ARTICLE IN PRESS

Journal of Great Lakes Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Great Lakes Research

journal homepage: www.elsevier.com/locate/ijglr

Enhancement of a global lake and reservoir database to aid climate studies and resource monitoring utilizing satellite radar altimetry

Charon M. Birkett [a,*], Katherine O'Brien [b], Scott Kinsey [b], Martina Ricko [c], Yuanjie Li [d]

[a] Geodesy and Geophysics Laboratory, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA
[b] ESSIC, University of Maryland, College Park, MD 20740-3823, USA
[c] KBR, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA
[d] Department of Geographical Sciences, University of Maryland, College Park, MD, USA

ARTICLE INFO

Article history:
Received 29 April 2021
Accepted 26 November 2021
Available online xxxx
Communicated by Anthony Vodacek

Keywords:
Database
Lakes
Reservoirs
Altimetry
Water level
Water resources
Anthropogenic influences

ABSTRACT

Here, we describe the expansion and enhancement of a large (surface area ≥100 km^2) lake and reservoir database (1.GREALD). These efforts have also resulted in the spin-off of two additional databases, one containing lakes and reservoirs (2.GREALD, area 10–99 km^2), and the other containing ephemeral lakes (3.GELD, area ≥100 km^2). These databases are unique sources for projects that utilize satellite radar altimeter data to monitor surface water levels. While 1.GREALD aims to be a complete catalog, 2. GREALD focuses on reservoirs in response to applied sciences programs that monitor water and energy resources. The creation of 3.GELD has climate change objectives as well as water resources and ecosystem conservation applications. The recording of information pertaining to the potential overpasses (water-body crossings) of the current and archive satellite altimeters is a primary objective as is the need to highlight any form of controlled water level variation. The permanent water databases now contain 6282 entries, half experience some form of anthropogenic influence and ~430 have been identified as potential climatically sensitive terminal lakes. The revised integral surface area distribution is a power law with exponent −1.016. Statistics reveal that with altimetric repeat visit times of 10-day to monthly, at least 80% of the permanent water bodies (≥10 km^2) have been overflown at some period since the 1990s. Current information on water use and reservoir formation date show that the primary use of the reservoir class is hydroelectric power, and that China, Brazil, India, Turkey, and Vietnam dominate the dam building in recent decades.

Introduction

A remote sensing program was initiated in the early 1990s with the objective of monitoring short and medium-term volume changes in large lakes with a view to interpreting the results in terms of "aridity variations", the associated "aridity index" being a measure of regional climate change (Mason et al., 1994). Closed lakes, which have no significant surface or subsurface outflow, are particularly climatically sensitive and have been used in the study of palaeoclimate (Street-Perrott et al., 1986). Such a climate study program considered the use of remote sensing techniques to derive surface area (multi-spectral imagery) and water levels (satellite radar altimetry). The program also required the global identification of potentially closed and climatically sensitive large lakes (≥100 km^2), and it needed a summary of which water bodies had satellite-based altimeter overpasses. The historical and current radar altimeters (Table 1) are profiling, not swath instruments, and are set in fixed repeat orbits, essentially tracing out a set of narrow

Abbreviations: AVISO, Archiving Validation and Interpretation of Satellite Oceanographic Data; DIFRA, Database on the Inland Fishery Resources of Africa; ENSO, El Niño Southern Oscillation; FAO, Food and Agriculture Organization; GIS, Geographic Information System; GDW, Global Dam Watch; GLWD, Global Lakes and Wetlands Database; GRanD, Global Reservoir and Dam Database; GREALM, Global Reservoir and Lake Monitor; GWM, Global Water Monitor; GWSP, Global Water Systems Project; HEP, Hydro Electric Power; ICOLD, International Commission of Large Dams; ILEC, International Lakes Environment Committee; MGLD, Mullard Global Lakes Database; MODIS, Moderate Resolution Imaging Spectroradiometer; NASA, National Aeronautical Space Administration; NGA, National Geospatial Intelligence Agency; ONC, Operational Navigation Charts; SMASH, SMall Altimetry Satellite for Hydrology; SWOT, Surface Water and Ocean Topography; USDA/FAS, United States Department of Agriculture Foreign Agricultural Service; WDBII, United States World Data Bank II; WGS, World Geodetic System; WLD, World Lake Database.

* Corresponding author.
E-mail address: Charon.M.Birkett@nasa.gov (C.M. Birkett).

https://doi.org/10.1016/j.jglr.2021.11.013
0380-1330/Published by Elsevier B.V. on behalf of International Association for Great Lakes Research.

Table 1
Historical, current, and future radar altimeter missions with short-repeat periods. These missions and mission phases are listed by nominal repeat period (i.e., the lake or reservoir re-visit time based on a single overpass), and this varies from 1 day to 35 days. The altimeter instruments are nadir-pointing (profiling) and separation of missions or phases into 12 sections in the table also denotes a change in ground track location and thus the specific water bodies (and the total number) that are overflown. Most instruments operate at Ku-band frequency, except those onboard SARAL and SWOT that operate at Ka-band. While SWOT will carry a profiling altimeter (this table), it will also operate the first swath-based altimeter. The 10-day and 27-day repeat missions will provide continuity of surface water level information to at least 2030. Those missions or phases labelled as "interleaved" have ground tracks that are shifted from a similar repeat cycle mission to provide additional spatial coverage.

Radar Altimeter Satellite Mission	Operating Agencies	Time Period	Latitude (degrees)	Repeat (days)
SWOT cal/val phase	Consortium 1	launch 2022	±78.0	1
ERS-1 commissioning phase	ESA	1991	±81.5	3
ERS-1 ice-1, ice-2 phases	ESA	1991–1992, 1993–1994	±81.5	3
TOPEX/Poseidon	NASA, CNES	1992–2002	±66.1	10
Jason-1	NASA, CNES	2002–2008	±66.1	10
Jason-2	Consortium 2	2008–2016	±66.1	10
Jason-3	Consortium 2	2016–2021	±66.1	10
Sentinel-6A Michael Freilich	Consortium 3	2020-present	±66.1	10
Jason-1 interleaved	NASA, CNES	2009–2012	±66.1	10
Jason-2 interleaved	Consortium 2	2016–2017	±66.1	10
HY-2A	CNSA	2011–2016	±80.7	14
Seasat	NASA	1978	±72.0	17
Geosat	US Navy	1985–1989	±72.0	17
GFO	US Navy	2000–2008	±72.0	17
SWOT	Consortium 1	launch 2022	±78.0	21
Sentinel-3A	ESA	2016-present	±81.4	27
Sentinel-3B interleaved	ESA	2018-present	±81.4	27
ENVISAT extension	ESA	2010–2012	±81.5	30
ERS-1	ESA	1992–1993, 1995–1996	±81.5	35
ERS-2	ESA	1996–2003	±81.5	35
ENVISAT	ESA	2002–2010	±81.5	35
SARAL	ISRO, CNES	2013–2016	±81.5	35

Agency Consortiums are,
(1) NASA, CNES, CSA, and UKSA.
(2) NASA, CNES, NOAA, and EUMETSAT.
(3) ESA, NASA, CNES, NOAA, and EUMESTSAT.

ground tracks over the Earth's surface. Hence a water body may have a satellite overpass or not. Early altimeter instruments were additionally restricted to observations of the largest water bodies. With limited published sources a lake and reservoir database (the Mullard Global Lakes Database or MGLD) was created based on visual inspection of a global set of 1:1,000,000 Operational Navigation Charts (ONC) dating between 1970s and 1990s and the manual overlaying of altimeter reference ground tracks derived from satellite orbit computation software. The resulting database (Birkett and Mason, 1995) contained details on lake name, geographical location, water body type, approximate lake surface area, ONC map number, altimetric-satellite crossing potential relating to the 1990s instruments, a rough guide to the severity of the terrain in terms of the approaching altimetric satellite overpass, and basic lake and reservoir background information. A few ephemeral lakes (endorheic basin or inland sink) and coastal lagoons were included, but, at the time, they were not the main research focus.

Later developments have seen other researchers and organizations exploit no-cost Landsat imagery, the availability of Google Earth (GE), and the dissemination of free information via the Internet, to compile their own lake and reservoir databases and on a much-expanded scale including the vast majority with surface area <10 km^2. The Global Lakes and Wetlands Database (GLWD, Lehner and Döll, 2004) is one such example where the earlier MGLD offered a limited cross-checking exercise for those $\geq$100 km^2. The original aridity index study program behind MGLD has since progressed and a variety of National Aeronautical Space Administration (NASA)-funded science and applied science programs now look to the monitoring of lake storage, via observation of surface area and levels. Altimetric-derived water levels, available over a 25 year period, are now an accepted climate index (Birkett and Crétaux, 2011), and stakeholders look to the operational monitoring of lakes and reservoirs to ascertain water and energy resources that are essential for municipal, industrial, agricultural, and fisheries uses. Anthropogenic effects and changes due to climate variability remain essential key influences, and regional stability studies are becoming more prevalent as potential threats to these resources are analyzed for national and international considerations. The ability to have archival and near real time measurements, and to highlight water-level controls, are essential requirements and so operational-based lake and reservoir monitoring programs are now running under multi-agency sponsorship.

With rapid reservoir development (Zarfl et al., 2015), increasing pressure on resources, recognition that even the largest open lakes can be influenced by large-scale El Niño Southern Oscillation (ENSO) or Indian Ocean events (e.g., Birkett et al., 1999), and an increase in the number of altimetric missions (Table 1), there was a need to expand and enhance the original MGLD database. This article describes this process and the various sources and methods utilized to enhance the original database. It also provides details of the creation of two spin-off databases. The contents of the revised and new databases are discussed in terms of highlighting the current number of anthropogenically influenced water bodies, the number overflown by the current and historical radar altimeter instruments, and the ability to meet the monitoring requirements of various end users. One primary stakeholder, the United States Department of Agriculture Foreign Agricultural Service (USDA/FAS) initiated the Global Reservoir and Lake Monitor (GREALM, Birkett et al., 2010), an on-line portal within their CropExplorer system, where both FAS and public users can access lake and reservoir altimetric-derived water level products. The NASA Global Water Monitor (GWM, see References for GREALM and GWM web site addresses) mirrors the same water level products as well as offering altimeter-derived water levels for river reaches, wetlands, and global sea level rise.

Methods

Data sources

Several information sources were utilized to revise the original MGLD (≥100 km^2) and to create two additional databases for the smaller (10 km^2 ≤ area ≤ 99 km^2) lakes and reservoirs, and for the ephemerals (≥100 km^2). The MGLD became the first new Global Reservoir and Lakes Database (1.GREALD), and the additional database for smaller water bodies became the second Global Reservoir and Lakes Database (2.GREALD). The third, the Global Ephemeral Lakes Database (3.GELD), was created for the ephemeral records. A brief overview of the sources is provided here noting that many are being upgraded on a continuing basis. While other databases are available, e.g., the ≥0.01 km^2 database of Sheng et al, (2016), and the ≥0.1 km^2 "HydroLAKES" database of Messager et al, (2016), small (<10 km^2) water bodies are not the focus here in consideration of the current satellite radar altimeter size limitations. Databases and catalogs containing the larger water bodies were given priority.

a) The Mullard Global Lakes Database

The original MGLD was specifically created for projects utilizing satellite radar altimeter data. It was constructed based on Operational Navigation Charts (ONC), limited published sources and catalogs. It was developed as a Filemaker Pro (Application) collection, one record for each lake or reservoir, with each record containing information fields (see Birkett and Mason, 1995 for details). Based on ONC depiction, water bodies roughly ≥100 km^2 were considered but it excluded most ephemerals and lagoons. In 1995, the number of water bodies held in the database was 1403 and each lake or reservoir entry was given a unique 4-digit Identification Number (ID). The database distinguished between reservoirs, open lakes, ephemeral lakes, lagoons, and closed lakes. The latter are terminal lakes and technically should have no surface or subsurface outflow, but the MGLD denoted a "closed" classification based on the absence of surface outflow according to the ONC chart depiction of drainage and contours, or if this designation was quoted in publications. If there was any ambiguity such lakes were classed as "closedx" (potentially closed) and any reports of ground seepage were noted in a general information field. Water bodies were classed as "res" (reservoir) if the ONC depicted the presence of a dam. The majority of MGLD lake area values were based on static estimates crudely and manually deduced from their representation on the ONC maps with no coastline digitization. The surface area was a first estimate used to define an approximate size-bracket for remote sensing instrument considerations. Satellite orbit prediction software and the United States World Data Bank II (WDBII) digital representation of the world were used to manually (via plotting) identify which water bodies had an altimetric satellite overpass (based on the then current missions), though in WDBII some water body perimeter polygons did not close and in complex regions some water bodies were indistinguishable from each other.

b) Technical Notes and Online Tools

Access to the MGLD enabled limnologists and other researchers to comment on the validity of its records. A collection of technical exchanges (S.V. Ryanzhin, private communications, 1995–1999) discussed comparisons of the MGLD to the Geographic Information System (GIS) WORLDLAKE database which contains geographic, morphometric, climatological, hydrological etc. data for 24,000 natural and 2300 manmade lakes based on historical literature sources (Ryanzhin and Straökraba, 1999). For water bodies >100 km^2, the WORLDLAKE contains data on 1600 natural lakes (including lagoons and ephemerals) and 300 reservoirs. Inaccuracies in the MGLD were later published (Ryanzhin and Geller, 2006) and mostly referred to name and location errors, duplicate records, or omissions (22 water bodies). This review greatly assisted with the identification of about a quarter of the MGLD lakes where no name could be deduced from the ONC and with discrepancies in water body type or lake surface area. Regarding the latter, roughly 5% of the MGLD lakes had lake areas that differed from published sources by more than the 20% error originally estimated by Birkett and Mason (1995), and these were investigated and corrected.

Considering online tools, the availability of the GE application based on Landsat imagery greatly assisted the upgrading and creation of the 1.GREALD and 2.GREALD databases. In addition, Google Maps, MapCarta and Wikipedia helped with water body names, type, approximate construction dates, general background on anthropogenic influences etc.

c) Source books for the Inland Fishery Resources of Africa

The Food and Agriculture Organization of the United Nations (FAO) originally published a three-volume source book set pertaining to African inland fishery resources (Vanden Bossche and Bernacsek, 1990). This became the on-line Database on the Inland Fishery Resources of Africa (DIFRA) data bank (see References section for web address). The source provides lake and reservoir details such as name, location, type, surface area, and bathymetry, and contains over 1000 records for the African continent.

d) The ILEC database

The original MGLD accessed information from the International Lakes Environment Committee (ILEC) 1988–1991 three data volumes which was in book form and contained only 183 lakes and reservoirs at the time. Now on-line (see References section for web address) and based on five volumes, the ILEC World Lake Database (WLD) contains over 750 water bodies (ILEC, 1993). As an ongoing project it provided much background information on the lakes and reservoirs.

e) Aquastat Dam Database

The FAO aims to provide a global information system on water and agriculture (AQUASTAT, see References section for web address) and within this there is a geo-referenced database on dams. Containing over 14,000 records, this source became an important input into the Global Reservoir and Dam Database (GRanD, see below), especially for the African dams. This dam database was used for the 1.GREALD and 2.GREALD African continent record entries in terms of locational, water body type (lake, reservoir, or lagoon) and general information.

f) The Global Lakes and Wetlands Database

The GLWD was one of the first large-scale global catalogs of lakes and reservoirs (Lehner and Döll, 2004, and see References section for web address). Based on satellite imagery, existing maps, freely available information sources, and GIS functionality, it distinguished lakes from reservoirs, and accepted saline lakes. It also offered lake area estimates based on a variety of methods and sources. The latter included vector-derived (polygon) published or reported maximum and minimum values, or a lake area derived from knowledge of volume and dam height. In general, the GLWD water body name could be associated with the reservoir or the dam, and the "dam year" could be any point in the history of the dam, e.g., commission date, date of dam construction start or completion, time of first turbine switch-on etc.

The GLWD-1 layer included shoreline polygons of 3067 lakes (surface area $\geq$50 km^2) and 654 reservoirs (surface area $\geq$35 km^2) while the GLWD-2 layer included 250,000 shoreline polygons of permanent open water bodies (surface area $\geq$0.1 km^2 but excluding those in GLWD-1, three quarters of these are North of 50° latitude). The GLWD-1 utilized multiple information sources and was cross-validated against documented data including the MGLD, retaining the MGLD type classification, ID number and water body area. Location, name, type, area (AREA_SKM i.e., polygon based), elevation, dam year, river source, and water use fields were utilized as input or for cross-validation efforts with respect to the 1.GREALD and 2.GREALD record fields.

g) The Global Reservoir and Dam Database

The 2004 GLWD reservoir information was based on sources from 1977 and 1998 publications and so the Global Reservoir and Dam database (GRanD) was created to catalog a more up to date reservoir inventory. Under the Global Water Systems Project (GWSP), GRanD Version 1.0 was made available in 2006, and was later followed by upgraded Versions 1.1 (6862 records) and finally Version 1.3 (in 2019) which offered an additional 458 reservoirs that had been (mostly) constructed between 2000 and 2016 (Lehner et al., 2011a,b, and see References section for web address). GRanD notably contains information on the dam location and year, river source, reservoir use, and reservoir surface area. The GRanD area value "Area_skm" (mostly based on the GRanD "Area_poly" i.e., based on the water body polygon) was entered into the 1.GREALD and 2.GREALD databases.

h) The International Commission of Large Dams Register

The International Commission of Large Dams (ICOLD) world register of dams (see References section for web address) lists many thousands of dams and reservoirs and is updated on an annual basis. Access is by annual license fee and the register omits specific location information, though one can search by country and confine the search engine by reservoir area. It was utilized as a check on reservoirs $\geq$10 km^2 for the USDA/FAS countries of interest (see the Discussion section).

i) The University of Minnesota Water Monitor

The University of Minnesota (UMN) Water Monitor is an online tool (see References section for web address) that aims to monitor the surface water area dynamics of all inland water bodies. It is based on the 8-day composite, 500 m pixel resolution, NASA Moderate Resolution Imaging Spectroradiometer (MODIS) image products from the year 2000. The system is backed by machine learning algorithm research (Khandelwal et al., 2017), and loading data by continent displays locations where there are "persistent water bodies" or "changing water regions". Water body area variations can then be displayed for a select location and in this manner the completion of a dam, i.e., the formation and expansion of a new reservoir, can be identified.

Enhancing the MGLD

With objectives focusing on a more up to date i) satellite radar altimetry overpass listing, ii) identification of water bodies which potentially have anthropogenic influences on their water level variations, and iii) a global lake and reservoir inventory, the original MGLD database was revised, ultimately translating into the new 1.GREALD database with two new spin-off databases, 2.GREALD and 3.GELD. The creation of 1.GREALD was done on a continuous basis across the 2000–2015 time frame. The creation of 2.GREALD was initiated in 2015 with revisions performed between 2015 and 2020. An early form of an ephemeral lake listing was put together at the same time as the MGLD (1995), but final construction and enhancement of 3.GELD was not undertaken until 2020.

While maintaining the original MGLD database Filemaker Pro format, the three new databases resulted in a more robust altimetric and waterbody inventory that contains additional, cross-validated records. In comparison with other databases and the original MGLD, one important note must be highlighted. The resulting 1.GREALD and 2.GREALD databases continue to use the "reservoir" type classification but it's meaning has expanded to represent all anthropogenically influenced water bodies and not just those with the presence of a dam. This modification allows investigators to note which water bodies may have artificial controls on the water level variations, and which still undergo natural variations.

The 1.GREALD Database

Using sources outlined above under Data Sources, MGLD became the new 1.GREALD, containing 2303 records ($\geq$100 km^2). Altimetric coverage fields now refer to many additional historical or current missions (Table 1) while remaining an inventory of permanent water bodies approximately $\geq$100 km^2, the number of entries has increased by two thirds. Name, location, and the water body type classification have been refined, and anthropogenic influences particularly noted (see the GREALD and GELD Information Fields section and "General Water Body Attributes"). Concerning the latter, any such water body potentially experiencing human controls on its water levels has been classed as a "reservoir". This is a very specific 1.GREALD and 2.GREALD type classification and it is not used in other databases.

The 1.GREALD entries were also updated to include impoundments created over the last three decades and a small number of bays, fjords, floodplain lakes and coastal lagoons (shallow water bodies separated from the ocean by an island, reef, or sand bank) were added. Some shallow waters that could be considered as wetlands, and a few river-widening reaches in delta regions were also included. Notably, a greater number of high latitude lakes (Canada, Russian Federation, United States of America (USA)) have been added, and a significant number of dammed reservoirs, especially in Brazil, India, and Kazakhstan. Each record now allows for the inclusion of multi-source surface area estimates, the pool creation dates behind a relatively new dam, and the water-uses of the lake or reservoir. Water bodies with significant agricultural or water resources importance have also been highlighted.

The 2.GREALD Database

The 2.GREALD contains 3979 records and focusses on water bodies 10 km^2 $\leq$ surface area $\leq$ 99 km^2, however its purpose is not to be a complete catalog for all water body types and regions but to focus on mid-low latitude open lakes, and especially reservoirs (again, defined here as any such water body potentially experiencing human controls) at all latitudes. The lower surface area limit here reflects improvements in altimetric instrumentation and data processing which is now enabling a smaller water body to be monitored. This database was primarily initialized using the layers GLWD.1 (records with surface area 11.6 km^2 to 99 km^2) and GLWD.2 (from record numbers 5028 to 18179, or 9.1 km^2 to 35.8 km^2), but GRanD and other sources were additionally used to construct its content.

The 3.GELD Database

The 3.GELD database (containing 213 records with potential extents $\geq$100 km^2) was created in response to several objectives; the need to i) assess potential water resources, ii) monitor basin ecology, and iii) study recharge frequency and location in response to climatic variations. Its original data source was the ONC map set which highlights ephemeral lakes and intermittent streams. While currently containing a minimum number of records, it also serves as a starting reference for those studying dust emission and transmission across arid lands (Prospero et al., 2002; Bryant 2003; Ginoux et al., 2012). The majority of 3.GELD records refer to large ephemeral lakes but several depression regions (where salt can form) are also included e.g., Qattara Depression. Note that water bodies within an ephemeral basin that have become notably permanent between 1984 and 2020 are entered into 1.GREALD or 2. GREALD.

GREALD and GELD Information Fields

Each of the three new databases contain records, one for each water body, and each record comprises a set of information fields (or field elements) as summarized in Table 2. The field elements (denoted in italics) can be grouped under general attribute labels and are discussed below. The record layout of 1.GREALD and 2. GREALD are identical except the more recently constructed 2. GREALD does not contain the *Map No.* field (see below) because the ONC are no longer considered a primary source of information. The 3.GELD ephemeral lake database only contains a subset of fields.

Location

The location field element set includes *Map No.*, *Region*, C*ountry, Lat deg* and *Lat* min (latitude), and *Long deg* and *Long* min (longitude). The *Map No*. field refers to the 1:1,000,000 scale ONC chart number. The *Region* field refers to continents (North, Central, and South America, Africa, Europe, Asia, and Australasia) with the Ural Mountains providing the boundary line between Europe and Asia. While many lakes and reservoirs are bounded by several countries, only one is entered into the *Country* field with updates to reflect the latest political changes. Geographical latitude and longitude co-ordinates (in degrees and minutes, with decimal places allowed in the 2.GREALD *Lat* min and *Long* min fields), refer to the central location of the main body of water, or a location immediately behind the dam for narrow reservoirs. Most original MGLD co-ordinates were revised using Google Earth (GE which is based on the World Geodetic System or WGS84). For the ephemeral lakes in 3.GELD the latitude and longitude fields refer to the central or deepest location based on a blend of GE imagery and ONC map contours. For complex reservoir cascade systems, evidence of white water (water outlet, or rapids) or elevation profiles often helped to detect water flow direction and so correctly place the latitude/longitude coordinates behind the dam.

Naming and Identification

This field element group comprises of *Name*, *Alt Name*, and *Lake ID*. While effort is made to store lake and reservoir names the task is complex due to source-based language or spelling variants as well as historical name changes. Overall, and for remotely monitoring purposes it is the latitude and longitude coordinates that uniquely identifies the water body in question. Nonetheless, effort has been made to insert names into the *Name* and *Alt Name* (Alternative names or spelling variants) fields, making occasional use of the name of a nearby village or landmark as a substitute if no name can be readily found. Regardless, some *Name* fields remain as "Unknown". The *Name* field takes the water body name as the primary entry, rather than the dam, barrage, gate, lock or electricity generating station name, though these are sometimes entered. Water bodies with the same name employ a numeral suffix in brackets (e.g., Victoria (1) in Tanzania, Victoria (2) in Australia, Victoria (3) in Canada) while barrier or generating station names employ Roman numerals (e.g., Grand Riviere I, Grand Riviere II in Canada). Water body names are often abbreviated if named after a regional or national person of importance (e.g., the Luiz E. Magalhaes reservoir in Brazil). A lake that contains enclosures (e.g., Lake Windsor, Bahamas) or a major causeway (e.g., Lake Urmia, Iran) where there may be little exchange of water, especially at low levels, also uses the numerals in brackets to mark different portions of the basin.

Water body names have been derived from the ONC, GE, Mapcarta, Wikipedia and other web sources, and the GLWD and GRanD catalogs. For many records, separate words in the *Name* field translating to mean "lake" or "reservoir" or "dam" have been omitted (e.g., Co, Danau, Embalse, Golu, Lac, Loch, Ozero, Repressa, Reservoir, Sap, Vodokhranilishche). For simplification, many apostrophes and hyphens have also been excluded. Many ephemeral lake names have been sourced from the ONC and translations of "salt" (e.g., Salar, Salina, Sabkha, Chott, Erg, Wadi, Pan) have also been removed. The *Lake ID* field is now a 5-digit integer field. It is a unique ID number allowed to take values 1–3000 (1.GREALD), 3003–7000 (2-GREALD), and 10,000–10,300 (3.GELD).

General water body attributes

There are six fields in this category. The first field, *Type*, is a general water body classification. The original MGLD classifications included "closed", "closedx", "open", "res", "ephem" and "lag". These denoted terminal lakes with no surface or subsurface outlet ("closed" or potentially closed "closedx"), those with an outlet ("open"), those with a dam or other man-made retention structure ("res"), ephemeral water bodies ("ephem"), and lagoons ("lag"). These are kept in the three new databases but as stated earlier, the definition of the "res" classification has been modified, and three additional water body types, "ephemx" (potentially ephemeral), "bay", and "fjord" have been introduced.

A task specific to interpreting surface water level variations required the identification of potential anthropogenic interference on all water bodies. This included the noting of man-made structures such as dams, dikes, embankments, weirs, locks, floodgates, canals, causeways (particularly if there is little or no exchange of water between the partitioned water bodies, e.g., Lake Urmia in Iran) and partitions or enclosures on the surface water (e.g., Lake Windsor in the Bahamas). The presence of agriculture around the basin or references to water being used for irrigation was noted, as were comments on other up- or down-stream water divergences or the presence of cooling towers. The *Type* "res" therefore expanded to represent both man-made (dammed) waterbodies and lakes undergoing any form of anthropogenic influence that might also affect their water level variations. This *Type* definition is very specific to 1.GREALD and 2.GREALD and is not used in other databases. The Caspian and fragmented Aral Seas for example, are thus included in the databases as "res" to signify controls on inflowing water. Several original "closed", "closedx", and "open" records were also re-classified to "res" if any potential anthropogenic influence was discovered, though the *Gen_Info* field (see Background Information section below) recorded the original classification and the reason for the modification. Examples of reclassification to "res" included Lake Ontario (previously "open") and Lake Turkana, Kenya (previously "closed"). Even in cases where dam regulations imitate natural outflows (e.g., Lake Victoria (1) in Tanzania) or there are run-of-the-river Hydro Electric Power (HEP) stations on the outflow (e.g., Lake Malawi in Malawi), the classification remains "res".

Table 2
The 1.GREALD, 2.GREALD and 3.GELD field elements. The fields (in italics) are broadly grouped under different attributes (in bold and underlined). A field is present (''yes'') or not (''no'') in each database, while ''x'' depicts a field that is present but not currently filled.

Field	Details	1.GREALD	2.GREALD	3.GELD
Location				
Map No.	ONC chart number	yes	no	yes
Region	Continent	yes	yes	yes
Country	Country	yes	yes	yes
Lat deg	Latitude (degrees)	yes	yes	yes
Lat min	Latitude (minutes)	yes	yes	yes
Long deg	Longitude (degrees)	yes	yes	yes
Long min	Longitude (minutes)	yes	yes	yes
Naming and Identification				
Name	Primary name of water body	yes	yes	yes
Alt Name	Alternate names, dam name	yes	yes	yes
Lake ID	5-digit unique water body identifier	yes	yes	yes
General Water Body Attributes				
Type	e.g., open, closed, res, ephem, lag, bay, fjord	yes	yes	yes
Area	Approximate surface area (km^2)	yes	yes	no
Area Poly	Google Earth polygon surface area (km^2)	yes	yes	no
Lake Elev	ONC lake elevation (feet)		yes	x
Terrain	Surrounding topography severity index (1 to 3)	yes	yes	yes
Res Year	Year when pool first formed behind dam	yes	yes	no
Radar Altimetry Coverage (presence of an overpass)				
SWOT 1 day	SWOT profiling altimeter 1-day cal/val phase	yes	x	x
ERS1 Ice 3 day	ERS-1 3-day ice-1 and ice-2 phases	yes	x	x
ERS1 Comm 3 day	ERS-1 3-day commissioning phase	yes	x	x
10 day	TOPEX/Poseidon, Jason, Sentinel-6A 10-day repeat	yes	yes	yes
IL 10 day	Jason-1, Jason-2 interleaved 10-day phase	yes	yes	yes
HY2A 14 day	HY-2A mission 14-day phase	yes	yes	yes
17 day	Seasat, Geosat, GFO 17-day repeat	yes	yes	yes
SWOT 21 day	SWOT profiling altimeter 21-day repeat	x	x	x
27 day	Sentinel-3A 27-day repeat	yes	yes	yes
IL 27 day	Sentinel-3B interleaved 27-day repeat	yes	yes	yes
ENVI 30 day	ENVISAT extended mission 30-day phase	yes	yes	yes
35 day	ERS-1, ERS-2, ENVISAT, SARAL, 35-day repeat	yes	yes	yes
Inclusion in Operational and Research Programs				
USDA/FAS	A USDA/FAS water body of interest	yes	yes	yes
GREALM	In the USDA/FAS GREALM program	yes	yes	yes
ESDR	In the NASA Earth Science Data Record program	yes	yes	no
Background Information – General and from other Databases				
Gen_Info	Background information from multiple sources	yes	yes	yes
GLWD_ID	GLWD lake ID	yes	yes	no
GLWD_Area	GLWD lake area (km^2)	yes	yes	no
GLWD_Elev	GLWD lake elevation (meters)	yes	yes	no
GLWD_River	Dammed River	yes	yes	no
GLWD_D_Yr	Reservoir Dam year	yes	yes	no
GRanD_ID	GRanD reservoir ID	yes	yes	no
GRanD_Area	GRanD reservoir area (km^2)	yes	yes	no
Area_Wiki	Wikipedia surface area (km^2)	yes	yes	no
Area_FAO	FAO surface area (km^2)	yes	yes	no
Area_ILEC	ILEC surface area (km^2)	yes	yes	no
WU_1 to WU_7	Water Usage (up to 7 entries)	yes	yes	no

For the ephemeral lakes, the *Type* ''ephemx'' was introduced if enclosures or other human activity was noted in the region, e.g., the Menindee lakes system along the Darling River, Australia, some of which are now being used for water conservation projects. Two other classifications, ''bay'' and ''fjord'', are also new though only a minimum number have been included in 1.GREALD for research purposes, e.g., the Shishmaref and Baird Inlets in the USA, the Limfjorden and Tyrifjorden fjords in Norway. Imagery within GE plus information on the world wide web also allowed greater checks on the terminal lakes, notably evidence of a surface outlet (using GE elevation profile to highlight up- and down-stream direction for example) or suspected subsurface outlet. As an example, Lake Naivasha (Kenya) has no surface outflow but is suspected to have groundwater seepage, so its classification was modified to ''closedx''. A greater number of large lagoons have been allowed into 1.GREALD and 2.GREALD, being distinguished from bays due to the non-influence of tides. Several floodplain lakes (e.g., along the Amazon River) have been included and classed as ''open'', as well as several shallow lakes that could also be classed as a wetland (e.g., Dipor Bil in India, a RAMSAR site and used for fisheries).

The second and third fields in the general water body attribute category are *Area* and *Area Poly* which refer to water body surface area in square kilometers. Due to seasonal fluctuations and problems in defining the perimeter limits of some water bodies, the *Area* field remains a first order approximation i.e., a rough guide to the number of water bodies in a given size category noting limitations of the satellite radar altimetry instruments such as footprint diameter (surface roughness dependent, but typically a few km over lakes) and the along track spatial resolution (~300–600 m) of the surface water levels within the altimeter data sets. The *Area* values are a mix of the original 1.GREALD values, and Wikipedia, ILEC, GLWD, GRanD values, plus new estimates based on the GE surface area measuring tool (which is also entered into the *Area Poly* field). Where the area estimates differed by ≥20%, the GE area determination tool was used as a check. The *Area* field therefore is a multi-source parameter with an unknown error estimate and is frequently rounded to the nearest 5 or 10 km^2 (for those $\geq$100 km^2). All area fields are not filled for the 3.GELD ephemeral database.

The fourth field, *Lake Elev*, retains lake height estimates in feet from the MGLD and these originated from ONC spot heights and contours. While retained for some records, in general it is not a filled field across the three new databases. The fifth field, *Terrain*, is a basic integer value that represents terrain severity (1 = plain or plateau, 2 = some mountainous terrain, 3 = enclosed by mountainous terrain). This also has MGLD heritage and is only available for a subset of records due to improvements in altimetric surface tracking and data processing methods that downplay terrain severity as a major hindrance. Both GLWD and GRanD attempt to provide a dam date but as explained in Lehner et al. (2011a,b) this can equate to many stages in the dam development, e.g., date of dam creation, completion, or last modification, the date of turbine installation or hydro-electric energy commissioning date etc. A reservoir may also be influenced by multiple dams of varying completion dates. With focus on highlighting when an altimetric satellite overpass first crosses the pool formed behind a dam, a new *Res Year* field was created to indicate the year in which the reservoir was first observed to form as per displayed in GE imagery or as per stated in Wikipedia. This field is not filled for all records (GE imagery dating back mostly to 1984), but it is helpful in identifying the most recent dam completions.

Radar altimetry coverage

To date, the radar altimetry instruments have been profiling instruments i.e., they only view the surface directly below the antenna (i.e., at nadir). As they orbit the Earth, they trace out a set of ground tracks which repeat at a set frequency. These are not exact repeat paths, the nadir view is allowed to drift, but no greater than ±250 m or ±1 km (depending on the mission) from a reference ground track. Each satellite also has geographical north/south latitude limits. With such constraints a given satellite therefore either has a lake or reservoir satellite overpass or not, and it is important for the various programs to identify which water body is crossed over by an altimeter and where within its perimeter the crossing takes place. Table 1 lists the historical and current nadir instruments in order of increasing temporal resolution. Separation of missions into sets in this table denotes a change in ground track location and thus the specific lakes or reservoirs that are overflown.

At the end of their main science phases both Jason-1 and Jason-2 were moved to an ''interleaved orbit'' for a short period such that their ground tracks interleaved with those of the original science phase (though still at 10-day resolution). Due to energy constraints, the ENVISAT mission orbit was also modified near the end of its lifetime and operated at 30-day resolution (extended mission phase) with a limited data download capability. The current Sentinel-3A and Sentinel-3B satellites operate simultaneously with the ground tracks of Sentinel-3B interleaving with those of Sentinel-3A. The future Surface Water and Ocean Topography (SWOT) mission (Fu et al., 2012) will carry both an onboard nadir altimeter, and a swath-based radar altimeter. The nominal SWOT repeat periods will be 1-day during the calibration and validation (cal/val) phase and 21-day during the main science phase, though the repeat frequency over a given lake or reservoir will be significantly better for the swath operation.

There are thus twelve radar altimetry coverage fields, the name of each includes a reference to the nominal orbit repeatability (in days) i.e., the repeatability of the ground tracks or the basic temporal resolution of a time series of water level variations derived from a single overpass. Larger bodies of water, or those at high latitude (a denser ground track coverage), may have more than one overpass within a given mission, however this is not distinguished within the three new databases, only the presence of (at least) one overpass is acknowledged. The original MGLD identified radar altimetry coverage (a simple ''yes'' or ''no'' in the fields) based on the then historical and current missions, i.e., the 10-day repeat (TOPEX/Poseidon), 17-day repeat (Seasat, Geosat) and 35-day repeat (ERS-1), whose orbits were later adopted by several follow-on missions (Table 1). The upgraded 1.GREALD additionally notes altimetric coverage for the missions or mission phases with 1-, 3-, 14-, 21-, 27-, and 30-day repeats, as well as the ten day repeat of the Jason-1 and Jason-2 ''interleaved'' mission phases. The resulting twelve fields, *SWOT 1 day*, *ERS1 Ice 3 day*, *ERS1 Comm 3 day*, *10 day*, *IL 10 day*, *HY2A 14* day, *17 day*, *SWOT 21 day*, *27 day*, *IL 27 day*, *ENVI 30 day, and 35 day*, (Table 2) contain a simple ''yes'' or ''no'' if an overpass from that mission/phase is present. Here, ''IL'' denotes an interleaved orbit and the six *10 day*, *IL 10 day*, *17 day*, *27 day*, *IL 27 day*, *35 day* fields do not include in their names a reference to a single mission because multiple missions have utilized, or are expected to utilize, these resolutions. The *SWOT 1 day* and *SWOT 21 day* fields refer to the on-board nadir-pointing (profiling) instrument and not the swath coverage. The other two new databases also contain these coverage fields, but the *SWOT 1 day*, *SWOT 21 day*, *ERS1 Comm 3 day*, and *ERS1 Ice 3 day* fields in 2.GREALD and 3.GELD, and the *SWOT 21 day* field in 1.GREALD, are not currently completed but will be filled via future efforts as the SWOT launch (2022) approaches. To note is that the historical 3-day, 17-day, and 35-day repeat orbit configurations have been discontinued, while the 10-day and 27-day repeat orbits have future continuity.

The approach to identifying which water bodies had an overpass initially considered an automated effort, comparing intersection of ground track locations with perimeter polygons. However, coastline complexity and the discovery of some perimeters that where not fully closed, had the project switch to a manual effort, which resulted in the inspection of ground track locations overlaid on GE imagery. This allowed a much closer examination, identifying radar altimetry crossings over small lake inlets and bays, and noting overpasses that are far from the dam location, or so close to a dam that the pass might clip the adjoining reservoir in a cascade. This manual approach helped to identify potential anthropogenic influences and allowed a check on the current water extent status of the lake or reservoir in terms of flood or drought conditions. The GE time slider tool was used to check the magnitude of seasonal and inter-annual changes in surface area.

There are several sets of mission reference ground tracks. This project utilized those from the Archiving Validation and Interpretation of Satellite Oceanographic Data (AVISO) which are based on an average of actual satellite location over a set number of repeat cycles, rather than a theoretical computation. The one exception was a ground track for the HY-2A mission which was based on mission repeat cycle 112 only (courtesy of Françoise Mertz, AVISO User Services). Because ground tracks can shift by up to ±1 km and the effective instrument footprint size can vary, the satellite coverage fields are set to ''yes'' if the reference ground tracks approach the water body coastline by up to 1 km providing the tracks are not in an adjacent water body. The perimeter of many lakes and reservoirs is often difficult to define, not the least for how to set a boundary limit across inflows and outflows. The surface extent and connectivity of many high-latitude lakes (e.g., Canadian Shield) is also problematic and without detailed hydraulic knowledge the satellite overpass fields are likely to contain a higher proportion of errors in such regions. Defining a perimeter in a cascade of dams was trivial (surface area defined as being from one dam location to the next), however defining the outer limits of a single dammed reservoir was more subjective. In many locations, particularly where very narrow reservoirs are present, flow direction and correct dam location had to be further investigated to define the perimeter. The GE time slider tool was also used to view a currently desiccated basin with the view that if no water has been present at the overpass location since 1984, then the radar altimetry coverage field was set to "no". For the ephemeral lakes, radar altimetry coverage field entries are based on the historical basin perimeter, though their perimeter and connectivity can be similarly complex if multiple small pools are present.

The *Gen_Info* field (see below) sometimes notes details of the altimetry ground track, such as if the track skims the lake (i.e., is roughly a few km inside or outside of the perimeter), or the track resides within, or is a few km distant from, an inflowing or outflowing river, or the track is far upstream from the dam. The first two cases are highlighted because the altimeter ground tracks do not exactly repeat and the instrument footprint is several km wide. A ground track which skims the lake may acquire the lake correctly, or it may suffer from land interference (degrading or resulting in loss of the altimetric lake level measurements). A ground track close to rivers (or deltas) may cause the altimeter to incorrectly monitor their seasonality rather than that of the lake, and a track far upstream from the dam may not capture the full reservoir seasonal variation. The distance of the track from the dam is not recorded and the "far from dam" phrase is rather arbitrary. However, it is applied when the GE time slider tool reveals the track may not always be crossing lake water (e.g., Rajghat reservoir, India), or when the track is more than 50 km (thalweg distance) upstream (e.g., Millenium reservoir, Ethiopia where the associated *10 day* field ground tracks are ~85 km upstream). The on-line GWM lake and reservoir altimetric water level variation products warn of these ground track situations within the GWM product "Advisories".

Inclusion in operational and research programs

The *USDA/FAS*, *GREALM*, and *ESDR* fields are a simple "yes" or "no" entry that highlights respectively, water bodies in the 40°S to 52°N latitude band (the main crop-producing regions), the availability of an altimetric surface water level product on the GREALM web site, and whether the water body was included in the NASA Making Earth System Data Records for use in Research Environments (MEaSUREs) "Development of pre-SWOT ESDRs for global surface water storage dynamics" project (Principal Investigator Dennis Lettenmaier, Tortini et al., 2020).

Background Information – general and other databases

This set of attributes contains 18 fields, all filled using externally sourced information. The *Gen_Info* field allows for brief background knowledge on diverse topics such as water depth, salinity, periods of winter freezing, current desiccation state, wind and tide effects, the location of the satellite overpasses in relation to the main body of water or the dam, anthropogenic influences etc. Five fields retain information from the GWLD database (Lehner and Döll, 2004), the water body ID number *GLWD_ID*, the surface area from polygon outlines *GLWD_Area*, the lake mean elevation *GLWD_Elev*, the name of the dammed river *GLWD_River*, and the dam completion year *GLWD_D_yr*. Both *GLWD_River* and *GLWD_D_yr* can be supplemented by information from the GRanD database (Lehner et al., 2011a,b), which also provides details for two additional fields, *GRanD_ID* (the GRanD ID number) and *GRanD_Area* (the GRanD reservoir area which prioritizes polygon-derived estimates). Lake area estimates can also be stored in three additional fields, *Area_Wiki*, *Area_FAO* and *Area_ILEC*, from the Wikpedia, FAO and ILEC sources. Because the lake and reservoir ID numbers differ between GRanD and GLWD, not all are inserted into 1.GREALD and 2.GREALD, noting again that a greater weight is placed on recording an accurate water body location.

The final seven fields, *WU_1* to *WU_7*, are for holding lake water use information i.e., the primary and secondary functions of the lake or reservoir. Both GLWD and GRanD supplied such information, but Web sources supplemented or provided these details in many records. The water-use fields allow up to seven entries with a single letter denoting the category. There is some departure from the GLWD and GRanD water-use definitions. Here we use, h = HEP, g = geothermal power, s = water supply, i = irrigation (direct or via inflows), f = local commercial fishing or aquaculture (but not sport fishing), l = livestock, c = flood control, n = navigation, p = industry (e.g., pollution control, pulp mill, logging, mining, salt production), r = recreation, and e = of ecological importance.

Results

This section discusses the current contents of the three databases. Here, the records associated with the minimum number of bay and fjord entries have been excluded, and again, specific to these databases, is that the "res" entry in the *Type* field is associated with any form of anthropogenic influence on water levels.

Cross-validation

In particular, the GLWD and GRanD databases proved excellent for original sources and for cross-validation efforts. Some errors were found in GLWD (e.g., lakes were not present at the given location, locations were partially incorrect, duplicate entries) and reported. Other GLWD records did not have a 1.GREALD or 2. GREALD match, e.g., GLWD record numbers 1520, 1466, 1838 are in fact a widening of a river in a delta region, an old dry lakebed full of agriculture, and a group of small lakes, respectively. However, as an early database attempt, GLWD inaccuracies were not unexpected and ongoing efforts aim to improve all catalog collections. Currently ~260 records in 1.GREALD do not have a current match with those in GLWD.

Upgrading efforts (MGLD to 1.GREALD) enabled the number of records to be increased by 65% adding in many water bodies of varying type, area, and location. Various sources helped to redefine high-latitude lakes where perimeter borders were unknown due to complexity and the possibility of connecting pools. For several water bodies a more up to date surface area (e.g., Aral Sea) was also entered and the provision of a water body name for the majority was achieved though 1.GREALD still contains 34 records, and 2. GREALD 386 records, where the name remains as "Unknown", the majority, not surprisingly, in Canada. The 1.GREALD database now contains records $\geq$60 km^2 but the majority are $\geq$95 km^2. The 2. GREALD database contains records with surface area 1–99 km^2 but the majority are 10–94 km^2.

Water body type – number and geographical distribution

Table 3 provides the number of records in 1.GREALD, 2.GREALD, and 3.GELD by *Type* and *continent*. Excluding the few bays and fjords, the combined databases now identify 6495 permanent and ephemeral water bodies, increasing the number of "res" as a *Type* field entry from 226 to 3278. However, the number of closed lakes only increased from 320 to 432. Over 50 lakes, originally classed as "closed", were identified as having anthropogenic influences on the water levels, and therefore *Type* was changed to "res". In addition, ~30% of the currently closed lakes are classed as "closed" i.e., potentially closed, with no obvious surface outflow, but there is evidence of agriculture or other human activity in the basin. The largest closed lake is the Great Salt Lake 6000 km^2 in Utah, USA, though the majority (75%) of closed lakes are $\leq$200 km^2. Similarly for the ephemeral lakes, ~10% have agriculture or human activity in their basin and as such they are classed as "ephemx".

Table 3
Distribution of water body type by continent. Numbers are taken from the 1.GREALD, 2.GREALD, and 3.GELD lake and reservoir databases (represented by superscripts 1, 2, or 3). Note that these represent those ≥100 km^2 (1.GREALD, 3.GELD) and 10–99 km^2 (2.GREALD). Water body *Type* classifications are "res", "open", "closed", and "lag" (or lagoon). The "res" (or reservoir) class is specific to these databases and is defined as any water body experiencing anthropogenic influences on the water levels. The "closed" and "closedx" (potentially closed) classes are combined here, as are "ephem" and "ephemx" (potentially ephemeral). The 2.GREALD database aims to be a complete record set for "res" only.

Continent	All^1	res^1	open^1	closed^1	lag^1	ephem^3	All^2	res^2	open^2	closed^2	lag^2
North America	966	274	627	31	34	17	1,551	703	828	7	13
Central America	19	6	2	3	8	0	14	10	4	0	0
South America	211	94	79	26	12	24	275	174	96	4	1
Europe	250	112	119	14	5	0	540	386	149	2	3
Africa	196	92	42	382	4	30	293	195	78	14	6
Asia	619	242	159	199	19	47	1,224	909	236	78	1
Australasia	42	24	7	8	3	95	82	59	15	8	0
TOTAL	2,303	844	1,035	319	105	213	3,979	2,436	1,406	113	24

In the original MGLD, 61% were classed as open lakes and 16% as dammed reservoirs but this ratio changed to 45% ("open") and 36% ("res") respectively in 1.GREALD as water body *Type* was redefined and new records added. Just over 125 lakes have "desiccated" entered in the *Gen_ Info* field based on exposure of the lakebed or coastline within current GE imagery. Half of these are "closed" which can experience large seasonal surface water area variations, the remainder are equally split between "open" and "res", noting that some dammed reservoirs can experience large fluctuations in extent. Currently, 40% of the water bodies reside in North America which has the majority (32%) of large (≥100 km^2) water bodies defined as "res" (Table 3). At the smaller surface area scale (10–99 km^2), Asia has the majority (37%) of "res", and it has the highest number of "closed" (277) in any size category. Looking at the distribution of "res" by country, Table 4 highlights the top 20 locations by number. The USA, Canada, People's Republic of China, Russian Federation, Brazil and India have the majority in both 1. GREALD and 2.GREALD.

The water bodies are distributed across the globe from −54.5k° to 81.5° latitude. Fig. 1 shows the global distribution of *Type* taken from the combined contents of 1.GREALD and 2.GREALD, and from the 3.GELD database. Clearly marked are the concentrations of open lakes (majority in North America) and the dense network bands of "res" (People's Republic of China, India, South Africa, etc.). Most of the closed and ephemeral lakes occupy the arid and semi-arid lands and high plateau regions which are typically low population regions, though it is clear from the "closedx" and "ephemx" classifications that their water resources may still be crucial. Fig. 1c depicts some closed lakes in exorheic basins (e.g., Amazon, Yangtze), but the majority of these have been defined as "closedx" (i.e., potentially closed, shown as yellow symbols). While the classification was based on the absence of an obvious surface outlet, it highlights the need for some additional cross-checking of this *Type* field entry in the future.

Water body area and integral size distribution

Noting that the *Area* field is an approximation and subject to seasonal and interannual variations as well as anthropogenic influences, the continental distribution of surface water area is provided in Table 5. The total area estimate of the permanent lakes in 1.GREALD is 1.77 million km^2 (3.5% of total land surface area 510 million km^2) with all anthropogenically influenced water bodies ("res") representing ~71% (decreasing to ~28%, if for example, the Caspian Sea and Lake Baikal in Asia, Lake Victoria in Africa, and the Great Lakes, Lake Athabasca, and Great Slave Lake in North America are excluded). By comparison the "res" in 2.GREALD represent an additional 5% of surface area.

As a complete representation of water bodies (and by type) down to ~100 km^2, Fig. 2(left) shows the integral size distribution

Table 4
Top 20 locations with respect to the number of reservoirs. In the table below, "Reservoirs" refer to the "res" *Type* classification in the 1.GREALD and 2.GREALD databases (represented by superscripts 1 and 2 respectively). The "res" class is specific to the databases and is defined as a water body experiencing anthropogenic influences on the water levels. Number and surface areas below may thus be dominated by water bodies that are newly classed as "res", e.g., the Great Lakes (USA), Lake Athabasca and the Great Slave Lake (Canada), Lake Victoria (Tanzania), the Caspian Sea (Kazakhstan), Lake Baikal (Russian Federation).

Country	Reservoirs^1		Country	Reservoirs^2	
Designation	(Number)	(Area, km^2)	Designation	(Number)	(Area, km^2)
United States America	154	278,515	United States America	536	14,829
Canada	104	110,690	People's Rep. China	368	9,650
People's Rep. China	61	24,254	India	220	5,276
Brazil	60	33,012	Canada	117	3,708
Russian Federation	57	98,705	Brazil	99	3,152
India	47	12,625	Spain	74	1,722
Australia	17	3,935	Turkey	69	1,670
Mexico	16	4,120	South Africa	61	1,287
Sweden	14	7,590	Australia	49	1,386
Turkey	14	3,805	Mexico	46	1,437
Iran	13	6,775	Norway	41	966
Iraq	13	7,190	Vietnam	40	837
Kazakhstan	13	404,105	Sweden	36	1,278
Argentina	11	4,232	Iran	34	592
Finland	11	8,815	Russian Federation	29	872
Thailand	11	2,920	Argentina/Thailand	24/24	699/649
Ethiopia	9	5,010	Romania/Venezuela	21/21	311/628
Turkmenistan	9	27,200	Portugal	20	362
Ukraine	9	7,170	Myanmar/N. Korea	18/18	779/358
New Zealand	7	1,550	Pakistan/S. Korea	16/16	375/426

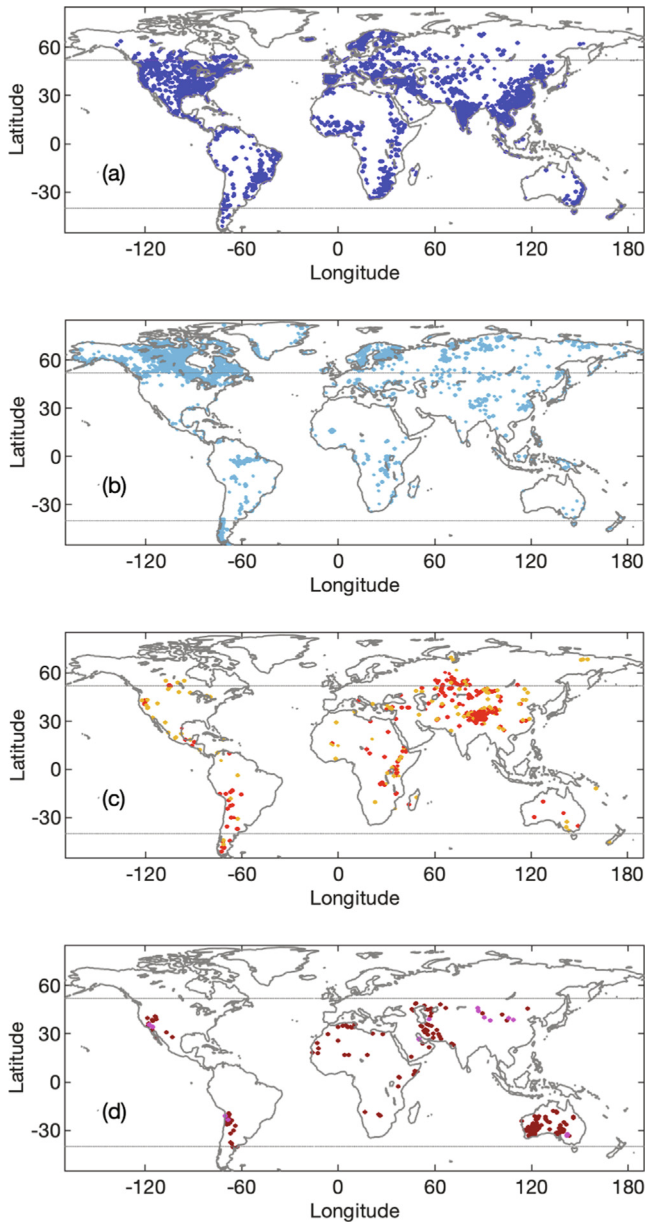

Fig. 1. The global distribution of water bodies according to the *Type* field in the combined 1.GREALD and 2.GREALD databases, and the 3.GELD database. The classes displayed below are (a) "res" (a specific reservoir definition defined as a water body that experiences anthropogenic influences on its water levels), (b) "open" lakes, (c) "closed" lakes, and (d) "ephem" or ephemeral lakes. Closed and ephemeral lakes are further divided into those potentially having some anthropogenic influence on their water levels and are depicted as yellow dots (c) and pink dots (d). Note that while 1. GREALD aims to be a complete database for those with surface area ≥ 100 km^2, the 2.GREALD (10–99 km^2) is only complete for the "res" classification. The dotted lines represent the main agriculture growing band of $-40°$ to $52°$ latitude. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the 1.GREALD water bodies, and Fig. 2(right) the combined 1. GREALD and 2.GREALD "res" records down to 10 km^2. Both graphs are logarithmic, plotting surface area (A) versus the number of water bodies (N) greater than that area value. The following equations describe the power relationships. Note again that the "res" classification is the combined representation of man-made (dammed) waterbodies and other lakes undergoing anthropogenic influences, and there is expectation of closed lakes being the most sensitive to seasonal area fluctuations.

all water bodies in 1.GREALD $100 - 10{,}000\ \mathrm{km}^2$

$$N = 188{,}702\,A^{-1.016} \quad (\text{Eq. }1)$$

"open" lakes in 1.GREALD $100 - 10{,}000\ \mathrm{km}^2$

$$N = 266{,}628\,A^{-1.275} \quad (\text{Eq. }2)$$

"closed" lakes in 1.GREALD $100 - 1{,}000\ \mathrm{km}^2$

$$N = 54{,}716\,A^{-1.187} \quad (\text{Eq. }3)$$

"lag" or lagoon type in 1.GREALD $100 - 1{,}000\ \mathrm{km}^2$

$$N = 4{,}802\,A^{-0.849} \quad (\text{Eq. }4)$$

"res" or reservoir type in 1.GREALD&2.GREALD $10 - 10{,}000\ \mathrm{km}^2$

$$N = 24{,}069\,A^{-0.794} \quad (\text{Eq. }5)$$

The power law form and its variations (e.g., based on the range of surface area, lake geomorphology, location), as well as the validity of extrapolating down to the smallest water body size, has been discussed by several authors (Halbfass 1922; Rapley et al., 1987; Wetzel 1990; Meybeck 1995, Birkett and Mason, 1995; Lehner and Döll, 2004; Downing et al., 2006; McDonald et al., 2012; Seekell et al., 2013; Verpoorter et al., 2014; Messager et al., 2016). Eq. (1) is a revised estimate of that quoted in Birkett and Mason (1995) of N = 84,353$A^{-0.95}$ for all water bodies 100–10,000 km^2. Comparing exponents, Eq. (1) is similar to the Lehner and Döll (2004) relationship of N = 155791$A^{-0.9926}$ (1–100,000 km^2, based on GLWD.1 and GLWD.2 but excluding reservoirs), and similar to the Messager et al., (2016) equation N = 164,698$A^{-1.05434}$ ($\geq$0.1 km^2, natural lakes in the HydroLAKES database). While the "res" type classification differs somewhat from the "reservoir" definition in other databases, the exponent of Eq. (5) is also very close to that quoted by Lehner et al. (2011a), N = 12604$A^{-0.7807}$ (10 km^2 $\leq$ area $\leq$ 1000 km^2, and reservoirs within the GRanD database).

Because the power law form is an integral distribution, the equations are expected to be reasonably accurate, despite the variability of lake area. Interestingly, as the power law index approaches a value of -1, the distribution becomes scale invariant. Thus, in any size interval the percentage of all lakes is approximately the same and as previously stated (Birkett and Mason, 1995) the distribution can be viewed as fractal in form.

Satellite overpass coverage

The databases reflect the profiling altimetry trade-off rule concerning mission temporal repeat and the density of ground tracks i.e., the faster the revisit time, the less dense the ground track coverage, and consequently the fewer water bodies with a satellite overpass. Hence the altimeter suite with a 10-day repeat-frequency crosses over fewer lakes and reservoirs than the instruments with a 35-day repeat. Fig. 3 shows this trade-off rule with respect to the current contents of 1.GREALD. Eleven of the twelve mission/phases (Table 1) are under consideration, with results from the future *SWOT 21 day* field excluded due to incompleteness. The resulting eight repeat periods are thus 1-, 3-, 10-, 14-, 17-, 27-, 30-, and 35-days. The 3-, 10-, and 27-day repeats are also employed by different missions/phases but on shifted ground track locations. These were additionally included as a cross-check and reveal very similar overpass numbers (Fig. 3). The number of water bodies with overpasses is biased towards the natural distribution of water bodies (greater at mid to high northern latitudes), the latitude coverage of the mission (Table 1), larger lakes (by virtue of

Table 5
Distribution of water body surface area by continent. The contents of the 1.GREALD and 2.GREALD databases are represented by superscripts 1 or 2 respectively. Water body *Type* classifications are "res", "open", "closed", and "lag" (or lagoon). The "closed" type includes both "closed" and "closedx" (potentially closed). The "res" (or reservoir) type is specific to the databases and is defined as any water body experiencing anthropogenic influences on the water levels. The surface areas below may thus be dominated by water bodies that are now classed as "res", e.g., in North America (the Great Lakes, Lake Athabasca, and the Great Slave Lake, a total of 279,100 km^2), in Africa (Lake Victoria, 68,800 km^2), and in Asia (the Caspian Sea and Lake Baikal, total 411,500 km^2). The 2.GREALD surface area values have been rounded to the nearest 1 km^2 and its records are not currently complete for the "open", "closed", and "lag" classes.

Continent	All^1	res^1	open^1	closed^1	lag^1	res^2	open^2	closed^2	lag^2
North America	595,408	393,839	172,569	12,360	16,640	20,061	45,511	193	815
Central America	11,465	1,195	9,000	1,270	0.0	262	282	0.0	0.0
South America	118,164	56,619	24,315	7,880	29,350	5,420	4,857	205	60
Europe	131,435	73,670	48,665	6,680	2,420	9,326	7,627	85	59
Africa	234,220	162,910	54,845	8,650	7,815	4,894	2,335	413	115
Asia	675,637	565,684	42,458	57,415	10,080	22,997	13,453	2,788	54
Australasia	9,340	5,485	1,500	1,865	490	1,588	894	81	0.0
TOTAL	1,775,669	1,259,402	353,352	96,120	66,795	64,548	74,959	3,765	1103

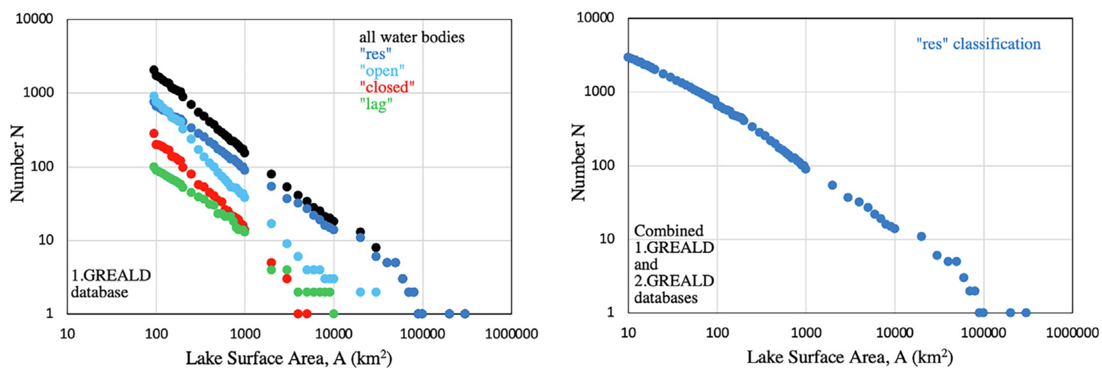

Fig. 2. The integral size distribution of water bodies. This is the number (N) with surface water area greater than a given area value (A), and both plotted on a logarithmic scale. (Left) Relationships based on those held within 1.GREALD and thus for those roughly $\geq$ 100 km^2. The chart depicts the total number of all water bodies, and then the "res", "open", "closed" and "lag" (or lagoon) *Type* classifications. (Right) The same relationship but for the combined 1.GREALD and 2.GREALD "res" classification. Note that "closed" and potentially closed ("closedx") classifications have been combined here and the "res" (or reservoir) class is specific to these databases and defined as any water body experiencing anthropogenic influences on the water levels. Note also that surface area is a seasonally and interannually variable parameter.

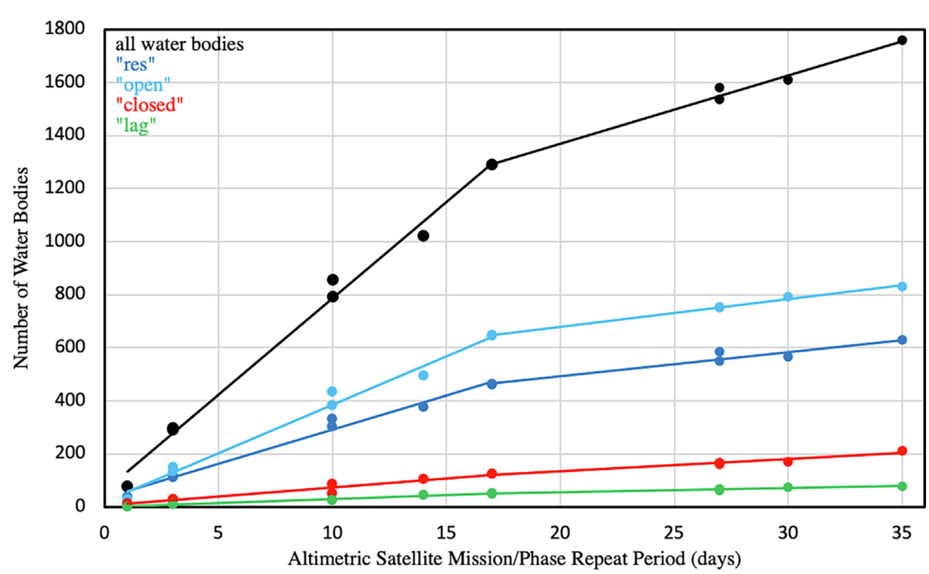

Fig. 3. The relationship between the altimetric repeat period and the number of water bodies with at least one overpass from that altimeter mission or mission phase. Numbers are based on the 1.GREALD database. Excluding the SWOT (21-day profiling) phase, 11 missions/phases are considered (Table 1) resulting in 8 different repeat periods (1-, 3-, 10-, 14-, 17-, 27-, 30-, and 35-days). The 3-, 10-, and 27-day repeats are also employed by different missions/phases but on shifted ground track locations. At these three resolutions there are thus pairs of water body numbers with very similar magnitude. The curves represent all water bodies, and then the "res", "open", "closed", and "lag" (or lagoon) *Type* classes. Note that "closed" and potentially closed ("closedx") lakes have been combined here and the "res" (or reservoir) class is specific to the databases and defined as any water body experiencing anthropogenic influences on the water levels. Solid lines are for visual guidance only.

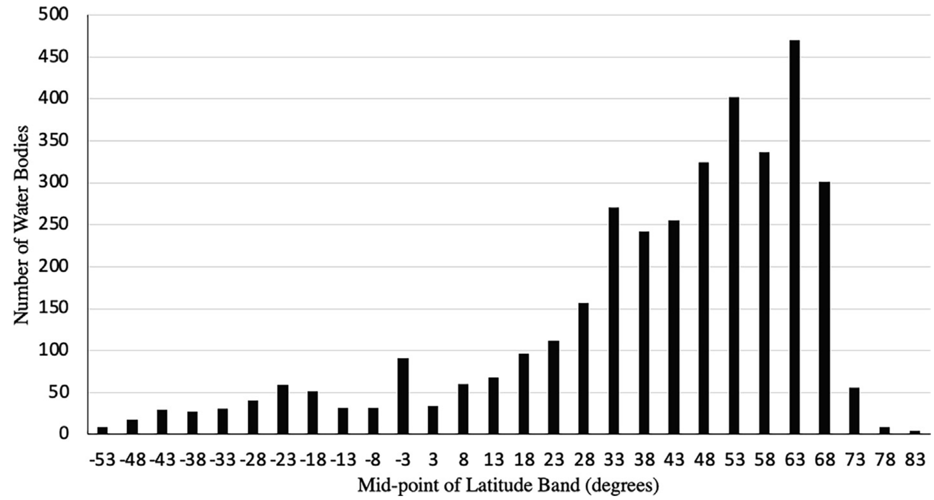

Fig. 4. Histogram of the number of water bodies within a given latitude band that have at least one satellite altimeter overpass. Example is for the 35-day repeat period suite of altimetric instruments (the *35 day* database field). Latitude bands are 5° wide with tick labels centered on the mid-band, e.g., −53° represents the latitude band −55 to −51°. Numbers are based on the combined 1.GREALD and 2.GREALD database records.

their size), those with an East-West aspect, and those at high latitudes where there is a greater ground track density. Fig. 4 looks at the latitude bias based on the total number of water bodies in 1. GREALD and 2.GREALD and the ground track overpasses associated with the 35-day repeat instrument suite. While there are small peaks in the −25 to −20°, and −5 to 0° latitude bands, by far the greater number of passes are between 30 and 70°.

Fig. 5 also reflects the trade-off between temporal repeat and water body coverage with respect to the contents of 1.GREALD and 2.GREALD and the repeat periods with long-term continuity i.e., 10-, 27-, and 35-days (Table 1). Each temporal resolution is associated with a sub-set of open and closed lakes, lagoons, and reservoirs (the "res" *Type* definition). However, as the Venn diagram shows (with respect to "res" as an example), some water bodies can have multiple observations at different temporal resolutions allowing an overall higher temporal resolution to be achieved in a resulting time series of water level variations. However, this is providing that the missions are operating simultaneously, and the level variations are not location dependent. For example, 256 large and 97 small reservoirs have potential for water level measurements at all three temporal resolutions over the mission operating periods.

Water body altimetric-coverage statistics can vary depending on which sets of missions or mission phases are considered and how they are prioritized (see also "scenarios" in the following Discussion section). On a global scale, and with emphasis on accepting a monthly resolution of altimetric water levels, a combined 35-day (*35 day* field), 27-day (Sentinel-3A, the *27 day* field), 27-day (Sentinel-3B, the *IL 27 day* field), and 10-day (*10 day* field) repeat set (in that priority order) could be considered. Statistics then reveal that ~6% of the large lakes in 1.GREALD and 23% of the smaller lakes in 2.GREALD will not have an overpass. Not surprisingly, the majority of these are the smaller water bodies ≤300 km^2 (1. GREALD) and ≤50 km^2 (2.GREALD), and mostly at northern lati-

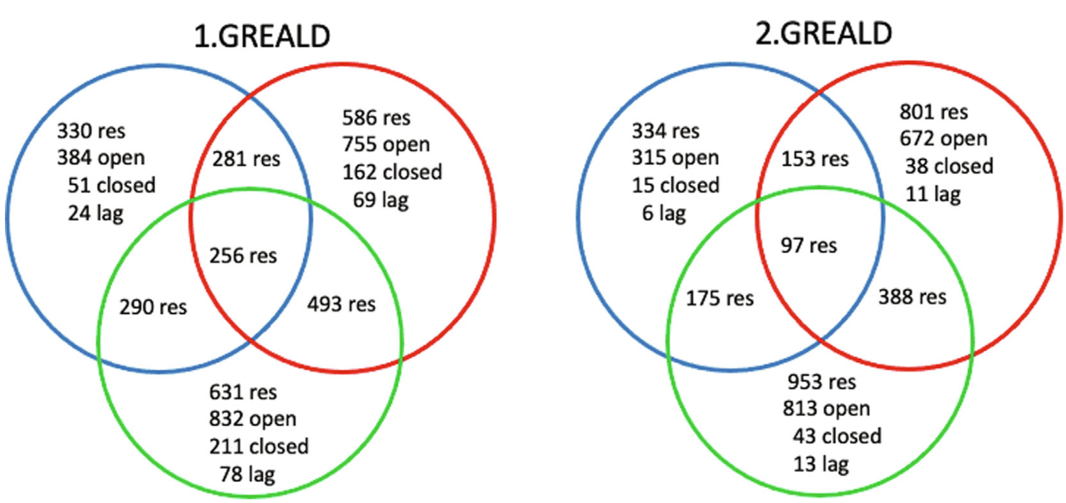

Fig. 5. Venn diagrams showing the trade-off between altimetric repeat period and the number of water bodies with at least one satellite overpass at that repeat period. These are based on the more complete 1.GREALD database records (left) and the current contents of the 2.GREALD database (right). The blue circle represents the 10-day repeat (the *10 day* database field) of the TOPEX/Poseidon, Jason, and Sentinel-6 instrument suite. The green circle relates to the 27-day repeat (the *27 day* field) of the Sentinel-3A mission and the red circle the 35-day repeat (the *35 day* field) of the historical ERS, ENVISAT, and SARAL mission suite (see Table 1). Water body *Type* classifications are "res", "open", "closed", and "lag" (lagoons). Note that "closed" and potentially closed ("closedx") lakes have been combined here and the "res" (or reservoir) class is specific to these databases and defined as any water body experiencing anthropogenic influences on the water levels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tudes. However, it does mean that ~2200 large lakes (1.GREALD) and ~3100 smaller lakes (2.GREALD) do have an overpass. Of these, the most northerly lakes are in Greenland e.g., Lake Romen So (811atitude, with 27- and 35-day repeats) and the most southerly is Lake Fagnano in Argentina (−54.5° latitude, with 27- and 35-day repeats). While many lakes are situated in mountainous terrain and plateaus, the lowest lake is the Dead Sea, Israel, with overpasses offered by the 27-day Sentinel-3B (the *IL 27 day* field) mission only.

Discussion

The contents of the three databases can be applied to both science and applied science programs that utilize satellite radar altimetry for surface water level measurements. For discussion, two altimetric spatial coverage scenarios are being considered here. The first assumes that long-term, multi-decadal datasets are required i.e., the 10-, 27-, 35-day repeat series. The second scenario puts greater emphasis on the current instruments (Jason-3, Sentinel-3A, Sentinel-3B) which are offering near real time, water level products. In either scenario, priority is on temporal resolution, so the 10-day repeat data has greater importance, followed by 27-day and then the 35-day repeat data.

With respect to science projects, researchers interested in climate variability and the utilization of altimetric water level as a proxy for precipitation or water storage can now use the new databases to better identify those water bodies that are expected to have natural variations, i.e., the open lakes, and especially the closed and ephemeral lakes which are known to be particularly climatically sensitive. The *Type* classification coupled with the *Gen_Info* field helps to highlight potential anthropogenic effects on water levels and assists the setting of the on-line GWM altimeter product "Advisories" in this respect. As a water-level product end-user this is relevant when evaluating long-term time series for climatic-related changes. Considering scenario 1, the databases (≥10 km^2) reveal that over 2000 open and over 370 closed lakes are available for climate interpretation of their water levels. Because closed lakes experience larger area fluctuations than open lakes, a surface area versus level (hypsometry) relationship can be utilized to determine water levels for the closed lake if altimetry data are only available for a short period. Other altimetric missions and mission phases then (e.g., when the *IL 10 day or ENVI 30 day* fields are "yes") may offer small time snapshots of water level observation to help build hypsometry, with surface area then being utilized as the main monitoring parameter. Observation of ephemeral lakes is more complex, due to water being temporally and spatially sporadic. The combined altimetric coverage may have to be considered to observe a single water in-fill. The 3.GELD database reveals that only 5 records (out of 213) are not covered by any short- or long-term altimetric mission overpass, with the majority of these situated in Australia.

The databases also aid those interested in the applied sciences particularly with respect to water resources (irrigation, storage) and energy resources (HEP), though lakes and reservoirs also have major roles in navigation, aquaculture, recreation etc. The primary stakeholder of the GREALM program, the USDA/FAS, is interested in all lakes and reservoirs (roughly) between −40° to 52° latitude i.e., the main agricultural growing regions. According to the combined contents of 1.GREALD and 2.GREALD, ~65% of the total number of records (over 4000 lakes and reservoirs) reside in this latitude band. Of these, 20% note irrigation as a water use, and ~3% have been marked as "desiccated" in their *Gen_Info* field to mark extreme conditions as currently revealed in GE imagery.

The current operational instruments (scenario 2) provide an overpass for two-thirds of the water bodies in the USDA latitude band and so near real time measurements (within hours to days) can be acquired. 1.GREALD and 2.GREALD inform us that these would consist of 521 reservoirs ("res" *Type* definition), 109 open lakes, and 52 closed lakes at 10-day resolution (the *10 day* field). An additional 860 "res", 216 open lakes, and 140 closed lakes could then be monitored with Sentinel-3A (the *27 day* field), and then Sentinel-3B (the *IL 27 day* field) could provide an additional 548 "res", 141 open lakes, and 79 closed lakes. In addition to the near real time measurements, those water bodies with a 10-day overpass currently have their 10-day archive data examined to form multi-decadal (from 1992) time series. In this way, USDA can compare near real time measurements with a long-term mean to formulate a "Status Index" to quickly highlight deviations from normal water levels.

The USDA/FAS have certain regions of interest which vary according to weather, climate, flood and drought hazards, water-sharing policies, and regional stability, but drought dominates issues relating to irrigation and crop production. Countries such as Iran, Iraq, Syria, Turkey, India, Pakistan, Brazil, Argentina, Mexico, Thailand, Australia, and all those on the African continent are currently of interest. Again, considering scenario 2, overpass coverage of these countries varies from 54% (India) to 95% (Iraq), though the average is 68%.

In general, the accuracy, delivery time, temporal resolution, and spatial coverage of the altimetric datasets do meet the original USDA/FAS stakeholder requirements regarding the larger (≥100 km^2) water bodies (Curt Reynolds, personal communication, Reynolds et al., 2003, 2007, 2009; Birkett et al., 2010). The latest altimetric instruments and enhanced data processing techniques though places emphasis on additionally monitoring the 10–99 km^2 lakes and reservoirs and in this regards the spatial coverage is compromised. For example, considering the "res" *Type* coverage within −40° to 52° latitude, only 90 "res" in 1.GREALD do not have an operational mission overpass (scenario 2) but as many as 951 "res" in 2.GREALD do not. However, these numbers can be halved (1.GREALD) or reduced to a third (2.GREALD) if the historical 35-day repeat and hypsometry method can be utilized to infer water levels.

While the formation of the reservoir behind the dam (*Res Year*) and water use (*WU_1 to WU_7*) fields continue to be filled for 2.GREALD, the databases do draw attention to dams that have been completed in recent decades and highlight water use priorities. Based on the *Res Year* field in the combined 1.GREALD and 2.GREALD databases, at least 323 dammed reservoirs have been created since 1992, and 67 dams have been built over the last 10 years (from 2011). The majority (77%) of these have smaller surface extents with their records held in 2.GREALD. Those formed since ~1992 are important to note as the dam completion coincides with the start of the long-term altimetric dataset timelines, thus the rate of reservoir in-fill can be captured as required by other stakeholders such as the United States National Geospatial Intelligence Agency (NGA). For the larger reservoirs (≥100 km^2), the People's Republic of China, Brazil and India dominate the number of new dam projects. These country's also lead the creation of smaller (10 km^2 ≤ area ≤ 99 km^2) reservoirs, with both Turkey and Vietnam additionally contributing a large proportion. While the lake databases allow the storing of multiple water uses, 1.GREALD reveals that the "res" *Type* records are dominated by requirements for HEP (65%), irrigation (30%), and basic water supply (15%), noting that the operational altimetric instruments overpass more than 90% of the reservoirs currently providing HEP, an important point noting the energy supply interests of NGA (global) and USDA/FAS (for Africa).

The question has also arisen as to whether radar altimetry can also serve the applied sciences in consideration of natural hazards and flooding. For example, considering the "open" *Type* classification, the altimetric missions with a monthly (27–35 days) repeat frequency (Table 1) do pass over many lakes at very high latitudes (e.g., over 1700 at ≥50° latitude) and will therefore monitor these lakes throughout the spring to fall (i.e., ice-off) seasons. While spatial coverage is good, the monthly resolution is poor compared to the daily observations of (say) Landsat imagery to precisely capture ice-jam release events. At lower latitudes, both the 10-day and monthly repeat altimeter instruments are excellent at capturing the seasonal and inter-annual hydrological variability although, again, any fast-changing variability such as a flood event may not be observed.

The lack of a satellite overpass for all water body types is more severe at low latitudes where the inter-track spacing between the satellite ground tracks is the widest. This calls for consideration of surface area to be a primary monitoring parameter in both the science and applied science programs. It also highlights the need for swath-based altimeters (such as SWOT) in the future.

Summary and future efforts

The upgrading and creation of water body records has resulted in three new databases, 1.GREALD, 2.GREALD, and 3.GELD. Combined these hold 6,495 records pertaining to open lakes, closed (terminal) lakes, lagoons, reservoirs, and ephemeral lakes. The open and closed lake, lagoon, and reservoir database 1.GREALD, and the ephemeral database 3.GELD, aim to be a complete catalog for those water bodies with surface area ≥100 km^2. The lake and reservoir database 2.GREALD (10–99 km^2), is only a complete catalog for reservoirs. All three databases were formed to assist science and applied science programs that look to climate change or resource monitoring via the use of satellite radar altimetry. This remote sensing technique has an ability to monitor surface water levels. The databases thus contain unique details on altimetric mission contributions to water level monitoring, and they also store general information that might assist with the interpretation of the level variations. Of particular note is that the usual definition of a reservoir is not employed, and here, "reservoir" (or the databases "res" *Type* field classification), is specifically defined as any water body where anthropogenic influences may affect water levels. The new database records note the presence of an overpass by the current and historical suite of profiling satellite radar altimeters. Each of these instruments has the potential to monitor surface level variations over the mission lifetime and so contribute to the useful monitoring of a hydrological parameter. While the satellite overpass statistics from each mission show a far from global coverage, and repeat visit times are 10-day to monthly, these profiling instruments do meet a number of end user or stakeholder requirements and between the various missions, at least 80% of the 6282 combined 1.GREALD/2.GREALD permanent water bodies (≥10 km^2) have been overflown since the 1990s. Future missions though must look to acquiring full global coverage especially over arid lands where water resources are critical, and there is the need for improved temporal resolution to capture fast hydraulic changes.

The new lake databases include 3280 records with "res" *Type* classification noting that the infill rates of those created over the last three decades can be captured by the long-term altimetric suites. The number of closed lakes held in the databases totals 432 with an additional 213 identified as ephemeral lakes for those seeking to examine natural variability. The new integral lake surface area distribution law still conforms to a power law observed by Birkett and Mason, (1995) but has a revised exponent of −1.016 in consideration of all water bodies in the area range 100–10,000 km^2 i.e., the distribution is scale invariant. The exponent decreases to −0.794 for the combined 1.GREALD and 2.GREALD "res" *Type* records in the 10–10,000 km^2 range.

Validation efforts will continue to refine the 1.GREALD, 2.GREALD and 3.GELD databases, and their contents will be made available at the GWM web site where comments and suggestions will be welcomed. Meanwhile, the project will make every effort to complete the reservoir year and water use (*Res Year* and *WU_1* to *WU_7*) fields for the 2.GREALD database, and to cross check the water body *Type* field label with additional sources. The *SWOT 1 day* and *SWOT 21 day* fields will also be completed in readiness for the SWOT launch in late 2022. Several new record fields are also being considered. One new field could be used to flag water bodies where surface area sources differ markedly leading to further investigation into potential causes (e.g., definition of perimeter, natural seasonality, ongoing in-filling, climate change). A winter freeze field flag and an average freeze duration field would also enhance the general background information, noting that radar altimetry derived water levels may be erroneous during the ice-on periods due to ice penetration effects. While water levels are useful as a stand-alone parameter, there are end user requirements to estimate and monitor the changing water storage (Gao et al., 2012; Zhou et al., 2016; Tortini et al., 2020), and for dammed reservoirs, to highlight when dead or flood water storage or levels are approaching. In this respect there are ongoing efforts to gather information on basic reservoir operating rules and bathymetry. New reference sources, especially on the location of large dams (e.g., Mulligan et al., 2020), small dams (the on-line Global Dam Watch or GDW initiative, see References section for web site address) and the presence of smaller-scale infrastructure on rivers (e.g., Whittemore et al., 2020) will also be investigated.

The current database project acknowledges the potential for many new dams to be completed over the forthcoming decade (e.g., Zarfl et al., 2015, with respect to those being built for HEP) and future tasks will seek the locations and expected surface area and mark satellite overpass coverage. Many new end users are requesting information relating to glacial lakes especially the emergence of new water bodies resulting from ice sheet or glacial melt. In this respect 96 lakes have been added, mainly to the 2.GREALD database, based on information from Shugar et al., (2020). Two thirds of these high-latitude lakes have altimetric overpasses from currently operating instruments. With refinements to radar and lidar altimetric techniques and instruments, the potential for including other lakes and reservoirs <10 km^2 is beginning to come to the forefront. Also, additional permanent or temporary water holdings such as melt ponds, storage holdings (small water tanks or enclosures), and even seasonally irrigated fields have been requested, providing the technology can meet the water level acquisition requirements.

While the Sentinel-6 Michael Freilich mission will take over the operational 10-day repeat monitoring of Jason-3 in November 2021, it is the launch of the SWOT mission in 2022 that will mark the beginning of swath based interferometric altimetry, i.e., a move away from profiling altimeters which lack a full global coverage. SWOT will in addition simultaneously measure both surface water level and area (Biancamaria et al., 2016). With a 21-day repeat frequency or better at high-latitudes, SWOT will satisfy most current stakeholder requirements though continuation of altimetric swath missions into the future is currently not assured, unlike the profiling altimeter suites which will offer continuity to at least 2030. Meanwhile, other missions are being proposed to improve on the current temporal resolution limit (10-day) of the profiling altimeters. For example, a constellation of ten profiling radar altimeters, the SMall Altimetry Satellite for Hydrology (SMASH) mission (Blumstein et al., 2019) has been proposed with inland water and estuary objectives as primary concerns and with an ultimate objective of providing 1-day sampling.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to particularly acknowledge Bernard Lehner for access to the GLWD and GRanD, lake and reservoir databases, and to acknowledge the Google Earth, Wikipedia, and Mapcarta on-line applications and tools. Thanks also goes to AVISO for provision of the satellite mission reference ground tracks, and to the three anonymous reviewers for comments on the manuscript. This research was supported by the NASA Applied Sciences/Water Resources and the NASA Ocean Surface Topography programs

References

Biancamaria, S., Lettenmaier, D.P., Pavelsky, T.M., 2016. The SWOT mission and its capabilities for Land Hydrology, Surveys in Geophysics. Springer Verlag (Germany) 37 (2), 307–337.
Birkett, C.M., Crétaux, J-F., 2011. 'Lake Levels', Chapter 2. Global Climate, Section 3 Hydrological cycle. In: Blunden, J., Arndt, D.S., Baringer, M.O. (Eds.), State of the Climate in 2010. 6th., 92 Bulletin of the American Met. Soc., pp. 53–54.
Birkett, C.M., Mason, I.M., 1995. A new Global Lakes Database for a remote sensing programme studying climatically sensitive large lakes. J. Great Lakes Research 21 (3), 307–318.
Birkett, C., Murtugudde, R., Allan, T., 1999. Indian Ocean climate event brings floods to East Africa's lakes and the Sudd Marsh. GRL 26 (8), 1031–1034.
Birkett, C.M., Reynolds, C., Beckley, B., Doorn, B., 2010. From Research to Operations: The USDA Global Reservoir and Lake Monitor, Chapter 2 In Coastal Altimetry, Springer Publications. 978-3-642-12795-3.
Blumstein, D., Guérin A., Lamy, A., Couhert, A., Piquereau, A., Palacin, B., Rouzies, F., Mallet, A., Biancamaria, S., le Gac, S., Amiot, T, Maisongrande, P., Cherchali, S., Coutin-Faye, S., 2019. SMASH: A constellation of Small Altimetry Satellites dedicated to Hydrology, Proc. 6th Workshop on Advanced RF Sensors and Remote Sensing Instruments, ARSI'19 and 4th Ka-band Earth Observation Radar Missions Workshop, KEO'19, 11-13 November, ESA/ESTEC, Noorwijk, The Netherlands, pp.5.
Bryant, R.G., 2003. Monitoring hydrological controls on dust emissions: preliminary observations from Etosha Pan, Namibia. Geograph. J. 169 (2), 131–141.
Downing, J.A., Prairie, Y.T., Cole, J.J., Duarte, C.M., Tranvik, L.J., Striegl, R.G., McDowell, W.H., Kortelainen, P., Caraco, N.F., Melack, J.M., Middelburg, J.J., 2006. The global abundance and size distribution of lakes, ponds, and impoundments. Limnol. Oceanogr. 51 (5), 2388–2397. https://doi.org/10.4319/lo.2006.51.5.2388.
Fu, L.-L., Alsdorf, A., Morrow, R., Rodriguez, E., Mognard, N., 2012. SWOT: The Surface Water and Oceanography Mission, Wide-swath Altimetric Measurement of Water Elevation on Earth, NASA JPL. Publication, 12–105.
Gao, H., Birkett, C., Lettenmaier, D.P., 2012. Global monitoring of large reservoir storage from satellite remote sensing. Water Resour. Res. 48 (9). https://doi.org/10.1029/2012WR012063.
Ginoux, P., Prospero, J.M., Gill, T.E., Hsu, N.C., Zhao, M., 2012. Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue Aerosol products. Rev. Geophys. 50, RG3005.
Halbfass, W., 1922. Die seen der Erde, Erganzungsheft, ed. Petermanns Mitteilungen, Gotha: Justus Perthes, No.185.
ILEC, International Lake Environment Committee 1988-1993, 1993. Survey of the State of the World's Lakes: Data Books of the World Lake Environments, Volumes 1-5, ILEC/UNEP publications, Otsu, Japan.
Khandelwal, A., Karpatne, A., Marlier, M., Kim, J., Lettenmaier, D., Kumar, V., 2017. An approach for global monitoring of surface water extent variations using MODIS data, Rem. Sens. Environ. 202, 113–128.
Lehner, B., Döll, P., 2004. Development and validation of a global database of lakes, reservoirs and wetlands. J. Hydrol. 296 (1-4), 1–22.
Lehner, B., Liermann, C.R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rödel, R., Sindorf, Wisser, D., 2011. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Front. Ecol. Environ. 9 (9), 494–502.
Lehner, B., Reidy Liermann C., Revenga C., Vörösmarty C., Fekete B., Crouzet P., Döll P., Endejan M., Frenken K., Magome J., Nilsson C., Robertson J.C., Rodel R., Sindorf N., Wisser, D., 2011b. Global Reservoir and Dam Database, Version 1 (GRanDv1): Dams, Revision 01. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC)
Mason, I.M., Guzkowska, M.A.J., Rapley, C.G., Street-Perrott, F.A., 1994. The response of lake levels and areas to climatic change. Clim. Change 27 (2), 161–197.
McDonald, C.P., Rover, J.A., Stets, E.G., Striegl, R.G., 2012. The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of the global lake extent. Limnol. Oceanogr. 57, 597–606. https://doi.org/10.4319/lo.2012.57.2.0597.
Messager, M.L., Lehner, B., Grill, G., Nedeva, I., Schmitt, O., 2016. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. Nat. Commun. 7 (1). https://doi.org/10.1038/ncomms13603.
Meybeck, M., 1995. Global distribution of lakes. In: Lerman, A., Imboden, D.M., Gat, J.R. (Eds.), Physics and Chemistry of Lakes. Springer, Berlin, pp. 1–36.
Mulligan, M., van Soesbergen, A., Sáenz, L., 2020. GODD, a global dataset of more than 38,000 georeferenced dams. Sci. Data 7 (31).
Prospero, J.M., Ginoux, P., Torres, O., Nicholson, S.E., Gill, T.E., 2002. Environmental characterization of global sources of atmospheric soil dust identified with the Nimbus 7 Total Ozone Mapping Spectrometer (TOMS) absorbing aerosol produc. Rev. Geophys. 40 (1). 2.1 – 2.31.
Rapley, C.G., Guzkowska, M.A.J., Cudlip, W., Mason, I.M., 1987. An exploratory study of inland water and land altimetry using Seasat data, ESA Report No. 6483/85/NL/BI.
Reynolds, C.A., Doorn, B., Birkett, C.M., Beckley, B., 2003. Near-Real Time Monitoring of Lake and Reservoir Surface Elevations. Africa GIS'03 Conference, Dakar, Senegal, November 4-7.
Reynolds, C.A, Doorn, B., Birkett, C., Beckley, B., 2007. Monitoring Reservoir and Lake Water Heights with Satellite Radar Altimeters, Africa GIS 2007 Conference, Ouagadougou, Burkina Faso, September 17-21.
Reynolds, C.A., Birkett, C.M., Beckley, B., 2009. Satellite Radar Altimeters Monitor Lake Water Heights in Africa. Proceedings of the 2009 Africa GIS Conference.
Ryanzhin, S.V., Straökraba, M., 1999. Geographical Information system for the world lakes limnological studies (GIS WORLDLAKE). SIL News 27 (May), 10 pp.
Ryanzhin, S.V., Geller, W., 2006. Review of the MGLD-MSSL Global Lakes Database. Int. J. Lakes & Rivers 1 (2), 50–60.
Seekell, D.A., Pace, M.L., Tranvik, L.J., Verpoorter, C., 2013. A fractal-based approach to lake size-distributions. Geophys. Res. Lett. 40 (3), 517–521. https://doi.org/10.1002/grl.50139.
Sheng, Y., Song, C., Wang, J., Lyons, E.A., Knox, B.R., Cox, J.S., Gao, F., 2016. Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. Remote Sens. Environ. 185, 129–141. https://doi.org/10.1016/j.rse.2015.12.041.
Shugar, D.H., Burr, A., Haritashya, U.K., Kargel, J.S., Watson, C.S., Kennedy, M.C., Bevington, A.R., Betts, R.A., Harrison, S., Strattman, K., 2020. Rapid worldwide growth of glacial lakes since 1990. Nat. Clim. Change 10 (10), 939–945.
Street-Perrott, F.A., Guzkowska, M.A.J., Mason, I.M., Rapley, C.G., 1986. Response of lake levels to climatic change - past, present and future. In, Effects of Changes in Stratospheric Ozone and Global Climate, Volume 3: Climate Change, pp. 211-216. United Nations Environment Programme.
Tortini, R., Noujdina, N., Yeo, S., Ricko, M., Birkett, C., Khandelwal, A., Kumar, V., Marlier, M.E., Lettenmaier, D.P., 2020. Satellite-based remote sensing data set of global surface water storage change from 1992 to 2018. J. Earth System Science Data (Copernicus) 12 (2), 1141–1151.
Vanden Bossche, J.-P., Bernacsek G.M., 1990. Source book for the inland fishery resources of Africa, Vol.1-3, CIFA Technical Papers 18/1 to 18/3, Rome, FAO.
Verpoorter, C., Kutser, T., Seekell, D.A., Tranvik, L.J., 2014. A global inventory of lakes based on high-resolution satellite imagery. Geophys. Res. Lett. https://doi.org/10.1002/2014GL060641, 6396-6402.
Wetzel, R.G., 1990. Land-water interfaces: Metabolic and limno- logical regulators. Verh. Int. Ver. Limnol. 24 (1), 6–24.
Whittemore, A., Ross, M.R.V., Dolan, W., Langhorst, T., Yang, X., Pawar, S., Jorissen, M., Lawton, E., Januchowski-Hartley, S., Pavelsky, T., 2020. A participatory science approach to expanding instream infrastructure inventories. Earth's Future 8 (11). https://doi.org/10.1029/2020EF001558.
Zarfl, C., Lumsdon, A.E., Berlekamp, J., Tydecks, L., Tockner, K., 2015. A global boom in hydropower dam construction. Aquat. Sci. 77 (1), 161–170.
Zhou, T., Nijssen, B., Gao, H., Lettenmaier, D.P., 2016. The contribution of reservoirs to global land surface water storage variations. J. Hydromet. 17 (1), 309–325. https://doi.org/10.1175/JHM-D-15-0002.110.1175/JHM-D-15-0002.s1.