Intercomparison and validation of continental water level products derived from satellite radar altimetry

Martina Ričko  
Charon M. Birkett  
James A. Carton  
Jean-François Crétaux
Intercomparison and validation of continental water level products derived from satellite radar altimetry

Martina Ričko, a Charon M. Birkett, b James A. Carton, a and Jean-François Crétaux c

aUniversity of Maryland, Department of Atmospheric and Oceanic Science, College Park, Maryland 20742
mricko@atmos.umd.edu
bUniversity of Maryland, Earth System Science Interdisciplinary Center, College Park, Maryland 20740
cCNES/LEGOS, 14 av. Edouard Belin, 31400 Toulouse, France

Abstract. Satellite radar altimeter measurements of lake and reservoir water levels complement in situ observations by providing stage information for ungauged basins and by filling data gaps in existing gauge records. Such additional measurements assist both research and operational programs. However, for a particular lake or reservoir, altimetric products offered to end-users may differ due to choice of employed instrument, processing technique, and applied geophysical corrections. To explore these differences, particularly with their potential impact on climate-based research, an intercomparison of three web-based water-level products (produced by Laboratoire d’Etudes en Géophysique et Océanographie Spatiale, National Aeronautics and Space Administration/United States Department of Agriculture, and European Space Agency/De Montfort University) has been undertaken based on 18 lakes and reservoirs. The products are well correlated with each other (r = 0.87 to 0.99) and where in situ data are available are quite well correlated with the gauge measurements (r = 0.73 to 0.99). Despite variations in data processing, the poorest root-mean-square differences between any altimeter product and gauge data (~0.20 to 1.41 m) occur for the narrow reservoirs and smaller lakes. The largest discrepancies between the altimeter products occur for the lakes that freeze (Lake Athabasca and Woods). The current altimeter products provide acceptable accuracy, long-term trends and seasonality for climate applications. We discuss the merits of each product system, but recommend further validations and the provision of ice-detection flags. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.JRS.6.061706]

Keywords: altimetric water level products; validation; lakes; reservoirs; climate.

1 Introduction

Measurement of surface water level for continental water bodies such as inland seas, lakes, and river systems presents multiple challenges, one of which is the need to compensate for the declining quantity of networks of in situ gauges. Recent improvements in satellite radar altimetry offers the possibility of compensation via the provision of near-global coverage of continental water levels with near-real time availability.1-4 As a result, a growing number of users are interested in applying altimetry water level estimates to a wide variety of studies related to climate change, prediction of natural disasters such as floods and droughts, water resource and fishery management, navigation, sediment transport etc. These users include the climate community, who are primarily interested in historic water level estimates; and the geodetic community, who are interested in near-real time water level estimates for current alerts and predictions of water levels. The water management community also has a great interest in these altimeter products.
There are several web-based services that offer altimeter-derived water level products. There are similarities and differences among techniques and products being offered, but all groups offer this service based on published research studies. However, end-users requiring lake and reservoir water-level information can be presented with a selection of product options for a given water body. Although each service offers absolute validation examples (comparison with in situ measurements), these have been limited in geographical scope and to present date, intercomparison between the products themselves (relative validations) has not yet been undertaken.

Here we undertake a more extensive absolute validation exercise and the first relative validation exercise based on continental water-level products from three on-line sources: the Hydrology from Space program run by the Centre National d’Etudes Spaciales (CNES) Laboratoire d’Etudes en Géophysique et Océanographie Spatiale (LEGOS), which offers products for lakes, reservoirs, rivers and wetlands; the “Global Reservoir and Lake Monitor” (GRLM) a combined United States Department of Agriculture’s Foreign Agricultural Service (USDA/FAS) and National Aeronautics and Space Administration (NASA) program; and the “River and Lake” European Space Agency/De Monfort University (ESA-DMU) program. The intention here is not to present a complete description of the technique behind each product type, but to note possible differences that may explain discrepancies, to examine the product validity, and to comment on the overall selection from the viewpoint of an end-user. In this respect, this study also explores the applicability of these products for climate-oriented research programs that utilize the products as a form of climatic index to detect inter-annual variability and long-term trends in water levels. The investigation is based on a reference set of 18 lakes and reservoirs with altimetric data products that span a 19-year observation period (1992 to 2011).

1.1 Satellite Radar Altimetry

Satellite radar altimeters are primarily designed to operate over oceans and ice sheets (for general information, see Fu and Cazenave). As an altimeter satellite orbits the earth, the nadir-viewing instruments continuously emit microwave pulses toward the surface. The altimetric range (the distance between the satellite antenna and surface) can be deduced from the two-way time delay between pulse emission and echo reception. This measurement, together with additional geophysical data and the known satellite orbital position, enables the topography of surface water level to be derived with respect to a single reference datum; for example, a reference ellipsoid. As each satellite is placed in a repeat orbit (accurate to within ±1 km), water level variations can be then constructed for a specific target (e.g., lake or reservoir) during the lifetime of the mission.

The main advantages of satellite radar altimetry are that it can be operational during day and night, with monitoring that is not hindered by clouds, vegetation, or canopy cover. The fact that surface water level heights are given with respect to a common reference datum is especially useful in forming a globally consistent data set. Altimeters can provide water level information for targets where in situ gauges have not been deployed or can provide measurements that contribute to an incomplete or potentially erroneous gauge data set. While radar altimeter height estimates have a number of contributing error sources (e.g., Birkett, Crétaux and Birkett), their advantages enable the ability to monitor monthly, seasonal, interannual and overall record trend variations in water level (over the lifetime of the mission) for a large number of lakes.

The orbit parameters of a satellite altimeter mission determine both the temporal resolution (typically 10 to 35 days) of the water level measurements, as well as the spatial density and location of the ground tracks. The larger lakes (e.g., Lake Malawi and the Great Lakes) are crossed by multiple ground tracks, while the smaller lakes (e.g., Lake Chad) only offer the option of height measurements from a single ground track. The spatial sampling along each ground track (∼350 to 660 m) also varies according to the mission. The accuracy of an individual lake height measurement and of a lake-level product (e.g., a time series of water-level variations) is controlled by a number of factors. One important element is the size of the lake (or the extent of water along the ground track) that determines the number of radar echoes collected and the ability to average measurements along the satellite track. Additional factors include the complexity of the terrain surrounding the lake, the ground-tracking logic, and the algorithms used for processing the echoes, all of which affect how quickly the lake surface is acquired, uniquely identified, and maintained. The surface roughness of the water (wave height), and various
atmospheric and geophysical influences such as water vapor, wind, rate of precipitation, presence of ice, tides, additionally play a role in product accuracy. Knowledge of the water vapor is particularly important and is dependent on the on-board microwave radiometer or global forecast model. A full explanation of the construction of altimetric lake height and the expected accuracy over lakes and reservoirs has been discussed by a number of authors [e.g., for the TOPEX/Poseidon (T/P) Geophysical Data Records (GDR), Birkett; e.g., for the Jason-2/OSTM Interim Geophysical Data Records (IGDR), Birkett and Beckley, and Dumont et al.].

A number of studies have also examined the complexity of the returned altimeter waveforms (radar echoes) over continental water targets and employed more sophisticated echo-shape analysis software (“retracking” algorithms) to improve altimetric height accuracy. While lake level products from the GRLM and LEGOS sites are based on a limited number of simple echo analysis algorithms (assuming waveform shapes from a typical ocean- or ice-based surface), the ESA-DMU products are based on different retracking algorithms that are adapted to many variations in surface type and waveform shape. The choice of retracking algorithm and the selection of only one to cover changing surface conditions may introduce a variable bias that folds into the height accuracy. While this topic is an ongoing research element, we note here that such bias effects may exist.

A great number of studies have used radar altimetry to monitor water level variations of individual lakes, rivers, wetlands, and floodplains, and a number of validation studies have been performed. For the large open North American Great Lakes with wind-roughened surfaces, comparison of T/P satellite radar altimetry with in situ gauge data provides a ~3 to 5 cm root-mean-square (RMS) difference estimate, which we interpret as an estimate of altimeter height error. Crétaux et al. found RMS errors < 10 cm for the largest lakes (Victoria, Superior, and Erie), while a study of Lake Issykkul (of intermediate size) gave an error estimate of 3 cm RMS. Smaller lakes and calm or sheltered waters have larger RMS errors due to fewer radar echoes for height averaging and typically have more narrow-peaked waveforms leading to a poorer range resolution. The latter occurs because the energy distribution within the waveform is spread across fewer range bins (in many cases, 1 to 2 range bins in the waveform window) than is typically seen from a wind-roughened surface, disabling range interpolation across multiple range-bins. In these cases, the RMS estimates have been observed to vary from tens of cm (e.g., Lake Chad, Lake Winnebago and Lake of the Woods, Lake Kariba, Lake Mar de Chiquita, and Lake Titicaca) to over a meter (e.g., narrow reservoirs such as Lake Powell). Dedicated retractors for smaller lakes and narrow reservoirs may improve the RMS error, but the current error over Lake Powell is still an order of magnitude lower than the water level variations of this lake and relevant for time series interpretation.

### 1.2 Study Regions

We focus on a sample of 18 lakes and reservoirs distributed across three continents: six in Africa, 10 in North America, one in South America, and one in Southeast Asia (Fig. 1), for which

![Fig. 1 Locations of selected lakes and reservoirs (stars). The Laurentian Great Lakes consist of five lakes: Erie, Ontario, Michigan, Huron, and Superior.](http://spiedigitallibrary.org/)

altimeter products and in many cases in situ observations are available. The lakes have been selected to sample a range of critical parameters: location, surface size, lake type (lake or reservoir; managed or unmanaged), freezing, and surrounding terrain. Lakes smaller than 100 km² are not included here due to altimetric technique limitations.

Lake Chad and the Volta Reservoir are located in central North Africa. The Volta Reservoir, located within the Ghana basin, is the world’s largest reservoir (by surface area), and is connected to the Atlantic Ocean via the Volta River. Lake Chad is shallow (<7 m deep) and experiences seasonal level fluctuations in its permanent waters and surrounding marsh, expanding in area from ~2000 km² to ~15,000 km² between dry and wet seasons. Most water loss is via evaporation and water extraction, though ~15% is believed lost due to ground seepage. Other African lakes include the smaller Lake Tana, which lies in the Ethiopian Rift Valley, and the Kainji Reservoir, located to the southwest of Lake Chad. Larger and deeper lakes, Lake Tanganyika and Lake Malawi, are also included, both situated along the Rift Valley of eastern and central Africa.

On the North American continent, we include the five Laurentian Great Lakes: Erie, Ontario, Michigan, Huron, and Superior, which are the largest group of freshwater lakes on Earth (containing roughly 22% of the world’s fresh surface water). The smallest by area is Lake Ontario, while the shallowest and smallest by volume is Lake Erie. The largest and deepest is Lake Superior. The Lake Michigan-Huron system, considered hydrologically as a single lake, exhibits the greatest range of level fluctuation of all the Great Lakes.44 Lake Winnebago in the eastern region of the Wisconsin State, Lake of the Woods in Minnesota, and the northern Canadian Lake Athabasca are also included. The two largest reservoirs in the United States—Mead and Powell—are examined as well. While Lake Mead is formed by water impounded by the Hoover Dam, Lake Powell is formed by the Glen Canyon Dam.

In Southeast Asia, we include Lake Tonle Sap in Cambodia. This is the largest lake in Southeast Asia and an important part of the Mekong hydrological system. With the start of the monsoon season in late May, Lake Tonle Sap begins to flood, and the shallow terrain means that its surface area rapidly expands at a rate of ~1000 km² per meter rise in water level.45,46 In South America, we consider Lake Guri, a reservoir whose dam and generator is a major source of hydroelectric power for Venezuela.

Eight lakes and reservoirs are located in the tropics and subtropics and do not experience the effects of an annual freeze and thaw cycle. In contrast, the North American lakes we consider, except Powell and Mead reservoirs, freeze at least partially during the northern winter season. The freezing period for the Great Lakes starts in December and lasts until March, with maximum freezing area during February and March. For some lakes, ice can last through whole northern spring season until May (especially smaller Great Lakes such as Erie and Ontario). Lakes Winnebago, Woods, and Athabasca can start freezing early in late November. The duration of the freezing period varies from year to year.

Most of the lakes and reservoirs are controlled to some extent. However, the impact of management is probably largest for the reservoirs: Guri, Mead, Powell, Volta, and Kainji. Most reservoirs are multipurpose; for irrigation, hydropower, flood control, water supply, navigation, fishing, and recreation.47 All lakes and reservoirs here are freshwater and are permanent.

2 Data

2.1 Web-Based Altimetric Data Products

A number of satellite radar altimetry missions have been used to construct water level products. The suite of NASA and CNES satellite radar altimetry missions consist of TOPEX/Poseidon (T/P) (1992 to 2002), Jason-1 (2002-present), and Jason-2/Ocean Surface Topography Mission (OSTM; 2008-present), all having a 9.92-day exact repeat cycle with a track spacing of 350 km at the equator. These satellites have near-polar orbits spanning ±66-deg latitude. The future Jason-3 satellite is scheduled to follow Jason-2 with an expected launch in 2013. This NASA/CNES suite is complemented by the ESA Earth Remote Sensing Satellite radar altimeters: ERS-1 (1991 to 1996) and ERS-2 (1995 to 2002), and the Environmental Satellite ENVISAT (2002 to 2012). These all have a 35-day exact repeat cycle with 70 km equatorial track spacing.
and an orbit spanning ±82-deg latitude. The Naval Research Laboratory (NRL) GeoSat Follow-On (GFO) Mission satellite (2000 to 2008) is somewhat different from either of these satellite suites in that it had a 17.05-day exact repeat cycle with 170-km equatorial track spacing, and an orbit spanning ±72-deg latitude. In general, most of these satellites require a minimum target area >100 km² or width >500 m to achieve sufficient data quality. The more recent Jason-2/OSTM and ENVISAT missions were specifically developed for monitoring continental water level as well as ocean and ice surfaces. Here we examine water-level products from three web-based sources sponsored or operated by: (i) LEGOS, (ii) NASA/USDA, and (iii) ESA-DMU. In this study, the downloaded products are referred to as the “raw” products. These have been constructed from the altimetric sensor and/or geophysical data records. Table 1 lists general product information. Specific information on mission and data set processing is summarized in Dumont et al. While the utilized data records are consistent across each web-based program, the altimetric products are created using slight variations in methodology.

2.1.1 LEGOS

The LEGOS product is available for 17 of the study lakes and spans the period late-September/early October 1992 until 2011. This product merges data from six satellite radar altimeters (T/P, Jason-1, Jason-2, ERS-2, ENVISAT, and GFO) to form one altimeter product, which is updated once or twice per year. The product has been created using all standard range/height corrections (orbit, ionospheric, wet and dry tropospheric corrections, polar and solid Earth tides, and sea state bias). Smaller lakes and reservoirs (such as Mead and Powell) are processed with 18- to 20-Hz data using the Ice-1 retracker on ENVISAT and Jason-2, while the majority lakes water levels (~80%) are derived with 1-Hz data using Ocean retracker (J.-F. Crétaux, by personal communication). Depending on the lake’s size, the satellite data may be averaged over very long distances (spanning the lake), and for large lakes the data is corrected for the slope of the geoid. Because original reference geoid provided with the altimetry measurements (e.g., EGM96 for T/P data) may not be accurate enough for lake level estimation, the LEGOS water-level products are converted to a geodetic frame based on the new and improved-resolution GRACE gravity model (GGM02C) and then distributed with an approximate monthly time resolution. Where multiple satellites cross over a lake surface (enhancing the product accuracy), each data set is processed independently, then inter-satellite range biases are removed using the T/P mission as a reference.

For each time series element (e.g., a given date and height) in the LEGOS product, there is an estimated height error (an associated standard deviation from the ensemble mean height), but there is no flag to indicate the possible presence of ice. Absolute validation exercises show LEGOS product RMS values; for example, of 3 cm (Lake Issykkul), 4 to 5 cm (Great Lakes: Superior and Erie), and 80 cm (Lake Powell). Additional points to note within the LEGOS system are a few shortened product time spans (commencing in the year 2000 for Bangweulu and Titicaca, ending in the year 2002 for Tonle Sap), possibly coinciding with mission starts/ends. The LEGOS system also covers 160 lakes and reservoirs (and about 1,300 river sites),

| Table 1 Satellite radar altimeter products. |
|-----------------|-----------------|-----------------|-----------------|
| Product          | ESA-DMU         | LEGOS           | GRLM            |
| Radar altimeter products | ERS-2 (personal communication), ENVISAT, Jason-2 | Multi-satellite product (T/P, Jason-1, OSTM/Jason-2, ERS-2, ENVISAT, and GFO) | TOPEX/Poseidon, Jason-1, OSTM/Jason-2 |
| Time resolution/repeat period (days) | 35 (ERS-2, ENVISAT), 10 (Jason-2) | ~30 | 10 |
| Retracking/processing system | Expert system approach (Berry et al. 1997) | Ocean (80%) and Ice-1 (20%) retrackers (Crétaux et al. 2011a) | Ocean and Ice retrackers (Birkett and Beckley 2010) |
classifying target type, and users can visualize the geographic location of the lake and the multi-
satellite ground track locations via Landsat imagery. A future updated version will additionally
provide surface area and volume products via imagery and hypsometry for users interested in
hydrological studies.

2.1.2 GRLM

The GRLM provides both archived and near real time water-level products for 16 of the study
lakes spanning the period late-September/early October 1992 until 2011 for the majority. The
primary product has a 10-day resolution and is a merged product (with applicable inter-mission
height bias) derived from the three NASA/CNES satellite radar altimeters: T/P, Jason-1, and
Jason-2/OSTM), which formed a continuous (satellite follow-on) record of observation from
1992. The Jason-2/OSTM near real-time observations are obtained from the IGDR, available
at a delay of one to three days after the satellite overpass and added to the existing time series
on a weekly basis. The GRLM system was the first to utilize near real-time radar altimetry data
over inland water bodies in such an operational manner.\(^5\) For some lakes, there are 17-day reso-
lution products derived from the NRL/GFO mission, and there are newly created beta-version
35-day resolution products derived from the ESA ENVISAT mission, but these were excluded in
our product comparison. The monitored lakes and reservoirs (currently \(\sim 75\) lake products of
10-day resolution, for lakes \(> 100\) km\(^2\)) are located within major agricultural regions. Because
the USDA/FAS specified relative water level variations as a product requirement, the GRLM
time series are not converted to geoid-based datum. Instead, these products remain in a reference
ellipsoid system, but are given a datum unique to each lake based on a mean water level derived
from the first nine years of T/P altimetry observations.

The GRLM products are available in graphical and text format, both offering a raw and
smoothed (standard IDL median-type filter) output. The latter is for visualization purposes
but helped in the removal of outliers, reducing high-frequency noise, and identifying significant
lower frequency features in the unsmoothed GRLM product. In addition to relative lake height
(see Dumont et al. 2006 for construction notes),\(^4\) these GRLM data products contain an error
estimate (discussed below) and a mean value of the radar backscatter coefficient. Large height
error estimates helped identify radar echo interference from land (coastlines, islands) or dry lake-
bed conditions. The backscatter coefficient serves to highlight particular dates when the lake
surface may be calm or frozen\(^17,18\) (in the latter case, the heights may be erroneous and can
be rejected) and is a somewhat general indicator, though it was not specifically used to remove
outliers or used to draw any conclusions related to freezing related issues in the time series.
However, future use of this coefficient together with knowledge of winter ice conditions, perhaps
utilizing synergistic lake imagery, could help identify problematic data for removal.

Due to an error in the on-board data processing software, many Jason-1 radar echoes were
lost over the surfaces of smaller lakes, so in these cases end-users need to rely on the GFO-based
product as a substitute. This alternative combined product (T/P, GFO, and Jason-2) has proved a
mixed success in that it allows many gaps to be filled, but adds additional errors due to differ-
ences in the location of the ground tracks and instrumentation of the different missions. In its
current form, the GRLM system does not allow for any detailed echo-shape interpretation and
must rely on standard retracking algorithm output provided within the GDR or IGDR. Like the
LEGOS system, this will either be Ocean- or Ice-1 retracker output. Absolute validation of the
GRLM products shows RMS errors \(<10\) cm for the North American Great Lakes, and the
African Lakes Victoria and Tanganyika.\(^4,41\) Smaller lakes or ribbon-like reservoirs with narrow
water extents have RMS accuracies \(\sim 20\) cm (Lake Chad)\(^50\) up to \(\sim 1.6\) m for Lake Powell.\(^4\)

2.1.3 ESA-DMU

The ESA-DMU River and Lake web site offers products for 14 of the study lakes. These products
are based on only two satellite radar altimeters: (1) the historical data from ENVISAT (2002 to
2010) available at approximately monthly resolution (35 days), and (2) the more recent near real
time data from Jason-2/OSTM (2009 to 2011) available at 10-day resolution. The products are
available with delays ranging from three days (e.g., near real time (NRT) Jason-2 product
delivered by subscription in three days) to up to four months (historical ENVISAT product with better orbit and more retrackers in the research expert system delivered upon request) after the satellite overpass. Additional ERS-2 data (1995 to 2003) (courtesy of Richard Smith, De Montfort University) was provided to the authors at 35-day resolution and has been utilized in this study for five of the African lakes (Chad, Kainji, Tana, Tanganyika, and Malawi) and for the Asian Lake Tonle Sap.

Processing of the ESA-DMU altimeter products starts by identifying ENVISAT and Jason-2 tracks that sample the chosen set of river and lake targets. Individual altimeter echoes are analyzed by an "expert system," which examines shape and complexity and selects the most appropriate retracking algorithm. To correctly identify the boundaries of continental water bodies, a box algorithm is used based on pre-defined geographical grids holding the allowable coordinates for a given altimetric pass. After all the standard range and height corrections are applied (e.g., orbit corrections, atmosphere delays of radar pulses, instrument, and surface-related corrections), the altimetry range data is referenced to the EGM96 geoid model to construct sea level orthometric heights. For an easier direct comparison between different targets, the orthometric height products are referenced to a climatological mean, which is computed by averaging over the full time period, thus creating a set of products with time series that are based on an orthometric relative mean.

ESA-DMU products also contain the mean latitude and longitude, the total number of data points, the number of good quality data, and the standard deviations of the samples in the crossing. Variations in the ground track location and the differences in the horizontal extent of the water at different times are given as well.

The ESA-DMU system monitors a large number of targets: 750 targets with 35-day sampling and 57 with 10-day sampling. Users need to register and access the River and Lake web site via a log-in process. It is interesting to note that for the largest lakes, the ESA-DMU product is available on a grid point system, with multiple product options over the large lake depending on the location. A user can therefore choose a record closest to the point of interest (e.g., close to a gauge station for validation) or take the mean of all given targets for a chosen lake or area. For example, the Great Lakes have 15 to 31 grid points to select from, while Lake Tana has only one grid point.

2.2 In Situ Gauge Data

Many in situ gauge networks have been reduced in density over the last few decades owing to geographical, political, or economic constraints, and data is not publically released for many that are still operating. Other limitations include: (1) they are generally located along the lake shore or near river outlets and thus may not be representative of a lake area average; (2) absolute location and sampling rates are sometimes unknown; (3) for the majority of gauges, no estimate of the measurement error is provided; and (4) gauge monitoring is often subject to interruption and occasional relocation. It is generally thought that gauge measurement error ranges from 0.3 cm for continuous-record gauging stations up to ~1 cm. However, as Crétaux et al. points out for Lake Issykkul, technical problems with gauges, data gaps, as well as unrepresentative siting, can easily increase the gauge error level to 4 to 5 cm, making them an unreliable source of ground truth if variations at the few cm level are required. Gauge time series are also generally not available for many lakes and reservoirs and, if available, are likely undocumented, especially their error estimates.

From online sources or personal contact, we have obtained gauge time series for 14 of the 18 case study lakes listed in Table 2. Even though the water levels measured by radar altimetry and in situ gauges are principally different (a weighted mean of reflecting waters over a wider surface area versus local water levels recorded at a specific gauge location), at times they may be equal in the specific case where the satellite track goes close by the gauge location. We present comparisons at crossovers of satellite tracks that are independent of in situ gauge measurements. The distance between the gauge locations and the center position of the altimeter crossings given by the database on average varies between 9 and 169 km (see Table 2). The gauge observations are daily (except for Lake Tana, for which we have monthly data) and each spans a portion or all of the period 1992 to 2011.
For the gauges in most lakes presented in this study, a reference datum is not known, while for some lakes such as Lake Mead and Powell, elevation is given above mean sea level, relative to the average sea level datum. For example, water level data for the North American Great Lakes are given in the International Great Lakes Datum (1985). For climate studies, this fact is not relevant, though it is important to maintain the same gauge datum for the duration of the data time series.

### 2.3 Summary of Product Errors and Differences in Processing

1. We are assuming that gauge errors are $<1\text{ cm}$. We are noting the discrepancy between a gauge measurement (single point) and an altimetric measurement (along track or grid/box averaged), and the fact that there may be considerable distances separating gauge location and satellite ground tracks (see Table 2).
2. The three programs offering altimetric products have their differences with respect to radar echo retracking (e.g., range determination), single or multi-satellite data mergers,

### Table 2 Case study lakes and reservoirs. Columns show gauge station name, source, time period of available data, and distances between the gauge station location and the middle position of the satellite radar altimeter track crossing over or near lake and reservoir (km) for three satellite radar altimeter products: LEGOS (a merged multi-satellite product), GRLM (T/P, Jason1, and Jason-2) and ESA-DMU (ENVISAT).

<table>
<thead>
<tr>
<th>Lake</th>
<th>Gauge station</th>
<th>Source</th>
<th>Time period</th>
<th>Distance to gauge:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LEGOS</td>
</tr>
<tr>
<td>Chad</td>
<td>Bol</td>
<td>Personal contact, F. Delclaux(^a)</td>
<td>1992 to 2008</td>
<td>9</td>
</tr>
<tr>
<td>Volta</td>
<td>Akosombo</td>
<td>Volta River Authority</td>
<td>1999 to 2010</td>
<td>164</td>
</tr>
<tr>
<td>Tana</td>
<td>unknown</td>
<td>Personal contact, J. Ratsey</td>
<td>1992 to 2006</td>
<td>—</td>
</tr>
<tr>
<td>Mead</td>
<td>Hoover Dam</td>
<td>Lake Mead Water Database(^b)</td>
<td>1992 to 2011</td>
<td>67</td>
</tr>
<tr>
<td>Powell</td>
<td>Glen Canyon Dam</td>
<td>Lake Powell Water Database(^d)</td>
<td>1992 to 2011</td>
<td>126</td>
</tr>
<tr>
<td>Erie</td>
<td>Cleveland</td>
<td>NOAA(^c)</td>
<td>1992 to 2011</td>
<td>82</td>
</tr>
<tr>
<td>Ontario</td>
<td>Oswego</td>
<td>NOAA(^c)</td>
<td>1992 to 2011</td>
<td>88</td>
</tr>
<tr>
<td>Michigan</td>
<td>Milwaukee</td>
<td>NOAA(^c)</td>
<td>1992 to 2011</td>
<td>119</td>
</tr>
<tr>
<td>Huron</td>
<td>Harbor Beach</td>
<td>NOAA(^c)</td>
<td>1992 to 2011</td>
<td>169</td>
</tr>
<tr>
<td>Superior</td>
<td>Marquette</td>
<td>NOAA(^c)</td>
<td>1992 to 2011</td>
<td>76</td>
</tr>
<tr>
<td>Winnebago</td>
<td>Oshkosh</td>
<td>USGS(^d)</td>
<td>1992 to 2011</td>
<td>—</td>
</tr>
<tr>
<td>Woods</td>
<td>Warroad</td>
<td>USGS(^d)</td>
<td>1992 to 2011</td>
<td>53</td>
</tr>
<tr>
<td>Athabasca</td>
<td>Crackingstone point</td>
<td>Environment Canada</td>
<td>1992 to 2009</td>
<td>89</td>
</tr>
<tr>
<td>Guri</td>
<td>Central Hidroelectrica Simon Bolivar</td>
<td>OPSIS(^e)</td>
<td>2002 to 2010</td>
<td>50</td>
</tr>
</tbody>
</table>

\(^a\)HydroSciences Montpellier (HSM).
\(^b\)U.S. Bureau of Reclamation.
\(^c\)National Oceanic and Atmospheric Administration (NOAA) Tides and Currents Database.
\(^d\)U.S. Geological Survey (USGS) Database.
\(^e\)Venezuelan electricity grid system operator (OPSIS), Boletin Estadistico Mensual del Sistema Electrico Nacional, 2008.

For the gauges in most lakes presented in this study, a reference datum is not known, while for some lakes such as Lake Mead and Powell, elevation is given above mean sea level, relative to the average sea level datum. For example, water level data for the North American Great Lakes are given in the International Great Lakes Datum (1985). For climate studies, this fact is not relevant, though it is important to maintain the same gauge datum for the duration of the data time series.
low versus high-resolution data rate usage, full along-track averaging versus grid/box averaging, and transformation to geodetic (geoid-based) reference systems versus non-transformation.

3. Expected altimetric product errors (e.g., absolute validation via comparison of a gauge time series of level variations versus an altimetric series) are going to be at best 3 to 5 cm, reducing to 10 or 10 s of centimeters according to lake size and surface roughness.

Gauge validation comparison by Frappart et al.\(^{13}\) has been performed earlier for the ENVISAT Radar Altimeter 2 (RA-2) for lake targets in the Amazon basin, showing RMS differences in the range of 25 to 53 cm depending on the location. A recent extensive validation for targets in the Amazon basin by Berry and Benveniste\(^ {37}\) has shown that ENVISAT consistently has the lowest RMS error among the radar altimeters (0.47 m versus: ERS-2 0.63 m, TOPEX 1.84 m, and Jason-1 1.22 m). It has also been shown that the use of retracking algorithms (especially ICE-1 that performs best for hydrological applications) have allowed refined data selection and, when combined with other corrections these changes, have lowered RMS errors given by ENVISAT over the Amazon basin.\(^ {13}\) More sophisticated ESA-DMU retrackers may offer improvements over ENVISAT. We note here that the choice of retrackers and specific locations will offer variable RMS errors.

3 Results

3.1 Altimetric Product Evaluation

End-users of the altimetric products have access to published materials that quote absolute accuracies of a given product (e.g., a time series of height values) for a set, and limited number, of lake targets. In addition, altimetric products offer error values for an individual date and height measurement. For a given lake or reservoir, important questions are how accurate is an altimetric product, what differences exist between the three products being offered, and do the differences affect climatic or hydrological interpretation of the water level time series in terms of accurately recognizing seasonal and inter-annual variations and long-term trends? Noting the subtleties and differences summarized in Sec. 2.3, we therefore investigate the absolute and relative accuracy of the altimetric products.

3.2 Product Filtering

Here we review the additional processing of the altimeter products and in situ time series. The following method was employed to adjust for differences in data product time spans and datums and data gaps. First, data outliers in the three altimeter products were removed with respect to in situ gauge data as a reference (accepting good altimeter data within one standard deviation of the original gauge data). For the periods when gauge data are missing, outliers in the GRLM time series were filtered out using the smoothed product version of the GRLM time series. For the other two altimeter products (LEGOS and ESA-DMU), these were filtered out with respect to the filtered GRLM product.

The number of outliers varies for each lake and satellite radar altimeter data product. For example, for Lake of the Woods the largest number of outliers that were removed from GRLM is 231, and only 1 from LEGOS and ESA-DMU time series. Due to the varying initial time resolution of the raw altimeter data (between 10 and 35 days), removed outliers need not be at the same dates, unless they are all classed as outliers during approximately the same time period (within five days) to keep the maximum number of good available data.

Overall, the GRLM data product experiences the most outliers and several gaps of missing data, especially during Jason-1 period (2002 to 2008). The GRLM product that consist of three separate satellites (T/P, Jason-1, and Jason-2/OSTM) has been combined into one time series, such that overlapped periods of Jason-1 data were removed. Similarly during overlapped periods with ENVISAT, the ERS-2 data were removed. The majority of outliers tend to occur during winter season for smaller lakes that freeze, such as Lake of the Woods. However, larger lakes, such as Lake Ontario, do not experience this kind of significant seasonal variation in the number of outliers, thus we do not exclude winter season data out in our study. The filtered altimeter time series

---

were then aligned (shifted vertically to overcome variations in height bias) with the gauge time series. For gauge data with large gaps, the filtered altimeter products were aligned with respect to the filtered LEGOS product (generally, the longest and cleanest merged product). After outliers have been removed, the altimeter products were interpolated to a uniform one-day time interval.

To quantify possible interpolation error arising from this method, pairing between filtered gauge and altimeter heights was allowed within a five-days interval. The two compared methods produce small differences when applied to a case study of the five Great Lakes; hence for the rest of this study we use the daily interpolated time series.

3.3 Product Errors

In order to estimate the error in the gauge data itself due, for example, to the representativeness of a single gauge station interpreted as a lake average, water level heights from four gauge stations located along the shores of Lake Ontario (at Eolcott, Cape Vincent, Oswego, and Rochester) have been compared for our base period 1992 to 2009. The average RMS difference among those four gauge stations is 3 cm (varying between 2 and 4 cm). The average RMS difference between gauge and altimetry height for this lake is 6 cm. If we assume the errors are uncorrelated and that the former provides an estimate of the gauge observation error (actually it should be a 40% overestimate if the errors are uncorrelated), then we may conclude that the RMS error in the altimetry itself is \((36 - 9)^{1/2} \approx 5.2\) cm. Thus in this case at least, it is evident that the RMS difference between altimeter and gauge heights is a reasonable estimate of the RMS error in the altimeter height itself.

To evaluate the consistency of error estimates, we compare the error estimates provided by the three altimeter products (LEGOS, GRLM, and ESA-DMU) with the RMS differences between the products and gauges for 14 lakes and reservoirs (Fig. 2). For LEGOS the largest median error occurs for Lake Tonle Sap (0.81 m) followed by reservoirs Powell, Kainji, Guri, Mead, and Volta (0.20 to 0.36 m). For GRLM the largest median error value is observed for Powell Reservoir (0.74 m), followed by Kainji, Guri, Tonle Sap, and Volta (0.19 to 0.29 m); while for ESA-DMU the largest median error value is seen for Kainji Reservoir (0.92 m) followed by Lake Woods, Tana, Tonle Sap, and Athabasca (0.21 to 0.38 m). The smallest median error estimates occur for the Great Lakes in all three products (0.04 to 0.05 m for GRLM, 0.06 to 0.07 m for LEGOS, and 0.08 to 0.11 m for ESA-DMU). We conclude that, although the order of which lakes have the lowest mean/median error varies among the altimeter products, reservoirs and smaller lakes tend to have the largest product internally estimated errors.

The observed RMS differences between the gauge and altimeter water level time series compare well with the median internal errors seen in the altimeter products (see Fig. 2), with the largest discrepancy occurring for Guri Reservoir (0.82 m versus 0.20 m for GRLM; 1.08 m versus 0.01 m for ESA-DMU) and Volta Reservoir (0.54 m versus 0.19 m for GRLM; 0.20 m versus 0.01 m for ESA-DMU). Overall, where available, they tend to be larger than the median error observed in the LEGOS product, but mostly lying in the upper 75th percentile of LEGOS internal errors for Mead, Powell, Guri, and Volta reservoirs. For Lake Chad, the observed RMS difference between the gauge and each altimeter product data is consistently higher than their median error seen in all three altimeter products (0.35 m versus 0.07 m for GRLM; 0.36 m versus 0.20 m for ESA-DMU; 0.30 m versus 0.13 m LEGOS).

3.4 Absolute Validation of Altimetric Products with In Situ Observations

Overall, the comparison of the filtered altimeter products with gauge data over time period when data are available shows good agreement for 14 lakes and reservoirs (Tables 3 and 4). Interestingly, all reservoirs show excellent correlations between the three altimeter product data and gauge data \((r\) between 0.98 and 0.99). All of the Great Lakes have low RMS differences \(\leq 11\) cm, with the smallest difference of only \(~5\) cm for Lake Superior. This result confirms previous studies reviewed above (e.g., 5 to 7 cm, \(\leq 10\) cm, 4 to 5 cm for Superior and Erie).\(^3\)\(^4\)\(^41\) We find the largest RMS differences for reservoirs, especially Lake Powell (1.41 m for LEGOS), Guri (1.08 m for ESA-DMU), and Mead (0.59 m for LEGOS). These RMS differences are also similar to the published estimates (e.g., for Lake Powell: 0.8 m by Crétaux et al.;\(^4\) 1.4 m by Ross and McKellip;\(^41\) \(~6\) m by Birkett et al.;\(^4\) and for Amazon basin: 1.1 m by Birkett et al.).
Figure 3 shows examples of 10 lakes and reservoirs for which gauge data is available, illustrating the agreement between the three altimeter products and gauge time series (for compactness we include only one of the Great Lakes, Lake Ontario, in this figure). At Lake Chad, the gauge data lag all three altimeter products by $\sim 40$ days. This is likely due to the fact that the gauge station at Bol is located in a seasonally inundated marsh region $\sim 50$ km northeast of the permanent waters of the lake. This time lag is somewhat comparable to a previous estimate of a 20 day lag.\textsuperscript{54} In contrast, at Lake Tana gauge data leads the altimeter products by $\sim 20$ days when computed during the interval from late 2002 to 2011. Tana experienced an increase in the amplitude of its seasonal cycle from 16 to 21 cm in 2002 as seen in the gauge observations. This sudden increase in amplitude raises the question of the validity of this gauge data; for example, whether gauge data before and after 2002 are from two different locations, or whether some

Fig. 2 Box plots of lake level error in meters given by: (a) LEGOS, (b) GRLM, and (c) ESA-DMU. Crosses show the mean of the errors for each lake or reservoir. The top and bottom of each box shows the 25th and 75th percentiles of the error, and the line in the middle of the box shows the median (50th percentile). The “whiskers” extend to the farthest outlying errors that are no more than 1.5 times the interquartile range about the median. The open circle symbols beyond the whiskers denote outliers that are farther than 1.5 times the interquartile range from the median. The black box symbols represent the observed RMS difference between the altimeter and gauge time series.
Table 3 Correlation among filtered altimeter and gauge time series of water level. Correlation has been computed over time period when both time series are available (95% confidence intervals are included).

<table>
<thead>
<tr>
<th>Lake</th>
<th>GRLM versus Gauge</th>
<th>LEGOS versus Gauge</th>
<th>ESA-DMU versus Gauge</th>
<th>LEGOS versus GRLM</th>
<th>LEGOS versus ESA-DMU</th>
<th>GRLM versus ESA-DMU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chad</td>
<td>0.90 ± 0.02</td>
<td>0.91 ± 0.02</td>
<td>0.91 ± 0.02</td>
<td>0.94 ± 0.02</td>
<td>0.95 ± 0.02</td>
<td>0.91 ± 0.02</td>
</tr>
<tr>
<td>Kainji</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.95 ± 0.09</td>
<td>0.98 ± 0.08</td>
<td>0.93 ± 0.09</td>
</tr>
<tr>
<td>Volta</td>
<td>0.99 ± 0.10</td>
<td>0.98 ± 0.09</td>
<td>0.99 ± 0.11</td>
<td>0.98 ± 0.07</td>
<td>0.97 ± 0.10</td>
<td>0.99 ± 0.12</td>
</tr>
<tr>
<td>Tana</td>
<td>0.97 ± 0.02</td>
<td>0.97 ± 0.02</td>
<td>0.96 ± 0.02</td>
<td>0.99 ± 0.02</td>
<td>0.97 ± 0.02</td>
<td>0.97 ± 0.02</td>
</tr>
<tr>
<td>Tanganyika</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.99 ± 0.01</td>
<td>0.96 ± 0.01</td>
<td>0.95 ± 0.01</td>
</tr>
<tr>
<td>Malawi</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.99 ± 0.02</td>
<td>0.97 ± 0.02</td>
<td>0.98 ± 0.02</td>
</tr>
<tr>
<td>Mead</td>
<td>—</td>
<td>0.99 ± 0.31</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Powell</td>
<td>0.99 ± 0.31</td>
<td>0.99 ± 0.26</td>
<td>—</td>
<td>0.98 ± 0.30</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Erie</td>
<td>0.97 ± 0.01</td>
<td>0.95 ± 0.01</td>
<td>0.86 ± 0.01</td>
<td>0.95 ± 0.01</td>
<td>0.91 ± 0.01</td>
<td>0.93 ± 0.01</td>
</tr>
<tr>
<td>Ontario</td>
<td>0.98 ± 0.01</td>
<td>0.98 ± 0.01</td>
<td>0.96 ± 0.01</td>
<td>0.96 ± 0.01</td>
<td>0.96 ± 0.01</td>
<td>0.96 ± 0.01</td>
</tr>
<tr>
<td>Michigan</td>
<td>0.98 ± 0.01</td>
<td>0.98 ± 0.01</td>
<td>0.93 ± 0.01</td>
<td>0.97 ± 0.01</td>
<td>0.91 ± 0.01</td>
<td>0.91 ± 0.01</td>
</tr>
<tr>
<td>Huron</td>
<td>0.99 ± 0.01</td>
<td>0.99 ± 0.01</td>
<td>0.93 ± 0.01</td>
<td>0.99 ± 0.01</td>
<td>0.96 ± 0.01</td>
<td>0.93 ± 0.01</td>
</tr>
<tr>
<td>Superior</td>
<td>0.97 ± 0.01</td>
<td>0.97 ± 0.01</td>
<td>0.95 ± 0.01</td>
<td>0.96 ± 0.01</td>
<td>0.97 ± 0.01</td>
<td>0.96 ± 0.01</td>
</tr>
<tr>
<td>Winnebago</td>
<td>0.73 ± 0.01</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Woods</td>
<td>0.86 ± 0.01</td>
<td>0.81 ± 0.01</td>
<td>0.81 ± 0.01</td>
<td>0.89 ± 0.01</td>
<td>0.93 ± 0.02</td>
<td>0.87 ± 0.01</td>
</tr>
<tr>
<td>Athabasca</td>
<td>—</td>
<td>0.91 ± 0.01</td>
<td>0.85 ± 0.02</td>
<td>—</td>
<td>0.92 ± 0.02</td>
<td>—</td>
</tr>
<tr>
<td>Guri</td>
<td>0.99 ± 0.20</td>
<td>0.99 ± 0.21</td>
<td>0.99 ± 0.23</td>
<td>0.99 ± 0.13</td>
<td>0.99 ± 0.21</td>
<td>0.99 ± 0.19</td>
</tr>
<tr>
<td>Tonle Sap</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.99 ± 0.09</td>
<td>0.98 ± 0.10</td>
<td>0.98 ± 0.10</td>
</tr>
</tbody>
</table>

Other change in instrumentation is influencing these differences. The Guri Reservoir, which appears to have a steady seasonal variation of ~6 m, reveals in its time series the effect of active water management (notice the seasonal maximum height is constant with time. Any additional water is spilled. When the water level drops below about ~24 m, the power-generation at Guri would stop, and when water levels approach this value, water conservation is imposed).

The lakes that freeze (Great Lakes, Lake of the Woods, and Lake Athabasca) show higher mean correlations (0.86 to 0.92) and lower RMS differences (13 to 16 cm) among the altimeter products during summer months (June to August), while during winter months (December to February) mean correlations are slightly smaller (0.81 to 0.90) with larger RMS differences (16 to 21 cm). Lake Athabasca shows erroneously large differences between the altimeter products and the gauge data during the seasonal minima of certain years (e.g., 1995, 2001, 2004, 2005, 2007, and 2008), where these differences for LEGOS can range up to 1.69 m in 1995. Similar differences between the altimeter products and the gauge data during seasonal minima are observed in the Lake of the Woods time series during several years as well (maximum of 1 m in 2007 for the ESA-DMU product). These differences are mostly due to the poor performance of radar altimeters during those periods when Lake Athabasca and Lake of the Woods freeze. The information about what the in situ gauge at the lake is measuring and where the gauge is located (for example, how deep under the ice) during a lake’s freezing period is unknown and could also help explain for some of the observed differences.

Even though many data outliers in the GRLM product water level time series for Lake of the Woods have been removed prior to the comparison with the gauge data, and the gauge data for
Table 4  RMS difference between filtered altimeter and gauge time series of water level variations, in meters. RMS difference has been computed over time period when both time series are available (95% confidence intervals are included).

<table>
<thead>
<tr>
<th>Lake</th>
<th>GRLM versus Gauge</th>
<th>LEGOS versus Gauge</th>
<th>ESA-DMU versus Gauge</th>
<th>GRLM versus GRLM</th>
<th>LEGOS versus ESA-DMU</th>
<th>LEGOS versus ESA-DMU</th>
<th>GRLM versus ESA-DMU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chad</td>
<td>0.29 ± 0.01</td>
<td>0.28 ± 0.01</td>
<td>0.29 ± 0.01</td>
<td>0.20 ± 0.01</td>
<td>0.19 ± 0.01</td>
<td>0.24 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Kainji</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.93 ± 0.04</td>
<td>0.73 ± 0.04</td>
<td>1.06 ± 0.05</td>
<td></td>
</tr>
<tr>
<td>Volta</td>
<td>0.54 ± 0.05</td>
<td>0.53 ± 0.04</td>
<td>0.20 ± 0.06</td>
<td>0.56 ± 0.03</td>
<td>0.61 ± 0.05</td>
<td>0.49 ± 0.06</td>
<td></td>
</tr>
<tr>
<td>Tana</td>
<td>0.18 ± 0.01</td>
<td>0.17 ± 0.01</td>
<td>0.21 ± 0.01</td>
<td>0.09 ± 0.01</td>
<td>0.17 ± 0.01</td>
<td>0.19 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Tanganyika</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.05 ± 0.01</td>
<td>0.14 ± 0.01</td>
<td>0.15 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Malawi</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.08 ± 0.01</td>
<td>0.17 ± 0.01</td>
<td>0.16 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Mead</td>
<td>—</td>
<td>0.59 ± 0.16</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Powell</td>
<td>0.82 ± 0.16</td>
<td>1.41 ± 0.13</td>
<td>—</td>
<td>1.41 ± 0.15</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Erie</td>
<td>0.06 ± 0</td>
<td>0.10 ± 0</td>
<td>0.10 ± 0</td>
<td>0.10 ± 0</td>
<td>0.10 ± 0.01</td>
<td>0.07 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td>0.06 ± 0</td>
<td>0.06 ± 0</td>
<td>0.07 ± 0</td>
<td>0.08 ± 0</td>
<td>0.07 ± 0.01</td>
<td>0.07 ± 0</td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td>0.08 ± 0</td>
<td>0.11 ± 0</td>
<td>0.07 ± 0.01</td>
<td>0.11 ± 0</td>
<td>0.08 ± 0.01</td>
<td>0.08 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Huron</td>
<td>0.06 ± 0</td>
<td>0.08 ± 0</td>
<td>0.07 ± 0.01</td>
<td>0.08 ± 0</td>
<td>0.05 ± 0.01</td>
<td>0.06 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Superior</td>
<td>0.05 ± 0</td>
<td>0.06 ± 0</td>
<td>0.05 ± 0</td>
<td>0.07 ± 0</td>
<td>0.04 ± 0</td>
<td>0.05 ± 0</td>
<td></td>
</tr>
<tr>
<td>Winnebago</td>
<td>0.11 ± 0.01</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Woods</td>
<td>0.19 ± 0</td>
<td>0.27 ± 0.01</td>
<td>0.24 ± 0.01</td>
<td>0.20 ± 0.01</td>
<td>0.16 ± 0.01</td>
<td>0.21 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>Athabasca</td>
<td>—</td>
<td>0.28 ± 0.01</td>
<td>0.28 ± 0.01</td>
<td>—</td>
<td>0.22 ± 0.01</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Guri</td>
<td>0.82 ± 0.10</td>
<td>0.87 ± 0.11</td>
<td>1.08 ± 0.12</td>
<td>0.63 ± 0.06</td>
<td>0.76 ± 0.11</td>
<td>0.66 ± 0.10</td>
<td></td>
</tr>
<tr>
<td>Tonle Sap</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.44 ± 0.05</td>
<td>0.58 ± 0.05</td>
<td>0.68 ± 0.05</td>
<td></td>
</tr>
</tbody>
</table>

Lake of the Woods experiences a three-year gap during 1994 to 1997, those two data sets correlate better during the shorter overlapped time period (1992 to 2002, \( r = 0.90 \)) than during the full time period (1992 to 2011, \( r = 0.86 \)). Only the GRLM (T/P) data complements the missing gauge observations during that time, as both the LEGOS and ESA-DMU products for Lake of the Woods begin from the end of 2002. RMS differences between the altimeter products and the gauge data range from 19 to 27 cm, comparable to Ross and McKellip’s value of 26 cm.\(^{41}\) Many data outliers in the GRLM product water level time series for Lake Winnebago also had to be removed prior to comparison with the gauge data. GRLM shows many data gaps, and the other two altimeter products cover this lake reflecting a lack of data from the ESA-DMU satellites, thus it is not surprising that the GRLM product for Lake Winnebago shows the lowest correlation with gauge time series (\( r = 0.73 \)). The RMS difference between the GRLM and gauge time series (11 cm) is significantly lower than Ross and McKellip’s value of 27 cm,\(^{31}\) probably due to the greater length of our data product and our removal of outliers.

Because the time range of the ESA-DMU data product is shorter than the other two products, comparisons among the three altimeter products are restricted to 2002 to 2010 when all three altimeter data are available. The comparison is limited to 10 lakes and reservoirs (Chad, Volta, Tana, Guri, Woods, and five Great Lakes) for which most products and gauge data are available (Fig. 4). We find that LEGOS has insignificantly higher median correlation between the gauge and altimetry data (0.97) than GRLM and ESA-DMU (0.96 and 0.94), with an insignificantly smaller spread of all correlation values (0.93 to 0.97) versus GRLM (0.93 to 0.98) and ESA-DMU (0.92 to...
0.98). The slightly larger spread of RMS differences (6 to 34 cm) is observed between the GRLM and gauge data, while LEGOS and ESA-DMU follow closely (7 to 28 cm and 7 to 23 cm).

We acknowledge that our sample of 10 lakes is small, and their choice likely has an impact on the results. However, the altimeter products (LEGOS, GRLM, and ESA-DMU) are in excellent agreement with the gauge observations (median correlations >0.95), and the median correlations among the altimeter products are even slightly higher (0.97). No single product has the best accuracy overall; product accuracies vary for a given lake and during any period of time.

3.5 Relative Validation of Altimetric Products

For our set of lakes (Chad, Tana, Kainji, Tanganyika, Malawi, and Tonle Sap), for which both the ENVISAT and ERS-2 data are available, a validation of the two filtered satellite data in the ESA-DMU product confirms that ENVISAT has up to a 33% lower error, consistent
with previous studies.\textsuperscript{37,55} For Lake Tana, for example, we find RMS differences with respect to the gauge heights of 9 cm for ENVISAT, but 24 cm for ERS-2 (although we need to bear in mind that the two instruments were not contemporaneous and conditions may have changed).

For 14 lakes and reservoirs (Tables 3 and 4), we observe high correlations among the three filtered altimeter products over time period when data are available ($r$ between 0.87 and 0.99). The largest RMS differences are consistently observed for reservoirs (e.g., Lake Powell 1.41 m between LEGOS/GRLM, Kainji 1.06 m between GRLM/ESA-DMU, and Guri 0.76 m between LEGOS/ESA-DMU).

A more complete comparison among the three filtered altimeter products for a subset of 10 lakes and reservoirs is restricted to 2002 to 2010 when all three altimeter data are available (Fig. 4). The median correlation between any two altimeter products show excellent agreement ($r = 0.95$ to 0.97). Slightly larger spread of correlation values (0.92 to 0.98) is seen between GRLM and ESA-DMU than between LEGOS/ESA-DMU (0.94 to 0.99) and LEGOS/GRLM (0.93 to 0.97). Somewhat larger RMS differences (7 to 25 cm) observed between LEGOS and GRLM indicate that these two altimeter products differ more than any other two (LEGOS/ESA-DMU 7 to 15 cm and GRLM/ESA-DMU 7 to 23 cm). Lower correlations are evident [outliers, Fig. 4(a)] for Lake of the Woods and Lake Chad, while the largest RMS differences [outliers, Fig. 4(b)] occur for the reservoirs Guri and Volta.

The comparisons above have concentrated on relative lake level, but the question arises: as datum requirements of the various end-users vary, which datum should be provided? Here, we...
consider a special case, Lake Ontario, which has accurate altimeter level estimates (RMS altimetry minus gauge difference is 6 to 7 cm). Figure 5 compares the orthometric heights directly. Since GRLM is not based on an orthometric system and uses a different reference system, an approximation is made via calculating the mean reference ellipsoid height (38.08 m) across the reference ground track and subtracting the mean geoid height (−35.95 m using the Earth Gravitational Model 1996) at this geographical location (courtesy of Brian Beckley, personal communication, 2011). The three altimeter products are offset from the International Great Lakes Datum: for 0.11 m for ESA-DMU (ENVISAT), 0.61 m for LEGOS (a merged multi-satellite product), and 0.81 m for GRLM (T/P, Jason-1, OSTM/Jason-2), respectively. For climate studies, this issue of the choice of datum does not have an impact on the types of studies that interest most users. For other users the choice may be important.

### 3.6 Identifying Seasonal Cycles and Trends

Lake and reservoir water levels generally have distinct seasonal cycles due to variations in net water flux, thermal expansion, and ice formation (Fig. 3). Because they are a periodic feature of time series records, the seasonal amplitude and phase is a readily identifiable feature even if time series from multiple sources do not completely overlap in time. Unexplained long-term changes or trends in the observed seasonal cycle, or systematic differences in seasonality between altimeter products and gauge data, may suggest a problem with the instrument measurements or observing strategy, while seasonal trends that appear in both data are clearly a valid phenomena.

For most lakes and reservoirs, the seasonal cycle, as depicted in both altimetric product and gauge data set, is well represented with good agreement between the three altimeter products and the in situ measurements. The two exceptions are lakes Powell and Guri, where for both reservoirs, the seasonal cycles have greater amplitude in the gauge time series than in the altimeter products (Table 5). This may be reflecting the large spatial displacement of gauge site and altimeter crossing location, or highlighting the failure of the altimeters to monitor seasonal lows in water level. In contrast, lakes Woods and Athabasca have lower amplitudes in the gauge time series than in any of the three altimeter product time series due to erroneous altimeter measurements during winter freezing.

Some lakes experience seasonal cycle variations in their water level records. Lake Tana’s seasonal amplitude has increased with time from 1.31 m estimated during 1992 to 2002 to 1.50 m in later years (2002 to 2006), an observation that is apparent in both the altimetric products and the gauge time series. For Lake Chad, the GRLM and LEGOS products confirm the earlier observation of Ricko et al. that the seasonal amplitude has decreased. A decrease from an average of 1.33 m estimated during 1992 to 2002 to an average of 0.55 m during 2002 to 2008 is seen in the GRLM product time series, while LEGOS shows a decrease on average from

---

**Fig. 5** Absolute water level orthometric height (m) for Lake Ontario during 1992 to 2011. Time series include gauge observations from Oswego station (red) and the three altimeter products: LEGOS (a merged product; black), GRLM (T/P, Jason-1, OSTM/Jason-2; blue), and ESA-DMU (ENVISAT; green). Dashed lines are linear trends computed from the filtered time series. The height offset with respect to gauge time series is 0.61 m for LEGOS, 0.81 m for GRLM, and 0.11 m for ESA-DMU.
1.42 m prior to 2002 to 1.12 m post 2002 during the same time periods. Interestingly, the ESA-DMU product reveals no such change in amplitude in the early 2000s. These differences in the mean amplitude are mostly due to different locations (lake versus marsh area) where the measurements of water level were taken. The marsh area can experience much larger seasonal cycle amplitude variations.\textsuperscript{18} Gauge data, in contrast, shows a slight increase in amplitude from 1.17 to 1.21 m during the same time periods.

Both \textit{in situ} and altimetric products show responses to distinct weather and climate events. For example, the Volta Reservoir experienced extreme minima during 1998, 2003 to 2004, and 2007, related to specific drought episodes. Heavy rainfall in 2010 also led to a record high water level and forced the opening of the Volta dam spillway resulting in downstream flooding events. Some lakes and reservoirs are also experiencing long-term trends (over the 19-year altimetric record). These include the reservoirs in the drought-stricken southwestern United States. Lake Powell has undergone a decline of 1.06 m/yr according to gauge records (1.13 m/yr according to the LEGOS product), and Lake Mead has undergone a 3.03 m/yr decline according to gauge records (2.70 m/yr according to the LEGOS) over a period of time when the data are available. For Lake Powell, the trend is somewhat misleading in that it is the result of a dramatic 35 m decrease in water level in the middle years of the record (1999 to 2005). There is a different story for the Volta Reservoir where flooding, mostly since 2007, has caused an increase of nearly 10 m, clearly evident in all the altimeter products. A significant positive trend is also confirmed in the water levels of Lake Malawi, which, due to an increase in rainfall,\textsuperscript{56} shows a 0.11 m/yr rise evident in all three altimeter products. Some of the variations in the trends observed in different altimeter products are mainly due to missing data gaps and possibly due to differences in ground track position.

### 4 Discussion and Conclusions

This study aimed to validate and intercompare altimeter-derived water level products for 18 test case lakes that are currently available from three different systems and to do so with relevance.
to climate-related research studies. The systems provided by LEGOS, NASA/USDA and ESA-DMU, offer products that span almost two decades (19 years) with either a ~10- or ~30-day temporal resolution. The study included a much larger absolute validation exercise that had been previously published and included the first relative validation study, which inter-compared the available system products.

The three product systems differ in methodology. The LEGOS methodology primarily utilizes low-resolution geophysical data and combines multi-temporal resolution data from various missions to form a merged product.\(^3\) The ESA-DMU system utilizes high-resolution sensor data, performs retracking of the radar echoes to achieve an improved height estimate, and merges a given temporal resolution data into a grid system on the lake surface.\(^1,11-13\) The GRLM system utilizes high-resolution geophysical data and keeps the 10- and 35-day resolution products separate. Despite these differences, and using gauge data as absolute truth, all three radar altimeter products perform well for our sample of 18 lakes and reservoirs of varying latitude, size, surface roughness, and surrounding terrain. Any differences resulting from merged versus single data sets, or using an expert retracking system versus using standard retracking options, have not shown to be very significant for the case study lakes.

The absolute validation exercises revealed similar RMS accuracies (when compared to gauge data) to what has been previously published.\(^3,4,41\) \([\text{e.g., } \leq 11 \text{ cm RMS for the North American Great Lakes, } > 1 \text{ m for the very narrow Powell and Mead reservoirs, and the largest errors 19 to 28 cm RMS for the lakes that freeze (Lake Athabasca and Lake of the Woods)}]\). For the latter lakes, the RMS differences between the three products and gauge data are \(\geq 50\%\) of the lake level variability itself, due to erroneous radar altimeter measurements when ice is present. However, even for Lake Athabasca and Lake of the Woods, these error levels are sufficiently low in comparison to seasonal variation to detect climate variability in the two-decade record. Overall, the lakes that freeze \(\text{e.g., Great Lakes, Lake of the Woods, and Lake Athabasca}\) show lower RMS differences (13 to 16 cm) among the altimeter products during summer season versus RMS differences (16 to 21 cm) during winter season. From the relative validation of the three filtered altimetric products, we conclude that the median correlations are comparable (0.95 to 0.97, with smaller spread of values) to those compared with the gauge data (0.94 to 0.97). However, the median values of RMS differences obtained from the relative validation are slightly smaller (0.09 to 0.12 m) than those obtained from the absolute validation (0.13 to 0.15 m).

From a climate analysis perspective, seasonal and interannual variability, and long-term trends must be represented accurately in the altimeter products. The seasonal cycle, which is a distinct and generally stable feature of most lakes and reservoirs, allows us to compare records that may not completely overlap in time. In doing so, we find slight disagreement among the products, mostly due to missing data, and possibly differences in data ground processing and ground track locations \(\text{e.g., Lake Chad}\). Thus we conclude that there is no overall winner in product system. The most accurate product can vary for a given lake and period of time.

Some lakes and reservoirs, such as Lake Mead, exhibited a clear water level trend over the study period, while Lake Powell displayed a multiyear nonlinear variability. These trends provide interesting insights into the changing climates and regional anthropogenic demands. The water level decrease for Mead reflects the persistent drought in the southwestern United States and growing demand for water. Positive trends are also evident for Volta Reservoir \(0.37 \text{ m/yr}\) and Lake Malawi \(0.11 \text{ m/yr}\), observed equally well in the altimetric products and gauge time series, due to an increase in frequency of rainfall events in these regions. The individual trend estimates as revealed in the three altimeter products and gauge time series agree well.

The ESA-DMU system offers multiple water level time series on a grid point system allowing multiple spatial averaging options for a single lake, so users can choose a time series closest to the geographical point of interest. The LEGOS system provides additional information in the form of lake surface extent and volume variation, with a time sequence of surface imagery for some lakes and reservoirs. The GRLM system provides a filtered version of the altimeter product, as well as a radar backscatter coefficient, which could both help end-users to identify poor or erroneous data. The fact that the three altimeter products use different reference datums is not a concern for climate studies.
5 Recommendations and Future Plans

In this study, we intercompared three water level products and validated them with respect to gauge observations. Our results lead us to recommend that end-users obtain all three altimetric products when and where available for cross-check purposes. For climatic interpretation, we encourage the product providers to include the earlier ERS-2 data in all three product systems.

We also encourage the LEGOS, ESA-DMU, and GRLM systems to consider, including ancillary information of interest to end-users such as ice presence, wind conditions, calm waters, storms/floods events, as well as additional information about lake basin parameters (e.g., surface area extent, volume, temperature and salinity, surrounding soil moisture, land cover, precipitation, etc.). Indeed, for lakes at higher latitudes with seasonal ice cover, it is essential that altimeter products provide information on ice related parameters (e.g., ice and snow cover time, ice thickness) to help reject erroneous data. Currently, of the three altimeter products we examine, only GRLM gives some information about the presence of ice obtained from the backscatter coefficient. However, we note that LEGOS is developing a similar data set and plans to apply the methodology to large seasonally frozen lakes in order to provide duration of ice appearance and specific dates of ice events (e.g., the first appearance of ice, the formation of stable ice cover, the first appearance of open water, and the complete disappearance of ice) for each lake or subregion of the lake that freezes.3 LEGOS plans to update its database for these products every year after winter time and give to users both remote sensing data (radar altimetry and SSM/I) and ice related in situ data.

The intercomparison method presented in this paper could be reused to qualify new and improved data sets and quantify their accuracy. Even though some recent studies have already shown considerable improvements in quality from current altimeters due to improvements in data processing,9 there is still additional room for improvement in waveform tracking logic and retracking methods to better identify the signal response of very small lakes in regions of highly varying terrain. The ESA-DMU system, for example, recommends a retracking system that adapts to each waveform shape.1,16 Studies of data filtering options and the estimation of time series error estimates, including further checks on range corrections, and validation of products for smaller lakes and reservoirs are still needed.

In addition, new geoid model or reference datum implementation and range/height corrections are under further development within each system. The ESA-DMU and LEGOS plans to replace the current model EGM96 or GGM02C by a newer version of EGM09 or GOCE derived model solution. GRLM will continue to incorporate ERS and ENVISAT data thus allowing expansion of coverage to ~600 lakes.

In the near future, the databases will integrate new data sets from the forthcoming Jason-3, SARAL (Satellite with ARgos and ALtika) and Sentinel-3 missions. Already in operation Cryosat-2 will also be included in the future processing of all the altimeter data products. Additional improvements will be gained from these next generation of satellite radar altimeters that will utilize enhanced technologies. The Indian Space Research Organization’s (ISRO)/CNES SARAL mission (launch December 2012) and the ESA’s Sentinel-3 (launch 2013) should improve tracking, have smaller footprints, and finer range precision. Satellite laser altimetry (Lidar) [e.g., the multi-beam Ice, Cloud, and Land Elevation Satellite (ICESat-2, launch 2016)],57 could offer water level information at better spatial resolution and accuracies under cloud-free conditions, with additional information on the surface gradient and extent of surface waters. A Ka-Band Radar Interferometer (KaRIn) has also been proposed by NASA/CNES for improved spatial observation of inlands waters. This will be carried onboard the Surface Water and Ocean Topography mission (launch 2019) with improved swath capabilities to cover more lakes.58,59

Both the World Meteorological Organization and the Global Climate Observing System identify lake and reservoir levels as forming a key climate reference data set. In particular, water level changes for a group of 79 lakes, including Chad and several Great Lakes, have been designated Essential Climate Variables (ECVs) as part of the Global Terrestrial Network for Lakes (GTN-L). A data center called HydroLare (www.hydrolare.ru) has been created by the State Hydrological Institute of St. Petersburg, Russia, under the sponsorship of Global Climate Observing System and Global Terrestrial Observing System. The main purpose of HydroLare is to develop web-based delivery of the ECVs for the GTN-L lakes based on a combination of in situ gauges and remote sensing data, principally radar altimetry. Comparison of the gauge
data and the three radar altimeter products given in this study is a step toward developing the regular comparisons required by Hydrolare to fulfill its mission.

Acknowledgments

The authors would like to thank the reviewers and Jérôme Benveniste for their valuable comments and suggestions. We also thank Richard Smith and Philippa Berry from EAPRS Laboratory, De Montfort University, UK, and Jérôme Benveniste from the European Space Agency for supplying the ERS-2 data products. Altimetric data products used in this study was provided by: LEGOS [www.legos.obs-mip.fr/soa/hydrologie/hydroweb]; NASA/USDA [www.pecad.fas.usda.gov/cropexplorer/global_reservoir]; ESA-DMU [tethys.eaprs.cse.dmu.ac.uk/RiverLake/shared/main]. Gauge data was provided by John Ratsey from the Ministry of Water Resources for Lake Tana; Francois Delelaux from HydroSciences Montpellier for Lake Chad; the Volta River Authority [www.vra.com]; the Lake Mead Water Database [lakemead.water-data.com] and the Lake Powell Water Database [lakepowell.water-data.com], the U.S. Bureau of Reclamation and NOAA [tidesandcurrents.noaa.gov] and the U.S. Geological Survey [waterdata.usgs.gov/nwis] for USA lakes; Environment Canada for Lake Athabasca; and the Venezuelan [opsis.org.ve/home3.html] for Lake Guri. Geoid height data for Lake Ontario was provided by Brian Beckley. Support for MR and JAC has been provided by NASA Grant NNX10AO99G, while support for CMB has been provided by NASA Grants NNX08AT886 and NNX08AM72G. This work forms part of the dissertation research of MR.

References


49. B. D. Tapley et al., “GGM02—an improved Earth gravity field model from GRACE,” J. Geod. 79(8), 467–478 (2005), http://dx.doi.org/10.1007/s00190-005-0480-z.


