This document was prepared by Fangfang Yao (Fangfang.yao@colorado.edu; yaoff.luke@gmail.com)

Overview: The global database of lake water storage (GLWS) v1.1 dataset contains time-varying lake areas, water levels, and storages for 1,972 Earth's large water bodies spanning 1992 to 2020 at a near monthly frequency. These water bodies represent 95% of the total global lake volume. The method is described in detail in Yao et al., Science (2023), but a quick overview is provided here.

1. We used a total of 248,649 satellite imagery, including both good and partially contaminated imagery, to map time-varying water areas following Yao et al. (2019).
2. We depended on elevation measurements from nine satellite altimeters to estimate water levels.
3. We combined near-monthly water areas with water levels to estimate lake volume changes over the period 1992-2020.
4. We accounted for sedimentation-induced storage losses in reservoirs filled prior to our study period (1992-2020) using sedimentation surveys and statistical methods.
5. We validated our lake water storage estimates against 26,052 monthly storage anomalies from in-situ gauging stations.
6. We used an ensemble of datasets and models to determine the drivers for natural lakes experiencing water losses or gains.

This dataset includes:

1. Global database of lake water storage GLWS time series v1.1: time-varying lake volume variables including area, level, and storage.
2. Global database of lake water storage GLWS shapefile v1.1: geographical locations of studied lakes, lake water storage trends and dominant drivers, and other information about lakes.
3. GLWS lake coordinates v1.1.csv: coordinates of studied lakes in a cvs format.
4. Validation_monthlystorageanamolies_against_insitu.csv: validation on estimated monthly storage anomalies against in-situ measurements.
5. Water masks: water masks for deriving water levels from ICESat and ICESat-2 satellites.

Citation: "Yao, F., Livneh, B., Rajagopalan, B., Wang, J., Crétaux, J.-F., Wada, Y., & Berge-Nguyen, M. (2023). Satellites reveal widespread decline in global lake water storage. *Science.* https://doi.org/10.1126/science.abo2812."

Global database of lake water storage GLWS time series v1.1

"Poly" folder: volume estimates using well-calibrated hypsometry (polynomial fitting) based on long-term water levels and areas

"ConstantArea" folder: volume estimates using time-varying water levels and a constant area

"LinearLi" folder: volume estimates using a hypsometry (linear fitting) from global reservoir bathymetry data (Li et al., 2020)

"SimpleLinear" folder: volume estimates using a simple linear hypsometry based on maximum and minimum values of levels and areas

Table 1. Attribute table of lake time series data (*.csv)

Column	*Description*
Time	Format: YYYYDDD, a combination of year plus a relative day number within the year Note that this should be interpreted as monthly data.
Mean_Levels	Mean water levels for that month [Time], unit: m Sources: "Poly" and "ConstantArea" folders: Hydroweb (Crétaux et al., 2011), DAHITI (Schwatke et al., 2015), and G-REALM (Birkett et al., 2011). "LinearLi" and "SimpleLinear": this study (Yao et al., 2023)
Mean_Level_Errors	Errors of the mean water levels, unit: m
Cleaned_Area	Lake areas mapped from Landsat satellite imagery using the method Yao et al. (2019), unit: km2
rws	Water storage anomaly relative to the first record, unit: Gigaton (Gt)
rws_err	Errors of water storage anomalies, unit: Gt
rws_wSediAdj	rws adjusted for sedimentation-induced loss, unit: Gt
rws_err_wSediAdj	Errors propagated from [rws_err] and errors of sedimentation-induced loss.
Hypsometry	Hypsometric function of the relationship between water areas and levels. "x" denotes water level (m) Area unit: [AreaunitinHypsometry_and_RMSE]

RMSE	Root-mean-square deviation of the hypsometric model unit: [AreaunitinHypsometry_and_RMSE]
R2	R2 of the hypsometric model
R2adj	Adjusted R2 of the hypsometric model
AreaunitinHypsometry_and_RMSE	Area unit for the hypsometry and its RMSE

Global database of lake water storage GLWS shapefile v1.1

Table 2. Attribute table of lake shapefile data

Column	*Description*
LakeID	Unique lake identifier
LakeName	Name of lake or reservoir
LakeArea	Area of lake or reservoir in km2
TypeName	Lake type as natural lake (non-regulated) or reservoir (regulated)
Dryland	Whether the lake is in arid regions 1: yes 0: no
Trend	Lake water storage trend rate based on the Mann-Kendall method, unit: Gigaton (Gt) per year This field was used to report lake water storage trends
TrendPval	p-value of the trend shown in the field “Trend”
TrendErr	Error (uncertainty) of the trend shown in the field “Trend”, unit: Gt per year
Notrend	Whether a lake has no significant trend ($p > 0.1$) in water storage 1: yes 0: no
Domidriver	Dominant driver of the lake volume variability
IsRunoff	Whether the dominant driver is runoff

	1: yes 0: no
Modelagree	Model agreement on the dominant driver
ModelR2	Mean R2 of the attribution models on lake volume variability
PrepSeason	Dominant precipitation season Apr: summer precipitation from April 1 to September 30th Oct: winter precipitation from October 1 to March 31
SediRate	Sedimentation rate applied for accounting for sediment-induced storage loss, unit: % For reservoirs with sedimentation surveys, this rate was calculated using survey data; For the remaining reservoirs, the mean rates from bootstrapping were applied. We note that sedimentation in each reservoir is only an approximation, which could be further improved by imposing additional constraints from local data.
SediLoss	Sediment-induced storage loss, unit Gt per year This field equals to the difference between [Trend] and [TrendOri]
SediLosErr	Error (uncertainty) of estimated sediment-induced storage loss, unit: Gt per year
NReservoir	Whether this lake is a newly filled reservoir during the studied period (1992 – 2020) 1: yes 0: no
TrendOri	Original lake water storage trend rate in Gt per year without sedimentation adjustment It has the same value as that in the field “Trend” for natural lakes and newly filled reservoirs as we only applied sedimentation adjustment in existing reservoirs.
TrendOriPv	p-value of the trend shown in the field “TrendOri”
Hydropower	Whether this lake is used for hydropower generation 1: yes 0: no

InsituVali	Whether this lake is included for validating estimated monthly storage anomalies 1: yes 0: no
ReguNatura	Whether this lake is a regulated natural lake 1: yes 0: no

Note a value of -99 indicates no data or not applicable.

GLWS lake coordinates v1.1.csv

Geographic locations of lakes in GLWS database. This is supplementary to Global database of lake water storage GLWS shapefile v1.1 in case readers and potential users do not have access to a Geographic Information System (GIS) software.

Validation_monthlystorageanamolies_against_insitu.csv

Validation data on 102 reservoirs: including 72 reservoirs from the United States, 18 from Australia, and 12 from Spain. A total of 26,052 monthly storage anomalies from in-situ gauging stations were used in the validation.

Water masks

Water masks delineated from Global Surface Water dataset (Pekel et al., 2016) were used to derive water levels from ICESat and ICESat-2.

Replicating our key analyses and figures

We encourage readers and potential users to visit our published code capsule at https://codeocean.com/capsule/0322198/tree/v1.

References

Birkett, C., Reynolds, C., Beckley, B., & Doorn, B. (2011). From Research to Operations: The USDA Global Reservoir and Lake Monitor. In Coastal Altimetry (pp. 19–50). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12796-0_2

Crétaux, J. F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M. C., Nino, F., Abarca Del Rio, R., Cazenave, A., & Maisongrande, P.

(2011). SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. Advances in Space Research, 47(9), 1497–1507. https://doi.org/10.1016/j.asr.2011.01.004

Li, Y., Gao, H., Zhao, G., & Tseng, K. H. (2020). A high-resolution bathymetry dataset for global reservoirs using multi-source satellite imagery and altimetry. *Remote Sensing of Environment*, 244, 111831. https://doi.org/10.1016/j.rse.2020.

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

Schwatke, C., Dettmering, D., Bosch, W., & Seitz, F. (2015). DAHITI – an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrology and Earth System Sciences*, 19(10), 4345–4364. https://doi.org/10.5194/hess-19-4345-2015.

Yao, F., Wang, J., Wang, C., & Crétaux, J.-F. (2019). Constructing long-term high-frequency time series of global lake and reservoir areas using Landsat imagery. *Remote Sensing of Environment*, 232, 111210. https://doi.org/10.1016/j.rse.2019.111210

Yao, F., Livneh, B., Rajagopalan, B., Wang, J., Crétaux, J.-F., Wada, Y., & Berge-Nguyen, M. (2023). Satellites reveal widespread decline in global lake water storage. *Science.* https://doi.org/10.1126/science.abo2812.