
	5 %/95 %	1 d	0.76	2.2	14.4

	5 %/95 %	3 d	0.76	4.5	15.2
	10 %/90 %	1 d	0.92	-0.5	9.0
BUE	**10 %/90 %**	**3 d**	**0.94**	**1.7**	**8.3**
	5 %/95 %	1 d	0.92	0.2	9.2
	5 %/95 %	3 d	0.94	2.2	8.1

Third, we expanded Sect. 5.1 to discuss classification uncertainties associated with thin ice and mixed ice–water surfaces, wet snow, turbid water and algal blooms, sun glint, terrain shadows, low solar illumination, and shoreline mixed pixels (Gupta et al., 2005; Kropáček et al., 2013; Warren, 2019; Riggs et al., 2019; Carrea et al., 2023). We also clarified that the lake-selection criteria, LIC filtering, temperature-based correction, and persistence criterion can mitigate, but cannot eliminate, these errors. In particular, spectrally ambiguous conditions cannot be fully resolved using MODIS optical observations alone. We therefore identify the integration of complementary thermal-infrared, radar, and microwave observations as an important direction for future improvement.

The corresponding methods, results, and limitations are now presented in Sects. 3.3, 3.5.2, 4.1.3, 4.2.2, and 5.1 and in Supplementary Table S1.

References cited in this response:

Xu, Y., Long, D., Li, X., Wang, Y., Zhao, F., and Cui, Y.: Unveiling lake ice phenology in Central Asia under climate change with MODIS data and a two-step classification approach, Remote Sensing of Environment, 301, 10.1016/j.rse.2023.113955, 2024.

Reed, B. C. et al. Integration of MODIS-derived metrics to assess interannual variability in snowpack, lake ice, and NDVI in southwest Alaska. Remote Sensing of Environment 113, 1443–1452 (2009). https://doi.org/10.1016/j.rse.2008.07.020

Kropáček, J. et al. Analysis of ice phenology of lakes on the Tibetan Plateau from MODIS data. The Cryosphere 7, 287–301 (2013). https://doi.org/10.5194/tc-7-287-2013

Hallerbäck, S. et al. Climate warming shortens ice durations and alters freeze and break-up patterns in Swedish water bodies. The Cryosphere, 16, 2493–2503 (2022). https://doi.org/10.5194/tc-16-2493-2022

Riggs, G. A., Hall, D. K., and Román, M. O.: MODIS Snow Products Collection 6.1 User Guide, Version 1.0, NASA, 2019.

Warren, S. G.: Optical properties of ice and snow, Philos. Trans. R. Soc. A, 377, 20180161, https://doi.org/10.1098/rsta.2018.0161, 2019.

Gupta, R. P., Haritashya, U. K., and Singh, P.: Mapping dry/wet snow cover in the Indian Himalayas using IRS multispectral imagery, *Remote Sens. Environ.*, **97**, 458–469, https://doi.org/10.1016/j.rse.2005.05.010, 2005.

Carrea, L., Cretaux, J.-F., Liu, X., Wu, Y., Calmettes, B., Duguay, C. R., Merchant, C. J., Selmes, N., Simis, S. G. H., Warren, M., Yesou, H., Mueller, D., Jiang, D., Embury, O., Berge-Nguyen, M., and Albergel, C.: Satellite-derived multivariate world-wide lake physical variable timeseries for climate studies, Scientific Data, 10, 10.1038/s41597-022-01889-z, 2023.

Korver, M. C., Lehner, B., Cardille, J. A., and Carrea, L.: Surface water temperature observations and ice phenology estimations for 1.4 million lakes globally, Remote Sensing of Environment, 308, 10.1016/j.rse.2024.114164, 2024.

Kropáček, J. et al. Analysis of ice phenology of lakes on the Tibetan Plateau from MODIS data. The Cryosphere 7, 287–301 (2013). https://doi.org/10.5194/tc-7-287-2013

Major comment 8:

Stratified validation should be conducted across different climate regions, latitudes, lake sizes and types, and freeze–thaw stages. Accuracy differences among seasons and years should also be evaluated to demonstrate the spatial and temporal consistency and reliability of the dataset.

Response:

We thank the reviewer for this comprehensive suggestion. We agree that spatially and temporally stratified evaluations are necessary to characterize both the reliability and limitations of a global dataset. Because this comment overlaps with Major Comments 2, 4-6, we summarize the complementary assessments below and provide the full experimental details in the corresponding responses.

First, we conducted a global assessment of data availability using daily Terra and Aqua observations throughout 2008. The proportion of remaining unavailable observations in the original data and after each processing step was stratified by continent, Köppen climate group, lake area, and freeze–thaw stage (Sect. 4.1.1 and Table 2). This analysis revealed substantial spatial and seasonal differences in the capacity of the workflow to fill observational gaps.

Second, we expanded the external cross-product comparison of MODIS-derived LIC classifications with the Sentinel-3-derived LIE-NH product. The overall consistency was 98.6 % for the clear-sky conditions and 98.1 % for the classifications from reconstructed cloudy MODIS pixels. The consistency remained high across all four seasons and showed no systematic deterioration for small lakes within the evaluated sample. Regional and climatic differences were more apparent for the reconstructed classifications, with the lowest consistency in polar regions. The comparison was stratified by season, continent, Köppen climate group, and lake area (Sects. 3.4.2 and 4.1.2 and Table 3).

Third, reconstruction consistency under cloudy conditions across freeze–thaw stages was assessed using the MODIS-internal Terra–Aqua paired evaluation described in Sect. 3.4.1. For Step 2, which contributed most to gap filling, conditional reconstruction consistency was approximately 96 % during the relatively stable ice-free and ice-cover stages but decreased to approximately 80 % during freeze-up and break-up. We therefore identified rapidly changing transition periods as having greater reconstruction uncertainty.

Fourth, we expanded the LIP evaluation by adding an independent comparison with GLRIPD observations and a large-sample, same-source consistency assessment against the MODIS-derived GLIP dataset (Sect. 4.1.3). The GLIP comparison included more than 25,000 matched lakes and was stratified by lake area, latitude, and Köppen climate group. Small and medium-sized lakes did not show systematically weaker consistency than large lakes. Differences were more strongly associated with latitude and climate, with higher consistency at 45–75° N and in cold-climate regions. The year-by-year comparison showed no progressive increase in bias or deterioration in consistency between LI-CCR and GLIP during 2002–2019 (Figure S1).

A separate lake-type stratification was not conducted because globally consistent lake-type information was unavailable for all evaluated lakes. We could only use lake area, latitude, geographic region, and climate group to characterize broad environmental differences among lakes.

Together, these analyses assessthe seasonal agreement, interannual consistency, and variation associated with lake area, latitude, climate, geographic region, and freeze–thaw stage. The revised manuscript also clearly distinguishes the external Sentinel-3 comparison, the MODIS-internal paired evaluation, the independent GLRIPD comparison, and the same-source GLIP consistency assessment. The corresponding revisions are presented in Sects. 3.4.1–3.4.3 and 4.1.1–4.1.3, Tables 2–4, and Supplementary Fig. S1.

Minor comment 1:

Section 3.5.1 is more appropriately described as an uncertainty estimate based on different data sources and processing steps, while Section 3.5.2 mainly provides data-source information and quality-control flags for LIP. Neither section constitutes a strict quality assessment. More accurate terminology should be used.

Response:

We thank the reviewer for this important clarification. We agree that the analyses originally presented in Sections 3.5.1 and 3.5.2 did not constitute a strict quality assessment. We have therefore renamed Sect. 3.5 "Uncertainty estimation and quality-control information." Section 3.5.1 now describes the component-based uncertainty estimation for daily LIC associated with different input sources and processing steps. Section 3.5.2 describes the data-source information and quality-control flags for LIP, together with the newly added propagation of component-based LIC uncertainty into LIP date. The corresponding terminology has also been revised in Sect. 4.2 and throughout the manuscript

Minor comment 2:

Line 330 states that the cross-sensor validation used MODIS data from 2008, whereas the caption of Figure 4 states 2015. Please verify and correct this inconsistency.

Response:

We thank the reviewer for identifying this inconsistency. After verification, we confirm that the cross-sensor validation and accuracy assessment of observations from 2008. We have therefore corrected the year in the Fig. 4 caption from 2015 to 2008 to ensure consistency with the main text.

Minor comment 3:

In Line 535, the lake-area threshold should be written as "2 km2" rather than "2 km." The manuscript also contains several errors in abbreviations and wording and should be carefully checked throughout. For example, Section 3.4.3 lists the four lake ice phenology events as "FUS, FUS, BUS and BUE," where the second FUS should be FUE. The use of PCIO and PICO should also be made consistent.

Response:

We thank the reviewer for identifying these inconsistencies. We corrected the lake-area threshold from "2 km" to "2 km^2" and revised the four lake ice phenology events in Section 3.4.3 to "FUS, FUE, BUS, and BUE." We also standardized the abbreviation throughout the manuscript as "PCIO" (probability of complete ice-cover occurrence), replacing the incorrect use of "PICO." In addition, we carefully rechecked the entire manuscript for similar inconsistencies in units, abbreviations, terminology, grammar, and wording and corrected the additional issues identified.

We sincerely thank the reviewer for the careful and constructive comments, which have helped us improve the clarity and consistency of the manuscript.