Supplement of Earth Syst. Sci. Data, 15, 2781–2808, 2023
https://doi.org/10.5194/essd-15-2781-2023-supplement

Supplement of

Res-CN (Reservoir dataset in China): hydrometeorological time series and landscape attributes across 3254 Chinese reservoirs

Youjiang Shen et al.

Correspondence to: Youjiang Shen (yjshen2022@rainbow.iis.u-tokyo.ac.jp)

Contents of this file

- Figs. S1-3 shows the methodologies of generate reservoir water level, storage anomaly and evaporation, respectively.
- Fig. S4 describes the distribution of reservoir water surface area values and Top 20 reservoirs based on area size in our data product.
- Figure S5 shows the cross validation of the areas of delineated catchments in this study.
- Figure S6 represents the performances of altimetry-derived water level from each satellite altimeters with different retracking algorithms.
- Fig. S7 illustrates the comparison between our water level time series and other existing similar databases.
- Fig. S8 shows the uncertainties for each value in the time series of reservoir water level at some selected reservoirs.
- Figure S9 shows the performance of the Standard-rate (SR) and High-rate (HR) products in terms of the RMSE and the relative RMSE values of the validated reservoirs. For detailed validation metrics, please refer to our data.
- Figure S10 draws the comparison of reservoir water surface area time series against in situ, altimetric water levels, and GRSAD and ReaLSAT area time series for a sample of reservoirs of varying areas.
- Figure S11 shows an example that illustrates how the uncertainties in satellite datasets propagate to storage anomalies.
- Figure S12 reports the validation the evaporation values.
- Figure S13 shows the long-term mean meteorological variables that were used to calculate the evaporation rates.
- Figure S14 describes the spatial distribution of the ratios of reservoir water surface area and storage to catchment area.

- Table S1 describes the providers of water level, water surface area, storage anomaly, and evaporation time series for Chinese reservoirs.
- Tables S2-8 describes the source datasets for generating reservoir water level, water surface area, storage anomaly, upstream catchment, catchment-level characteristics products as well as the buffer distance for calculating reservoir area series in this study.
- Tables S9-16 describes the attributes provided in our datasets.

- Text S1 details the procedures of correcting the errors of delineated catchments and removed unrealistic or incorrectly catchments.

All validation reports and datasets, please find them at Zenodo link: https://doi.org/10.5281/zenodo.7664489

Figures

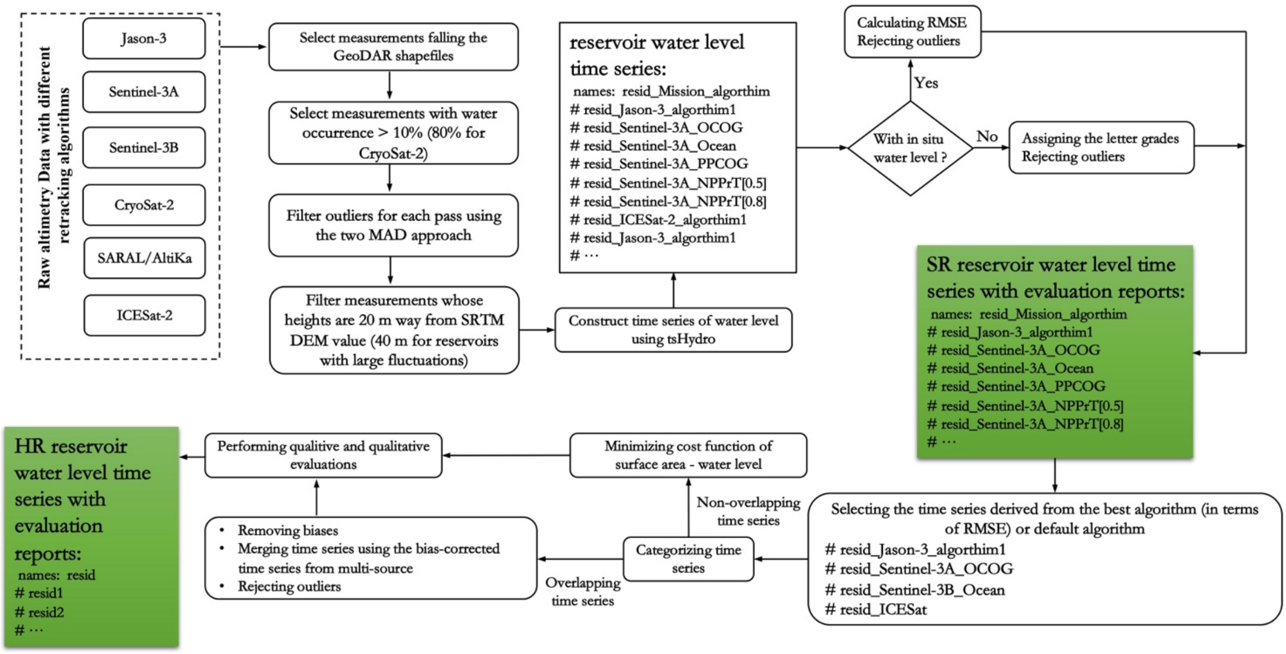


Figure S1. Flowchart of bias correction for obtaining SR and HR altimetric water level time series over reservoirs (Shen et al., 2022b).

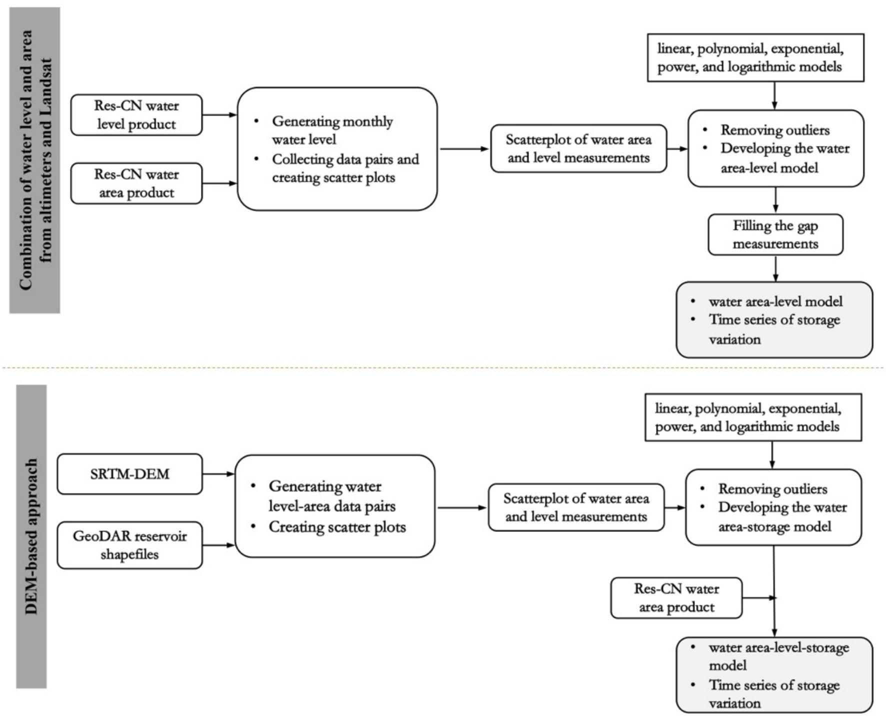


Figure S2. Flowchart of obtaining reservoir storage anomaly: 1) using water levels from satellite altimetry and water surface areas from satellite imagery (top panel); and 2) using imagery-based water areas and SRTM-DEM (bottom panel). (Shen et al., 2022b).

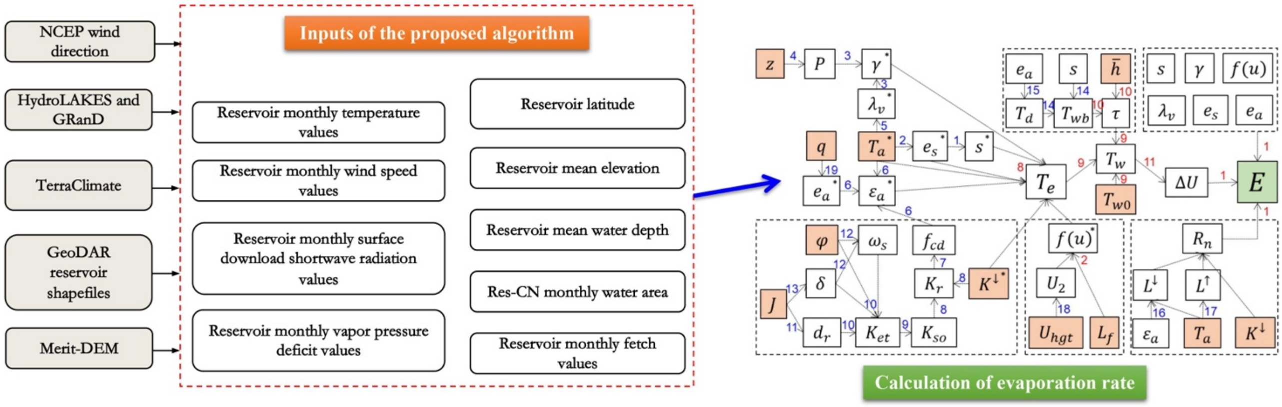


Figure S3. Flowchart of the proposed algorithm for generating time series of reservoir evaporation. For more details about the algorithm, evaporation calculation example, please find the https://ars.els-cdn.com/content/image/1-s2.0-S0034425719301063-mmc1.pdf.

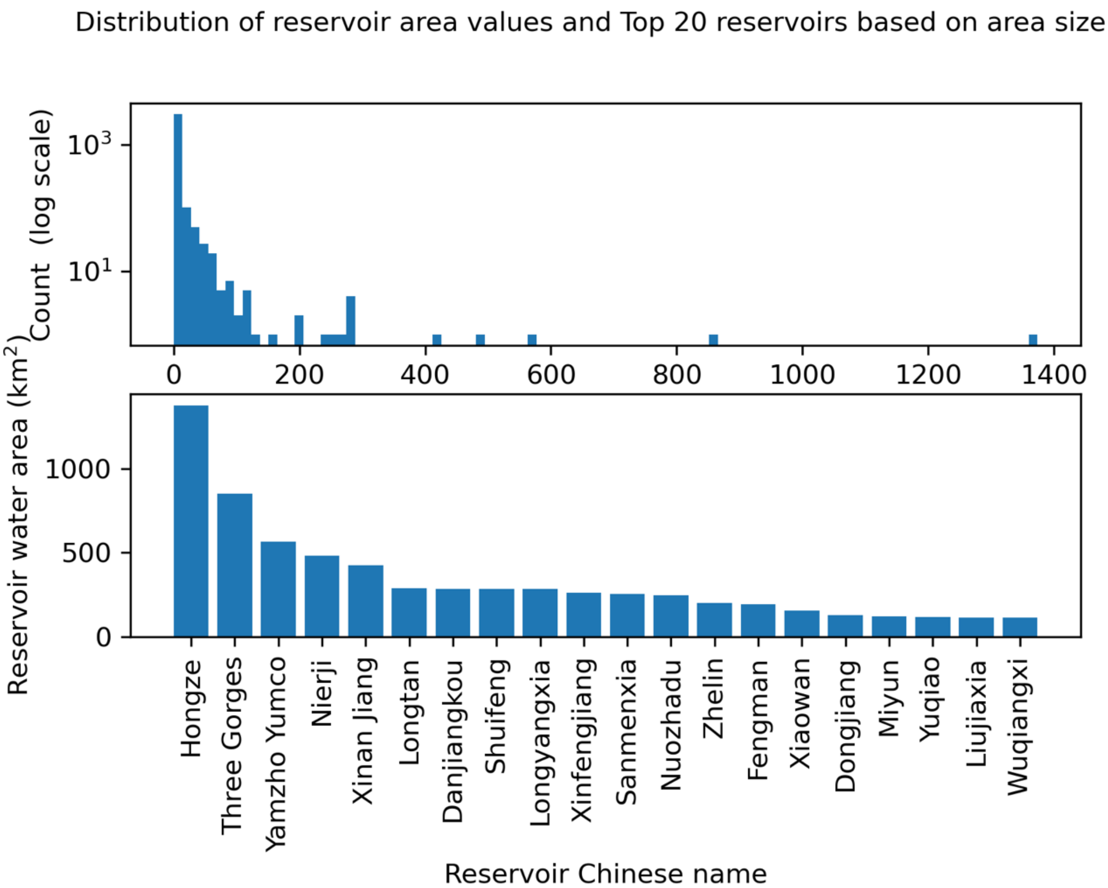


Figure S4. Distribution of reservoir water surface area values and Top 20 reservoirs based on area size in our data product. For more information such as area, name, and ID of all reservoirs, please refer to our data product.

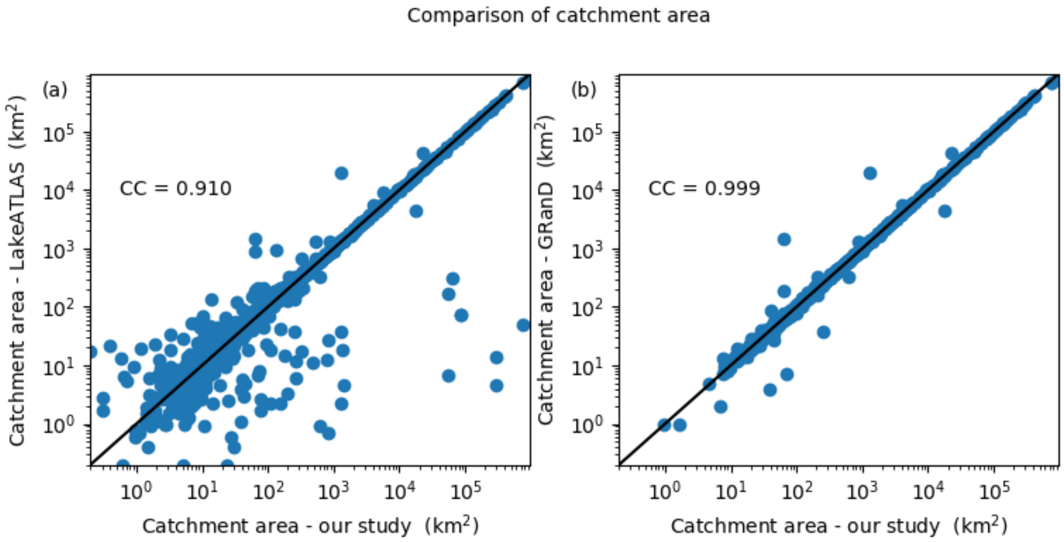


Figure S5. Comparison of the areas of delineated catchments in this study with those of LakeATLAS (Lehner et al., 2022), and those of GRanD reported value (Lehner et al., 2011).

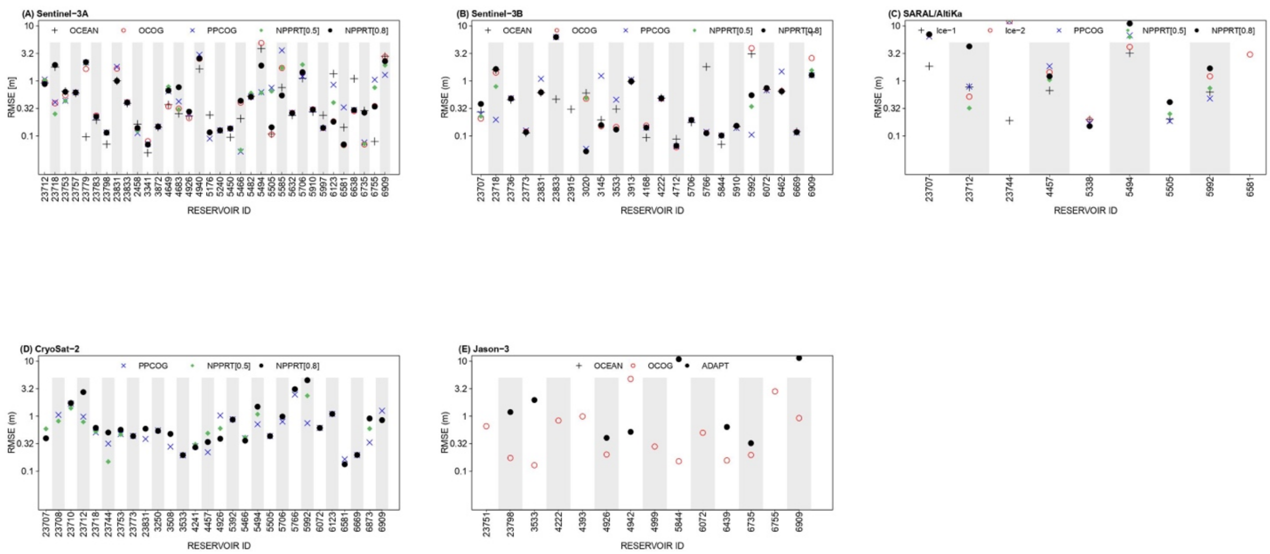


Figure S6. Performances of altimetry-derived water level from each satellite altimeters with different retracking algorithms.

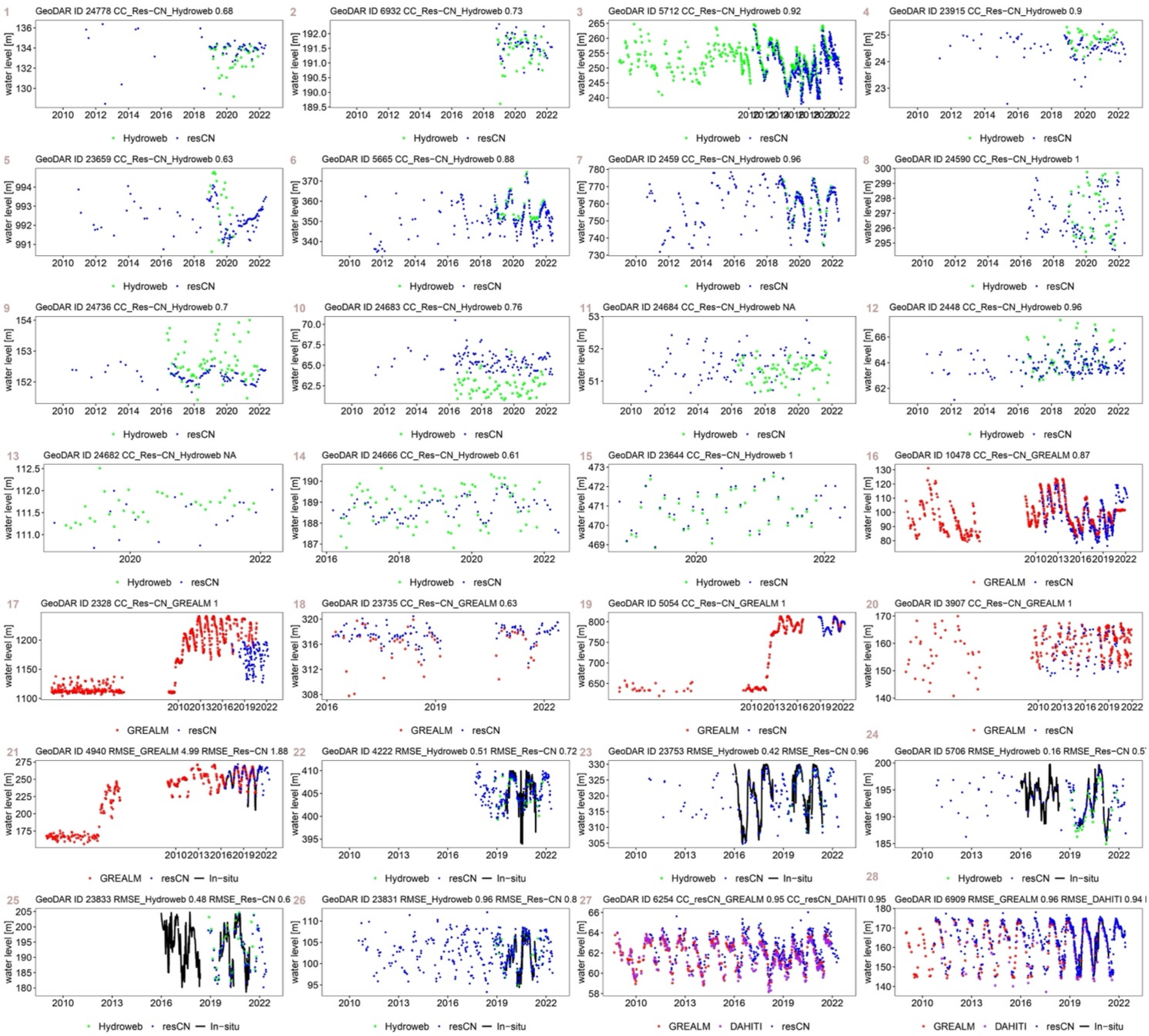


Figure S7. Comparison between our water level time series and other existing similar databases.

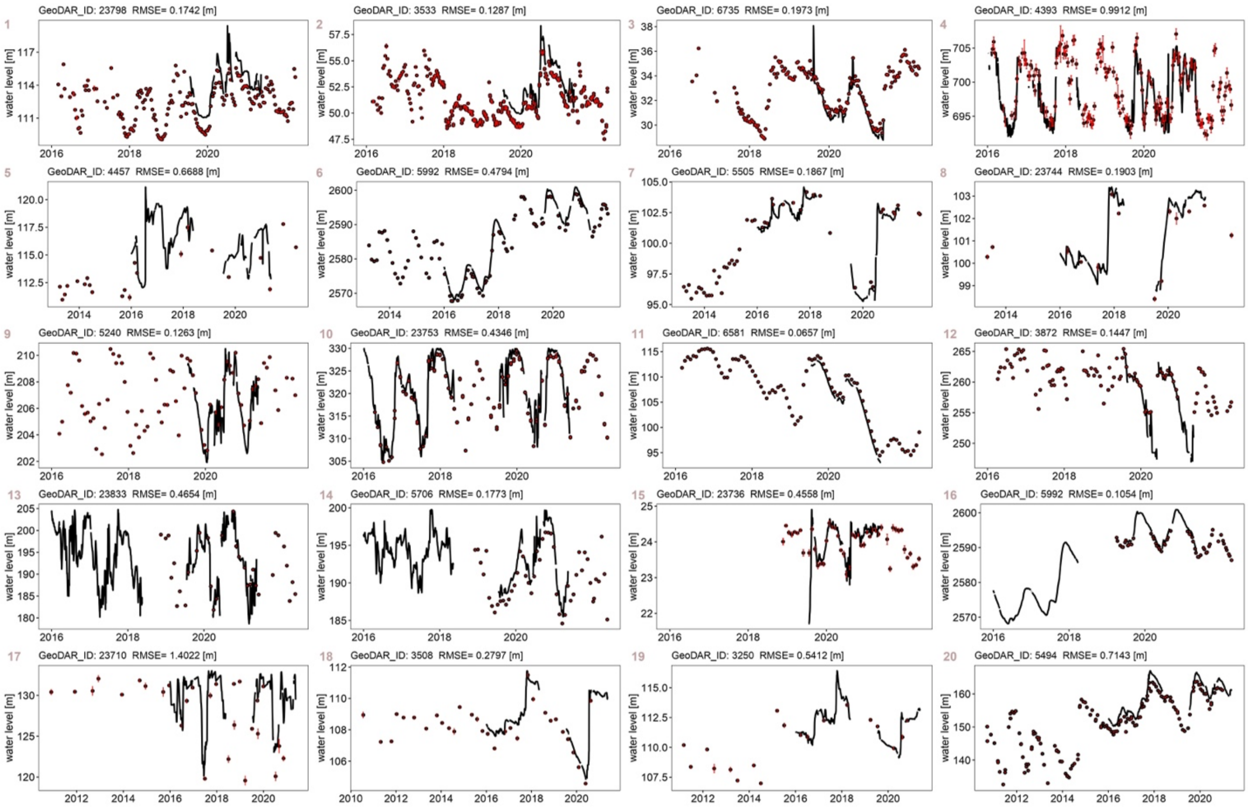


Figure S8. Uncertainties for each value in the time series of reservoir water level. In the figure, black line refers to the observed water level, dot is the altimetric water level, error bar quantifies the uncertainty of each value. Taking our 20 reservoirs in the standard rate product as an example, 1-4 are taken from Jason-3 mission, 5-8 are from SARAL/AltiKa mission, 9-12 are from Sentinel-3A mission, 13-16 are from Sentinel-3B mission, 17-20 are from CryoSat-2 mission. All uncertainties values are available in our product.

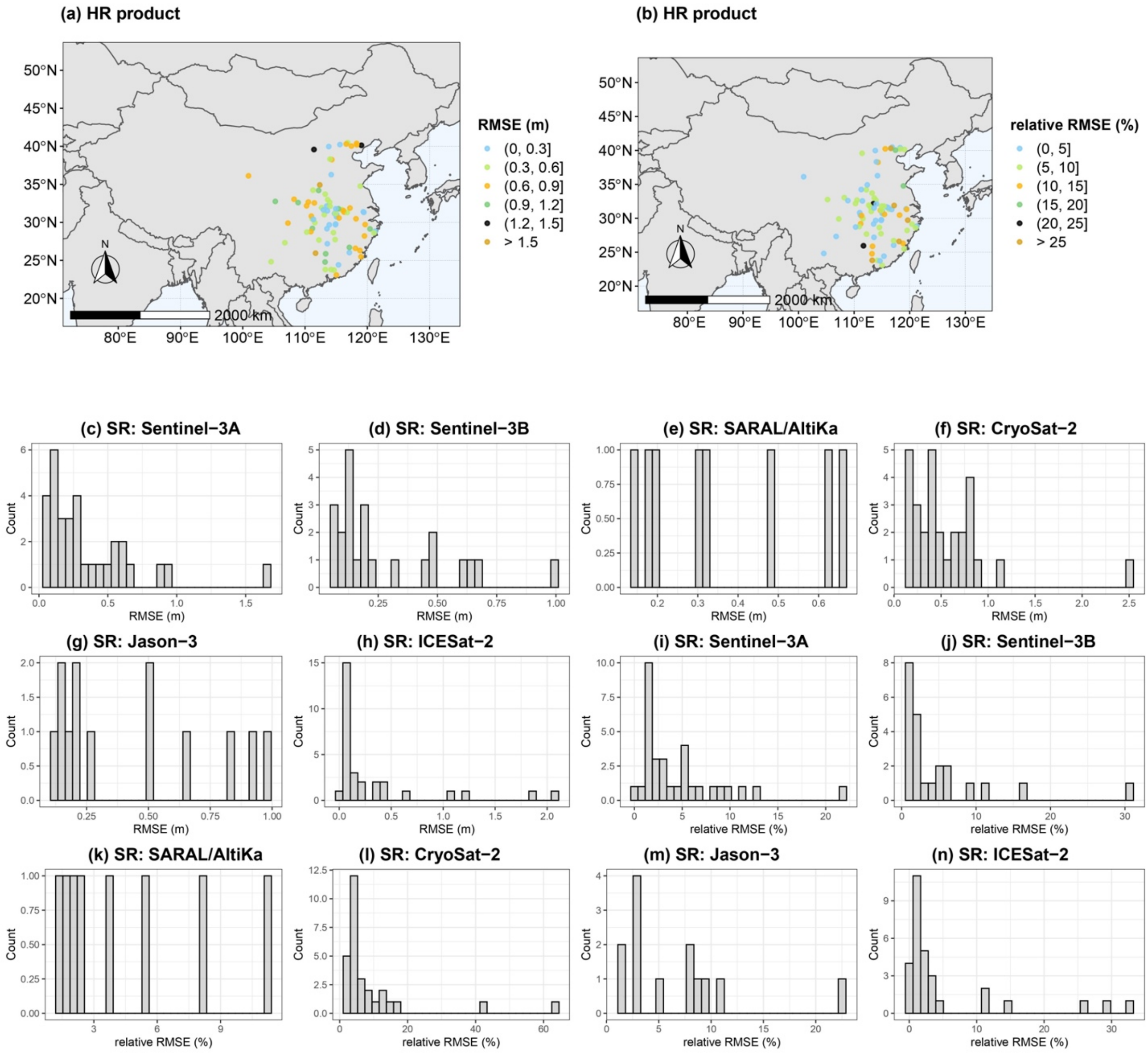


Figure S9. Performance of the Standard-rate (SR) and High-rate (HR) products in terms of the RMSE and the relative RMSE values of the validated reservoirs. For detailed validation metrics, please refer to our data.

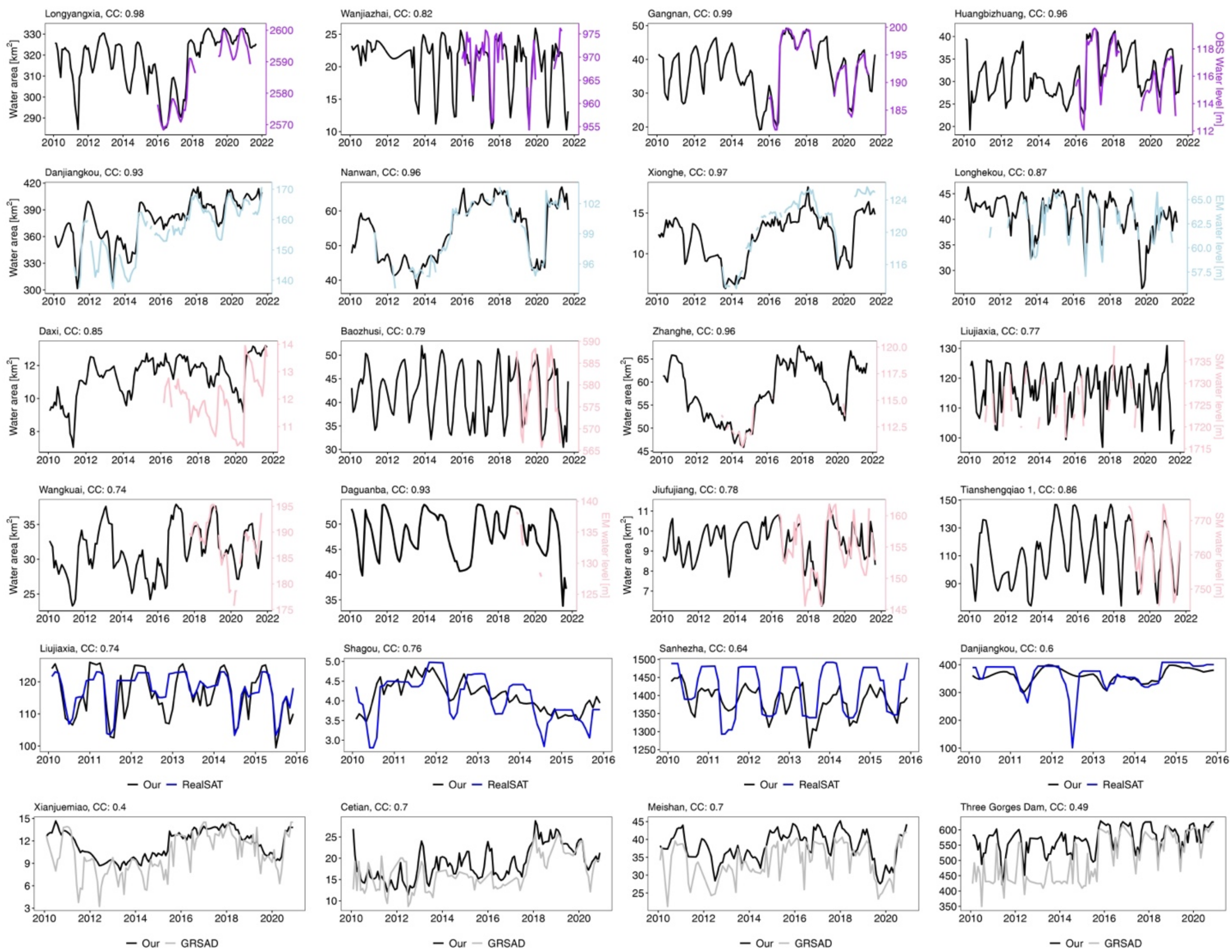


Figure S10. Graphs showing reservoir water area time series against in situ water levels, altimetric water levels from high and standard rates, and GRSAD and ReaLSAT area time series for a sample of reservoirs of varying areas (Shen et al., 2022b).

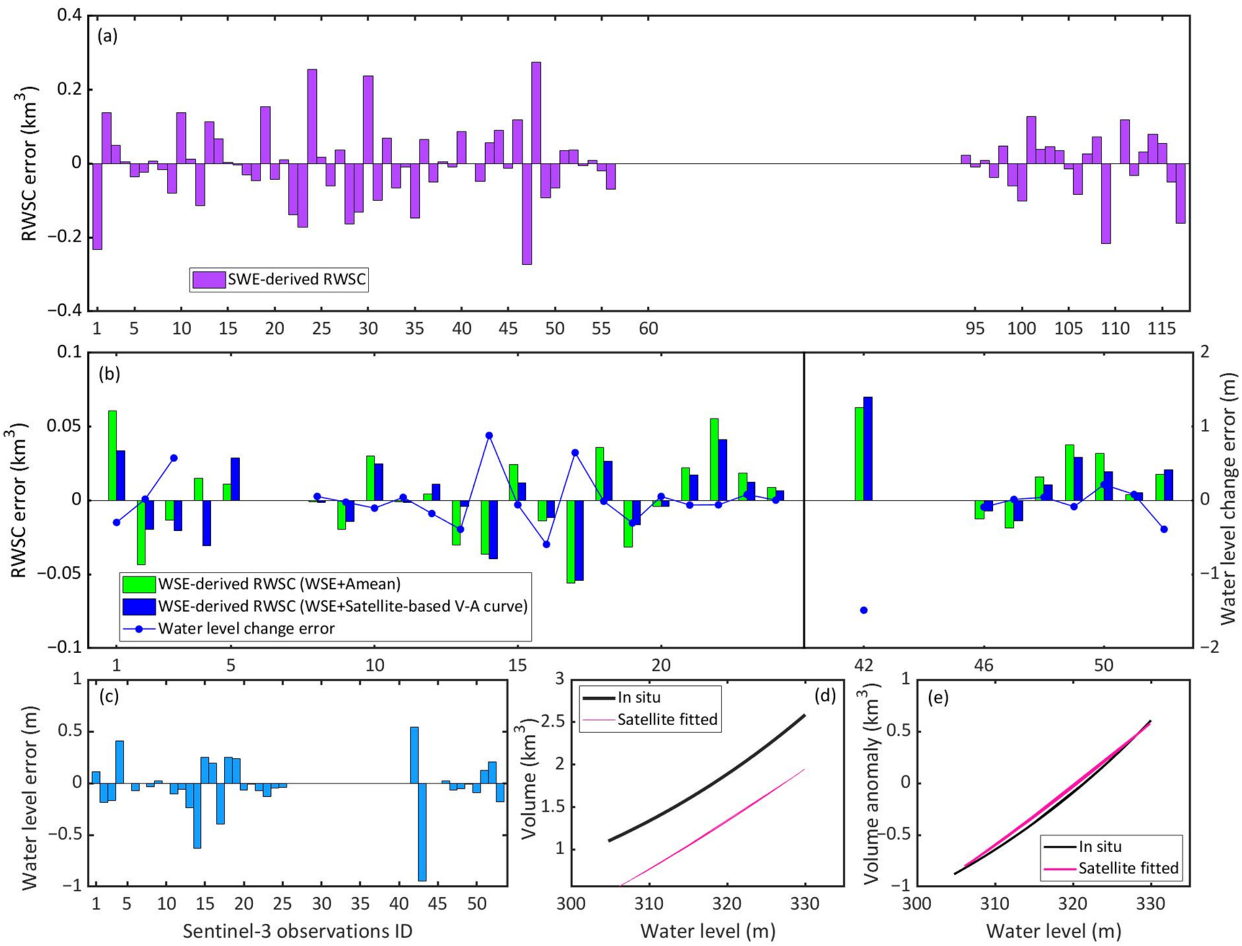


Figure S11. Graphs showing an example that illustrates how the uncertainties in satellite datasets propagate to storage anomalies. Error series and relationships of reservoir elevation-storage. Error series of (a) SWE-derived RWSC (i.e., storage anomaly), (b) WSE-derived RWSC and water level change, (c) WSE (i.e., water level). (d) and (e) Relationships of elevation-storage. The numbers on the x-axis (a, b, c) refer to the IDs of SWE, WSE, and WSE change observations, respectively. For more details about the propagation process, please find the reference Shen et al., (2020): https://doi.org/10.3390/rs14040815.

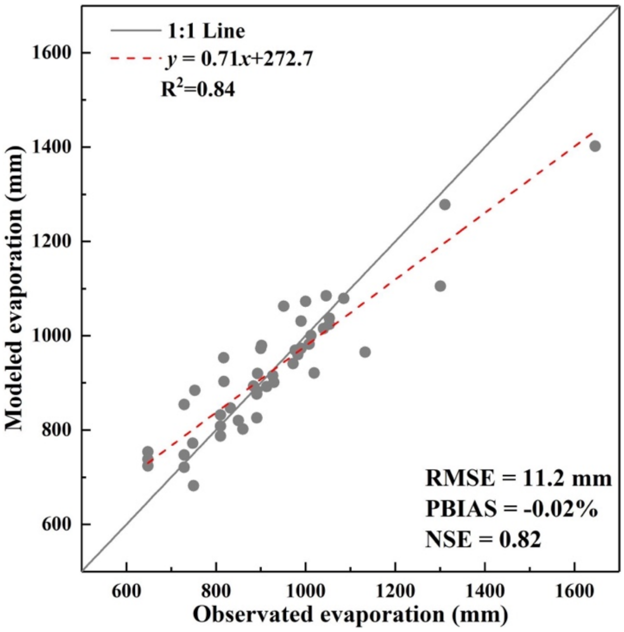


Figure S12. Observed and modeled average annual evaporation for 47 reservoirs (Tian et al., 2021).

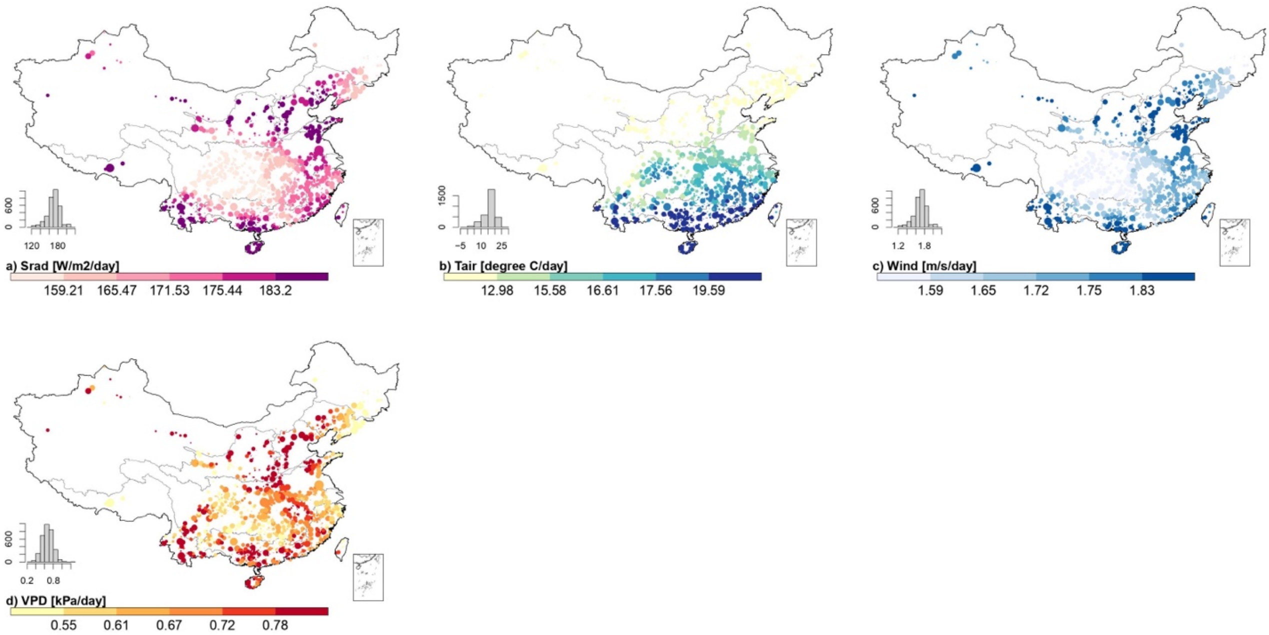


Figure S13. Long-term mean meteorological variables that were used to calculate the evaporation rates.

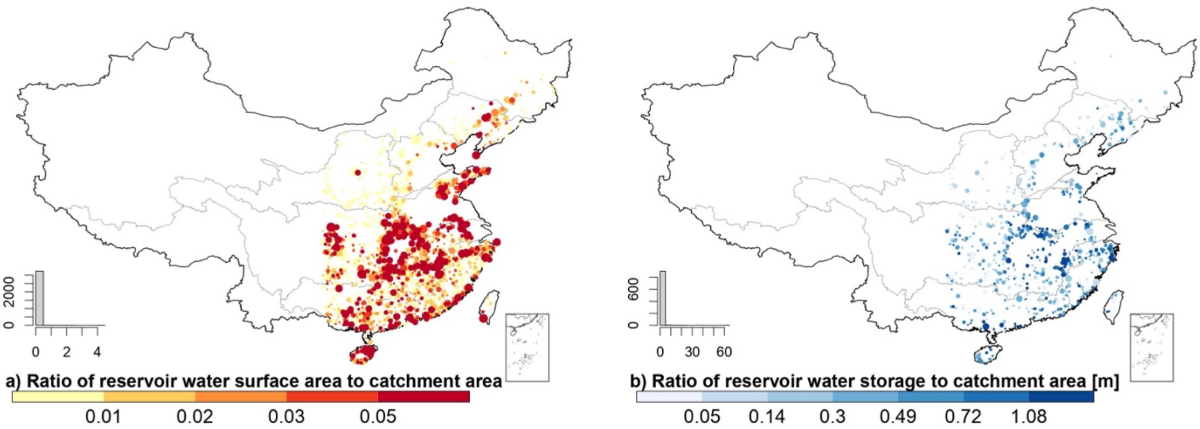


Figure S14. Spatial distribution of the ratios of reservoir water surface area and storage to catchment area. Note: not all reservoir water storage data are available from the GeoDAR database (Wang et al., 2022).

Tables

Table S1. Providers of water level, water surface area, storage anomaly, and evaporation for Chinese reservoirs.

Water levels:

Data sources	No. of reservoirs	Time and temporal resolution	Download link
Hydroweb	32	1992–2023, 10–35 day	http://hydroweb.theia-land.fr/
DAHITI	8	2002–2023, 10–35 day	https://dahiti.dgfi.tum.de/en/
G-REALM	~30	1992–2023, 10–35 day	https://ipad.fas.usda.gov/cropexplorer/global_reservoir
Tortini et al. (2020)	<10	1992–2018, sub-monthly	https://doi.org/10.5067/UCLRS-GREV2
Shen et al. (2022b)	338	2010-2021, monthly	https://doi.org/10.5281/zenodo.7251283

*** Last access: 15 October 2022.**

Water surface areas:

Data sources	Number of reservoirs	Time and temporal resolution	Download link
Bluedot observatory	Not clear	2016–2021, sub-monthly	https://blue-dot-observatory.com/*
GRASD	923	1984-2018, monthly	https://doi.org/10.18738/T8/DF80WG* (Zhao and Gao, 2018; Gao and Zhao, 2019)
Khandelwal and Kumar (2019)	<10	1992–2018, sub-monthly	https://doi.org/10.5067/UCLRS-AREV2*
ReaLSAT	85522 (lakes and reservoirs)	1984-2015, monthly	https://doi.org/10.5281/zenodo.6468209* (Khandelwal et al., 2022)
Donchyts et al. (2022)	9418	1985-2021, monthly	https://doi.org/10.6084/m9.figshare.20359860.v1*
Yao et al. (2019)	~8	1992–2018, sub-monthly	https://lakewatch.users.earthengine.app/view/glats*
Shen et al. (2022b)	338	2010-2021, monthly	https://doi.o https://doi.org/10.5281/zenodo.7251283*

Storage anomalies:

Data sources	Number of reservoirs	Time and temporal resolution	Download link
Vu et al. (2022)	10	2008-2020, monthly	https://doi.org/10.5281/zenodo.6299041*
Tortini et al. (2019)	<10	1992–2018, sub-monthly	https://doi.org/10.5067/UCLRS-STOV2*
Hou et al. (2022)	923	1984-2015, monthly	Not publicly accessible
Shen et al. (2021)	337	2010-2021, monthly	https://doi.org/10.5281/zenodo.7251283*
Zhao et al. 2022	>4,000	1984-2020, monthly	https://doi.org/10.5281/zenodo.4646621*

Evaporation:

Data sources	Number of reservoirs	Time and temporal resolution	Download link
Zhao et al. 2022	>4,000	1984-2020, monthly	https://doi.org/10.5281/zenodo.4646621*
Tian et al., 2022	908	1984-2016, monthly	https://doi.org/10.5281/zenodo.6042127*

*** Last access: 24 Feb 2023.**

Table S2. Source datasets for generating reservoir water level product in this study.

Satellite	Data period	Retracking algorithms	Repeat cycle
CryoSat-2	2010-2022.5	PPCOG, NPPTr[0.5], NPPTr[0.8]	369 days
SARAL/AltiKa	2016-2022.5	PPCOG, NPPTr[0.5], NPPTr[0.8], ICE-1, ICE-2	35 days
Sentinel-3A	2016.2-2022.5	PPCOG, NPPTr[0.5], NPPTr[0.8], OCOG, Ocean	27 days
Sentinel-3B	2018.4-2022.5	PPCOG, NPPTr[0.5], NPPTr[0.8], OCOG, Ocean	27 days
Jason-3	2016-2022.5	Adapt, OCOG, Ocean	10 days
ICESat-2	2018.10-2022.5	Official	91 days
In situ	2015-2021.5	- (99 reservoirs with observational water level records)	1 day
Global Surface Water Explorer	1984-2020	- (water occurrence product version 1.3)	-
SRTM-DEM	2000	-	-

Source download links:

- SARAL/AltiKa Geophysical Data Records (GDRs) from CNES AVISO+ at ftp://avisoftp.cnes.fr/AVISO/pub/
- CryoSat-2 baseline C level 1b dataset from ESA at https://science-pds.cryosat.esa.int/
- Sentinel-3 level 2 "Enhanced measurements" datasets from Copernicus Open Access Hub at https://scihub.copernicus.eu/dhus/.
- Jason-3 from ftp://avisoftp.cnes.fr/AVISO/pub/
- ICESat-2 ATL13 product from https://icesat-2.gsfc.nasa.gov/science/data-products
- In situ water level for 99 reservoirs from the local governments at http://113.57.190.228:8001/web/Report/BigMSKReport and National Hydrological Information Centre at http://xxfb.mwr.cn/index.html
- Global Surface Water Explorer from the https://global-surface-water.appspot.com/.
- SRTM-DEM from the https://srtm.csi.cgiar.org/srtmdata/.

Last access: 15 October 2022

Table S3. Source datasets for generating reservoir water surface area product in this study.

Product	Data period	Remark	Download link
GeoDAR	-	Reservoir shapefiles	https://doi.org/10.5281/zenodo.6163413
JRC-GWSD	1984-2020	Two products are used: monthly history and water occurrence products	https://global-surface-water.appspot.com/

Last access: 15 October 2022

Table S4. Source datasets for generating reservoir storage anomaly product in this study.

Product	Data period	Remark	Download link
Res-CN water surface area	1984-2021	Generated in in this study	https://doi.org/10.5281/zenodo.7664489
SRTM-DEM	2000	-	https://srtm.csi.cgiar.org/srtmdata/
In situ	2015-2021.5	139 reservoirs with daily water level and storage data	http://113.57.190.228:8001/web/Report/BigMSKReport http://xxfb.mwr.cn/index.html

Last access: 15 October 2022

Table S5. Source datasets for generating reservoir evaporation product in this study.

Satellite	Data period	Remark	Download link
GeoDAR	-	Reservoir shapefiles	https://doi.org/10.5281/zenodo.6163413
TerraClimate	1984-2020	Four variables: air temperature, wind speed, vapor pressure deficit, and surface downward shortwave radiation	https://climate.northwestknowledge.net/TERRACLIMATE/
Res-CN water surface area	1984-2021	Generated in in this study	https://global-surface-water.appspot.com/
NCEP	1984-2020	wind direction data from National Centers for Environmental Prediction	https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html
Merit-DEM	2019	To derive mean elevation	http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/
HydroLAKES	-	To derive mean water depth	https://www.hydrosheds.org/products/hydrolakes
GRanD	-	To derive mean water depth	https://globaldamwatch.org/grand/

Last access: 15 October 2022

Table S6. Source datasets for generating reservoir upstream catchment product in this study.

Satellite	Data period	Remark	Download link
GeoDAR	-	Reservoirs and dams shapefiles	https://doi.org/10.5281/zenodo.6163413
Merit-Hydro	2019	Flow directions	http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_Hydro/

Last access: 15 October 2022

Table S7. Source datasets for generating catchment-level characteristics product in this study.

Category	Source data	Remark	Download link
Topography	Merit-dem and Merit-Hydro	Yamazaki et al. (2017, 2019)	http://hydro.iis.u-tokyo.ac.jp/~yamadai/*
	GeoDAR	Reservoir shapefiles	https://doi.org/10.5281/zenodo.6163413*
	Our study	Catchment shapefiles	Provided in the main text.
Climate	National Station-based Climatic Data set V3	800 gauges offering 10 variables	closed
	Global Aridity Index and Global Reference Evapo-Transpiration	Zomer, R. J. (2022)	https://cgiarcsi.community/2019/01/24/global- aridity-index-and-potential-evapotranspiration-climate-database-v3/*
Land cover	NPP, GPP, NDVI, LAI, EVI	Myneni et al., 2015; Didan, 2021; Running et al., 202a, b;	Available at GEE (original datasets and references are in the main texts)
	IGBP, Zeng, X. (2001)	Rooting depth	https://lpdaac.usgs.gov/products/mcd12q1v006/*
	ESA WorldCover 2020, 10m	10 Land cover types	https://worldcover2020.esa.int/*
Soil	SoilGrids250m	Hengl et al. (2017)	https://files.isric.org/soilgrids/former/2017-03-10/data/*
	Dai et al. (2019)	Soil hydraulic and thermal properties	http://globalchange.bnu.edu.cn/research/soil5.jsp*
	Shangguan et al. (2013)	Soil property data	http://globalchange.bnu.edu.cn/research/soil2*
Geology	GLHYMPS (Global HYdrogeology MaPS)	Gleeson et al. (2014)	https://borealisdata.ca/dataset.xhtml?persistentId=doi:10.5683/SP2/DLGXYO*
	GliM (Global Lithological Map)	Hartmann and Moosdorf, 2012	https://doi.pangaea.de/10.1594/PANGAEA.788537*
Anthropogenic activity characteristics	Gridded Population of the World (GPW)	Population amount in 2000, 2005, 2010, 2015, and 2020	https://sedac.ciesin.columbia.edu/*
	Global Roads Inventory Project (GRIP) dataset	Road density	https://www.globio.info/download-grip-dataset*
	Global Human Footprint v2	Human Footprint	https://sedac.ciesin.columbia.edu/data/set/wildareas-v2-human-footprint-geographic/data-download*
	DMSP-OLS Nighttime Lights v4 dataset	Doll, 2008	https://eogdata.mines.edu/products/dmsp/*

*Last access: 15 October 2022.

Table S8. Buffering distance for GeoDAR reservoir datasets.

Buffered distance (m)	Number of reservoirs	Reservoir water surface area (km^2)
150	650	A<=0.1
250	1515	0.1<=A<=1
500	795	1<=A<=10
750	273	10<=A<=100
1000	21	100<=A<=1000

Table S9. Attributes of reservoirs provided in the Res-CN.

Attribute	Unit	Description and values
id_v11	-	Dam ID of GeoDAR version 1.1
id_grd_v13	-	GRanD ID of this dam if also included in GRanD v1.3
lat	degree	Latitude of the dam point
lon	degree	Longitude of the dam point
rv_mcm_v11	millions of m^3	Reservoir storage capacity taken from Wada et al. (2017) and GRanD v1.3
plg_a_km2	km^2	Area of the retrieved reservoir polygon
length	km	Perimeter of the reservoir
elev	m	Mean reservoir elevation from Merit-DEM (Yamazaki et al., 2019)
region	-	Reservoir-based River basins
reservoir_development_index	-	Ratio: perimeter of the reservoir / perimeter of the circle whose area is that of the reservoir

Note: missing or inapplicable values are filled by "-999".

Table S10. Attributes of topography provided in the Res-CN.

Attribute	Unit	Description	Data source and reference
length	m	The length of the main stream measured from the basin outlet to the remotest point on the basin boundary. The main stream is identified by starting from the basin outlet and moving up the catchment.	Subramanya (2013)
area	km^2	Calculated catchment area	Merri-Hydro (Yamazaki et al., 2019), GeoDAR (Wang et al., 2022)
elev	m	Mean catchment elevation	Merit-DEM (Yamazaki et al., 2019)
elev_max	m	Maximum catchment elevation	See above
elev_min	m	Minimum catchment elevation	See above
elev_std	m	Standard deviation of elevation in catchment	See above
elev_range	m	Range of catchment elevation (maximum minus minimum elevation)	See above
slope	m km^{-1}	Mean catchment slope, Horn (1981)	See above
mvert_dist	km	Horizontal distance from the farthest point of the catchment to the corresponding gauge (length axis)	Merri-Hydro (Yamazaki et al., 2019)
mvert_ang	degree	Angle between the north direction and connection from farthest point of catchment to the corresponding gauge (length axis); e.g., direction from north (farthest catchment point) to south (gauge):180 degree, direction from east to west: 270 degree	See above
elongation_ratio	-	Ratio: elongation ratio, i.e., ratio between the diameter of an equivalent circle and the area of the catchment area to its length, Schumm (1956)	Subramanya, K. (2013)
strm_dens	km km^{-2}	Ratio: stream density, i.e., ratio of lengths of streams and the catchment area	See above
resArea	km^2	reservoir area.	Wang et al. (2022)
form_factor	-	Ratio: catchment area / (length)^2	Subramanya, K. (2013)
shape_factor	-	Ratio: (catchment length)^2 / catchment area	See above
circulatory_ratio	-	Ratio: perimeter of the catchment / perimeter of the circle whose area is that of the basin	See above
compactness_coefficient	-	Ratio: perimeter of the catchment / perimeter of the circle whose area is that of the basin	See above
resArearatio	-	Ratio: reservoir area / catchment area	Merri-Hydro (Yamazaki et al., 2019), GeoDAR (Wang et al., 2022)
relief	-	Ratio: mean catchment elevation / Maximum catchment elevation	See above

Table S11. Meteorological data provided in the Res-CN.

Variable	Unit	Description	Source
DOY	-	Day of year	-
pet_mean	mm/d	Mean daily PET (Penman–Monteith equation)	Subramanya (2013), Nation station-based climate data v3
ep	mm/d	Mean daily evaporation (observations)	Nation station-based climate data v3, 1980-2020
ts	◦C	Mean, maximum, and minimum daily ground surface temperature	See above
prec_nscd	mm/d	Mean daily precipitation	See above
prs	hpa	Mean, maximum, and minimum daily ground surface pressure	See above
humid	-	Mean daily relative humidity	See above
sun	h	Mean daily sunshine duration	See above
ta	◦C	Mean, maximum, and minimum daily temperature	See above
wind	m/s	Mean, and maximum daily wind speed	See above

Table S12. Climate indices provided in the Res-CN.

Attribute	Unit	Description	Data source
p_mean	mm/d	Mean daily precipitation[a]	Nation station-based climate data v3
et0_mean	mm/d	Mean daily reference evapotranspiration ET0[b]	Global Evapo-Transpiration (ET0) Database v3 (Zomer and Trabucco, 2022)
arid	-	Global Aridity Index Database v3[b]	Global Aridity Index Database v3 (Zomer and Trabucco, 2022)
p_season	-	Seasonality and timing of precipitation (estimated using sine curves) to represent the annual precipitation cycles; positive (negative) values indicate that precipitation sums are higher during summer (winter) months; values close to 0 indicate uniform precipitation throughout the year; Eq. (14) in Woods (2009)[a]	Nation station-based climate data v3
frac_snow	-	Fraction of precipitation falling as snow, i.e., falling on days with mean temperature below 0 ◦C[a]	See above
hi_prec_fr	d yr^{-1}	Frequency of high-precipitation days (≥ 5 times mean daily precipitation)[a]	See above
hi_prec_du	d	Mean duration of high-precipitation events (number of consecutive days with ≥ 5 times mean daily precipitation)[a]	See above
hi_prec_ti	season[c]	Season during which most high-precipitation days (≥ 5 times mean daily precipitation) occur[a]	See above
lo_prec_fr	d yr^{-1}	Frequency of dry days (< 1 mm/d precipitation)[a]	See above
lo_prec_du	d	Mean duration of dry precipitation events (number of consecutive days with < 1 mm/d precipitation)[a]	See above
lo_prec_ti	season[c]	Season during which most dry days (< 1 mm/d precipitation) occur[a]	See above

[a] Period 1 October 1990 to 30 September 2019. [b] Period 1970 to 2000. [c] List of abbreviations for seasons: djf – December–January–February, mam – March–April–May, jja – June–July–August, son – September–October–November.

Table S13. Attributes of land cover provided in the Res-CN.

Attribute	Unit	Description	Data source
lai_max	-	Maximum monthly mean of one-sided leaf area index (based on 12-monthly means)	MODIS MCD15A3H version 6.1
lai_min	-	Minimum monthly mean of one-sided leaf area index (based on 12-monthly means)	See above
lai_diff	-	Difference between maximum and minimum monthly mean of one-sided leaf area index (based on 12-monthly means)	See above
ndvi_max	-	Maximum monthly mean of NDVI (based on 12-monthly means)	MODIS MOD13Q1 version 6.1
ndvi_mean	-	Average monthly mean of NDVI (based on 12-monthly means)	See above
EVI	-	Mean enhanced vegetation index	See above
GPP	kg c m^{-2}	Mean gross primary production	MODIS MOD17A2H version 6
NPP	kg c m^{-2}	Mean net primary production	MODIS MOD17A3HGF version 6
root_depth_50	m	Root depth (percentiles = 50 % extracted from a root depth distribution based on IGBP land cover)	Eq. (2) and Table 2 in Zeng (2001)
root_depth_99	m	Root depth (percentiles = 99 % extracted from a root depth distribution based on IGBP land cover)	See above
first_type	-	Dominant Land class	ESA WorldCover 10 m
first_frac	-	Fraction of dominant class	See above
Trees	-	Fraction of “trees” (trees)	See above
Grassland	-	Fraction of “grassland” (grassland)	See above
Cropland	-	Fraction of “cropland” (cropland)	See above
Shrubland	-	Fraction of “shrubland” (shrubland)	See above
BU	-	Fraction of “built-up” (bu)	See above
BSV	-	Fraction of “barren sparse vegetation” (bsv)	See above
SI	-	Fraction of “snow and ice” (si)	See above
OC	-	Fraction of “open wate” (oc)	See above
HW	-	Fraction of “herbaceous wetland” (hw)	See above
Mangroves	-	Fraction of “mangroves” (mangroves)	See above
ML	-	Fraction of “moss and lichen” (ml)	See above

Table S14. Attributes of soil provided in the Res-CN.

Attribute	Unit	Description	Data source
bdod*	kg dm-3	Bulk density of the fine earth fraction	SoilGrids250 m (Hengl et al., 2017)[a]
cec*	cmol kg^{-1}	Cation exchange capacity of the soil	See above
soc*	g kg^{-1}	Soil organic carbon content in the fine earth fraction	See above
phh2o*	10	Soil pH	See above
pdep	cm	Soil profile depth	Shangguan et al. (2013)
cl	%	Percentage of clay content of the soil material	See above
sa	%	Percentage of sand content of the soil material	See above
por	cm^3 cm^{-3}	Porosity	See above
si	%	Percentage of silt content of the soil material	See above
grav	%	Rock fragment content	See above
som	%	Soil organic carbon content	See above
log_k_s*	cm d^{-1}	Log-10 transformation of saturated hydraulic conductivity	Soil hydraulic and thermal parameters (Dai et al., 2019)[a]
theta_s*	cm^3 cm^{-3}	Saturated water content	See above
tksatu*	W m^{-1} K^{-1}	Thermal conductivity of unfrozen saturated soils	See above
csol*	J/(m^3K)	Volumetric heat capacity of soil solids in a unit soil volume	See above
lambda*	-	Pore size distribution index for the Campbell model	See above
log_vgm_n*	-	Log-10 transformation of a shape parameter for the VG model	See above
psi_s*	cm	Saturated suction for the Campbell model	See above
tkdry*	W m^{-1} K^{-1}	Thermal conductivity of dry soils	See above
tksatf*	W m^{-1} K^{-1}	Thermal conductivity of frozen saturated soils	See above
vf_clay_s*	cm^3 cm^{-3}	Volumetric fration of clay	See above
vf_gravels_s*	cm^3 cm^{-3}	Volumetric fration of gravel	See above
vf_om_s*	cm^3 cm^{-3}	Volumetric fration of SOM	See above
vf_quartz_mineral_s*	cm^3 cm^{-3}	Volumetric fration of quartz within mineral soils	See above
vf_sand_s*	cm^3 cm^{-3}	Volumetric fration of sand	See above
vf_silt_s*	cm^3 cm^{-3}	Volumetric fration of silt	See above
vgm_alpha*	cm^{-1}	The inverse of the air-entry value for the VG model	See above
vgm_theta_r*	cm^3 cm^{-3}	Residual moisture content for the VG model	See above

* Within the aforementioned 28 soil variables, 21 variables marked with * are represented across 7 levels encompassing six soil layers as well as the entire soil layer. An instance of this is the cation exchange capacity (CEC) of the soil, which has 7 associated attributes denoted as cec_1, cec_2, …, cec_6, and cec, indicating the CEC of the first to sixth soil layers and the entire soil layer, i.e., at six layers of 0–0.05, 0.05–0.15, 0.15–0.30, 0.30–0.60, 0.60–1.00, and 1.00–2.00m, as well as the whole soil layer. In this sense, we provided 154 soil attributes.

Table S15. Attributes of geology provided in the Res-CN.

Attribute	Unit	Description	Data source
permeability	m^2	Subsurface porosity	GLHYMPS (Gleeson, 2018)
porosity	-	Subsurface permeability (log-10)	See above
first_type	-	Dominant geological class	GLiM (Hartmann and Moosdorf, 2012)
first_frac	%	Fraction of “first_type”	See above
ig	%	Fraction of “ice and glacier” (ig)	See above
mt	%	Fraction of “metamorphic” (mt)	See above
pa	%	Fraction of “acid plutonic rocks” (pa)	See above
pb	%	Fraction of “basic plutonic rocks” (pb)	See above
pi	%	Fraction of “intermediate plutonic rocks” (pi)	See above
py	%	Fraction of “pyroclastics” (py)	See above
sc	%	Fraction of “carbonate sedimentary rocks” (sc)	See above
sm	%	Fraction of “mixed sedimentary rocks” (sm)	See above
ss	%	Fraction of “siliciclastic sedimentary rocks” (ss)	See above
su	%	Fraction of “unconsolidated sediments” (su)	See above
va	%	Fraction of “acid volcanic rocks” (va)	See above
vb	%	Fraction of “basic volcanic rocks” (vb)	See above
vi	%	Fraction of “intermediate volcanic rocks” (vi)	See above
nd	%	Fraction of “no data” (nd)	See above
wb	%	Fraction of “water bodies” (wb)	See above

Table S16. Attributes of anthropogenic activity provided in the Res-CN.

Attribute	Unit	Description	Data source
population*	-	Population for the years 2000, 2005, 2010, 2015, and 2020	Gridded Population of the World (GPW) database v4.11
avg_vis*	-	The average of the visible band digital number values with no further filtering	DMSP-OLS Nighttime Lights v4 dataset (Doll, 2008)
stable_lights*	-	The cleaned up avg_vis contains the lights from cities, towns, and other sites with persistent lighting, including gas flares. Ephemeral events, such as fires, have been discarded. The background noise was identified and replaced with values of zero	See above
cf_cvg*	-	Cloud-free coverages tally the total number of observations that went into each 30-arc second grid cell. This band can be used to identify areas with low numbers of observations where the quality is reduced.	See above
avg_lights_x_pct*	-	The average visible band digital number (DN) of cloud-free light detections multiplied by the percent frequency of light detection. The inclusion of the percent frequency of detection term normalizes the resulting digital values for variations in the persistence of lighting. For instance, the value for a light only detected half the time is discounted by 50%. Note that this product contains detections from fires and a variable amount of background noise	See above
reproject_grip4_total_dens_m_km2	$m\ km^{-2}$	Road density	Global Roads Inventory Project (GRIP) dataset (Meijer et al., 2018)
reproject_hfp2009	-	The Human Footprint camp of cumulative pressures on the environment in 2009	Global Human Footprint v2 dataset (Venter et al., 2016)
reproject_hfp1993		The Human Footprint camp of cumulative pressures on the environment in 1993	Global Human Footprint v2 dataset (Venter et al., 2016)

* Within the population category, there are five included attributes, i.e., population_2000, population_2005, population_2010, population_2015, and population_2020. As for the Nighttime light category, which comprises of avg_vis, stable_lights, cf_cvg, and avg_lights_x_pct, both the mean and sum values for each variable are provided for all available time frames. To illustrate, the variable mean_cf_cvg_101994 denotes the mean value of cf_cvg for the month of October in 1994. Accordingly, a total of 288 anthropogenic attributes have been provided.

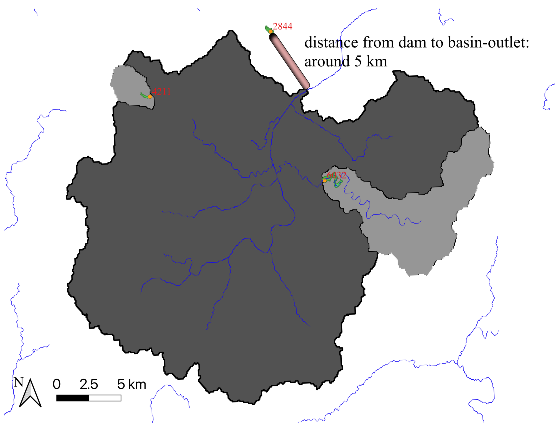
Texts

Text S1

The following steps details the procedures for automatically delineating reservoir upstream catchments.

1. Using the MERIT Hydro flow directions and GeoDAR dam locations to drive the outlet relocation algorithm, resulting the raw full catchment of each reservoir (using the watershed .exe).
2. Clearing the holes to remove topology errors across full catchments using QGIS 3.24.
3. Checking all full catchments and removing/modifying unrealistic or incorrectly catchments. Most of these incorrectly catchments are assigned to small reservoirs located near the confluences of rivers of different sizes. See the example below where the reservoir "2844" is assigned to a large catchment of the mainstream (dark) rather than a small catchment of the tributary. We remove these small reservoirs that are assigned by a large catchment considering three aspects. Firstly, these reservoirs are relatively small, with a median size of 0.06 km^2 and 63% of them are smaller than 0.01 km^2. Secondly, the reservoirs have little regulation impacts and are typically removed from the literature (see CAMELS datasets). Thirdly, no rivers intersect the reservoirs, even the latest small rivers provided by Merit-Hydro, which means that the delineated catchment area is extremely small.
4. Generating intermediate catchments by removing the overlapping areas of upstream reservoirs from the full catchment of the current reservoir using QGIS 3.24.
5. Fixing the invalid geometry of intermediate catchments by eliminating geometry errors.

Based on the literature, and our discussion – cross validation (See main text Section 3.2), verifying the reliability of our delineated catchments.



Example.