Manuscript prepared for Hydrol. Earth Syst. Sci.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 25 August 2015

~~Kalman Filter Approach for Estimating Water Level Time Series over Inland Water using Multi-Mission Satellite Altimetry~~DAHITI - An Innovative Approach for Estimating Water Level Time Series over Inland Water using Multi-Mission Satellite Altimetry

Christian Schwatke, Denise Dettmering, Wolfgang Bosch, and Florian Seitz

Deutsches Geodätisches Forschungsinstitut der Technischen Universität München (DGFI-TUM),
Arcisstr. 21, 80333 München, Germany

Correspondence to: Christian Schwatke
(christian.schwatke@tum.de)

[In this version only changes with respect to the reviewer comments are highlighted. This paper was also improved by a professional proofreading service whose changes are not highlighted in this version.]

Abstract. Satellite altimetry has been designed for sea level monitoring over open ocean areas. ~~However, since some years, this technology is also used for observing inland water levels of lakes and rivers.~~However, for some years, this technology has also been used to retrieve water levels from reservoirs, wetlands and in general any inland water body, although the radar altimetry technique has been especially applied to rivers and lakes. In this paper, a new approach for the estimation of inland water level time series is described. It is used for the computation of time series of rivers and lakes available through the web service 'Database for Hydrological Time Series over Inland Water' (DAHITI). ~~The method is based on a Kalman filter approach incorporating multi-mission altimeter observations and their uncertainties. As input data, cross-calibrated altimeter data from Envisat, ERS-2, Jason-1, Jason-2, Topex/Poseidon, and SARAL/AltiKA are used~~The new method is based on an extended outlier rejection and a Kalman filter approach incorporating cross-calibrated multi-mission altimeter data from Envisat, ERS-2, Jason-1, Jason-2, Topex/Poseidon, and SARAL/AltiKa, including their uncertainties. The paper presents water level time series for a variety of lakes and rivers in North and South America featuring different characteristics such as shape, lake extent, river width, and data coverage. A comprehensive validation is performed by comparisons with in situ gauge data and results from external inland altimeter databases. The new approach yields RMS differences with respect to in situ data between 4 cm and ~~38~~36 cm for lakes and ~~12~~8 cm and ~~139~~114 cm for rivers, respectively. For most study cases, more accurate height information than from ~~available other~~other available altimeter data bases can be achieved.

Keywords. Satellite Altimetry; Inland Water; Kalman Filter; DAHITI; Envisat; ERS-2; Topex/Poseidon; Jason-1; Jason-2; SARAL/AltiKa

1 Introduction

Since the 1990s, monitoring and modelling the water cycle of the system Earth have become a very important task (Stakhiv and Stewart, 2010). In particular, the knowledge of regional changes of water storage in rivers and lakes is fundamental for the risk assessment of natural disasters such as the droughts and floods which have been increasing over the last few decades (Guha-Sapir and Vos, 2011). Despite the growing importance of measurements, the number of in situ stations monitoring river discharge is globally declining. The number of river discharge time series provided by the Global Runoff Data Center (GRDC) decreased from about 7,300 to 1,000 stations between 1978 and 2013 (Global Runoff Data Center, 2013). In order to make a statement about the development of water level gauging stations an equivalent database such as the GRDC is required. In general, in situ water level data are managed by federal institutions which make data access very difficult. However, many remote sensing satellites have been launched in the last few years measuring parameters relevant for the investigation of the water cycle, e.g. precipitation, water level, and gravity.

~~One of~~Among these remote sensing techniques is satellite altimetry. Besides its main design goal of measuring ~~water level heights~~water levels in the ocean, satellite altimetry can also be used for deriving ~~water level heights~~water levels of inland water bodies, i.e. lakes, reservoirs, rivers, and wetlands (e.g. Birkett 1995, Crétaux and Birkett 2006 and Crétaux et al. 2011). The advantage of satellite altimetry is its global availability, which allows for estimation of water level time series even in remote areas without local infrastructure. Satellite altimetry can provide water level time series longer ~~then~~than two decades.

~~However, due to its measurement geometry providing measurements along separate ground tracks with distances between about 80 km (ERS-2, Envisat, SARAL/AltiKa) and 300 km (Topex/Poseidon, Jason-1, Jason-2) at the equator not all water bodies can be captured.~~However, because its measurement geometry provides observations along specific ground tracks touching water bodies is by chance. Thereby, big water bodies have a higher probability to be passed than smaller ones. In addition, because of a repeat orbit configuration the temporal resolution is limited to 35 (~~Envisat~~ERS-2, Envisat, SARAL/AltiKa) or 10 (~~Topex/Jason~~Topex/Poseidon, Jason-1, Jason-2) days when only single altimeter missions are used. Thus, the combination of different altimeter systems plays a key role in increasing the temporal and spatial resolution as well as the length of the time series. Satellite altimetry has to cope with different problems over inland water which are mainly caused by the large pulse-limited footprint of radar altimeters. For altimeter missions using Ku-band such as Envisat, the resulting footprint varies between 2 km over the ocean and up to 16 km over the land (Chelton et al., 2001). Even for SARAL/AltiKa, measuring in Ka-band, the footprint size ~~still is~~is still about 8 km (Schwatke et al., 2015). The majority of problems in the field of inland satellite altimetry are the result of land contamination. This effect is ~~twofold~~two-fold: on the one hand the contamination of the radar echo leads to degraded range quality or even to unusable data sets, on the other hand so-called ‘hooking’ or ‘off-nadir’ effects occur. The second effect arises from off-nadir radar returns when the satellite is still/already over land but receives the main reflection from the off-nadir water areas. This leads to longer ranges visible in a parabolic shape of the resulting height sequence. This effect can be ~~easily~~ corrected by fitting curves ~~on~~to the resulting ~~water level heights~~water levels (da Silva et al. 2010, Maillard et al. 2015). For each land-water transition a parabola can be fitted to the measurements that can be used to correct the off-nadir effect. In this paper, the off-nadir data are discarded since for all targets enough reliable nadir-measurements are available. The first effect is more challenging. The contamination of radar measurements by land causes a degeneration of ocean-like waveform shapes (Brown, 1977). ~~The affected waveforms are more peaky and reliable heights cannot be derived using ocean waveform retrackers~~The affected waveforms do not have typical brown-like shapes and cannot be retracked by using ocean waveform retrackers (MLE (Challenor and Srokosz, 1989), NASA β (Martin et al., 1983), etc). Therefore, additional retracking ~~has to~~can be applied with ~~robust~~ retracking algorithms such as OCOG (Wingham et al., 1986), Improved Threshold (Hwang et al., 2006), etc. ~~in order to achieve reliable heights.~~which are more robust with respect to the geometry of the waveforms and can achieve reliable heights. The choice of retracker depends on the quality of existing altimeter measurements which varies between investigated inland water bodies because of their extent, shape or ambient topography.

~~Despite of the aforementioned challenges, satellite altimetry has been successfully used for the estimation of water levels of lakes and rivers by different groups during the last years.~~ The potential of satellite altimetry for the estimation of water level time series and for understanding the terrestrial water cycle was shown by Birkett (1995), Crétaux and Birkett (2006) and Crétaux et al. (2011). In most studies, only single satellite tracks were used for the computation of water level time series. The most popular study areas were the Great Lakes (Ponchaut and Cazenave (1998) used Topex/Poseidon) and the Amazon basin. For the latter, investigations were based on different missions: e.g. Topex/Poseidon (de Oliveira Campos et al., 2001; Zakharova et al., 2006), Topex/Jason-1/Jason-2 (Seyler et al., 2013) and ERS-2/Envisat (da Silva et al., 2010). In addition to these individual investigations four global databases have been developed that provide the international community water level time series over inland. The different processing strategies of these four databases are described as follows.

The *Hydroweb database*[c17] was developed by the Laboratoire d’Etudes en Géophysique et Océanographie Spatiales (LEGOS). For the estimation of water level time series over lakes and rivers, a multi-mission approach using satellite altimeter data of Topex/Poseidon, ERS-1, ERS-2, Envisat, Jason-1, and GFO is applied. The physical heights are estimated in a track-wise manner and are corrected by the slope of the geoid or mean lake level and by range biases with respect to Topex/Poseidon. The final time series are computed by merging the altimeter data on a monthly basis. The applied approaches ~~used~~ are published in Crétaux et al. (2011) and da Silva et al. (2010).

The *River & Lakes database*[c18] was developed by the European Space Agency and the De Montfort University (ESA-DMU). It provides track-wise time series derived from Jason-2 and Envisat for a variety of inland waters. For each track crossing the water body of interest a single time series is processed. The methodology for the estimation uses an ex-

---

[c17] http://www.legos.obs-mip.fr/soa/hydrologie/hydroweb/

[c18] http://tethys.eaprs.cse.dmu.ac.uk/RiverLake/shared/main

pert system which is based on neural networks (Berry et al., 1997).

The *Global Reservoir and Lake Monitor (GRLM)*[c0] is maintained by the Foreign Agricultural Service of the United States Department of Agriculture (USDA). ~~Time series~~Water level time series of lakes and reservoirs are estimated by using a segment of one single altimeter track over the investigated target. The time series are composed of data from consecutive altimeter missions measured along the same ground track. A combination of contemporaneous missions is not performed. The method for the estimation of water level time series is described in Birkett et al. (2011).

The *Database for Hydrological Time Series over Inland Water (DAHITI)*[c1] was launched by the Deutsches Geodätisches Forschungsinstitut (DGFI, now DGFI-TUM) in 2013. Currently, DAHITI provides about 250 time series of rivers, lakes, reservoirs, and wetlands. The methodology for the estimation of water level time series in DAHITI is based on an extended outlier rejection and a Kalman filter approach described in detail in the article at hand.

In contrast to the methods already published in the literature, our approach is based on a rigorous combination of a variety of altimeter missions. In addition, extended outlier detection is applied and optional waveform retracking is implemented. Moreover, the processing contains a full error propagation and provides accuracies for each height measurement. ~~This will be discussed in a follow-on paper in which uncertainties of the applied geophysical corrections and models will be taken into account.~~ Furthermore, correlations between altimeter measurements are considered in order to achieve more reliable errors for each water level height. The current paper provides detailed information on the estimation of water level time series and performs a comprehensive validation by comparing the results with in situ gauging data and time series from other databases (~~LEGOS~~Hydroweb, ~~ESA-DMU~~River & Lakes, and GRLM).

The article is structured as follows: In Section 2 the altimeter data that serve as input for the ~~Kalman filter approach as well as the preprocessing of the data~~processing are described. In Section 3 the methodology for the estimation of water level time series from satellite altimeter data using a Kalman filter approach is explained. Section 4 starts with the introduction of the validation areas and data before the resulting water level time series and validation results are presented. ~~The paper finishes with a conclusion.~~The paper concludes with a summary of the results and outlook.

2 Altimeter Data and ~~preprocessing~~Height Estimation

~~For more than two decades, satellite altimetry has been provding data for various applications over ocean and inland waters. The approach presented in this paper combines as many as possible altimeter tracks from different missions over an investigated water body in order to increase the temporal resolution of the final water level time series, to maximize the probability to cover smaller inland waters, and to increase the accuracy.~~

In this paper, altimeter measurements from Topex, Jason-1, Jason-2, ERS-2, Envisat, and SARAL/~~AltiKA~~AltiKa are used depending on the data coverage for the inland water ~~body~~bodies under investigation. In principle, data from Geosat, ERS-1, HY-2A, IceSAT, and Cryosat-2 can be used. However, these missions are neglected in the current investigations for a number of reasons, i.e. lack of data over land, non/long-repeat cycle, bad data quality, or missing waveform information. The applied missions can be separated into two groups according to their orbit characteristics. Topex/Poseidon was launched in 1992 into an orbit with a repeat cycle of 9.9156 days and a track separation at the equator of about 300 km. The mission was followed by its successors, Jason-1 and Jason-2. These three altimeter satellites can be used for estimating continuous time series over more than two decades. The second group starts with ERS-2 (launched in 1995), followed by Envisat and SARAL/~~AltiKA~~AltiKa. The orbit of these missions is defined by a repeat cycle of 35 days and a track separation of about 80 km at the equator. The data are available for almost two decades with a data gap between 10/2010 (end of Envisat core mission) and 03/2013 (launch of SARAL/AltiKa). ~~due to the shift of Envisat to a drifting orbit that lasted until the launch of SARAL/AltiKA~~The data for Envisat on its drifting orbit (10/2010-04/2012) are not used. ERS-1 is not yet ready for use in DAHITI but will be integrated in the near future. This will enable extensions of the time series back to 1991.

For the estimation of ~~water level heights~~water levels, Sensor Geophysical Data Records (SGDR) altimeter products are used which provide ~~1 Hz and~~ high-frequent ranges as well as altimeter waveforms. The altimeter waveforms allow~~ for~~s individual retracking in order to achieve more reliable altimeter ranges, especially for smaller inland water bodies. Table 1 shows a list of the altimeter missions used and provides information about the product, cycle length, frequency, ~~cross~~along-track distance between altimeter measurements on the ground, time period, and mean range bias with respect to Topex.

Depending on the investigated inland water body the original ocean ranges in the SGDR are very often corrupted. Especially over small lakes and rivers the altimeter waveforms do not exhibit the typical ocean-like shape because of land contamination. Land-contaminated altimeter waveforms are usually more peaky and noisy which leads to flat-patched, quasi-specular and complex waveforms (Berry et al., 2005). The quality of the ranges can be improved by retracking these waveforms. In this study, the 'Improved Threshold Retracker' (Hwang et al., 2006) with a threshold of 10% is applied if additional retracking is necessary. In general, all altimeter measurements of smaller lakes and rivers are retracked if the ocean product does not lead to reliable time se-

---

[c0] http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/

[c1] http://dahiti.dgfi.tum.de

Table 1: List of all altimeter missions used in this study together with their main characteristics.

Mission	Product	Cycle length	Data rate	Along-track distance	Time period	Mean range bias
Envisat[1]	SGDR (v2.1)	35 d	18 Hz	~374 m	2002 - 2010	450.8 ± 7.9 mm
Envisat (EM)[1]	SGDR (v2.1)	35 d	18 Hz	~374 m	2010 - 2011	441.2 ± 2.6 mm
ERS-2[1]	SGDR (REAPER)	35 d	18 Hz	~374 m	1995 - 2007	71.2 ± 6.9 mm
Jason-1[2,3]	SGDR-C	9.9156 d	20 Hz	~294 m	2002 - 2009	97.3 ± 1.3 mm
Jason-1 (EM)[2,3]	SGDR-C	9.9156 d	20 Hz	~294 m	2009 - 2012	97.2 ± 2.6 mm
Jason-1 (GM)[2,3]	SGDR-C	9.9156 d	20 Hz	~294 m	2012 - 2013	103.1 ± 1.7 mm
Jason-2[2,3,4]	SGDR-D	9.9156 d	20 Hz	~294 m	2008 - active	-4.7 ± 1.0 mm
Poseidon[2,3]	ALT SDR (L1B)	9.9156 d	10 Hz	~620 m	1992 - 2002	-1.1 ± 7.2 mm
Topex[2,3]	ALT SDR (L1B)	9.9156 d	10 Hz	~620 m	1992 - 2002	-0.2 ± 1.2 mm
Topex (EM)[2,3]	ALT SDR (L1B)	9.9156 d	10 Hz	~620 m	2002 - 2005	-0.0 ± 2.5 mm
SARAL/~~AltiKA~~AltiKa[3,5]	SGDR-T (patch 2)	35 d	40 Hz	~173 m	2013 - active	-67.5 ± 1.7 mm

operated by: [1]ESA, [2]NASA, [3]CNES, [4]EUMETSAT, [5]ISRO

ries because of land contamination. Testing different thresholds for the retracking of altimeter measurements showed that a threshold of 10 % gives slightly better results for smaller lakes and rivers. In our implementation of the ‘Improved Threshold Retracker’ the first sub-waveform is always chosen. Nor do we use a reference height for choosing the sub-waveform such as the last range over ocean as described in (Hwang et al., 2006) since this is difficult in the case of small lakes and rivers. This algorithm is very robust and delivers ranges for all surface types~~. They~~ which are more reliable than the original ranges over small inland waters. However, over open water (i.e. larger lakes) the resulting ranges are less precise than ranges derived from retracking algorithms for ocean applications. It is known that switching retracking algorithms along a single satellite track leads to height offsets (Crétaux et al., 2009). To avoid those offsets, all altimeter measurements of an investigated inland water body are retracked with the same algorithm.

In order to convert the range measurements (original or retracked) to ~~water level heights~~water levels serving as input for our Kalman filter approach numerous preprocessing steps are necessary. Eq. 1 summarizes the height computation from altimeter products (orbit height h_{sat} and (retracked) altimeter range r_{alt}). These processing steps have to be performed for each individual altimeter measurement. The derived normal heights h_{normal} serve as input for the ~~Kalman filter~~DAHITI approach described in Section 3.

$$\begin{aligned} h_{normal} &= h_{sat} - r_{alt} - \\ &\quad \Delta h_{wet} - \Delta h_{dry} - \Delta h_{iono} - \\ &\quad \Delta h_{etide} - \Delta h_{ptide} - \\ &\quad \Delta h_{rad} - N \end{aligned} \tag{1}$$

First, the range has to be corrected for geophysical effects. For this purpose, the models and ~~conventions~~corrections given in Table 2 are applied. It is important to apply identical geophysical corrections for all missions and over the whole time period in order to avoid inconsistencies in the resulting multi-mission time series. To correct the wet (Δh_{wet}) and dry (Δh_{dry}) tropospheric delay, products of ECMWF for Vienna Mapping Function 1 (VMF1) (Boehm et al., 2009) are used. The ionospheric delay Δh_{iono} is corrected by using the NOAA Ionospheric Climatology 2009 (NIC09) (Scharroo and Smith, 2010) model. The solid Earth tide and pole tide corrections (Δh_{etide},Δh_{ptide}) are applied according to the IERS Conventions 2003 (McCarthy and Petit, 2004). Finally, each single altimeter measurement is corrected for its radial error Δh_{rad} in order to account for inter-mission~~ range~~ biases~~ is corrected~~. Radial errors are derived from a global multi-mission crossover analysis as described by Bosch et al. (2014). They are computed with the ocean products. Radial errors were interpolated over land to provide range bias corrections for each altimeter measurement over land. This approach works quite well as long as the ocean product is used for the computation of inland water levels. However, as soon as retracking is involved additional retracker offsets will occur. In order to minimize the relative offsets between different altimeter tracks, we use the same retracker for all measurements over one target. That minimizes the inter-mission biases which are shown later for selected results in Section 4.3 and allow us to use different altimeter missions as a single virtual altimeter system. The average values of the applied range errors are given in Table 1 for each altimeter mission. All data used in this study (the altimeter data as well as all corrections) are extracted from OpenADB[c5], the open altimeter ~~data base~~database of DGFI-TUM. More information on OpenADB is given in Section 3.1. The quality of extracted geophysical corrections is checked and altimeter measurements are rejected if they do not comply ~~with certain thresholds.~~with the valid ranges given in the mission handbooks.

[c5] http://openadb.dgfi.tum.de

Table 2: List of applied models and geophysical corrections

Correction	Source/Model	Reference
Wet troposphere	ECMWF (2.5° x 2.0°) for Vienna Mapping Functions 1 (VMF1)	Boehm et al. (2009)
Dry troposphere	ECMWF (2.5° x 2.0°) for Vienna Mapping Functions 1 (VMF1)	Boehm et al. (2009)
Ionosphere	NOAA Ionosphere Climatology 2009 (NIC09)	Scharroo and Smith (2010)
Solid earth tide	IERS Convention 2003	McCarthy and Petit (2004)
Pole tide	IERS Convention 2003	McCarthy and Petit (2004)
Range Bias	MMXO14	Bosch et al. (2014)
Geoid	EIGEN-6C3stat	Förste et al. (2012)

For the computation of water level time series within the Kalman filter approach normal heights h_{normal} are used as input data whereas altimetry provides ellipsoidal heights. However, ellipsoidal heights are purely geometrical and do not allow us to predict where the water will flow. We compute normal heights by subtracting a (quasi-)geoid model (N) from the ellipsoidal heights. For this purpose, the EIGEN-6c3stat (Förste et al., 2012) model is used which supplements the EGM2008 geoid model with additional GOCE gravity data. ~~The derived water levels are assumed to be constant over lakes since in general, the water is in balance with gravity and hydrodynamics of lakes is small compared to open ocean conditions.~~

3 ~~Kalmen Filter~~DAHITI Approach

In order to use altimeter measurements from different tracks and missions a consistent and reliable combination strategy is important. The irregular spaced observations from different locations must be merged into one time series per target and the optimal combination of measurements with different uncertainties must be ensured. This requirement is fulfilled by our DAHITI approach which is based on an extended outlier rejection and a Kalman filter for the estimation of water level time series. ~~This requirement is fulfilled by a Kalman filter that updates a model by measurement data of different accuracies and predicts the current state to the next time epoch Kalman,1960. In contrast to the common least-squares adjustment the Kalman filter works recursively and the amount of input observations per processing step is significantly reduced due to its sequential integration. This also enables real-time applicability in future.~~

The processing strategy for the estimation of water level time series over inland waters using ~~a Kalman filter~~the DAHITI approach is separated into three steps: preprocessing, Kalman filtering and postprocessing (cf. ~~Fig.~~Figure 1). The preprocessing step includes all necessary tasks for the preparation of the input altimeter heights such as waveform retracking, applying range corrections, calculation of ~~standard deviations (SDs) of~~ height~~s~~ errors, and rejection of outliers. ~~The Kalman filtering step starts with the definition and creation of a hexagonal computation grid covering the inland water body. This is followed by the Kalman filtering itself, estimating water levels for each grid point.~~In the Kalman filtering step, the computation of the water levels of the investigated water body is performed. In this paper, we apply Kalman filtering in a single location centred on the investigated water body and obtain one computed water level for each epoch. However, there is also an option for performing Kalman filtering on a grid which can be used for investigation of the surface variability of larger lakes.

In the postprocessing step, all ~~water level heights~~water levels from the previous step are merged to form a single water level time series referring to one reference location if the Kalman filtering was performed on a grid. Subsequent outlier detection ~~is conducted~~can be conducted if necessary. The final time series is stored in ~~the "Database for Hydrological Time Series of Inland Water" (DAHITI)~~DAHITI, accessible via the website.

3.1 Preprocessing

The Open Altimeter Database (OpenADB) holds satellite altimeter data and derived high-level products. OpenADB provides satellite altimeter data, geophysical corrections, models, etc. which are also accessible via the website. All information is stored in separate parameter files named Multi-Version-Altimetry (MVA) and can be individually combined to achieve water heights. ~~In OpenADB satellite altimeter data are stored in the Multi-Version-Altimetry structure which is designed to allow fast parameter updates and data base extractions with user-defined formats and parameters. This data structure allows for an easy extraction of the required altimeter measurements for an inland water body of interest. Furthermore, the desired geophysical corrections can be selected individually. Users can choose between different geoid models, wet troposphere models, etc. according to their individual purpose.~~ The data sets from OpenADB used for this study and the methodology used to derive individual ~~water level heights~~water levels are described in Section 2.

In addition to the normal heights of the water levels the Kalman filter requires information on the quality of each measurement. This information is used for the weighting of the individual data sets as well as for the error estimation of water level products. Because of ~~lacking~~the lack of ab-

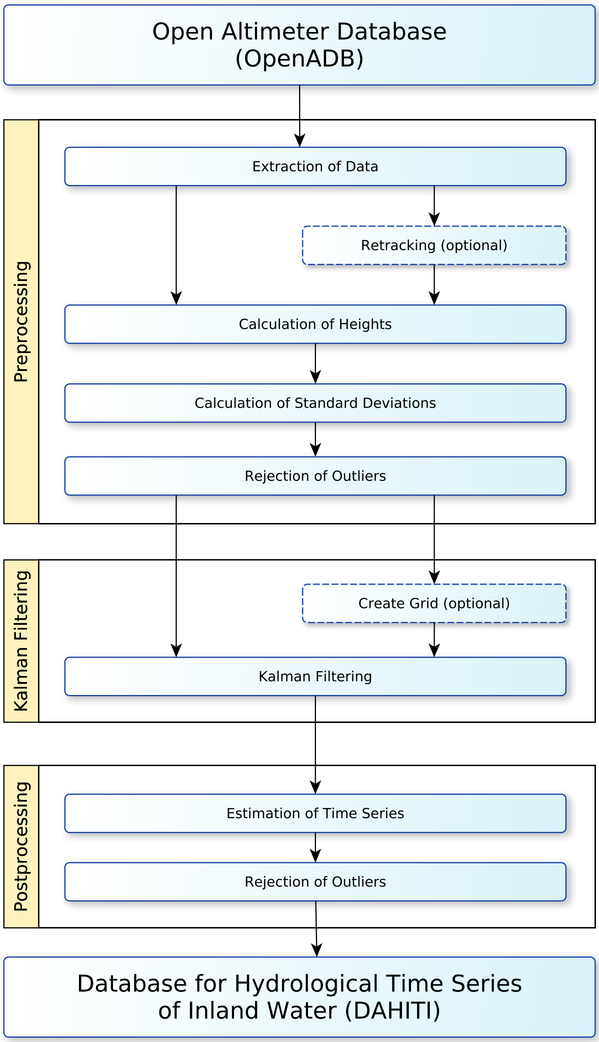


Figure 1: Processing strategy for the computation of water level time series for inland waters in DAHITI in three main steps: preprocessing, Kalman filtering, and postprocessing.

solute accuracy, the precision of the heights is computed by analysing the along-track scatter of the measurements.

~~For this purpose, a standard deviation for each water level height using a floating box of 5 data points along the altimeter track is estimated. Lower standard deviations imply higher accuracies of the water level heights and vice versa. This approach assumes a constant water level along the satellite track and is only valid for lakes and small river crossings without significant slopes. In general, the standard deviations increase when the measurements locations approach the shore.~~For this purpose, an ‘absolute deviation around the median’ (ADM) is estimated by using a sliding box along the altimeter track. The size of the sliding box varies for large lakes (±3.5km), small lakes/large rivers (±1.5km) and smaller rivers (±0.5km). The definition of the sliding box in kilometres instead of number of points allows consistent handling of missions with different data rates (10 Hz, 20 Hz, or 40 Hz) and ensures correct inter-mission weighting. The ADM is calculated by estimating a median of the water heights within the box. Then the median height is subtracted from the current water height and the absolute value of the difference is used as the ‘error’ of the altimeter measurement. Compared with estimated standard deviations, the ADM method is more robust against corrupted water heights and topography near shores and leads to more reliable errors as long as more than half of the altimeter measurements are over water.

~~Each water level height has to pass an outlier test before it is included in the Kalman filter. Different user-defined criteria can be selected for track-wise outlier detection.~~Before Kalman filtering is performed various user-defined outlier rejections can be applied. Inaccurate ~~water level heights~~water levels must be rejected before Kalman filtering; precise ones are used for the estimation of the resulting ~~water level heights~~water levels. The following outlier criteria can be applied in the preprocessing step.

- Latitude thresholds
- Water height thresholds
- Height error (ADM) threshold
- Backscatter coefficient (Sigma0) thresholds
- Along-track outlier test (SVR)

It is important to note that the criteria for the outlier detection are very flexible and the optimal configuration strongly depends on the investigated water body. As a consequence, the parameters for outlier rejection vary with the study areas. First, three outlier criteria (latitude thresholds, water height thresholds and height error threshold) are applied. ~~Different thresholds for water level height, standard deviation, or latitudes can be selected.~~ The backscatter coefficients of altimeter measurements provide information about the reflectance of the surface. This information can be used to reject altimeter measurements affected by ice. Moreover, outlier detection with Support Vector Regression (SVR) (Smola and Schölkopf, 2004) is implemented. This method applies linear regression to each altimeter track to reject altimeter measurements that do not represent the flat water level of the inland water target. SVR is similar to common regression but is more flexible and robust. SVR is an advancement of the Support Vector Machine (SVM) (Boser et al., 1992) which is used as a classification algorithm for applications such as pattern recognition and machine learning. Depending on the mathematical problem, the kernel for the regression varies. One can use linear, polynomial or radial base functions (Smola and Schölkopf, 2004). In our case, SVR is applied on single altimeter tracks over an inland water body using a linear kernel and zero-slope constraint~~ is applied~~. Based on the constant representing the flat water level, an interval is defined which separates into valid and invalid data. ~~Fig.~~Figure 2 shows an example of an altimeter track (Envisat, Pass 80, Cycle 007) crossing ~~a lake~~Lake Erie with an island in the middle. Blue dots indicate valid measurements, red dots indicate rejected data that exceed the ~~standard deviation~~ADM threshold of 5 cm (black dotted line), and green dots mean outliers detected by SVR (with rejection interval of ~~10~~±5 cm). The threshold of the SVR should be in the order of the noise of high-frequent altimeter measurements. One can see that all heights influenced by land contamination are detected as outliers and the remaining heights represent a flat surface.

~~It is important to note that the criteria for the outlier detection are very flexible and the optimal configuration strongly depends on the investigated water body. As a consequence, the parameters for outlier rejection vary with the study areas.~~

3.2 Kalman Filtering

~~The Kalman filtering is the most important step in the computation of water level time series and the heart of DAHITI.~~Kalman filtering is the central part of the computation of water level time series in DAHITI. It updates a model by measurement data of different accuracies and predicts the current state to the next time epoch (Kalman, 1960). In contrast to the common least-squares adjustment the Kalman filter works recursively and the number of input observations per processing step is significantly reduced because of its sequential integration. This also enables real-time applicability.

~~It describes~~The Kalman filter performs the estimation of water level time series from the track-wise input heights ~~. The combination of the~~by combining time-dependent input data available at irregular intervals and –in the case of larger lakes– at different locations ~~is realized by a Kalman filter approach~~. Different modified Kalman filter approaches have been used for geodetic applications (e.g. Yang and Gao 2006, Eicker et al. 2014 and Gruber et al. 2014) In principle, this algorithm realizes a sequential least squares adjustment by taking into account the accuracies of the input data as well as the deterministic and stochastic behaviour of the system and produces a statistically optimal estimate of the water level time series.

3.2.1 Update Interval

The Kalman filter uses input observations to update the current state of the system and predict the model of the following time epoch. This is performed in a continuous loop consisting of two steps (an update and a prediction step) running consecutively for every period of time t_k. ~~In~~At the beginning , an initialization is necessary in order to set the starting conditions. The work flow is illustrated in ~~Fig.~~Figure 3. The time increment of the Kalman filter can be defined arbitrarily. In our case an observation-based update interval instead of a ~~contant~~constant one is used. That means that our system ~~each time is updated if~~is updated each time a new altimeter track is available. Thus, the update interval strongly depends on the size and the data coverage of the investigated water body. It can vary between 35 days (if only an Envisat track crosses the target area) and one day (in the case of large lakes covered by different altimeter missions). ~~Shorter time intervals are precluded~~Time intervals shorter than one day are precluded by assigning the individual measurements to full days. The use of an adaptive update interval avoids smoothing effects in the case of data gaps that may occur when a fixed time increment is selected.

3.2.2 Computation Grid

~~Our approach includes the location of each altimeter observation by performing all computations on a regular grid over the water body. A hexagonal grid is selected in order to ensure equidistant grid nodes. The grid is created automatically by a recursive algorithm using one initial node over water as reference point for each water body. A landwater mask provides information about the extent of the grid. For smaller lakes and rivers, the grid is extended by a transient region in order to take into account uncertainties of the landwater mask and occurring flood events. Figure ?? shows an example of a grid used for the Lake Erie. The resolution of the grid and the number of nodes (n) can be defined individually depending on the extent of the inland water body. For rivers a small grid with a very high spatial resolution is selected in order to avoid errors due to river slopes.~~All computations can be referred to one location (centre of the target) or performed on a computation grid. The latter is optional and can be applied for special investigations on surface variability of larger lakes. The standard solution -also used for all computations within this study- assumes uniform lake surfaces in balance with gravity and merges all water heights of one update step to one location. Surface differences owed to systematic height or geoid errors or hydrodynamic effects from wind and waves are neglected. In practice, our approach automatically creates a grid by means of a recursive algorithm used on an initial grid node as reference point. A land water mask provides information on the extent of the water body and the grid. The grid node separation can be chosen manually depending on the extent of the investigated inland water. Thus, normally we define only one grid node over the target. However, in cases where surface differences are expected, a smaller grid node distance can be chosen. The computations will then be performed for all grid nodes and different water levels for the whole lake surface.

[Image of computation grid was removed]

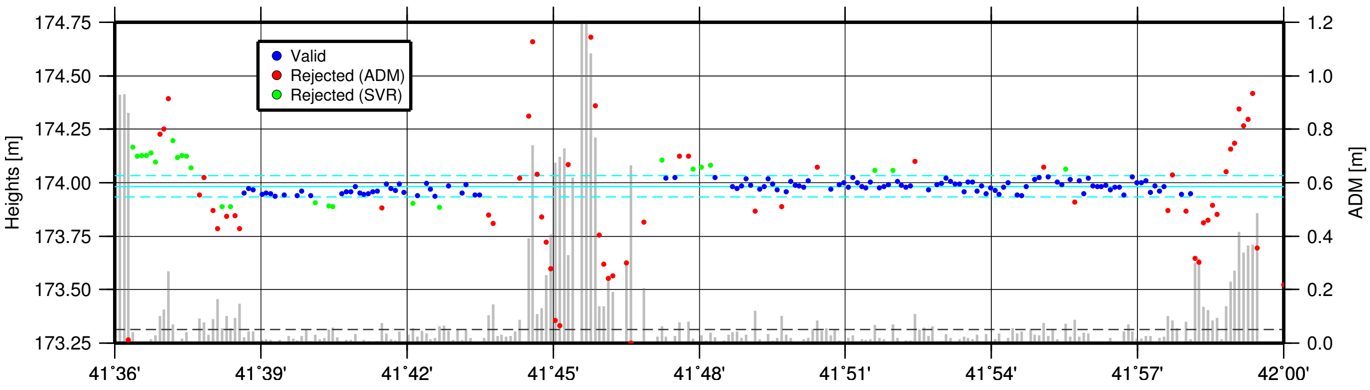


Figure 2: Example of an outlier detection using ~~standard deviations~~error threshold and SVR along a single satellite track over ~~a lake~~Lake Erie containing an island (between approx. 41°44’ and 41°47’). The result of the regression shows valid (blue) and rejected (red, green) ~~altimeter~~water heights. The ~~standard deviation~~height errors based on ADM are plotted as grey bars. Thresholds for ~~standard deviation~~height errors and SVR are marked by dashed lines (black and cyan respectively). [Image was updated]

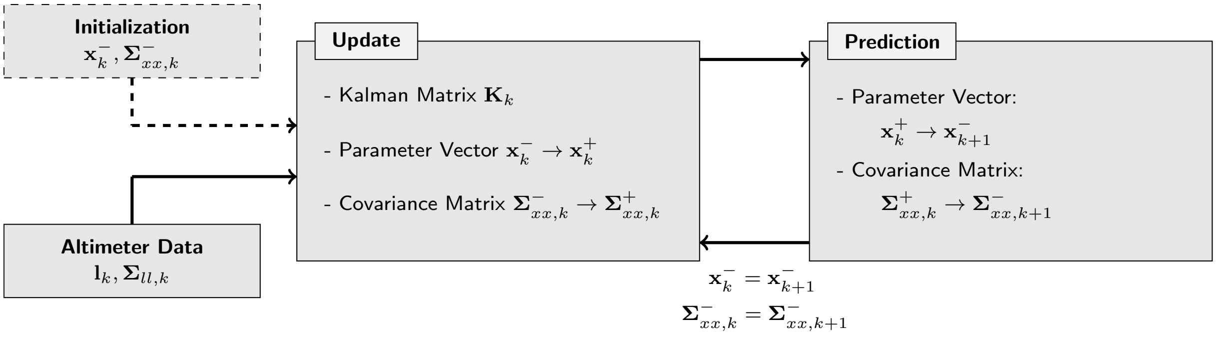


Figure 3: Procedure of Kalman filtering starting with an initialization step followed by a progressive loop containing one update and one prediction step.

3.2.3 Kalman Filter Equations

In the following, the basic equations of the Kalman filter are introduced. The algorithm consists of an observation model and a dynamic model.

The observations for each step k corresponding to epoch t_k are given in vector $\mathbf{l}_k$ and its co-variances in matrix $\mathbf{\Sigma}_{ll,k}$.

$$\underset{(m_k,1)}{\mathbf{l}_k} = \underset{(m_k,n)}{\mathbf{A}_k} \cdot \underset{(n,1)}{\mathbf{x}_k} - \underset{(m_k,1)}{\mathbf{v}_k} \qquad (2)$$

$$\underset{(m_k,m_k)}{\mathbf{\Sigma}_{ll,k}} = \underset{(m_k,m_k)}{\mathbf{I}} \cdot \underset{(m_k,1)}{\mathbf{s}_{l,k}} \qquad (3)$$

The vector length of $\mathbf{l}_k$ depends on the number of ~~water level heights~~water levels m_k available at each epoch t_k. The unknown grid node heights are compiled in vector $\mathbf{x}_k$. The $m_k \times n$ design matrix $\mathbf{A}_k$ is the core of the observation model and connects the ~~water level heights~~water levels with the computation grid consisting of n grid points ($n = 1$ using only a single grid point). $\mathbf{A}_k$ has a dimension of $m_k \times n$ and contains ones for those grid nodes where ~~water level heights~~water levels are available. Hereby, each water level height is assigned to the nearest grid node. In the case when the computation is performed on a single grid node all water level heights are assigned to them. The vector $\mathbf{v}_k$ absorbs the residuals of the observation model.

The uncertainties of the ~~water level heights~~water levels are described in $\mathbf{\Sigma}_{ll,k}$. Since there is no information on correlation between individual ~~water level heights~~water levels the matrix is defined as a diagonal matrix with variances σ_l^2 from ADM (computed in the preprocessing step) on the mean diagonal. These are collected in vector $\mathbf{s}_{l,k}$.

The dynamic model of the Kalman filter approach describes the transition of the system state from epoch t_k to

t_{k+1}.

$$\underset{(n,1)}{\mathbf{x}_{k+1}^{-}} = \underset{(n,n)}{\mathbf{\Phi}_k} \cdot \underset{(n,1)}{\mathbf{x}_k^{+}} + \underset{(n,n)}{\mathbf{\Lambda}_k} \cdot \underset{(n,1)}{\mathbf{q}_k} \quad (4)$$

$$\underset{(n,n)}{\mathbf{\Sigma}_{xx,k+1}^{-}} = \underset{(n,n)}{\mathbf{\Phi}_k} \cdot \underset{(n,n)}{\mathbf{\Sigma}_{xx,k}^{+}} \cdot \underset{(n,n)}{\mathbf{\Phi}_k^{T}} + \underset{(n,n)}{\mathbf{\Lambda}_k} \cdot \underset{(n,n)}{\mathbf{Q}_k} \cdot \underset{(n,n)}{\mathbf{\Lambda}_k^{T}} \quad (5)$$

This includes the prediction step (cf. ~~Fig.~~Figure 3) for the parameter vector $\mathbf{x}_k^{+}$ as well as for its covariance matrix $\mathbf{\Sigma}_{xx,k}^{+}$. The prediction of the grid node heights is done by the transition matrix $\mathbf{\Phi}_k$. In addition, system noise $\mathbf{q}_k$ is taken into account and mapped to the grid node heights by $\mathbf{\Lambda}_k$. The model uncertainties are predicted by Eq. (5) where the covariance matrix $\mathbf{Q}_k$ contains the uncertainties of the system disturbance, i.e. the system noise. Since no information on the temporal evolution of the water level is known in advance, the prediction is based purely on stochastic information ~~(transition matrices are identity matrices)~~. Moreover, the (deterministic) system disturbances in $\mathbf{q}_k$ are set to zero. The system noise σ_q^2 in matrix $\mathbf{Q}_k$ is assumed to yield 5 cm² for each grid node (without correlations).

~~In the following, the applied Kalman filter procedure is described in detail.~~The applied Kalman filter procedure as used in the DAHITI approach is described in detail below.

Initialization

The Kalman filter approach begins with an initialization step which is necessary before starting the recursive loop. The initial state vector $\mathbf{x}_k^{-}$ is filled by setting all elements to the observed water level with the smallest ~~standard deviation~~height error in the first epoch t_k. The covariance matrix $\mathbf{\Sigma}_{xx,k}^{-}$ is initialized by an identity matrix of size $n \times n$.

Update

In the update step, new altimeter ~~water level heights~~water levels are introduced in order to update the parameters of the ~~actual~~current state $\mathbf{x}_k^{-}$ to a new state $\mathbf{x}_k^{+}$. The update is done by comparing the estimated observations (based on the current model, cf. Eq. 2) with the ~~water level heights~~water levels. The weighting of this so-called innovation is described by matrix $\mathbf{K}_k$. It can be computed based on the design matrix and the covariance matrices of observations and parameters using

$$\underset{(n,m_k)}{\mathbf{K}_k} = \underset{(n,n)}{\mathbf{\Sigma}_{xx,k}^{-}} \cdot \underset{(n,m)}{\mathbf{A}_k^{T}} \cdot \Big(\underset{(m_k,m_k)}{\mathbf{\Sigma}_{ll,k}^{-}} + \underset{(m_k,n)}{\mathbf{A}_k} \cdot \underset{(n,n)}{\mathbf{\Sigma}_{xx,k}^{-}} \cdot \underset{(n,m_k)}{\mathbf{A}_k^{T}} \Big)^{-1} \quad (6)$$

The parameter update of vector $\mathbf{x}_k^{+}$ describes the updated ~~water level heights~~water levels for each grid node at the current epoch t_k.

$$\underset{(n,1)}{\mathbf{x}_k^{+}} = \underset{(n,1)}{\mathbf{x}_k^{-}} + \underset{(n,m_k)}{\mathbf{K}_k} \cdot \Big(\underset{(m_k,1)}{\mathbf{l}_k} - \underset{(m_k,n)}{\mathbf{A}_k} \cdot \underset{(n,1)}{\mathbf{x}_k^{-}} \Big) \quad (7)$$

$$\underset{(n,n)}{\mathbf{\Sigma}_{xx,k}^{+}} = \Big(\underset{(n,n)}{\mathbf{I}} - \underset{(n,m_k)}{\mathbf{K}_k} \cdot \underset{(m_k,n)}{\mathbf{A}_k} \Big) \cdot \underset{(n,n)}{\mathbf{\Sigma}_{xx,k}^{-}} \quad (8)$$

In parallel the corresponding covariance matrix $\mathbf{\Sigma}_{xx,k}^{+}$ of the height estimates is updated using Eq. (8). The uncertainties of new altimeter data are taken into account by applying the Kalman matrix as weighting matrix. It can easily be seen that the parameter accuracies will ~~be reduced within the updating step~~become smaller within the updating step.

Prediction

After the parameter vector and the covariance matrix of the current epoch t_k have been updated, the prediction of $\mathbf{x}_k^{+}$ and $\mathbf{\Sigma}_{xx,k}^{+}$ to the next epoch t_{k+1} is performed and $\mathbf{x}_{k+1}^{-}$ and $\mathbf{\Sigma}_{xx,k+1}^{-}$ are computed. The predictions are used as start parameters for the next update step, and the computation loop then continues until all ~~water level heights~~water levels have been processed. In our case, no additional information about the temporal propagation of the parameter vector and the covariance matrix is introduced. Therefore, no deterministic model is applied and the transition matrices $\mathbf{\Phi}_k$ for data and $\mathbf{\Lambda}_k$ for disturbances in Eq. (4) and (5) can be identity matrices. Furthermore, only system noise is taken into account by setting the disturbance value $\mathbf{q}_k$ equal to zero and its uncertainties $\mathbf{Q}_k$ to variances of 5 cm² for each grid node without any correlations.

3.3 Post-processing

The Kalman filter provides water heights $\boldsymbol{x}_k$ and their formal errors $\mathbf{\Sigma}_{xx,k}$ for each epoch t_k and grid node.

If Kalman filtering is performed on a single grid node, the final water level and error are immediately available. ~~Since we assume the water level to be constant over the grid area for each time step, the surface information shall be concentrated to one references point~~ ~~Thus,~~If it is computed on a grid, a ‘mean’ one-dimensional time series is computed. Instead of simply averaging all grid node heights, we select only the best water levels per epoch. Only ~~water level heights~~water levels are selected that which fulfill certain error criteria. In general, the limit for the ~~standard deviation~~maximum height error is set to values between 5 and 10 cm. The selected limit depends on the resulting height errors. Therefore, the limit is selected manually in such a manner that only reliable heights are used for the final time series. The remaining ~~water level heights~~water levels are averaged for each epoch by using the formal errors for the ~~resulting water heights~~ weighting factors. Finally, a time series of ~~water level heights~~water levels and their formal errors over the entire period of time are obtained.

In a last step an outlier rejection is performed. The water level time series can still contain outliers because of bad quality of data, ice coverage, orbit manoeuvres, etc. For the detection of those outliers, SVR can be applied again - now on the full time series. Complete tracks showing significant differences with respect to the other points of the water level time series can be rejected. This time, radial base functions instead of a linear kernel are used to perform the regression since a constant water level over time cannot be assumed. The radial base function as kernel of the SVR allows us to fit the time series including seasonal variations and trends. Figure 4 shows the results of an applied SVR on a six-year subset of the time series of Lake Erie. The fitted model is plotted as a cyan line together with its manually defined confidence interval. The confidence interval is selected depending on the noise of the water level time series which varies between 7.5 cm and 100 cm. ~~Water level heights~~Water levels which fulfill the limit of the SVR are kept (blue) whereas outliers are rejected (red).

4 Results and Validation

In this chapter, water level time series resulting from the Kalman approach are presented and validated. Since it is not possible to show results for all inland water bodies we focus on the selected study areas introduced in Section 4.1. Three inland water targets are described in more detail. They represent different target types, i.e. large lakes, small lakes, and rivers. Moreover, results from 16 lakes and 20 river crossings are validated by comparison with in situ data and altimeter time series provided by other groups.

4.1 Study Areas

For altimetry-derived water level time series, in situ measurements from gauging stations are the most important validation data sets. In order to perform reliable comparisons, only those inland water bodies are selected as study areas for which in situ data are available. Since we have access to many gauging stations in North and South America we focus our study on these two continents.

Another criterion for the selection of inland water bodies is the availability of external altimetry-derived time series to demonstrate the performance of our Kalman filter method compared with other approaches. Each study case is observed by at least one other group (i.e. ~~LEGOS~~Hydroweb, ~~ESA-DMU~~River & Lakes or GRLM). Thus, those targets in North and South America are selected which are best represented by other inland altimetry databases for a time period as long as possible. ~~Moreover, different water types should be covered, such as large lakes, small lakes, and rivers with different width.~~ We end up with the 16 lakes and 20 river crossings illustrated in ~~Fig.~~Figure 5. For almost all investigated inland water bodies at least one in situ gauging station and one external altimetry-derived time series is available.

~~In the following all investigated inland water bodies located in North America and South America and their corresponding in-situ data are introduced.~~

The first study areas are the Great Lakes of North America comprising Lake Superior (82,000km²), Lake Huron (59,000km²), Lake Michigan (58,000km²), Lake Erie (25,000km²), and Lake Ontario (19,000km²). The size of the these lakes leads to ocean-like conditions which means that the altimeter measurements are not disturbed by land. Only a few altimeter measurements near the lake shore are contaminated by land. The Great Lakes show seasonal variations of about 1 m. They are well-observed inland waters with many in situ stations provided by the 'Tides & Currents' platform of the National Oceanic and Atmospheric Administration (NOAA)[c9]. For the validation of Lake Superior, in situ stations of Duluth, Grand Marais, Marquette, Ontonagon and Point Iroqouis are used. Lake Huron has five stations for validation which are Essexville, Harbor Beach, Lakeport, Mackinaw City, and de Tour Village. The stations Calumet Harbor, Holland, Kewaunee, Ludington, Milwaukee, and Port Inland are used for Lake Michigan. Lake Erie has seven stations for validation which are Buffalo, Cleveland, Fairport, Fermi Power Plant, Marblehead, Sturgeon Point, and Toledo. For the validation of Lake Ontario, the in situ stations of Cape Vincent, Olcott, Oswego, and Rochester are used.

In addition to the Great Lakes, the Great Slave Lake (27,200 km²), Lake Winnipeg (24,000 km²), Lake Athabasca (7,800 km²), Lake Winnipegosis (5,100km²), Lake Manitoba (4,600 km²), Lake of the Woods (4,300 km²), Great Salt Lake (4,000 km²), Lake Claire (1,400 km²), and Cedar Lake (1,300 km²), which are located in Canada and the United States, are investigated. These lakes differ significantly in surface extent by up to a factor of 20. Estimation of water level time series in the Canadian lakes is made difficult by the winter conditions. Several lakes are frozen for several months which makes the water level computation challenging (Table 3). For the validation of the water level time series, in situ data provided by the government of Canada[c9] and the U.S. Geological Survey (USGS)[c9] are used.

In addition to the lakes in North America two lakes in the very south of South America are selected for validating our approach. Lake Argentino (1,466 km²) and Lake Buenos Aires (1,850 km²) are located in Argentina next to the Andes. The lakes are party surrounded by mountains, which can affect the altimeter measurements. The lakes have a similar shape with largest extent in across-track direction of the satellites ground track. This leads to rather short track crossings varying between 10 and 15 km. Despite their location in a temperate zone near high mountains the lakes are not frozen in winter. The seasonal variations of both lakes vary between 2.5 m and 3.5 m. For the validation of ~~Lago~~Lake Argentino and Lake Buenos Aires in situ data from the Ministerio de Planificación Federal, República Argentina[c10] are used.

[c9] http://tidesandcurrents.noaa.gov/

[c9] http://wateroffice.ec.gc.ca/

[c9] http://waterdata.usgs.gov/

[c10] http://www.hidricosargentina.gov.ar/

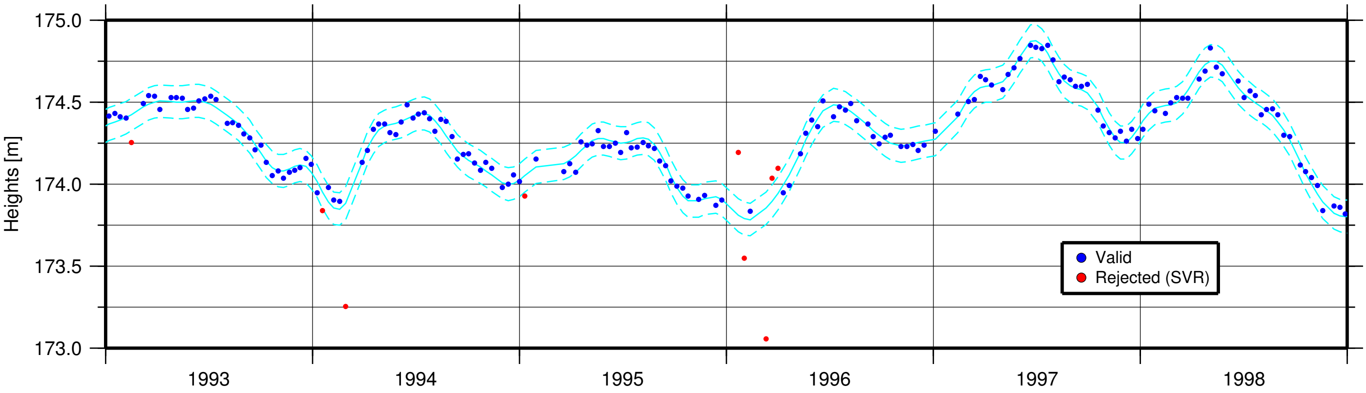


Figure 4: Example of applied SVR using radial base functions for outlier rejection on a resulting water level time series (Lake Erie) of the Kalman filtering step. ~~The estimated regression function and its error limits are plotted in cyan.~~The estimated regression function (cyan) and its confidence intervals (dotted cyan) are plotted. The result of the regression shows valid (blue) and rejected (red) altimeter heights. Each rejected water level height represents one complete satellite overflight.

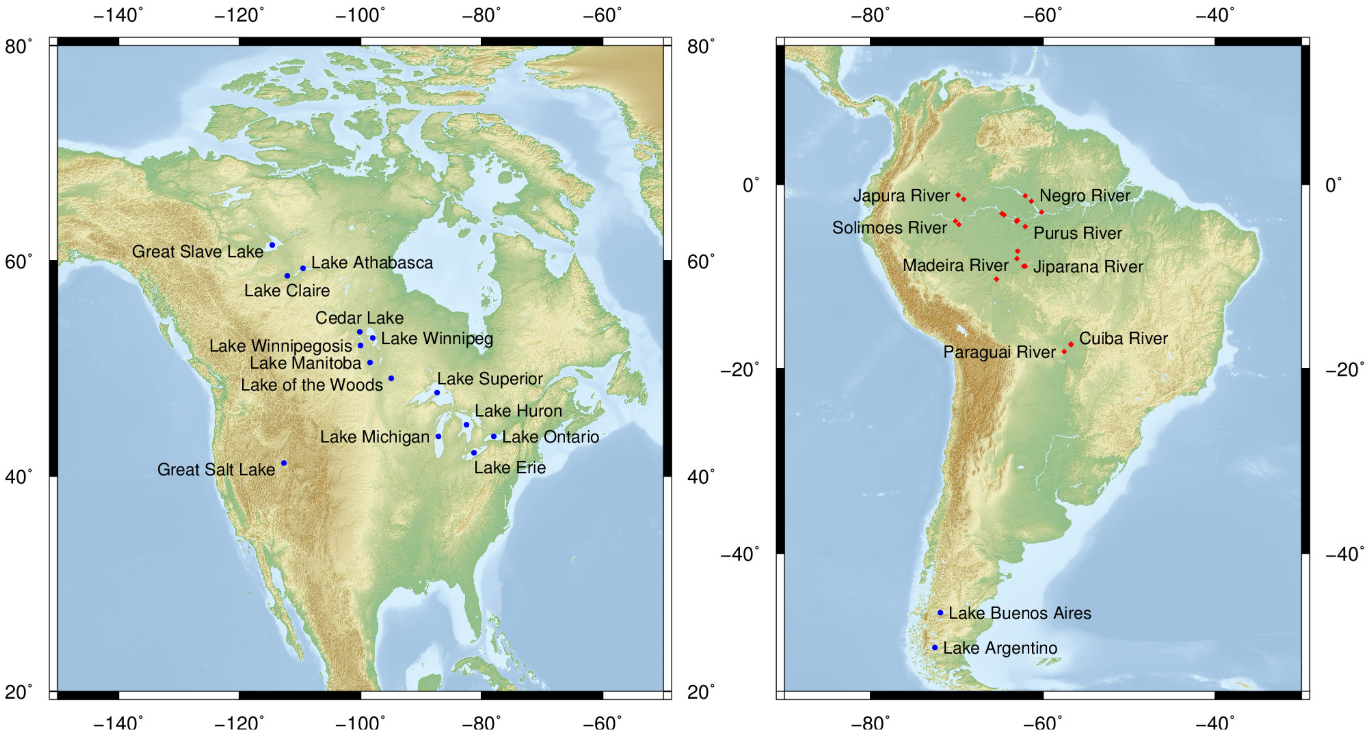


Figure 5: Map of selected study areas of lakes (blue) and rivers (red) in North America (left) and South America (right) [Image was updated]

For the analysis of rivers, the Amazon basin is selected as the study area; it is the largest basin in the world and covers about 7,000,000 km^2. The region is located in the tropics, and the climate is hot and ~~wet~~humid throughout the year. Because of the strong precipitation, the resulting seasonal variations of the water level ~~show amplitudes~~reach peak-to-peak variations up to 15 m. The Amazon basin consists of countless rivers which differ in terms of length, width, meanders, and seasonal variations. This ~~diversification~~variety is very useful for the quality assessment of water level time series from altimetry. For example, the river widths vary from up to 10 km for the Amazon river and a few hundred metres for the

~~Rio Jiparaná~~Jiparaná River. Moreover, the Amazon basin is a well-observed area since the Agência Nacional de Águas (ANA)[c14] provides data for numerous in situ gauging stations. For the validation, water level time series of gauges at the ~~Rio Japurá~~Japurá River, the ~~Rio Solimoes~~Solimões River, the ~~Rio Negro~~Negro River, the ~~Rio Purus~~Purus River, ~~Rio Jiparaná~~Jiparaná River, ~~Rio Paraguai~~Paraguai River, and the ~~Rio Cuiba~~São Lourenc̃o River are used. Another reason why we chose the Amazon basin is that other groups such as LEGOS and ESA-DMU have also investigated this area.

4.2 Validation data sets

~~For a validation of the Kalman filter results in-situ data of gauging stations provided by different institutions named in Sect. 4.1 are used.~~ Water level time series from gauges have a high relative accuracy, but some points must be borne in mind in the use of in situ data. The absolute comparison of heights from gauges and satellite altimetry is often very difficult since location, reference height and vertical datum of gauges are not always precisely known or may even be unknown. This leads to height offsets between water level time series from gauge and altimetry which must be considered in the validation step. In particular, the comparison between ~~water level heights~~water levels from altimetry and in situ data over rivers ~~show~~shows in most cases remaining offsets. In general, almost no altimeter satellite track crosses the river at the location of a gauging station, which leads to additional offsets because of the ~~slope of the river~~river slope. To avoid handling the uncertainties of in situ data only relative comparisons with water level time series from altimetry are ~~preformed~~performed.

In order to rank our results with respect to other time series derived from altimeter data, we download water levels from three external inland altimeter data bases, namely ~~LEGOS~~Hydroweb, ~~ESA-DMU~~River & Lakes, and GRLM. These results are based on ~~different~~various altimeter missions and ~~the groups perform different~~diverse approaches were performed to compute the water level time series. As a consequence, these external time series cover different time periods ~~and feature different temporal resolution~~with temporal resolutions between 10 and 35 days. This has to be kept in mind when the ~~different ~~time series of the four databases are compared.

4.3 Selected Results

We choose three of the aforementioned water bodies in order to present detailed results of our Kalman filter approach. The targets are selected to represent three ~~disparate~~diverse inland water body types featuring different characteristics. Lake Superior (~~Fig.~~Figure 6) is selected as representative of larger lakes with ocean-like conditions. Lake Athabasca (~~Fig.~~Figure 7) is a smaller lake which has to cope with ice coverage in winter, which is the case for most lakes in North America. Finally, the ~~Rio Madeira~~Madeira River (~~Fig.~~Figure 8) in the Amazon basin is selected to show the potential of the Kalman filter approach for river monitoring. For ~~those~~all examples, the ~~Kalman filter based ~~time series from DAHITI is compared with in situ data and results from ~~LEGOS~~Hydroweb, ~~ESA-DMU~~River & Lakes, and GRLM.

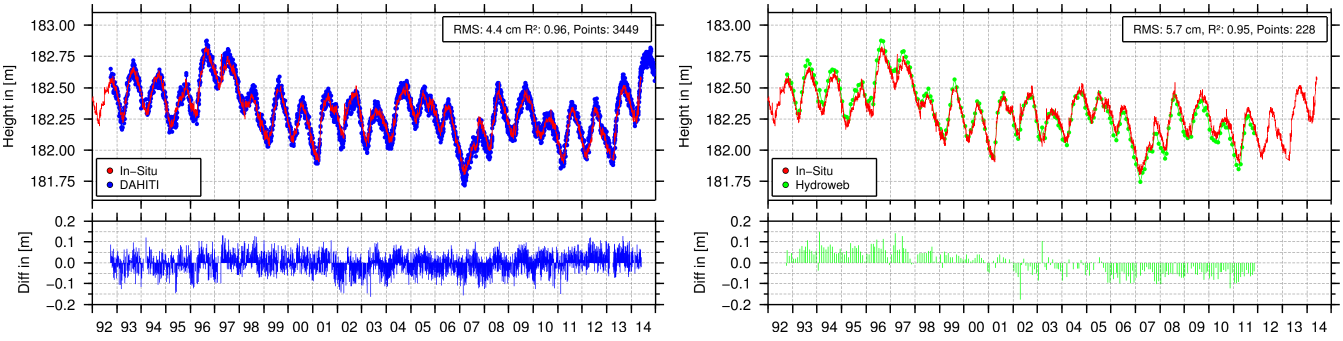
4.3.1 Lake Superior

Figure 6 shows the water level time series of Lake Superior between 1992 and 2014: the DAHITI result is plotted in blue (subplot a), the in situ data of station Ontonagon in red, and external altimetry-derived water levels in green (~~LEGOS~~Hydroweb, subplot b), light blue (~~ESA-DMU~~River & Lakes, subplot c), and orange (GRLM, subplot c). ~~For a detailed few, results from year 2004 are highlighted in the upper right corner.~~ In order to neglect constant offsets between the different solutions, all time series are shifted to the level of ~~DAHITI~~in situ data, and only water level changes are compared. The applied offset is estimated by using the average of height differences at all days in which in situ data and time series from altimetry are available. Additionally, differences between water levels from altimetry and in situ data are plotted for each time series. For the DAHITI computation, ~~1 Hz~~non-retracked high-frequent altimeter data of Topex, Jason-1, Jason-2, Envisat, ERS-2, and SARAL/~~AltiKA~~AltiKa are used. The Kalman filter provides a continuous time series with an irregular near-daily resolution which shows neither outliers nor inter-mission inconsistencies. In order to achieve reliable water level time series different outlier criteria are applied. Initially, the number of invalid water levels is reduced by using thresholds for latitude (depending on track length over Lake Superior), height (180 m to 185 m) and height error (10 cm). Furthermore, only backscatter coefficients between 10 db and 18 db are selected in order to reject data affected by ice coverage. Then, an SVR using a confidence limit of ±5 cm is applied along the crossing altimeter track to reject water levels near the shore which are affected by land contamination. Finally, an SVR using a confidence limit of ±7.5 cm is applied along the final water level time series to reject remaining outliers. Altogether, the time series is composed of ~~2779~~3449 single points, each representing one day with at least one altimeter track crossing the lake. During computation of the final water level time series 24 % of the data are rejected, mostly because of ice coverage.

The DAHITI water levels coincide very well with the daily in situ data of Ontonagon. The correlation coefficient R^2 is 0.96 and the RMS difference ~~show ~~is ~~4.1 cm~~4.4 cm. The alternative computation of the water level time series using a median filter instead of the Kalman filter leads to a slightly worse RMS difference of 4.5 cm (see Section 4.3.4). In comparison with the DAHITI time series, the other altimetry-derived water levels show significantly reduced temporal resolutions. In addition, the lengths of the time series differ, depending on the missions used by the

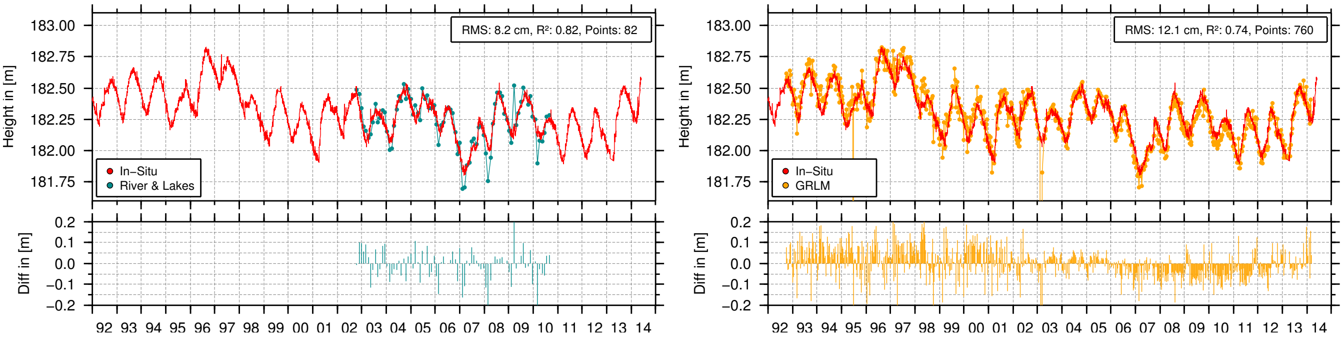
---

[c14] http://ana.gov.br/



(a) DAHITI vs. In situ data

(b) Hydroweb vs. In situ data



(c) River & Lakes vs. In situ data

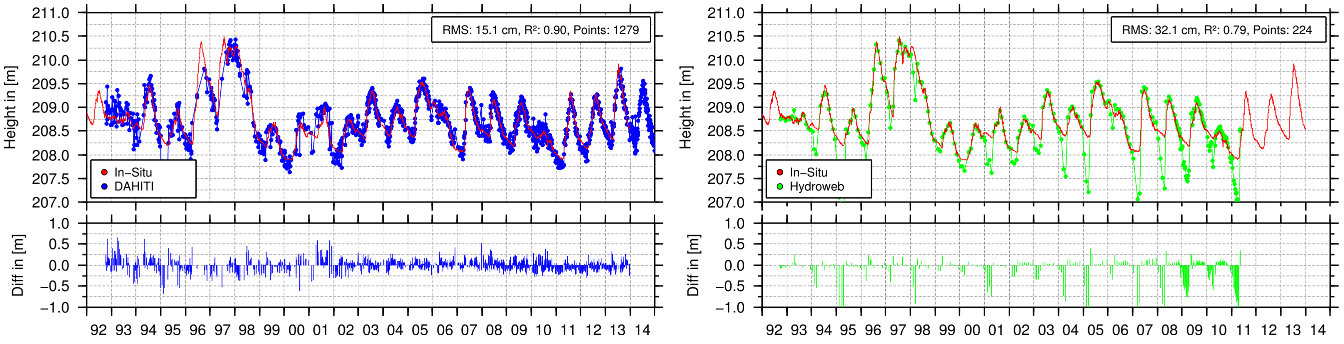
(d) GRLM vs. In situ data

Figure 6: ~~Water level time series of Lake Superior from DAHITI (1992-2014), LEGOS (1992-2011), ESA-DMU (2002-2010) and GRLM (1992-2014) compared with in-situ data (Ontonagon, 1992-2014) and shifted to the water level height of DAHITI.~~Water level time series of Lake Superior from DAHITI (1992-2014), Hydroweb (1992-2011), River & Lakes (2002-2010) and GRLM (1992-2014) compared with in situ data (Ontonagon, 1992-2014) and shifted to the water level height of the in situ data. Additionally, differences between heights from altimetry and in situ data are plotted for periods in which both data sets are available. [Figure was updated]

different groups. In order to rank the DAHITI result compared with other altimetry-derived water levels, we also compare the three external time series with in situ gauging data within the corresponding time intervals. For all three databases this gives smaller correlations and higher RMS (~~LEGOS~~Hydroweb: RMS=~~6.1~~5.7 cm, R²=~~0.94~~0.95, ~~278~~228 points, ~~ESA-DMU~~River & Lakes: RMS=8.2 cm, R²=0.82, 82 points, and GRLM: RMS=12.1 cm, R²=0.74, 760 points). For the validation, the water level time series of the other altimetry-derived water levels are used as they are, without any additional outlier rejection. This leads to higher RMS differences as published in Ričko et al. (2013), who applied an additional outlier rejection based on in situ data. The altimetry derived-solutions differ because of varying input data sets and the different approaches. ~~LEGOS~~Hydroweb uses a multi-mission approach with a merged monthly resolution whereas ~~ESA-DMU~~River & Lakes relies purely on Envisat with a temporal resolution of 35 days. GRLM applies a multi-mission approach ~~reaching~~providing a temporal resolution of about 10 days. The time series of Hydroweb and GRLM still show mission-dependent offsets which can be seen in the differences from the in situ data (mainly positive for ERS-2, mainly negative for Envisat). In contrast, mission-dependent offsets are quite small in the water level time series of DAHITI.

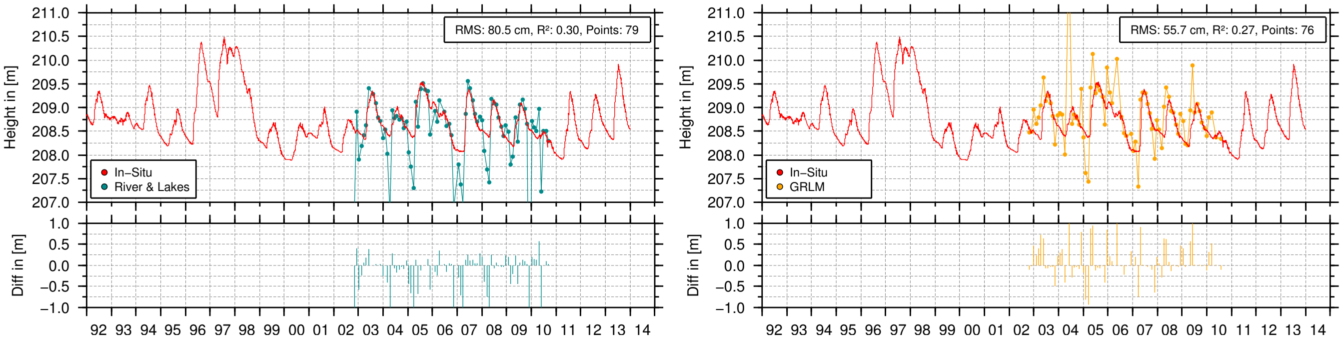
4.3.2 Lake Athabasca

Figure 7 shows the water level time series of Lake Athabasca between 1992 and 2014. Once again, water levels from DAHITI (blue), in situ data of Crackingstone Point (red), ~~LEGOS~~Hydroweb (green), ~~ESA-DMU~~River & Lakes (light blue), and GRLM (orange) are plotted. The time series of the four altimeter data bases are shifted to the level of the in situ data. ~~Now, the year 2010 is highlighted.~~ In principle, Lake Athabasca, whose surface covers 7,800 km², should be large enough to provide reliable altimetry-derived water level time series. However, different problems such as ice coverage because of regular freezing in winter, land contamination and off-nadir effects near lake shores have to be considered. For the estimation of the water level time series in DAHITI retracked altimeter data are used, with a 10% Improved



(a) DAHITI vs. In situ data

(b) Hydroweb vs. In situ data



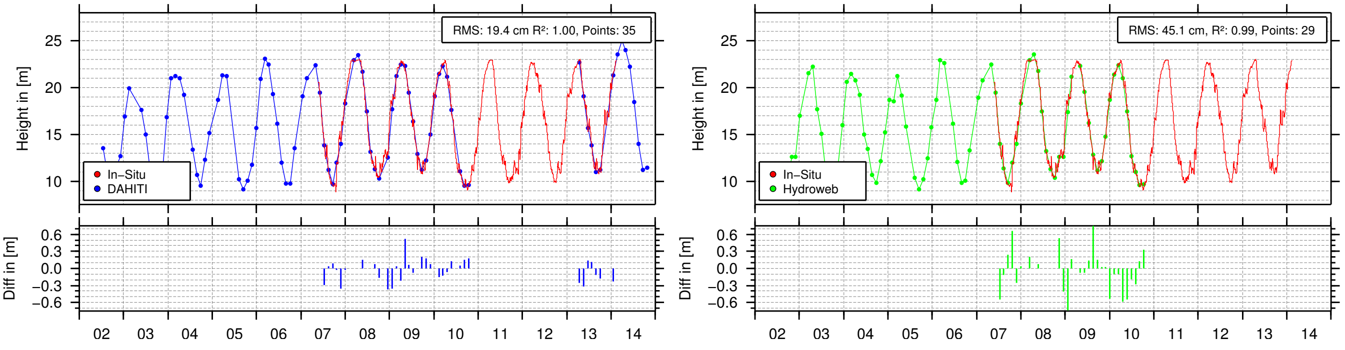
(c) River & Lakes vs. In situ data

(d) GRLM vs. In situ data

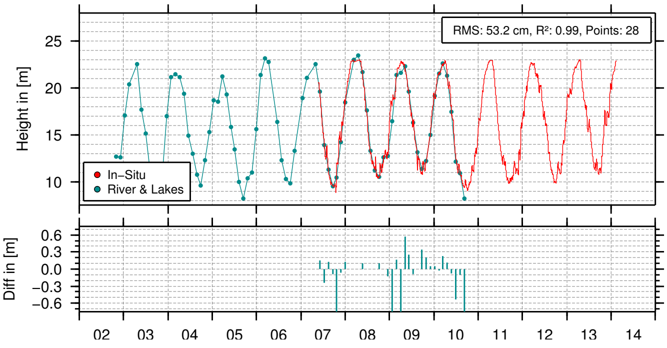
Figure 7: ~~Water level time series of Lake Athabasca from DAHITI (1992-2014), LEGOS (1992-2011), ESA-DMU (2002-2010) and GRLM (2002-2010) compared with in-situ data (Lake Athabasca, 1992-2013) and shifted to the water level height of DAHITI.~~Water level time series of Lake Athabasca from DAHITI (1992-2014), Hydroweb (1992-2011), River & Lakes (2002-2010) and GRLM (2002-2014) compared with in situ data (Lake Athabasca, 1992-2013) and shifted to the water level height of the in situ data. Additionally, differences between heights from altimetry and in situ data are plotted for periods in which both data sets are available. [Figure was updated]

Threshold retracker (Hwang et al., 2006). For the computation, altimeter data of Topex, Jason-1, Jason-2, Envisat, ERS-2 and SARAL/AltiKa are used. In order to achieve reliable water level time series the same outlier criteria as for Lake Superior but different thresholds are applied. First, outliers are rejected by using thresholds for latitude (depending on track length over Lake Athabasca), height (208 m to 212 m) and height error (50 cm). Furthermore, water levels affected by ice coverage are rejected if the valid backscatter coefficients are not between 10 db and 18 db. To reject water levels near the shore which are affected by land contamination, an SVR along the crossing altimeter track using a confidence limit of ±5 cm is applied. Finally, an SVR along the final water level time series using a confidence limit of ±50 cm is applied to reject remaining outliers.

The DAHITI water level shows a very good agreement with in situ data in summer ~~,a few outliers~~and almost no outliers owed to ice coverage are visible in winter compared with time series from Hydroweb and River & Lakes. The overall consistency with the gauge data yields a correlation coefficient of ~~0.88~~0.90 and an RMS difference of ~~17.0~~15.1 cm using ~~1337~~1279 points in the period between 1992 and 2014. The usage of a median filter leads to slightly worse RMS differences of 15.3 cm for Lake Athabasca. The differences between in situ data and ~~LEGOS~~Hydroweb (RMS=~~33.7~~32.1 cm, R^2=0.79, ~~272~~224 points), ~~ESA-DMU~~River & Lakes (RMS=80.5 cm, R^2=0.30, 79 points) and GRLM (RMS=55.7 cm, R^2=0.27, 76 points) show higher RMS values and smaller correlations. One can clearly see that the problems of altimeter time series occur mostly in winter because of ice coverage. In particular, water level time series of Hydroweb and River & Lakes show strong outliers in winter which are not contained in the time series of DAHITI because of the applied outlier rejection. A new problem with retracker biases arises for time series based on retracked altimeter data. To minimize those effects all altimeter measurements are retracked using with 10% Improved Threshold retracker. However, small retracker biases can also occur if identical retracking algorithms are applied on altimeter missions measuring in different bands such as Ku-band (Envisat) and Ka-band (SARAL/AltiKa).



(a) DAHITI vs. In situ data (b) Hydroweb vs. In situ data



(c) River & Lakes vs. In situ data

Figure 8: ~~Water level time series of Rio Madeira from DAHITI (2002-2014) and ESA-DMU (2002-2010) compared with in-situ data (Humaitá, 2007-2014) and shifted to the water level height of DAHITI.~~Water level time series of the Madeira River from DAHITI (2002-2014), and River & Lakes (2002-2010) compared with in situ data (Humaitá, 2007-2014) and shifted to the water level height of the in situ data. Additionally, differences between heights from altimetry and in situ data are plotted for periods in which both data sets are available. [Figure was updated]

4.3.3 Madeira River

~~As last example, we choose a river crossing in the Amazon basin.~~ Figure 8 shows the resulting water level derived from an Envisat and SARAL/~~AltiKA~~AltiKa crossing over the ~~Rio Madeira~~Madeira River. The water level time series from DAHITI (blue), Hydroweb (green) and River & Lakes (light blue) are compared with the in situ station Humaitá (red), which is located about 27.6 km upstream. All time series from altimetry are shifted to the water level of the in situ station. At this location the ~~Rio Madeira~~Madeira River is about 2.5 km wide. ~~The in-situ station Humaitá is located about 27.6 km upstream.~~ In order to achieve reliable water level time series over the Madeira River different outlier criteria are applied. First, thresholds for latitude (depending on track length over the Madeira River), height (30 m to 50 m) and height error (100 cm) are applied to reduce the number of invalid water levels. Finally, an SVR along the crossing altimeter track using a confidence limit of ±10 cm and an SVR along the final water level time series using a confidence limit of ±100 cm are applied to reject remaining outliers. In this case, no limit for the backscatter coefficients is applied because no ice coverage exists in the Amazon basin. In principle, the backscatter coefficient can also be used to distinguish between water and land but this is not considered here. All altimeter time series reach a temporal resolution of about one month since there is only one mission with 35-day temporal resolution at the same time. Altimeter data ~~is~~are available between 2002 and 2014 with a data gap ~~in 2011 and 2012~~between October 2010 and March 2013. The altimeter data from Envisat on the shifted orbit can not be used between October 2010 and April 2012 for the current water level time series. Gauging information does not start before 2007. Thus, the comparison with in situ data only comprises a time period of about 3.5 years. For DAHITI another year of SARAL/~~AltiKA~~AltiKa data is available. The Kalman filter result (blue) shows an RMS difference of ~~21.6~~19.4 cm and a correlation coefficient of 1.00 by using 35 points. The estimation of the water level time series using a median filter leads to RMS difference of 19.6 cm. The RMS is comparable to the result for Lake Athabasca, which is even more satisfactory when we take into account the seasonal variations of about 15 m of the ~~Rio Madeira~~Madeira River~~ into account~~. The high amplitude is also the reason for the extremely high correlation, which should not be overvalued. The RMS differences of ~~LEGOS~~Hydroweb and ~~ESA-DMU~~River & Lakes with respect to the gauge are twice as great, at 45.1 cm (~~LEGOS~~Hydroweb, 29 points) and 53.2 cm (~~ESA-DMU~~River & Lakes, 28 points) respectively. GRLM does not provide information for this virtual station.

4.3.4 Discussion

The DAHTIT time series show good consistency with in situ observations and clear advances over established approaches. However, some problems remain, especially for smaller lakes and rivers. For larger lakes, the assumption of a uniform surface level may no longer be justified. In addition to height differences owed to systematic errors in geophysical corrections or the geoid, hydrodynamic effects caused by wind and waves can cause horizontal lake level differences. Currently, these are neglected when combine observations from diverse parts of the lake. Moreover, measurements (altimetry as well as in situ) feature unequal accuracies observed over areas with different surface conditions. This effect can be seen when we compare the DAHITI water level time series of Lake Superior with additional gauging stations. The five possible comparisons lead to RMS differences varying by 2 cm (between 4.4 cm and 6.6 cm; Table 4). The two stations Duluth and Point Iroqouis show reduced consistency with altimetry. Both stations are located in smaller bays of the lake and are more affected by wind and waves than the other stations, which leads to more noisy in situ time series.

For small lakes and rivers, land contamination of waveforms is the largest problem because nearly all altimeter measurements are affected. For rivers, almost no nadir measurements may occur and even these can originate from river branches and distort the water level time series from the investigated target. Moreover, the river slope can influence the time series as well as the comparison with in situ data. The crossings between river and altimeter track can vary slightly (up to 1 km) because of orbit instabilities so that the reflections originate from different areas which do not exhibit the same water level. The most important challenge remaining is the handling of inter-mission biases and retracker biases. The usage of radial errors from a global crossover analysis and the restriction to one common retracker works reasonably well; however, small discrepancies remain in the time series. Moreover, the quality of the single altimeter measurements could surely be further improved by combining different retracking algorithms depending on the waveform shapes. This remains a major challenge and offers enormous potential for future work.

The validation of water level time series of DAHITI for Lake Superior, Lake Athabasca, and the Madeira River compared with in situ data and time series from Hydroweb, River & Lakes, and GRML showed clear improvements. To evaluate the impact of the outlier rejection and Kalman filtering on the improvements of the DAHITI time series, an alternative approach using a simple median filter instead of a Kalman filter was applied.

The resulting RMS differences for three inland waters decreased slightly by 0.1 cm to 0.2 cm, which indicates that the combination strategy has only a moderate effect on the overall accuracy. The strongest improvements are currently owed to rigorous outlier detection and data retracking. However, the Kalman filter has a considerable potential when upgraded by dynamic modelling and used for real-time applications.

4.4 Quality Assessment

The results for Lake Superior, Lake Athabasca, and the ~~Rio Madeira~~Madeira River presented in Section 4.3 already show the ability of the Kalman filter approach to provide reliable and highly accurate time series of inland ~~water level heights~~water levels. Since three results – even if they do represent different inland water types – are not enough to perform a reliable quality assessment of the method, we extend the validation to a larger sample and include all study targets (16 lakes and 20 river crossings) described in Section 4.1 in the comparison.

Table 3 gives an overview of the different parameters used for the estimation of water level time series in DAHITI. This information is provided for all investigated lakes and rivers. First, the used altimeter missions are shown, followed by the retracking flag which indicates if additional retracking is applied. Then the ice flag shows if the water body is affected by ice coverage in winter. This information originates from external sources, e.g. National Snow and Ice Data Center (*http://nsidc.org/*) for Lake Superior. Table 3 also shows which outlier criteria were applied for the different inland water targets to reject erroneous water levels. Consequently, appropriate thresholds for latitude, height, backscatter coefficient, height error, SVR along the pass and along the final time series can be selected. Finally, the number of data points of the water level time series are shown which is equal to the number of days in which altimeter data are available. The last column describes the percentage of outliers which were rejected during the computation of the water level time series.

Table 4 summarizes the comparisons of lake level time series from DAHITI, ~~LEGOS~~Hydroweb, ~~ESA-DMU~~River & Lakes, and GRLM with in situ gauge data. ~~Additional information such as winter ice coverage (Ice) or data retracking (Retr., for DAHITI only) is provided.~~ For each target, RMS difference, squared correlation coefficient and the number of points (No) used for the validation are provided. Depending on the availability of in situ time series of the investigated water body, more than one comparison is performed for the larger lakes. The smallest RMS difference for each target ~~is marked in green~~is highlighted in bold, the largest one in ~~red~~italics.

DAHITI results show RMS differences with respect to the gauge data between 4 cm and ~~38~~36 cm. It is obvious that accuracy declines with lake extent and ice coverage. For some lakes, the differences between DAHITI and in situ data vary

Table 3: Relevant parameters for the estimation of the water level time series for DAHITI. This table shows information about the used altimeter missions for selected lakes and rivers, applied retracking (Retr.), ice coverage (Ice), and applied outlier criteria which are used for the processing of water level time series. Finally, the number of data points and percentage of outliers of the final water level time series are given.

Target name (DAHITI Id)	Used Missions						Retr.	Ice	Applied Outlier Criteria						Data Points	Outlier
	TP	J1	J2	E2	EN	SA			Lat.	Height	Sig.0	Error	SVR Pass	SVR Series		
Superior, Lake (3)	√	√	√	√	√	√	–	√	√	√	√	√	√	√	3614	24 %
Huron, Lake (33)	√	√	√	√	√	√	–	√	√	√	√	√	√	√	2132	23 %
Michigan, Lake (11)	√	√	√	√	√	√	–	√	√	√	√	√	√	√	2132	37 %
Erie, Lake (6)	√	√	√	√	√	√	–	√	√	√	√	√	√	√	1968	41 %
Ontario, Lake (35)	√	√	√	√	√	√	–	√	√	√	√	√	√	√	2174	29 %
Athabasca, Lake (100)	√	√	√	√	√	√	√	√	√	√	√	√	√	√	1398	45 %
Great Slave, Lake (99)	√	√	√	√	√	√	–	√	√	√	√	√	√	√	1396	39 %
Claire, Lake (578)	–	√	√	–	√	√	√	√	√	√	√	√	√	√	593	23 %
Winnipeg Lake (101)	–	–	√	–	√	√	√	√	√	√	√	√	√	√	816	13 %
Manitoba, Lake (191)	–	√	√	–	√	√	√	√	√	√	√	√	√	–	535	4 %
Cedar, Lake (200)	√	√	√	–	√	√	–	√	√	√	√	√	√	√	647	18 %
Winnipegosis, Lake (281)	–	√	√	–	√	√	√	√	√	√	√	√	√	√	529	20 %
Lake of the Woods (73)	√	√	√	–	√	√	√	√	√	√	√	√	√	√	687	49 %
Great Salt, Lake (72)	–	–	–	√	√	√	√	–	√	√	–	√	√	√	147	9 %
Argentino, Lake (182)	√	√	–	√	√	√	–	–	√	√	–	√	√	√	880	49 %
Buenos Aires, Lake (139)	–	–	–	–	√	√	–	–	√	√	–	√	√	–	116	3 %
Solimões, River (405)	–	–	√	–	√	√	√	–	√	√	–	√	√	–	323	3 %
Solimões, River (406)	–	–	–	–	√	√	√	–	√	√	–	√	√	–	90	12 %
Solimões, River (389)	–	–	–	–	√	√	√	–	√	√	–	√	√	–	86	12 %
Solimões, River (581)	–	–	√	–	–	–	√	–	√	√	–	√	√	√	178	26 %
Solimões, River (384)	–	–	–	–	√	–	√	–	√	√	–	√	√	–	82	0 %
Solimões, River (582)	–	–	√	–	–	–	√	–	√	√	–	√	√	–	198	17 %
Purus, River (583)	–	–	–	–	√	√	√	–	√	√	–	√	√	–	86	10 %
Jiparaná, River (584)	–	–	–	–	√	√	√	–	√	√	–	√	–	–	91	9 %
Jiparaná, River (585)	–	–	√	–	–	–	√	–	√	√	–	√	–	√	235	2 %
Japurá, River (579)	–	–	–	–	√	√	√	–	√	√	–	√	–	–	99	0 %
Japurá, River (580)	–	–	–	–	√	–	√	–	√	√	–	√	–	–	81	2 %
São Lourenčo, River (1093)	–	–	√	–	–	–	√	–	√	√	–	√	√	√	233	3 %
São Lourenčo, River (1094)	–	–	√	–	–	–	√	–	√	√	–	√	√	√	232	3 %
Madeira, River (371)	–	–	–	–	√	√	√	–	√	√	–	√	√	√	90	10 %
Madeira, River (360)	–	–	√	–	–	–	√	–	√	√	–	√	√	√	227	5 %
Madeira, River (575)	–	–	–	–	√	–	√	–	√	√	–	√	√	–	81	4 %
Negro, River (161)	–	–	–	–	√	√	√	–	√	√	–	√	√	–	89	13 %
Negro, River (352)	–	–	–	–	√	√	√	–	√	√	–	√	–	–	99	2 %
Negro, River (346)	–	–	–	–	√	–	√	–	√	√	–	√	√	–	81	2 %
Paraguai, River (1095)	–	–	–	–	√	√	√	–	√	√	–	√	–	–	99	0 %

Missions: Topex/Poseidon (TP), Jason-1 (J1), Jason-2 (J2), ERS-2 (E2), Envisat (EV), SARAL/AltiKa (SA) [New Table added]

by more than a factor of two with different lake gauges. Especially for Lake Erie ~~and Lake Winnipeg~~ the difference between the RMS values can reach up to ~~10~~8.1 cm. ~~Since only one DAHITI time series is computed per lake, these variations demonstrate uncertainties of the in-situ data sets.~~ For most lakes, the relations between the different RMS values are similar for the different altimeter products.

For most lakes the DAHITI water levels are more consistent with in situ data than the results from external altimeter data bases. In addition, the temporal resolutions of the time series are significantly higher, as indicated by the number of ~~used points~~ points used for the validation. Of course, the different time periods of the other altimeter data sets have to be taken into account, too. The most notable improvements through the DAHITI Kalman approach with respect to the existing databases can be seen for smaller lakes. For example, for the Lake of the Woods, the DAHITI consistency with in situ data is more than twice as good as the other altimeter products, improving the RMS differences from about ~~40~~36 cm to approximately ~~15~~16 cm.

The validation results for different rivers in the Amazon basin are summarized in Table 5. We study eight different rivers with 20 virtual stations altogether. For the computation, data from ~~Jason-1,~~ Jason-2, Envisat, and SARAL/~~AltiKA~~AltiKa are used. Most of the time series are based on only one altimeter track (sometimes from consecutive missions, e.g. ~~Jason-1 and Jason-2~~Envisat and SARAL/AltiKa). Few locations allow use of more than one track in case of a crossover point between different altimeter tracks. Table 5 shows the comparison results of three altimeter products (DAHITI, ~~LEGOS~~Hydroweb, and ~~ESA-DMU~~River & Lakes) with different in situ stations. GRLM does not provide river level time series and is excluded from this investigation. In addition to RMS differences with respect to the gauging time series and correlation coefficients, the number of data points, river width and distance between altimeter crossing and gauge are given. The river width corresponds

Table 4: ~~Water level time series of DAHITI for selected lakes compared with in situ data and results from Hydroweb, River & Lakes and GRLM~~Water level time series of selected lakes and reservoirs from DAHITI, Hydroweb, River & Lakes and GRLM compared with in situ data. For each comparison of water level time series from altimetry with in situ data an RMS difference and squared correlation is computed. The number of points from the final water level time series which were used for the validation is given in the third column (No) of each altimeter dataset.

	DAHITI			Hydroweb			River & Lakes			GRLM		
Lake name - Station name - (DAHITI Id)	RMS [cm]	R²	No	RMS [cm]	R²	No	RMS [cm]	R²	No	RMS [cm]	R²	
Superior - Duluth^1 (3)	**~~5.2~~5.4**	~~0.93~~0.92	~~2810~~3485	~~5.6~~6.6	~~0.95~~0.92	~~279~~229	~~8.5~~9.3	~~0.80~~0.73	82	*~~11.8~~12.4*	~~0.75~~0.72	
Superior - Grand Marais^1 (3)	**~~4.0~~4.4**	~~0.96~~0.95	~~2808~~3483	~~5.6~~5.2	0.95	279	8.5	0.80	82	*11.8*	0.75	
Superior - Marquette^1 (3)	**~~4.5~~4.8**	0.94	~~2810~~3485	~~5.9~~5.5	~~0.94~~0.95	~~279~~229	8.5	0.80	82	*11.9*	0.75	
Superior - Ontonagon^1 (3)	**~~4.1~~4.4**	~~0.96~~0.95	~~2779~~3449	~~6.1~~5.7	~~0.94~~0.95	~~278~~228	8.2	0.82	82	*12.1*	0.74	
Superior - Point Iroqouis^1 (3)	~~6.4~~6.6	~~0.85~~0.84	~~2156~~2619	**6.5**	0.84	~~191~~141	9.5	0.75	82	*12.4*	0.62	
Huron - Essexville^1 (33)	~~8.9~~9.1	~~0.92~~0.93	~~1542~~2048	~~10.9~~11.2	~~0.92~~0.93	~~277~~230	**8.9**	0.80	76	*11.5*	0.90	
Huron - Harbor Beach^1 (33)	**~~4.9~~5.2**	0.98	~~1543~~2049	*~~7.5~~7.7*	0.97	~~277~~230	6.4	0.89	76	6.8	0.96	
Huron - Lakeport^1 (33)	**~~6.0~~6.3**	0.96	~~1482~~1960	*~~8.6~~8.4*	0.96	~~263~~215	7.2	0.86	75	7.6	0.95	
Huron - Mackinaw City^1 (33)	**~~4.3~~4.9**	~~0.98~~0.97	~~1446~~1925	~~5.9~~6.2	0.98	~~255~~208	6.7	0.88	75	*7.7*	0.94	
Huron - De Tour Village^1 (33)	**~~4.3~~4.7**	0.98	~~1510~~2007	*~~6.3~~6.9*	0.98	~~269~~222	6.2	0.89	76	*6.9*	0.96	
Michigan - Calumet Harbor^1 (11)	~~7.3~~7.6	0.95	~~1390~~2045	*~~11.7~~10.5*	~~0.92~~0.94	~~277~~228	**7.5**	0.87	76	8.7	0.94	
Michigan - Holland^1 (11)	~~5.5~~5.6	~~0.92~~0.91	~~1058~~1464	~~7.8~~7.2	~~0.82~~0.84	~~181~~131	**5.3**	0.93	76	*8.3*	0.82	
Michigan - Kewaunee^1 (11)	~~5.2~~5.4	~~0.93~~0.92	~~1013~~1403	~~7.3~~6.7	~~0.85~~0.86	~~176~~124	**5.0**	0.94	73	*8.7*	0.80	
Michigan - Ludington^1 (11)	~~5.2~~5.4	0.93	~~1013~~1448	~~7.3~~6.8	0.85	124	**5.0**	0.94	73	*8.7*	0.80	
Michigan - Milwaukee^1 (11)	~~5.7~~5.9	0.97	~~1409~~2075	*~~10.0~~8.8*	~~0.95~~0.96	~~279~~230	**5.3**	0.93	78	8.4	0.95	
Michigan - Port Inland^1 (11)	**~~5.5~~5.8**	~~0.92~~0.91	~~1034~~1448	~~8.2~~7.4	~~0.82~~0.83	~~181~~130	6.1	0.91	69	*10.1*	0.74	
Erie - Buffalo^1 (6)	**~~11.8~~11.5**	~~0.78~~0.79	~~1315~~1892	~~17.0~~16.4	~~0.70~~0.72	~~265~~221	17.2	0.50	81	*18.7*	0.61	
Erie - Cleveland^1 (6)	**~~5.1~~5.9**	~~0.95~~0.94	~~1294~~1859	~~9.4~~9.0	~~0.91~~0.92	~~257~~213	13.2	0.70	81	*13.6*	0.78	
Erie - Fairport^1 (6)	**~~4.7~~5.2**	~~0.96~~0.95	~~1293~~1858	~~9.5~~8.6	~~0.92~~0.93	~~257~~213	12.6	0.74	81	*13.5*	0.79	
Erie - Fermi Power Plant^1 (6)	**~~10.1~~10.5**	~~0.86~~0.84	~~1258~~1819	~~15.4~~14.9	~~0.76~~0.78	~~255~~212	15.3	0.63	81	*17.0*	0.70	
Erie - Marblehead^1 (6)	**~~8.3~~8.8**	~~0.90~~0.88	~~1260~~1827	~~12.9~~13.4	~~0.83~~0.81	~~251~~210	15.3	0.62	80	*15.9*	0.73	
Erie - Sturgeon Point^1 (6)	**~~10.5~~10.3**	~~0.82~~0.83	~~1294~~1870	~~15.4~~14.8	~~0.75~~0.77	~~262~~218	16.3	0.55	80	*17.5*	0.65	
Erie - Toledo^1 (6)	**~~13.1~~13.3**	~~0.79~~0.77	~~1314~~1888	~~18.4~~18.6	0.69	~~265~~220	17.2	0.57	81	*19.5*	0.64	
Ontario - Cape Vincent^1 (35)	~~5.0~~5.3	~~0.94~~0.96	~~1389~~2089	~~6.8~~6.5	~~0.94~~0.95	~~282~~227	**4.5**	0.97	75	*10.8*	0.85	
Ontario - Olcott^1 (35)	**~~3.9~~4.5**	~~0.98~~0.97	~~1312~~1976	~~6.2~~6.1	~~0.95~~0.96	~~266~~210	4.9	0.96	72	*11.0*	0.85	
Ontario - Oswego ^{1} (35)	~~4.8~~5.2	~~0.97~~0.96	~~1390~~2098	~~6.7~~6.6	~~0.94~~0.95	~~284~~229	**4.6**	0.97	75	*10.8*	0.85	
Ontario - Rochester^1 (35)	~~4.2~~4.6	0.97	~~1391~~2099	~~6.2~~6.1	~~0.95~~0.96	~~284~~229	**4.4**	0.97	75	*10.8*	0.85	
Athabasca - Crackingstone Point^2 (100)	**~~17.0~~15.1**	~~0.88~~0.90	~~1337~~1279	~~33.7~~32.1	0.79	~~272~~224	*80.5*	0.30	79	55.7	0.27	
Great Slave - Hay River^2 (99)	**~~15.9~~13.3**	~~0.61~~0.68	~~1221~~1209	*31.2*	0.37	246	-	-	-	-	-	
Claire - Prairie Point^2 (578)	**~~19.9~~19.6**	~~0.53~~0.37	~~141~~404	-	-	-	*37.9*	0.25	70	-	-	
Winnipeg - George Island^2 (101)	**~~14.6~~11.8**	~~0.81~~0.87	~~522~~778	~~27.7~~28.6	~~0.64~~0.66	~~257~~146	*41.9*	0.49	77	~~30.9~~33.0	~~0.56~~0.59	~~703~~
Winnipeg - Gimli^2 (101)	**~~24.4~~15.6**	~~0.49~~0.79	~~518~~758	~~30.0~~30.1	~~0.55~~0.61	~~264~~147	*~~41.9~~42.4*	0.48	76	~~35.9~~36.2	~~0.42~~0.50	~~722~~
Winnipeg - Pine Dock^2 (101)	**~~19.1~~12.6**	~~0.65~~0.86	~~499~~694	~~28.3~~29.7	~~0.63~~0.67	~~254~~139	*42.7*	0.51	74	~~33.8~~34.6	~~0.50~~0.56	~~702~~
Manitoba - Westbourne^2 (191)	**~~20.0~~13.4**	~~0.77~~0.85	~~318~~499	34.2	0.42	100	34.2	~~0.42~~0.33	73	*46.0*	0.11	
Manitoba - Steep Rock^2 (191)	**~~20.0~~13.3**	~~0.77~~0.85	~~315~~499	36.4	0.40	101	35.5	0.33	75	*47.3*	0.11	
Cedar - Oleson Point ^{2} (200)	**~~37.8~~35.9**	~~0.83~~0.86	~~530~~545	*76.7*	0.20	252	54.8	0.54	78	-	-	
Winnipegosis - Winnipegosis^2 (281)	**~~16.2~~16.5**	0.91	~~426~~469	*36.7*	0.63	136	34.2	0.61	70	36.2	0.53	
Lake of the Woods - Clearwater Bay^2 (73)	**~~16.3~~16.8**	~~0.75~~0.72	~~631~~648	~~44.4~~32.3	~~0.58~~0.46	~~143~~206	*36.6*	0.40	77	-	-	
Lake of the Woods - Cyclone Island^2 (73)	**~~15.0~~16.1**	~~0.78~~0.74	~~631~~642	~~43.3~~31.0	~~0.61~~0.51	~~143~~206	*35.9*	0.41	77	-	-	
Lake of the Woods - Hanson Bay^2 (73)	**~~14.7~~16.3**	~~0.79~~0.73	~~631~~642	~~42.6~~30.4	~~0.63~~0.53	~~143~~207	*36.0*	0.40	77	-	-	
Great Salt - Saltair Boat Harbor^3 (72)	**~~7.1~~7.4**	~~0.89~~0.91	~~27~~44	~~20.2~~20.0	~~0.36~~0.38	35	-	-	-	*29.4*	0.21	
Argentino - Calafate^4 (182)	**~~12.2~~14.6**	~~0.98~~0.97	~~747~~856	*~~23.5~~21.9*	0.93	185	-	-	-	-	-	
Buenos Aires - Los Antiguos^4 (139)	**19.0**	~~0.77~~0.73	~~33~~47	*~~29.2~~29.4*	0.70	19	-	-	-	-	-	

Source of in-situ data: ^{1}NOAA Tides and Currents, ^{2}Canada Wateroffice, ^{3}U.S. Geological Survey (USGS), ^{4}Ministerio de Planificación Federal, República Argentina

to altimeter track length crossing the river based on satellite images from Google Maps. Positive distances indicates downstream gauges, negative differences indicate upstream gauges.

Table 5: ~~Validation of water level time series of DAHITI for selected rivers in the Amazon basin compared with in-situ data and results from LEGOS, and ESA-DMU~~Water level time series of selected rivers of the Amazon basin from DAHITI, Hydroweb and River & Lakes compared with in situ data. For each comparison of water level time series from altimetry with in situ data an RMS difference and squared correlation is computed. The number of points from the final water level time series which were used for the validation is given in the third column (No) of each altimeter dataset. Additionally, the distance to the nearest in-situ station (upstream (+), downstream (-)) and the river width at the crossing altimeter track is shown.

Target name - Station name (DAHITI-ID)	Distance [km]	River width [km]	DAHITI RMS [cm]	DAHITI R²	DAHITI No	Hydroweb RMS [cm]	Hydroweb R²	Hydroweb No	River & Lakes RMS [cm]	River & Lakes R²
Solimões, River - Tabatinga[1] (405)	+ 28.8	~3.8	~~30.6~~39.6	~~1.00~~0.99	~~143~~222	*39.9*	0.99	86	**29.5**	1.00
Solimões, River - Tabatinga[1] (406)	- 23.2	~2.8	**~~28.0~~17.4**	1.00	~~40~~48	-	-	-	*~~1214.2~~119.9*	~~0.45~~0.88
Solimões, River - Tefé[1] (389)	- 14.0	~2.6	**~~11.6~~12.3**	1.00	~~32~~35	-	-	-	*~~173.3~~14.8*	~~0.76~~1.00
Solimões, River - Tefé[1] (581)	+ 23.1	~3.7	**~~21.7~~24.5**	0.99	~~74~~95	*53.9*	0.98	84	-	-
Solimões, River - Itapéua[1] (384)	- 13.9	~4.4	**~~49.9~~33.9**	~~0.98~~0.99	~~39~~40	*~~120.4~~110.9*	~~0.90~~0.91	4439	-	-
Solimões, River - Itapéua[1] (582)	+ 8.9	~2.6	**~~46.0~~31.3**	0.99	~~117~~137	*61.2*	0.97	97	-	-
Purus, River - Aruma-Jusante[1] (583)	- 12.5	~1.4	**~~23.3~~20.0**	1.00	16	24.1	1.00	7	*318.9*	0.61
Jiparaná, River - Tabajara[1] (584)	- 14.3	~0.4	**~~139.2~~113.8**	~~0.88~~0.87	~~43~~47	*335.5*	0.29	33	-	-
Jiparaná, River - Tabajara[1] (585)	+ 2.4	~0.3	**~~49.4~~46.7**	~~0.96~~0.97	~~84~~93	-	-	-	-	-
Japurá, River - Vila Bittencourt[1] (579)	- 40.1	~2.6	~~29.3~~34.0	0.99	24	*67.2*	0.90	24	**31.2**	0.99
Japurá, River- Vila Bittencourt[1] (580)	+ 47.5	~1.9	**~~40.2~~41.0**	~~0.97~~0.98	26	61.3	0.93	25	*115.1*	0.80
São Lourenčo, River - Posada Taiama[1] (1093)	+ ~~3.1~~0.8	~0.3	**~~31.5~~25.4**	~~0.86~~0.91	~~149~~160	-	-	-	-	-
São Lourenčo, River - Posada Taiama[1] (1094)	~~- 9.8~~+ 4.8	~0.3	**~~22.4~~21.7**	~~0.93~~0.94	~~110~~157	-	-	-	-	-
Madeira, River - Humaitá[1] (371)	- 27.6	~2.5	**~~21.6~~19.4**	1.00	35	45.1	0.99	29	*53.2*	0.99
Madeira, River - Humaitá[1] (360)	+ 70.5	~1.5	**~~29.0~~36.3**	0.99	~~93~~173	*50.2*	0.99	91	-	-
Madeira, River - Guajará-Mirim[1] (575)	- 48.7	~2.4	**~~78.4~~75.5**	~~0.88~~0.91	~~44~~36	87.4	0.88	35	*134.3*	0.77
Negro, River - Porto de Manaus[1] (161)	+ 15.5	~10.0	**~~12.7~~7.6**	1.00	~~74~~88	~~21.8~~25.2	1.00	~~72~~79	*~~67.5~~72.0*	0.96
Negro, River - Moura[1] (352)	- 64.8	~4.5	**~~65.3~~60.0**	0.97	62	*70.9*	0.96	43	-	-
Negro, River - Moura[1] (346)	+ 51.8	~18.5	**~~46.0~~43.6**	0.98	42	*46.3*	0.97	45	44.1	0.98
Paraguai, River - Sao Francisco[1] (1095)	- 32.3	~~~6.3~~0.8	**~~26.4~~22.5**	0.96	~~44~~46	-	-	-	-	-

Source of in-situ data: [1] Agência Nacional de Águas (ANA)

The RMS differences between altimeter time series and in situ data vary between ~~12~~8 cm and ~~139~~114 cm in the case of DAHITI. For most virtual stations, the consistency with the gauge is considerably lower than for lakes. It is not possible to prove a dependence between river width and distance to the gauge, not only because of the altimeter time series but also because of the accuracies of the in situ data which also contain measurement errors. Also, the angle in which the satellite track crosses the river has a strong impact on the quality of the water level time series. Furthermore, distances of tenths of kilometres between the in situ station and the nearest crossing altimeter track make it more difficult to prove dependences owed to unpredictable river flow effects.

Compared with time series from ~~LEGOS~~Hydroweb and ~~ESA-DMU~~River & Lakes, the new DAHITI approach can improve the gauge consistency for most of the targets. The improvement can reach several decimetres. Many correlation coefficients in Table 5 are close to one. This is not necessarily an indication of optimal consistency between altimeter water level and gauging observations but is significantly influenced by the large absolute water level variations (more than 10 m).

5 Conclusions

This paper presents a new method for estimating water level time series over inland water using multi-mission satellite altimetry data. It is based on careful data preprocessing (including waveform retracking), a Kalman filter approach, and a rigorous outlier detection. The introduced method is the basis of ~~the "Database for Hydrological Time Series over Inland Water" (DAHITI)~~DAHITI, an online database for inland water level time series from satellite altimetry observations operated by the Deutsches Geodätische Forschungsinstitut der Technischen Universität München (DGFI-TUM).

The study demonstrates the performance of the new method for numerous lakes and rivers in North and South America. A comprehensive validation is performed by comparison with time series of water level variations from in situ gauging stations. Moreover, a comparison with external altimetry-derived water level variations is presented based on data from Hydroweb (LEGOS), the River & Lakes database (ESA-DMU), and the Global Reservoir and Lake Monitor (GRLM).

The lake level data sets computed with the presented approach yield accuracies between 4 cm and ~~38~~36 cm depending on the surface extent of the lake and climate conditions (i.e. ice coverage). For rivers, the performance is considerably lower with RMS differences varying between ~~12~~8 cm and ~~139~~114 cm. Here the accuracy mainly depends on the crossing angle of the altimeter track~~ but also the river width plays a minor role.~~ and the surrounding conditions. Also other surrounding conditions such as topography, quality of waveforms and their retracked water heights can influence the resulting water level time series. Especially in the Amazon basin the river meander can also change over the years because of strong seasonal variations. ~~River width only plays a minor role.~~

For most study cases, the new approach yields significant accuracy improvements compared with water level variations provided by established inland altimeter databases, especially for smaller lakes and rivers. In addition, the temporal resolution of the DAHITI lake time series is significantly improved compared with other data sets, allowing for the detection of sub-monthly temporal changes.

The reasons for the improved performance of the presented approach are multiple: first, a larger observation data set is used as input as a multi-mission concept is realized. All available altimeter missions are cross-calibrated and incorporated in the computations. Second, the applied preprocessing consists ~~of a careful retracking and a robust outlier elimination~~of a robust outlier elimination and optional retracking. This ensures that only highly accurate data will be used. Moreover, the Kalman filter approach permits the optimal combination of all data sets and also includes the accuracies of the input data for weighting. This also enables rigorous error propagation and the computation of formal errors for each water level height. ~~This aspect will be highlighted in a follow-on paper.~~ Further comparisons for the three selected areas show that using the Kalman filter approach instead of a median approach leads to slightly decreased RMS differences. This indicates that the major improvements in the water level times of DAHITI are owed to the extended outlier rejection. In ~~principle~~future, the Kalman filter approach ~~also allows~~will also be used for (near) real-time analysis and integration of altimeter data (with the so-called Operational Geophysical Data Record, OGDR). This enables daily actualization of the water level time series and may also be used for short-time predictions. Furthermore, the introduction of a dynamic model in the Kalman filter will cause an increase in the temporal resolution of the water level time series. For the development of the dynamic model external data sets such as GRACE, precipitation, etc. can be used.

In spite of the improved water level time series of DAHITI compared with results from Hydroweb, River & Lakes and GRLM, there are still some challenging tasks which have to be taken into account to make further improvements. Retracking is the most challenging task in using altimeter data for smaller water bodies. The mixture of different waveform shapes such as ocean-like, specular, etc. makes it difficult to choose a suitable retracking algorithm. Each retracker is optimized for special waveform shapes, but switching the retracking algorithm to achieve the best ranges will lead to retracker biases which have to be taken into account. Furthermore, inter-mission offsets can also arise because of the different characteristics of the measurement systems (e.g. Ku-band (Envisat) and Ka-band (SARAL/AltiKa)).

All presented water level time series as well as results for many additional targets are freely available in the ‘Database for Hydrological Time Series over Inland Water’ (DAHITI) at http://dahiti.dgfi.tum.de.

Acknowledgements. The altimeter observations and geophysical corrections are taken from OpenADB (*http://openadb.dgfi.tum.de*). The altimeter missions are operated and maintained by CNES/NASA (Topex, Jason-1), ESA (ERS-1/2, Envisat), CNES/NASA/Eumetsat/NOAA (Jason-2), and ISRO/CNES (SARAL/~~AltiKA~~AltiKa) and the original data sets are disseminated by AVISO, ESA, and PODAAC.

We thank Agência Nacional de Águas (ANA), National Oceanic and Atmospheric Administration (NOAA), government of Canada, U.S. Geological Survey (USGS) and Ministerio de Planificación Federal, República Argentina for providing the in situ data.

We also thank the reviewers for helpful comments and suggestions which helped us to improve the quality of this paper.

References

Berry, P., Jasper, A., and Bracke, H.: Retracking ERS-1 altimeter waveforms over land for topographic height determination: an expert system approach, ESA Pub. SP-414, 1, 403–408, 1997.

Berry, P. A. M., Garlick, J. D., Freeman, J. A., and Mathers, E. L.: Global inland water monitoring from multi-mission altimetry, Geophysical Research Letters, 32, n/a–n/a, doi:10.1029/2005GL022814, http://dx.doi.org/10.1029/2005GL022814, l16401, 2005.

Birkett, C., Reynolds, C., Beckley, B., and Doorn, B.: From Research to Operations: The USDA Global Reservoir and Lake Monitor, in: Coastal Altimetry, edited by Vignudelli, S., Kostianoy, A. G., Cipollini, P., and Benveniste, J., pp. 19–50, Springer Berlin Heidelberg, doi:10.1007/978-3-642-12796-0_2, 2011.

Birkett, C. M.: The contribution of TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes, Journal of Geophysical Research: Oceans, 100, 25 179–25 204, doi:10.1029/95JC02125, 1995.

Boehm, J., Kouba, J., and Schuh, H.: Forecast Vienna Mapping Functions 1 for real-time analysis of space geodetic observations, Journal of Geodesy, 83, 397–401, doi:10.1007/s00190-008-0216-y, 2009.

Bosch, W., Dettmering, D., and Schwatke, C.: Multi-Mission Cross-Calibration of Satellite Altimeters: Constructing a Long-Term Data Record for Global and Regional Sea Level Change Studies, Remote Sensing, 6, 2255–2281, doi:10.3390/rs6032255, 2014.

Boser, B. E., Guyon, I. M., and Vapnik, V. N.: A Training Algorithm for Optimal Margin Classifiers, in: Proceedings of the Fifth Annual Workshop on Computational Learning Theory, COLT ’92, pp. 144–152, ACM, New York, NY, USA, doi:10.1145/130385.130401, 1992.

Brown, G.: The average impulse response of a rough surface and its applications, Antennas and Propagation, IEEE Transactions on, 25, 67–74, doi:10.1109/TAP.1977.1141536, 1977.

Challenor, P. and Srokosz, M.: The extraction of geophysical parameters from radar altimeter return from a non-linear sea surface, Mathematics in remote sensing, pp. 257–268, 1989.

Chelton, B., Ries, J., Haines, B., Fu, L., and P.S., C.: Satellite Altimetry, in: Satellite Altimetry and Earth Sciences: A Handbook of Techniques and Applications, edited by Fu, L. and Cazenave, A., pp. 1–131, Academic Press, 2001.

Crétaux, J.-F. and Birkett, C.: Lake studies from satellite radar altimetry, Comptes Rendus Geoscience, 338, 1098 – 1112, doi:10.1016/j.crte.2006.08.002, la Terre observée depuis l'espace Observing the Earth from space, 2006.

Crétaux, J.-F., Calmant, S., Romanovski, V., Shabunin, A., Lyard, F., Bergé-Nguyen, M., Cazenave, A., Hernandez, F., and Perosanz, F.: An absolute calibration site for radar altimeters in the continental domain: Lake Issykkul in Central Asia, Journal of Geodesy, 83, 723–735, doi:10.1007/s00190-008-0289-7, http://dx.doi.org/10.1007/s00190-008-0289-7, 2009.

Crétaux, J.-F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.-C., Nino, F., Rio, R. A. D., Cazenave, A., and Maisongrande, P.: SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, Advances in Space Research, 47, 1497 – 1507, doi:10.1016/j.asr.2011.01.004, 2011.

da Silva, J. S., Calmant, S., Seyler, F., Filho, O. C. R., Cochonneau, G., and Mansur, W. J.: Water levels in the Amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions, Remote Sensing of Environment, 114, 2160 – 2181, doi:10.1016/j.rse.2010.04.020, 2010.

de Oliveira Campos, I., Mercier, F., Maheu, C., Cochonneau, G., Kosuth, P., Blitzkow, D., and Cazenave, A.: Temporal variations of river basin waters from Topex/Poseidon satellite altimetry. Application to the Amazon basin, Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science, 333, 633 – 643, doi:10.1016/S1251-8050(01)01688-3, 2001.

Eicker, A., Schumacher, M., Kusche, J., Döll, P., and Schmied, H.: Calibration/Data Assimilation Approach for Integrating GRACE Data into the WaterGAP Global Hydrology Model (WGHM) Using an Ensemble Kalman Filter: First Results, Surveys in Geophysics, 35, 1285–1309, doi:10.1007/s10712-014-9309-8, http://dx.doi.org/10.1007/s10712-014-9309-8, 2014.

Förste, C., Bruinsma, S., Flechtner, F., Marty, J., Lemoine, J., Dahle, C., Abrikosov, O., Neumayer, H., Biancale, R., Barthelmes, F., and Balmino, G.: A new release of EIGEN-6C, presented AGU Fall Meeting 2012, San Francisco 2012, 2012.

Global Runoff Data Center: Long-Term Mean Monthly Discharges and Annual Characteristics of GRDC Station / Global Runoff Data Center, Koblenz, Germany: Federal Institute of Hydrology (BfG), 2013.

Gruber, C., Moon, Y., Flechtner, F., Dahle, C., Novák, P., König, R., and Neumayer, H.: Submonthly GRACE Solutions from Localizing Integral Equations and Kalman Filtering, in: Earth on the Edge: Science for a Sustainable Planet, edited by Rizos, C. and Willis, P., vol. 139 of *International Association of Geodesy Symposia*, pp. 383–389, Springer Berlin Heidelberg, doi:10.1007/978-3-642-37222-3_51, http://dx.doi.org/10.1007/978-3-642-37222-3_51, 2014.

Guha-Sapir, D. and Vos, F.: Quantifying Global Environmental Change Impacts: Methods, Criteria and Definitions for Compiling Data on Hydro-meteorological Disasters, in: Coping with Global Environmental Change, Disasters and Security, edited by Brauch, H. G., Oswald Spring, U., Mesjasz, C., Grin, J., Kameri-Mbote, P., Chourou, B., Dunay, P., and Birkmann, J., vol. 5 of *Hexagon Series on Human and Environmental Security and Peace*, pp. 693–717, Springer Berlin Heidelberg, doi:10.1007/978-3-642-17776-7_40, 2011.

Hwang, C., Guo, J., Deng, X., Hsu, H.-Y., and Liu, Y.: Coastal gravity anomalies from retracked Geosat/GM altimetry: improvement, limitation and the role of airborne gravity data, Journal of Geodesy, 80, 204–216, 2006.

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, Transactions of the ASME–Journal of Basic Engineering, 82, 35–45, 1960.

Maillard, P., Bercher, N., and Calmant, S.: New processing approaches on the retrieval of water levels in Envisat and SARAL radar altimetry over rivers: A case study of the São Francisco River, Brazil, Remote Sensing of Environment, 156, 226 – 241, doi:10.1016/j.rse.2014.09.027, 2015.

Martin, T. V., Zwally, H. J., Brenner, A. C., and Bindschadler, R. A.: Analysis and retracking of continental ice sheet radar altimeter waveforms, Journal of Geophysical Research: Oceans, 88, 1608–1616, doi:10.1029/JC088iC03p01608, 1983.

McCarthy, D. and Petit, G.: IERS Conventions 2003 (IERS Technical Note ; 32), Frankfurt am Main: Verlag des Bundesamtes für Kartographie und Geodäsie, p. 127 pp., 2004.

Ponchaut, F. and Cazenave, A.: Continental lake level variations from Topex/Poseidon (1993–1996), Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science, 326, 13 – 20, doi:10.1016/S1251-8050(97)83198-9, 1998.

Ričko, M., Birkett, C. M., Carton, J. A., and Crétaux, J.-F.: Publisher's Note: Intercomparison and validation of continental water level products derived from satellite radar altimetry, Journal of Applied Remote Sensing, 6, 060 103–060 103, doi:10.1117/1.JRS.6.060103, http://dx.doi.org/10.1117/1.JRS.6.060103, 2013.

Scharroo, R. and Smith, W. H. F.: A global positioning system-based climatology for the total electron content in the ionosphere, Journal of Geophysical Research: Space Physics, 115, n/a–n/a, doi:10.1029/2009JA014719, 2010.

Schwatke, C., Dettmering, D., and Boergens, E.: Potential of SARAL/AltiKa for inland water applications, Marine Geodesy, n/a, n/a–n/a, doi:10.1080/01490419.2015.1008710, 2015.

Seyler, F., Calmant, S., da Silva, J. S., Moreira, D. M., Mercier, F., and Shum, C. K.: From TOPEX/Poseidon to Jason-2/OSTM in the Amazon basin, Advances in Space Research, 51, 1542 – 1550, doi:10.1016/j.asr.2012.11.002, satellite Altimetry Calibration and Deformation Monitoring using GNSS, 2013.

Smola, A. J. and Schölkopf, B.: A Tutorial on Support Vector Regression, Statistics and Computing, 14, 199–222, doi:10.1023/B:STCO.0000035301.49549.88, 2004.

Stakhiv, E. and Stewart, B.: Needs for Climate Information in Support of Decision-Making in the Water Sector, Procedia Environmental Sciences, 1, 102 – 119, doi:http://dx.doi.org/10.1016/j.proenv.2010.09.008, http://www.sciencedirect.com/science/article/pii/S1878029610000095, world Climate Conference - 3, 2010.

Wingham, D., Rapley, C., and Griffiths, H.: New techniques in satellite altimeter tracking systems, in: ESA Proceedings of the 1986 International Geoscience and Remote Sensing Symposium(IGARSS'86) on Remote Sensing: Today's Solutions for Tomorrow's Information Needs,, vol. 3, 1986.

Yang, Y. and Gao, W.: An Optimal Adaptive Kalman Filter, Journal of Geodesy, 80, 177–183, doi:10.1007/s00190-006-0041-0, http://dx.doi.org/10.1007/s00190-006-0041-0, 2006.

Zakharova, E. A., Kouraev, A. V., Cazenave, A., and Seyler, F.: Amazon River discharge estimated from TOPEX/Poseidon altimetry, Comptes Rendus Geoscience, 338, 188 – 196, doi:10.1016/j.crte.2005.10.003, 2006.

Anonymous Referee #1
Received and published: 25 May 2015

MAJOR COMMENTS

This manuscript is about retrieving inland water level information using radar altimeters from satellites. A new processing schema is proposed based on 1) a careful data pre-processing (including waveform retracking), 2) a Kalman filter approach that incorporates cross-calibrated multi-mission altimeter observations with their uncertainties and 3) a rigorous outlier detection strategy. This processing schema is contributing to populate the new archive called “Database for Hydrological Time Series of Inland Water” (DAHITI) at DGFI-TUM. The performances of the new processing are here assessed in a number of lakes and rivers in North and South America. The authors compare their water level time series with available ground truth and other similar altimeter-derived water level products (e.g., Hydroweb, River & Lake, GRLM). The results show that with the new processing inland water height information is more accurate than that available from the other established inland altimeter services (i.e., Hydroweb, River & Lake, GRLM).

Overall, this manuscript presents a novel method to process altimeter data in inland waters that appears to be very effective for smaller lakes and rivers. It is clear that exploiting all available satellite missions is crucial to construct accurate water level time series, although inter-mission biases must be carefully taken into account.

The new method is clearly documented and data analysis is sufficiently complete. The results from the comparison statistics made at a good number of water targets are well discussed. Three case-studies (Lake Superior, Lake Athabasca and Rio Madeira) are also commented in detail.

What follows are some remarks:

- The authors try to explain the observed disagreements in the various comparisons, however, the discussion about the possible causes is not sufficiently supported by rigorous explanations. A strong recommendation for the authors is to better interpret the results in the three case-studies with reference to all possible reasons that might explain disagreements with ground truth and with the other remote sensing products.

Planned updates in revised version:

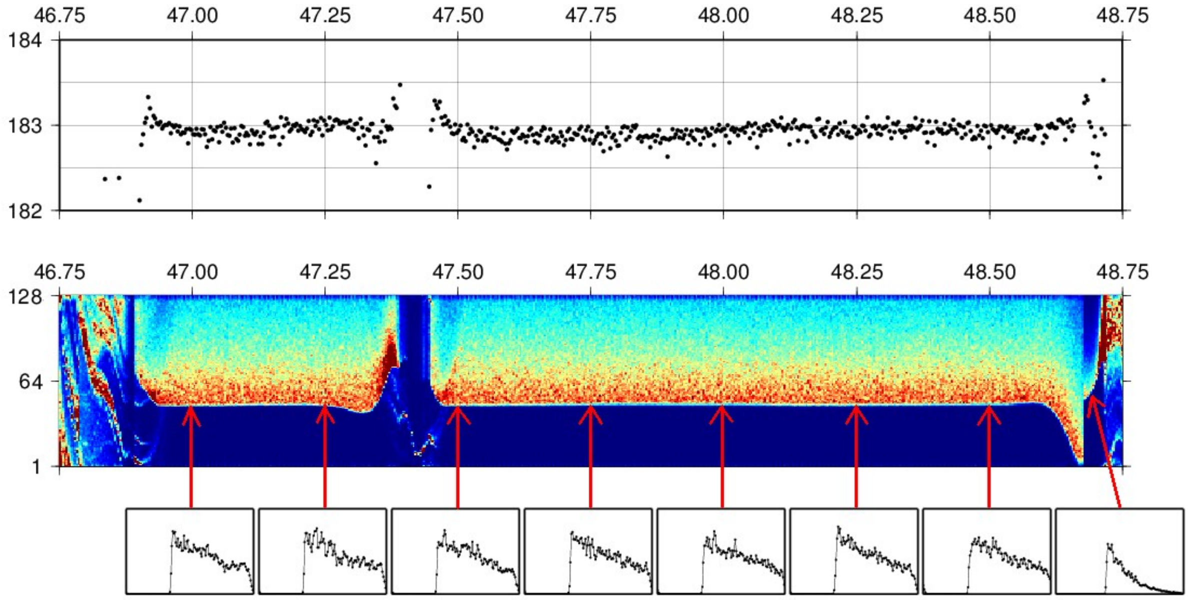
- The discussion of the observed disagreements will be extended for the three selected results in Figure 7,8,9.
- Lake Superior:
  - Ice coverage, wind & wave, geoid errors
- Lake Athabasca:
  - Ice coverage, land contamination, impact of mountains near the lake shore, retracker biases, hooking effect
- Amazon River:
  - land contamination, river slope, hooking effect, large footprints can lead to measuring river branches, retracker biases

- The effect of wind and wave fields in the case of Lake Superior should be investigated in order to prove the assumption that the level of the lake is constant. I am also not convinced that all waveforms in this big lake when the satellite is far from land conform to Brown model. I would expect to see Brown, specular and mixed. The authors should provide some figures about this classification from which follows the choice of the appropriate retracker.

We agree that larger lakes do not have a constant surface but hydrodynamic effects from wind and waves have to be taken into account. The sentence “The derived water levels are assumed to be constant over lakes since in general, the water is in balance with gravity and hydrodynamics of lakes is small compared to open ocean conditions” is part of the manuscript since we compute only one height per time step which is assumed to represent the whole water surface. However, in our computations, we take different water levels into account by defining a grid over the surface and estimate heights for each grid point. Thus, different hydrodynamic conditions are considered in the processing.

For larger lakes, we are using the ocean products of the different altimeter mission instead of retracking the data by ourselfs since the data qualitiy is sufficient. . Therefore, retracking is applied only for smaller targets in case the ocean product does not have a good quality.

The image shows that most waveforms are brown-like. The figure shows an altimeter track of Envisat (Pass 0510, Cycle 040) crossing the Lake Superior. The water heights (top) and the corresponding waveforms (bottom) are plotted. Most waveforms over the lake have brown-like shapes. Near land water transitions the trailing edge changes to an exponential shape which can be seen in the last waveform near 48.625°E



Planned updates in revised version:

- We will further explain and discuss the statement/assumption of the “constant” lake surface

The possible presence of ice in the case of Lake Athabasca could be verified looking at backscatter observations.

We will do this. The backscatter coefficient will be implemented as additional outlier criteria for measurements affected by ice coverage in the processing strategy of DAHITI.

Planned updates in revised version:

- We will explain the usage of the backscatter coefficient as additional outlier criteria to reject measurements affected by ice coverage

- The ground tracks are generally not located near water gauge. This means that the two systems observe different water dynamics. This is especially true in rivers where there is some regulation of the flow. There is a need to better characterize the observational context that in the paper is not done. Auxiliary data sources (e.g., optical imagery, meteo data, etc.) could help in this exercise.

We agree that distances between gauging station and altimeter track can influence the river level agreement at different locations. This is already discussed in the manuscript in chapter “4.2 Validation data sets”. These discrepancies will of course map in the validation results. The interpretation of additional optical images might help in interpreting the influence of these effects but are beyond the scope of the paper.

- The bibliography is well cited. Some new manuscripts of interest are suggested, e.g., Surajit Ghosh, Praveen Thakur, Vaibhav Garg, Subrata Nandy, Shivprasad Aggarwal, Sudip Kumar Saha, Rashmi Sharma & S. Bhattacharyya (2015): SARAL/AltiKa Waveform Analysis to Monitor Inland Water Levels: A Case Study of Maithon Reservoir, Jharkhand, India, Marine Geodesy, DOI: 10.1080/01490419.2015.1039680; Jean-François, Crétaux, et al. "Global surveys of reservoirs and lakes from satellites and regional application to the Syrdarya river basin." Environmental Research Letters 10.1(2015): 015002.

Crétaux, et al. (2015) is more related to volume estimation and will not be added as additional reference.

Planned updates in revised version:

- Sharma & S. Bhattacharyya (2015) will be added to our paper

- The title does not clearly reflect the content of the paper. It seems that the only improvement is due to Kalman filter, while it was clear from the text that there are other two important processing steps (waveform retracking and outlier detection). Maybe the author can make an attempt to modify a bit the title to reflect the content of the paper as a whole.

After some investigation we can say that the current implementation of the Kalman Filter without any dynamic model is not the major source for the improvements but the applied outlier detection. Therefore, we decided to change the title of the paper.

Planned updates in revised version:

- The title will be updated, probably to “DAHITI – an innovative approach for estimating water level time series over inland water using multi-mission satellite altimetry”

- Figures 7, 8, 9 are difficult to interpret without zooming out. Author are aware and in fact one year is shown separately. The variability in Lake Superior is around 50-60 cm over the selected time period. It is higher in Lake Athabasca, with some evident inter-annuality (2 meters around 1996-1998). The variability in the river is very high (15m) and no inter-annuality is observed. Is this behaviour realistic, even though the ground truth confirms?

We think that this behavior is realistic. In the case of the Amazon river, an inter-annuality can also be detected (e.g. Between 2009 and 2010). However, it is quite small compared to the large variations of almost 15m and can hardly be seen in the figure.

Planned updates in revised version:

- The figures 7,8,9 will be improved to show the results more clearly

- The manuscript is written with understandable English and very few typos, however, the fluency of the text should be improved with the help from a native English.

Planned updates in revised version:

- We will try to further improve the use of English language.

In summary, I invite the authors to follow the above recommendations and expand the discussion of results, especially in the three case-studies, also where possible with the support of auxiliary information (bibliography, other data sources, etc.).

This study is certainly of interest to the inland water altimetry community with reference to the new processing method, but also to hydrology scientists that could exploit the water level time series in their research studies.

Therefore, the manuscript calls for some revision before to become publishable. I like having a look at the revised manuscript and authors' answers.

MINOR COMMENTS

We will not comment on each point here. All minor corrections will be addressed.

Pg. 2, row 2, "
...
for observing inland water levels of lakes and rivers" – I suggest to rephrase as water levels can be retrieved from reservoirs, wetlands and in general any inland water body, although the radar altimetry technique has been especially applied to rivers and lakes
Update from "However, since some years, this technology is also used for observing inland water levels of lakes and rivers." to "However, for some years, this technology is also used for observing inland water levels of lakes, rivers, and other inland water bodies."

Pg. 2, row 16, "height", I suggest to be homogeneous in the text in using "water height" or "water level"
All "water level heights" will be changed to "water levels"

Pg. 2, row 19, "important task", Who states that ? please refer to bibliography
The following bibliography will be added:

- E. Stakhiv, B. Stewart, Needs for Climate Information in Support of Decision-Making in the Water Sector, Procedia Environmental Sciences, Volume 1, 2010, Pages 102-119, ISSN 1878-0296, http://dx.doi.org/10.1016/j.proenv.2010.09.008.

Pg. 2, row 23-24, "the number of in-situ stations monitoring river discharge is globally declining", Is the number stable or decreasing for lakes too? If you provide the infor for rivers you need also to say something for lakes, otherwise if lakes are well monitored with ground truth there is no need to use altimetry.

Information about the status for in-situ data of lakes will be added

pg. 3, row 4, "water level heights" – it doesn't make sense to say at same time level and height. Again, please be uniform in the text using "water height" or "water level"
and hereinafter correct all occurrences
All "water level heights" will be changed to "water levels"

pg. 3, row 9, "then", change to "than"

pg.3, row 10, "its measurement geometry providing measurements", please rephrase, it is not a problem of geometry but rather than configuration and trade-off between revisiting and coverage

pg. 3, row 12, "not all water bodies can be captured" – I suggest to explain that the touching is by chance, however, big water bodies have more probabilities to be passed.

pg. 3, row 20, "still is", change to "is still"

pg. 3, row 22, “twofold”, change to “two-fold”. A third effect might be due to specular returns with non-brownian response of the target. This may happen frequently in small rivers.
We agree that there are specular waveforms over smaller rivers but the reason for these waveforms is land contamination.

pg. 4, row 2, “The affected waveforms are more peaky” – This statement could confuse the reader. The land can interfere in early or late gates depending on the position of track with respect to the water target (see Abileah, R., et al. "Coherent ranging with Envisat radar altimeter: A new perspective in analyzing altimeter data using Doppler processing." Remote Sensing of Environment 139 (2013): 271-276.).
Phrasing will be updated from “The affected waveforms are more peaky and reliable heights cannot be derived using ocean waveform retrackers” to “The affected waveforms do not have brown-like shapes and cannot be retracked by using ocean waveform retrackers.”

pg. 5, rows 9-10, “Time series of lakes and reservoirs” – add “water level” before “time”

pg. 5, rows 20-26, please rephrase the aim of this paper (that is a new approach to retrieve water heights) that has to be clear to the reader. Comparisons with approaches used in other hydrospace services have to be discussed later. Follow-up work has to be mentioned in the conclusions
In our opinion, the comparison with the other approach should remain in the introduction because it reflects the current state of the art.

Planned updates in revised version:

- We will rephrase to aim of the paper in more detail
- Follow-up work will be moved into the conclusion/outlook

pg. 6, row 8, “Section”, please be homogenous writing always “Section” or “Sect.”

pg. 6, row 10, “The paper finishes with a conclusion”, I don’t like this sentence, please rephrase

pg. 6, row 12-13, “For more than two decades, satellite altimetry has been providing data for various applications over ocean and inland waters”, please remove as already stated in the introduction

pg. 6, row 13-16, “The approach presented in this paper combines as many as possible altimeter tracks from different missions over an investigated water body in order to increase the temporal resolution of the final water level time series, to maximize the probability to cover smaller inland waters, and to increase the accuracy” – In this section the reader expects description of data used and then processing applied. The previous sentence is somewhat to be placed in the introduction (what, why).
We think that the information about our multi-mission approach and the resulting increase of the temporal resolution should be highlighted (repeated)also in the data section because it is strongly related to the data which is introduced afterward.

pg. 6, row 19, “body”, change to “bodies”

pg. 6, row 23-27 and pg. 7, rows 1.3, Info about revisiting time a cross track separation was already provided in the introduction. I recommend to avoid repeating things, suggesting removal from introduction, where are not key to understand (one can simply say that revisiting is of order of ten days and more, along track coverage is dense and there gaps between tracks)

pg. 7, row 15-17, can you show some examples of waveforms with reference to the areas of investigation ?

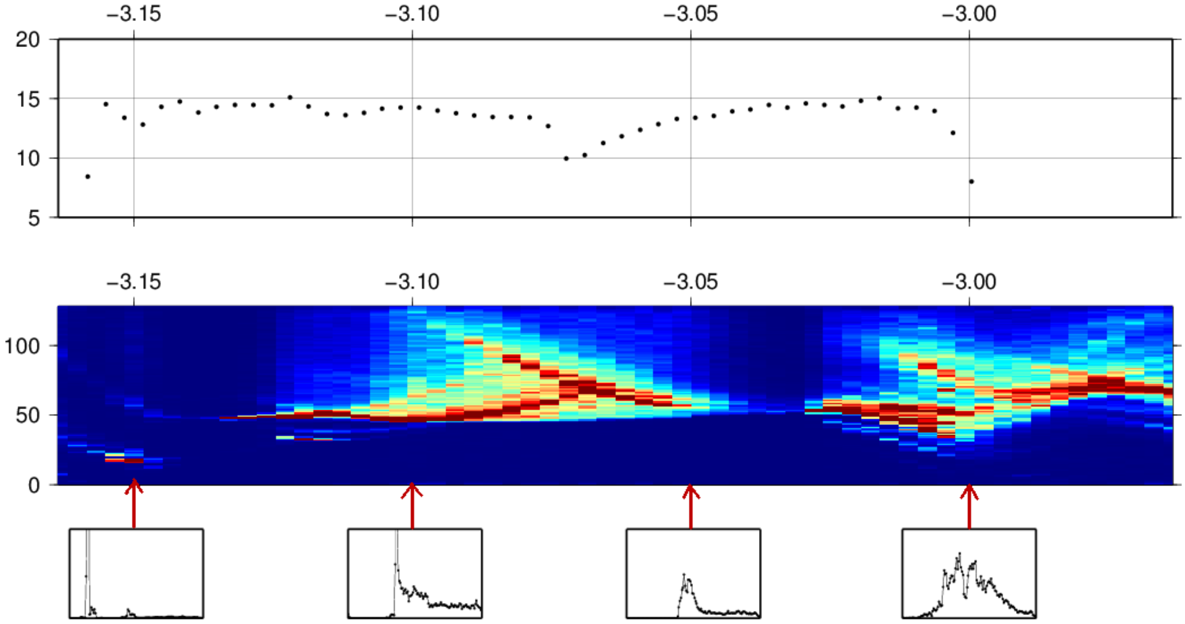


Figure 2. Example of a crossing satellite track from Envisat (Cycle 040, Passnr 0564) over the Amazon river with water heights (top) and waveforms (middle). Additionally four waveforms are shown in detail.

Planned updates in revised version:

- We don’t want to add additional images about waveforms in the paper because this is well-known in the community. Instead, we will add a reference which shows different waveforms over inland water

pg. 8, row 6, “conventions”, maybe you wanted to say “corrections”

pg. 9, rows 2-3, “The derived water levels are assumed to be constant over lakes since in general, the water is in balance with gravity and hydrodynamics of lakes”. This can be a reasonable assumption for small water bodies. The wind set-up can pile-up water somewhere in lakes and the water level cannot be assumed constant. I think the point needs to be explained here and then discussed if supposed relevant in selected water targets (e.g., Lake Superior).
The impact of wind and currents on the lake surface will be discussed.

pg.9, row 24, “water level height”, see previous note

pg. 10, row 16, “a floating box of 5 data points” – why defining a window of 5 points ? is there any reason ? why not using consecutive points to measure the noise ?
We improved our “error” estimation for each measurement. Now we, are using sliding boxes along the satellite track of +-3.5 km for large lakes, +-1.5 km small lakes/large rivers and +-0.5 km for small rivers. The definition of the floating box in km instead of number of points makes this results more comparable with other missions which are measuring in 10Hz, 20Hz, or 40Hz. Because the number of measurements can vary within a box but the used track length is constant.
Moreover, instead of calculating standard deviations within each box, we are estimating a median of the water heights. Then the median height is subtracted from the current water height. The absolute

value of the difference is then used as “error” of the measurement. These method lead to reliable errors than the former approach if more than the half of the used point are over water.

Planned updates in revised version:

- The new error estimation method will be explained

pg. 12, row 11, “Figure”, please be homogeneous in the text (Fig. or Figure)

pg. 12, row 5, Which land mask are you using ? how much is the resolution ?
The used land mask is provided within the Generic Mapping Tool (GMT). GMT uses the Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG). For cases that investigated inland bodies are not included in the dataset or have an unuseful shape, no mask is applied for the computation. I don’t know the resolution but the data is provided as shape file which will be rastered for our purpose.

pg. 13, row 15, missing “.”

pg. 14, row 20, remove “.”

pg. 17, row 12, “LEGOS, ESA-DMU or GRLM”, please use the exact acronyms of these hydrospace services and not the institutions who developed them.

Pg. 17, row 25, “The Great Lakes show seasonal variations of about 1m” – please provide reference

Pg. 18, row 24, “extend”, change to “extent”

pg. 19, row 17, “LEGOS, ESA-DMU, see previous note

pg. 20, row 3, “show”, change to “shows”

pg. 20, row 7, “preformed”, change to “performed”

pg. 20, row 10, “LEGOS, ESA-DMU, see previous note

pg. 20, row 24, “LEGOS, ESA-DMU, see previous note and correct all occurrences hereinafter

pg. 20, row 28, “For a detailed few, results”, please rephrase

pg. 21, row 9, “RMS differences show is”, please rephrase

pg. 23, rows 16-18, “Since only one DAHITI time series is computed per lake, these variations demonstrate uncertainties of the in-situ data sets” – I am not convinced about this statement. For big lakes we can have discrepancies due to metocean effects. This is a key point to be investigated.

I agree that the different RMS differences between the single DAHITI time series and different in-situ data can not only be explained by uncertainties of in-situ data. For large lakes effects such as wind can influence the consistency to used in-situ station.

Planned updates in revised version:

- We will explain that the discrepancies of RMS from different in-situ stations are not only due to the in-situ data but also to wind or currents which can lead to a rougher surface over larger lakes.

Pg. 24, row 16, “Certainly this is not only due to the altimeter time series but also caused by the accuracies of the in-situ data” - The authors here are a bit speculating as there is no proof that ground truth is not accurate.

All data which are measured in any way contain measurement errors. Also in-situ data contain measurement errors which are mostly unknown (since gauging stations do not provide any error estimates). Therefore, we can not suppose that in-situ are perfect.

The sentence will be updated from “Certainly this is not only due to the altimeter time series but also caused by the accuracies of the in-situ data.” to “Certainly this is not only due to the altimeter time series but also caused by the accuracies of the in-situ data which also contain measurement errors.”

Anonymous Referee #2

Received and published: 1 June 2015

The manuscript by Schwatke et al. is of major importance since it presents new results for the monitoring of lakes and lakes that look to be a dramatic improvement (much lower rms in comparison with in-situ series) with respect to existing databases. Consequently, the paper deserve publication but it is not acceptable in the present state. Major points have to be improved.

The presentation of the results is not clear in the sense that their methodology merges three aspects:
1 - the use of a "homemade" retracking of the radar echoes
2 - replacement of the true data value by one predicted through a Kalman filtering.
3 - fine rejection of the outliers

And It is absolutely necessary that the part played by each of these three points in the final improvement is better stated since these 3 points more or less make the difference between the existing website cited in the paper (some use GDR retracking when other use their own retracking, some have a refined detection of the outliers in the raw data when others are more loose on this question, some publish all the values obtained after the post-processing step when other reject dubious values, ...). In other word, Authors must show how much improvement is due in the present study to the retracking of the radar waveform, how much is due to the Kalman filter ?, and how important is the choice of the "valid" points ?

We agree that analyzing the importance of the different aspects of our approach will improve the paper. Thus, we will show the impact of the different processing steps in more detail. We will separately analyze the impact of the pre-processing (i.e. retracking and outlier detection) and the kalman filter approach. We will not handle the retracking separately since we did no innovative work here: we simply apply the existing “Improved Threshold Retracker” to achieve reliable heights for missions such as Jason-1 and smaller water bodies. Furthermore, it is shown in Table 3 that the usage of the ocean product (e.g. Lake Argentino, Lake Buenos Aires, etc) will also improve the resulting RMS and R² compared to the other databases. This shows that the retracking has only a small impact on the improvements of the time series.

Planned updates in revised version:

- Height estimation by using a median approach instead of the and Kalman Filter will be performed for the three selected study cases to demonstrate the impact of the Kalman Filter on the resulting time series
- Additional explanation of the impact of different applied outlier criteria on the resulting time series will be added.
- Additional information which outlier criteria was applied will be added to Table 3 and 4.

According to the Title, the Kalman filter seems to be the major source of improvement. If it happens that the retracking algo and/or outlier detection play an important part in the improvement, maybe the title should be changed to take it into account.

After some investigation we can say that the current implementation of the Kalman Filter without any dynamic model is not the major source for the improvements but the applied outlier detection. Therefore, we decided to change the title of the paper.

Planned updates in revised version:

- The title will be probably updated to "*DAHITI – an innovative approach for estimating water level time series over inland water using multi-mission satellite altimetry"*

The question of the biases is not clear. Authors should -as least briefly- state how they evaluated the biases for the tracker that they use, for each mission (were they estimated globally prior the computation of the series?, are they evaluated separately for each series merging several missions?), And -in the case of global values- publish the values.

We apply range biases which are derived from a global multi-mission crossover analysis described in Bosch et al. (2014). For the multi-mission crossover analysis, the ocean products from the different altimeter missions were used. For our application, the radial errors were interpolated (due to missing crossover points) over land to provide an individual range bias correction for each altimeter measurement. The global mean range biases are already published in our manuscript (in Table 1).

This approach works quite well (not only over ocean but also over land) as long as the ocean product is used for the computation of water levels. However as soon as retracking is involved retracker offsets will occur. In order to minimize the relative offsets between different missions and altimeter tracks we use the same retracker (with identical parameters) for all measurements over one target. That minimizes the inter-mission biases. Remaining differences may still be present in the time series (e.g. due to mission-dependent or location-dependent effects such as waves). These inconsistencies will map into the quality of the time series and will show up by increased RMSE when comparing the altimetry product with in-situ data. This is a point which needs additional investigation in the future.

Planned updates in revised version:

- The explanation on handling range biases will be extended
- The problem of remaining inter-mission offset will be discussed additionally in the outlook

The authors mention in the Introduction the key point of the slant measurements (off-nadir measurements). But they do not explain how they deal with it. Do they ignore it ? If not, where is it corrected for? In the pre-processing step ? How is it modelled (best fitting parabola, parabola constrained by geometrical considerations?, other analytical expression?, etc ..)

In the current processing strategy, the hooking effect is not corrected but the effected measurements are excluded from further processing. The restriction of measurement latitude within the preprocessing step limits the usage of off-nadir observations. Furthermore, the additional computed input uncertainty for each altimeter measurement will increase for off-nadir observations. Thus, only few off-nadir measurements will be used within the approach and these will have larger uncertainties.

In addition, we are about to publish an advanced approach for correcting the hooking effect. We plan to integrate this approach in the DAHITI software in the future.

Planned updates in revised version:

- An explanation on the handling of the hooking effect will be added

Legends in the Tables are not complete. For example, in Table 4, what does the N stand for (number of cycles ?) ? Besides, it would be useful to indicate the mission (Jason-2, Envisat, ...). Also, the river widths do not correspond to actual widths of the reaches. Is it the length of the track segment ? Figures showing comparison of time series are not easy to read. For exp, in Figure 9, it is not possible to see if points are missing or hidden by others. Authors should seek for better way to show all the series in a single view (points in the background larger than the points in the foreground, or use different symbols ....)

Planned updates in revised version:

- The legends will be extended
- The used altimeter mission will be added to Table 3 and 4.
- A definition of the river width will be added (it is the length of the track segment based on Google Maps)
- The Figures 7, 8 and 9 will be updated to make them easier to read and more understandable

I don’t comment the English which is better than mine .

Anonymous Referee #3
Received and published: 3 June 2015

The authors present an interesting approach for deriving water heights over inland water bodies utilizing the Kalman filter method and increasing the temporal resolution of inland water time series by combining multiple altimeter missions that cover a target of interest. The resulting time series for various inland water bodies are made available in the Database for Hydrological Time Series over Inland Water (DAHITI), maintained by the Deutsches Geodätisches Forschungsinstitut (DGFI). Generally, I like the idea to employ a Kalman filter technique in the processing of inland water heights and the increase in temporal coverage by combining different missions. However, I think that the proposed Kalman filter approach is quite simple and some details of the methodology should be made clearer, especially why a detailed hexagonal grid is needed for the whole water body instead of just utilizing the measurement positions (or some fixed bins along the groundtrack) of the altimeter itself. Therefore, I generally recommend the paper for publication since the DAHITI database is an interesting addition to the available range of databases, but after addressing the major issues below.

Thanks a lot for your very valuable comments. Based on this, we did some important changes – not only of the manuscript but also of our approach and software. Thus, your review really helped to improve the results of our work.
For detailed answers to your comments and the updates planned for the revised manuscript see the green parts (and especially the bullets) below.

GENERAL COMMENTS

The focus of the manuscript is unclear. The title suggests that the Kalman filter approach is new and the core topic of the paper, but in fact the authors describe their own database and its features all quite wordy while the core of the KF method is described relatively sketchy (apart from the general Kalman filter equations, which can also be found in a textbook). If the authors feel their method is innovative, they should put a stronger focus on it and provide more details on the method, motivations for choosing certain parameters and validations (and maybe shorten some other topics). See also below.

Since the focus of our paper is the whole processing strategy and not only the Kalman filter we will change the title of the paper accordingly, probable to "DAHITI – an innovative approach for estimating water level time series over inland water using multi-mission satellite altimetry".
Indeed, detailed investigations show that the main improvement of our approach is not due to the Kalman Filter approach but mainly from the preprocessing step with outlier detection.

- We will describe this in more detail in order to make the focus of the paper more clear.

The core of every Kalman filter is the dynamic model, the propagation of which is merged with the data in a least squares sense. The authors assume an identity matrix for their dynamic model and introduce an a-priori error of 5cm. This basically reduces the Kalman filter to a recursive least squares method. The choices for the model and the a-priori error, but also for the (optional) retracking procedure should be properly motivated in the text. Additionally, it could be validated whether the result is sensitive against e.g. the choice for the a-priori error or the choice of retracking algorithm, or the grid spacing, etc. . This would make the choices, which might also vary with respect to the target water body, more plausible.

It is true that the current implementation of the Kalman Filter is simple but it has a big potential in the future if a dynamic model is applied. We will try to make this more clear in the text and motivate and describe the chosen parameters in more detail. A sensitive analysis was made in the evaluation step of our approach by testing different grid sizes, a-priori values, etc. Based on this analysis the setting parameters were fixed.

- We will describe the choices for the model, the a-priori error, and the retracking procedure in more detail
- We will provide more information about the applied outlier criteria and thresholds

The authors introduce their Kalman filter method utilizing a precomputed, hexagonal grid while assuming a constant lake level. I do like the idea of employing a Kalman filter, but I do not really see the necessity of the grid when all the grid points are assumed to be the same height (P4828, line 1). In the end all grid points over the lake are averaged into one final height output. I suppose, utilizing the actual altimeter locations (with the same assumption for constant heights and averaging) would yield nearly the same result? The authors state on P4825, line 8-10 that the observations are linked by A_k only to the nearest grid points. When computing the average, only the "best water levels" of all grid points of each epoch are selected (P4828, line 3-4). I assume, these should in general be the grid points very close to the actual altimeter groundtrack since only these will be influenced by the observations?

We changed our approach and do not use a grid for the standard computation anymore. This is no longer necessary due to improved “errors” of the altimeter measurements which are now more realistic. Now, a more reliable weighting of the input observations in the Kalman Filter is realized and makes the usage of a grid obsolete. However, we will keep the option for calculations on a grid because it is very useful for investigations regarding the surface variability of larger lakes.

- We will remove the grid for our standard computations

Furthermore, when using the grid, the individual grid points would be correlated, yet the authors assume uncorrelated system noise (P4826, line 5) and initialize their co-variance matrix E_xx,k using an identity matrix (P4825, line 12). On P4824, line 10-11 the authors say that the grid is used to account for river slopes, but I think referencing the individual heights to a fixed point (e.g. by correcting for the differences in geoid height between the measurement point and the reference point) might also do the trick.The authors should better motivate these choices to make it more comprehensible.

- As described in the point before, the grid will be removed.

Another point would be the applied retacking. Why are not all the lakes (rivers) retracked. This would be more consistent. Additionally, why use 1 Hz heights instead of 20Hz (or 18Hz or 40Hz) heights all the time over inland water bodies? Furthermore, the small threshold of the applied Improved Threshold Retracker of 10% might lead to the detection of small bumps in front of the desired leading edge in a sub-waveform, depending on the number of additional range gates considered before and after the detected sub-waveform.

Over large lakes the ocean product of the different altimeter mission provides better water heights than any retracking algorithm which is not optimized for ocean applications. Moreover, using the ocean product enable us to provide absolute water heights since we can apply the range corrections from global ocean analysis and can neglect retracking biases. Therefore, we apply only an additional retracking when it is really necessary: over smaller lakes and rivers.
In future, we will switch now completely to high-frequent altimeter data for consistency in our DAHITI approach even if 1Hz data are sufficient for lager lakes.
We agree that a 10% Improved Threshold Retracker is not always optimal, but we don't want to mix retrackers and tests show that smaller thresholds in most cases (especially over rivers) lead to better results in the resulting water heights.

Maybe the biggest problem is that no systematic analysis of the many choices is made in which the results differ. It is difficult for the reader to find out which of the choices actually leads to improvements and which is not really required. What I mean is something like, first, testing the effect of the SVR outlier detection (see also “specific comments” on this). Then investigate the effect of the additional threshold outlier correction, the Kalman filter, etc.

We agree with your statement that it is not really clear what part of the processing leads to the improved results. Therefore, the following updates will be made in the revised version.

- We will add information in Table 3 and 4 which outlier criteria are applied for the different inland water bodies.
- Detailed information about applied outlier criteria and thresholds will be added for the three study cases
- Furthermore, we will demonstrate the impact of the preprocessing and the Kalman filtering. Hereby, we replace the Kalman filter step by an Median filter to demonstrate the impact of both parts. This will be done for the three investigated water bodies. Investigations show that the major improvements are due to the preprocessing and not to the Kalman Filter. Therefore, we will updated the title of the paper (see above)

As mentioned, parts of the manuscript are quite wordy, e.g. the description of the possible choices in the openADB database in the beginning of 3.1 (P4822). Additionally, some information is provided several times throughout the manuscript (see also the “specific comments” part). These could be shortened and more focused.

- We will update/shorten certain parts (see specific comments, below)

Command of language is a problem. I am not a native English speaker, but there are, e.g., a lot of commas missing in some parts of the text. Someone with better knowledge of this than me should proofread the manuscript later.

- We will try to further improve the use of English language.

Why are there two sections containing pre-processing? (Sect. 2 and Sect 3.1)

I agree that this is a bit confusing.

- Therefore, the title of Sect. 2 will be changed from “Altimeter Data and preprocessing” to “Altimeter Data and Height Estimation”.

I like the selection of the investigated water bodies, as they cover large lakes, smaller lakes, as well as rivers. However, some additional information on the chosen criteria for outlier detection etc. would be helpful.

Additional columns of the applied outlier criteria will be added in Table 3 and 4.

- Furthermore, additional information such as used outlier thresholds will be shown in detail for the three selected study cases in ‘4.3 Selected results’

For the results, the DAHITI time series are corrected for possible outliers while time series from other groups which might still contain outliers (e.g. GRLM) are not corrected. I think, it would be more consistent when the same outlier detection procedure is applied to the other time series, too. Additionally in my opinion, the correlation with an in-situ time series is more meaningful in combination with mentioning the number (or percentage) of removed outliers from the total number of points; after removing a lot of outliers, the correlation will always become better.

We do not modify the external time series before the validation process and also do not use the in-situ data for an outlier detection of the DAHITI time series since we want to document the quality of the time series without further modification. We agree that doing so would lead to a decrease of RMS but each altimeter database has the possibility to add own outlier rejection before releasing the time series.

- We will add additional information such as number of days when altimeter data was available as input data and the resulting number of days of the final time series (thus: number of outliers)
- We will add a link to Ricko et al. in which outlier rejection was applied with respect to in-situ data

In the introduction, the authors mention that the determination of the accuracies will be covered in another paper. Why is this not included here, since the outlier detection and Kalman filtering are depending on the derived accuracies? Generally, I think this is an interesting and important part of the time series provided by DAHITI.

In general, the whole method of error propagation is already implemented and explained in this paper.
However, in the present state, the provided “accuracies” are only formal errors strongly depending on the errors of the input heights which are not perfect. In addition, detailed investigations on the resulting errors are necessary covering uncertainties of the applied geophysical corrections or geoid model. And the most difficult part is the determination of an realistic error from the applied retracking. All these points will be highlighted in another paper.

SPECIFIC COMMENTS

DAHITI The name of the database is introduced several times. At first in the abstract and again on page 4817, line 15, which is fine. But also on page 4821, line 26 or page 4837, line 2 or page 4838, line 8-9. I think one (or two) time(s) is fine.

- DAHITI will be introduced once

DGFI Similar to “DAHITI“, I think that introducing it once on page 4817, line 16 is sufficient.

- DGFI will be introduced once

— Abstract

P4814, line 7-9: I think both sentences basically providing the same information. I would suggest writing something like “... approach incorporating cross-calibrated altimeter data from Envisat ...”

- Phrasing will be updated

— Introduction

P4815, line 18: The definition of “footprint” is not precise. The pulse limited footprint (which the authors talk about) is much smaller than the beam-limited footprint defined by the antenna beam-width.

- For clarification we will change “footprint” to “pulse-limited footprint”.

P4815, line 24: How are the hooking effects treated? This is important, especially when small lakes or rivers are investigated.

In the current processing strategy, the hooking effect is not corrected but the affected measurements are excluded from further processing. The restriction of measurement latitude within the preprocessing step limits the usage of off-nadir observations. Furthermore, the additional computed input uncertainty for each altimeter measurement will increase for off-nadir observations. Thus, only few off-nadir measurements will be used within the approach and these will have larger uncertainties. In addition, we are about to publish an advanced approach for correcting the hooking effect. We plan to integrate this approach in the DAHITI software in the future.

- An explanation on the handling of the hooking effect will be added

P4815, line 27-28: I think this sentence might be misunderstood in the way that the curves are fitted to the final heights instead of the heights of one overpass over the target of interest.

- The statement will be explained in more detail for clarity

P4816, line 1: When talking about ocean-like waveforms, one might refer to the Brown model (Brown, 1977). Similar on P4819, line 16.

- Reference will be added

P4816, line 4: Generally, all threshold retrackers are sensitive to the geometrical waveform shape, e.g. large peaks will lead to a larger amplitude etc. Therefore, I think “robust” should be rephrased. Same for P4819, line 20.

The statement that we indicate retackers such as OCOG, Improved Threshold as robust means that every waveform can be retracked compared to retracking algorithms such as MLE, Beta5, etc. which try to fit form function to the waveform.

- Phrasing will be updated

P4816, line 7-11: The first two sentences of this paragraph basically provide the same information.

Phrasing was updated from *“Despite of the aforementioned challenges, satellite altimetry has been successfully used for the estimation of water levels of lakes and rivers by different groups during the last years. The potential of using satellite altimetry for the estimation of water level time series and for understanding the terrestrial water cycle was already shown e.g. in Birkett (1995), Crétaux and Birkett (2006), and Crétaux et al. (2011).” to “*The potential of using satellite altimetry for the estimation of water level time series of lakes and rivers and for understanding the terrestrial water cycle was already shown e.g. in Birkett (1995), Crétaux and Birkett (2006), and Crétaux et al. (2011).”

P4817, line 22: “optional retracking”. Why is the retracking optional and not mandatory for all inland water targets. See also general comments.
See answer in general comments

— Altimeter data and preprocessing

P4818, line 12-13: The first sentence repeats information already provided in the introduction. I’d suggest to simply remove this sentence.

- Sentence will be removed

P4819, line 3: “about 80km” → “about 80km at the equator”

- Phrasing will be updated

P4819, line 3-5: Since Envisat stopped working in April 2012 and SARAL was launched in March 2013, there is a gap in the coverage.

That is true, but in line 3-5 we are talking about the time series using the same orbit as ERS-2 and SARAL/AltiKa to achieve a time series of about two decades. In 10/2010 the Envisat mission was shifted to an interleaved orbit which means that these measurements are no longer on the same ERS-2/SARAL/AltiKa orbit. This will lead to a gap between 10/2010 and 03/2013 exist.

Envisat : 05/2002 - 10/2010
Envisat extended : 10/2010 - 04/2012
SARAL/AltiKa : 03/2013 – active

- We will update this part from "The orbit of these missions is defined by a repeat cycle of 35 days and a track separation of about 80 km. The data is available for almost two decades with a data gap between 2010 and 2013 due to the shift of Envisat to a drifting orbit that lasted until the launch of SARAL/AltiKa." to "The orbit of these missions is defined by a repeat cycle of 35 days and a track separation of about 80 km at the equator. The data is available for almost two decades with a data gap between 2010 (end of Envisat core mission) and 2013 (launch of SARAL/AltiKa). The data of Envisat on its drifting orbit (10/2010-04/2012) is not used."

P4819, line 18-20: Why use a threshold of 10%? Additionally, the algorithm might provide more than one sub-waveform and corresponding water height. Is the "correct" height selected by utilizing a reference height like in Hwang et al, 2006 or Guo et al., 2009?

We tested different values of thresholds but in our opinion the 10% threshold leads in the most cases to better results as using a 50% threshold. But the the difference are not quite large. Furthermore, we are using only the first sub-waveform for our retracking. We don't use a reference height such as last SSH over ocean (Hwang et al.) since this is difficult over small lakes and rivers.

- We will explain the used version of the Improved Threshold Retracker in more detail

P4819, line 20: How is it decided whether retracking is necessary or not? Is there a reference for the statement, that ranges from the 10% Improved Threshold Retracker are more reliable?

There is no certain criteria for applying an additional retracking. In general, all altimeter measurements of smaller lakes and rivers are retracked if the ocean product does not lead to reliable time series due to land contamination. There is no reference for selecting the 10% Improved Threshold Retracker but we have implemented several retracking algorithm and found out that this leads to the best results for our application.

P4820, line 1: Suggestion: "altimeter range" could be renamed to "retracked altimeter range" to make it clear that the ranges are always retracked and not tracker ranges.

We keep "altimeter range" because not all altimeter measurements are retracked. It can also happen that the original ocean product is used.

P4820, line 22: It would be nice if the authors could give more detail on the "certain thresholds" that were used.

- The different applied outlier criteria for the different inland water targets will be added in Table 3 and 4.
- Furthermore, detailed information about "certain thresholds" will be added for three selected results

P4820, line 25: "... and do allow to predict where ...". I think "... and do not allow to predict where ..." was intended.

- Phrasing will be updated

P4821, line 1-4: Remaining uncertainties in the geoid might lead to significant height differences depending on the location over the lake, especially for larger lakes. So the assumption of a constant height level of the derived heights might not be satisfied.

We agree that in reality the resulting surface of larger lakes are not constant because geoid errors or wind and waves lead to height differences. Our statement was based on the assumptions without that errors (geoid error, retracking error) and hydrodynamics (wind, waves) are neglected. But remaining height difference is no problem for our approach.

- We will explain our assumption in more detail

— Kalman filter approach

P4821, line 22: Rephrase “hexagonal computation”.

- Phrasing will be updated

P4822, line 14: Accuracy is not the same as precision.
Yes, of course. But since we have no information on the accuracy we use the precision as an alternative.

- Phrasing will be updated

P4822, line 15: The abbreviation SD for presumably “standard deviation” has not been introduced at this point.

- SD will be removed due to the new approach described in the next comment

P4822, line 16: The authors should elaborate why they use a “floating” box (and maybe rephrase it to “sliding”) with the size of 5 instead of just all the valid heights over the target of interest. Especially, since deriving the standard deviation reduces the degrees of freedom (dof) by two, leaving only 5-2=3 dof, which is quite a small number.

The reason for using a sliding box and not all “valid” heights is that we do not know which values are valid, and which are invalid due to topography or outliers.

We improved our “error” estimation for each measurement. Now, we are using sliding boxes along the satellite track of +-3.5 km for large lakes, +-1.5 km small lakes/large rivers and +-0.5 km for small rivers. The definition of the box in km instead of number of points makes the results more comparable between different missions measuring with 10Hz, 20Hz, or 40Hz. Moreover, instead of calculating standard deviations within each box, we use the “absolute deviation around the median” (ADM) as “error” of the measurement. For that purpose, in each box, we are estimating a median of the water heights. Then the absolute difference between the median height and the current water height is computed. This method leads especially near shores to more reliable errors than the former approach..

- The new error estimation method will be explained

P4822, line 19: “without significant slopes”. As mentioned before, there might be a significant slope from uncertainties in the geoid correction. Therefore, what will happen in case there is a significant slope and the assumption is not satisfied?

- This sentence will be removed due to the new approach of estimating “errors”

P4822, line 21- P4823, line 4: I think it is not necessary to explain all possible choices for outlier detection in the software package, but rather focus on what was chosen for the water height derivation in DAHITI.

- The different applied outlier criteria for the different inland water targets will be added in Table 3 and 4.
- Furthermore, detailed information about certain thresholds will be added for three selected results

P4823, line 6-7: “ the SVR on […] is applied.” → “SVR is applied to ...”

- Phrasing will be updated

P4823, line 9: “a lake” → Which lake?

It is Lake Erie which is crossed by Pass 80 of SARAL/AltiKa

- Information will be updated

P4823, line 11-12: How have these thresholds been selected? Are there different thresholds for different targets (large lakes, smaller rivers, etc.)? When looking at Fig. 2, only a small number of points (green) are actually rejected by the outlier detection which would probably be also rejected by the selected threshold criterion. Maybe the authors could elaborate on why they need both stages of outlier detection?

In general, a threshold of 0.05 m is selected for lakes when this outlier criteria is applied. This threshold was selected because high-frequent data have a noise within +-5cm which can be seen in Figure 2. It is often not necessary to apply all introduced outlier criteria.
Planned updates in revised version:

- Table 3 and 4 will be updated by adding additional information about applied outlier criteria

P4823, line 14-16: The authors should mention this a little bit earlier in the before the example with the lake. Additionally, it is not mentioned in the results section which criteria for outlier detection were chosen for the different inland water bodies.

- This information will be moved before the example of the lake
- Table 3 and 4 will be updated with additional information about applied outlier criteria

P4823, line 19: Suggestion: Maybe rephrase “heart of DAHITI”.

- The role of the Kalman Filter in DAHITI will be rephrased

P4823, line 24-25: The Kalman filter only gives an optimal estimate for linear models and Gaussian distributed model states and observations.

We agree!

P4824, line 23-24: I think the sentence might be a bit unclear. Suggestion: “Time intervals shorter than one day are precluded by assigning the individual measurements to full days”

- Phrasing will be updated

P4825, Eq 2: I think the x_k vector is missing a minus sign in the exponent

In this part the basic equation especially the observation model is introduced which is not time dependent itself. Therefore, the x_k vector has no minus sign.

P4826, line 5-6: Why 5cmˆ2? Why are the grid nodes uncorrelated (see also general comments)?

The 5cm² was selected based on personal experience.
We did not apply correlations because they are unknown (see before)

P4827, line 7-8: I suggest to rephrase it to “ … accuracies will become smaller within the updating step” since “reduced accuracies” implies that the accuracy becomes worse.

- Phrasing will be updated

P4827, line 15-20: The information on the transition matrix and the assumed variances has already been provided on P4826, line 2-6.

Yes, in P4827, line 15-20 the general Kalman Filter is explained. But P4826, line 2-6 in we focus on our implementation in DAHITI.

- Phrasing in P4827, line 15-20 will be updated

P4828, line 6: Why 5-10cm? Is this value larger for smaller lakes or rivers? How is it chosen?

The resulting errors depend on the input errors of the altimeter measurement. By using the standard deviation which was implemented before the errors were quite small and increase for small lakes and rivers. Our new error estimation leads to more realistic and larger errors which increases also the resulting errors of the Kalman Filter approach. Therefore we choose an error limit for larger lakes which is selected manually depending on the resulting errors. For rivers and small lakes we are using the height which has the lowest error.

- We will explain in more detail how the error limit is selected

P4828, line 17-21 (Figure 5): How is the confidence interval “manually defined”? Is it different for each target? The dashed lines in Figure 5 are not described in the caption of the figure. Which lake is used for the example in Figure 5?

The confidence interval is defined manually because it depends on the scatter of the different water bodies
Lake Erie is used in Figure 5 which is already mentioned in P4828, line 18

- Dashed lines in Figure 5 will be described
- Name of the lake will be added to the description of Figure 5
- Selection of confidence interval will be explained in more detail

— Results and validation

P4829, line 14-15: The different target types have already been mentioned in the beginning of Sect. 4.

- We will remove this sentence

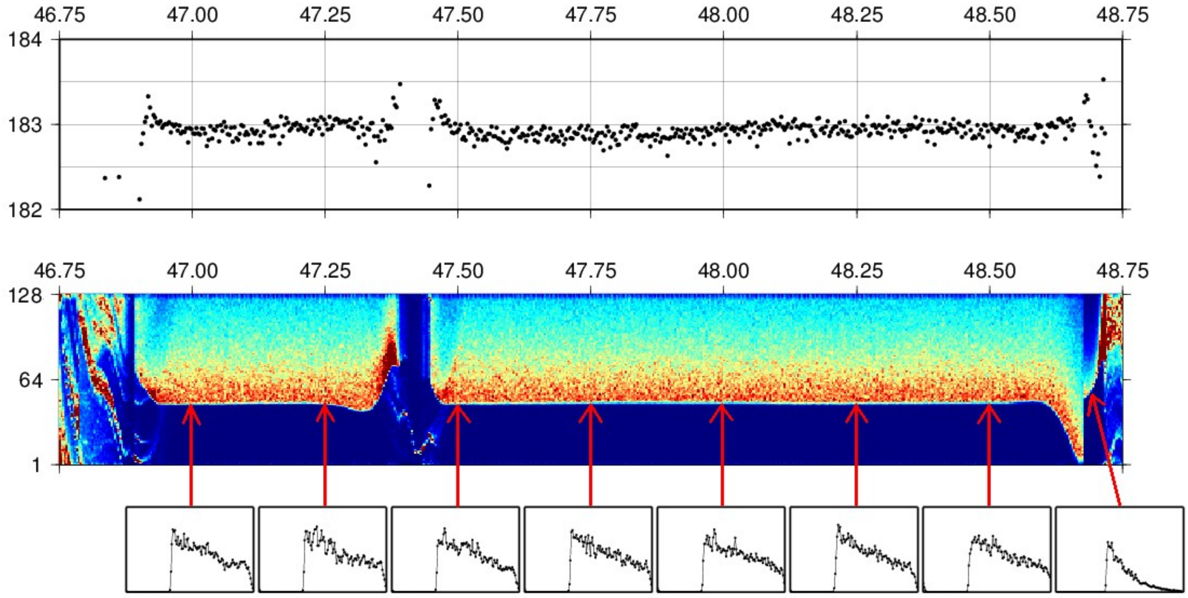
P4829, line 19-20: Suggestion: Remove that sentence, as it provides no new information.

- The sentence was removed

P4829, line 23-25: This may be correct for the central parts of the lake but closer to the land influences will impact the waveform shape.

Yes, that is true, but there is only a small band of few kilometers along the lake shore which is affected by land contamination. For the Great Lakes, the waveforms have almost always ocean-like shapes which is shown in the image below for a crossing track over Lake Superior. For those lakes the corrupted waveforms near the lake shore will be rejected by the different outlier criteria.

- We will add information about corrupted waveforms near the lake shore



P4830, line 16: “several lakes are frozen for several month” Which ones? Are all measurements during these month treated as outliers?

Sentence was updated from “Several lakes are frozen for several months which makes the water level computation challenging” to “Several lakes (indicated in Table 3) are frozen for several months which makes the water level computation challenging” .

- We will explain the usage of the backscatter coefficient as additional outlier criteria to reject measurements affected by ice coverage

P4831, line 20-21: I’d suggest to remove the first sentence, as it repeats information that has been mentioned before.

- Sentence will be removed

P4832, line 5-6: Computing the relative height changes does not avoid all the uncertainties in in-situ gauge data. E.g. there could still be gross reading errors in manually operated gauges.

That is true. The objective of the comparing relative heights is that the different altimeter time series have different absolute heights. These offsets are now removed. Other uncertainties in in-situ data will show up as increased RMS within the validation process. In our case, only clear outliers of several meters are rejected for few stations in the Amazon Basin.

P4832, line 28 (Figure 7) Why is the year 2004 highlighted?

There is no certain reason why year 2004 was highlighted. The objective is to show the data distribution of the different time series in detail. Especially the increased number of points of the DAHITI time series.

P4833, line 2: How have the time series been shifted? With respect to the mean or one (or more) selected cycles?
All external time series from altimetry and in-situ stations are shifted with respect to the DAHITI time series. Hereby, differences at all common epochs are computed and averaged.

- The method about estimating the applied offset will be explained in more detail

P4833, line 4: After applying two outlier corrections during the processing, one should in general not expect any more outliers.

That`s true. So, this result proofs the efficiency of our outlier rejection.

P4833, line 5: Why use the 1Hz ranges instead of consistent 20Hz (or 18Hz or 40Hz) ranges for all targets? Small land influences on the trailing edge of the waveforms might still influence the results of the standard ocean retracker ranges, available in the GDRs. (See also general comments).

In general, 1Hz altimeter data are sufficient for large lakes. But we changed our approach in DAHITI and we are now using only high-frequent data for all investigated water bodies.

- 1Hz altimeter data will not longer be used in the DAHITI approach

P4833, line 8-9: From Table 1, one can see that other stations showed larger RMS compared to Ontonagon and Grand Marais. Is this connected to the distance of the in-situ stations to the altimeter ground tracks?

Probably, there is a small impact of the distance to the gauging station. However, the in-situ time series of Duluth and Point Iroqouis look more noisy than the station Ontonagon and Grand Marais.
This is probably caused by wind and waves because the station Duluth and Point Iroqouis are located in brights in the most west and south-east. Thus, effects due to the quality of these stations are more likely. The distances between in-situ station and nearest altimeter track for Ontonagon and Point Iroqouis vary between 5-10 km. We think that the distance has not a major impact on the RMS differences.

P4833, line 9: In my opinion the correlation coefficient alone is not a good measure for the quality, since removing a large number of “outliers” automatically improves the correlation drastically. So additional information on the total number of measurements and the number of removed outliers from the final time series might be helpful here.

- We will add an information on the number of outliers (number of days when altimeter data was available as input data and the resulting number of days of the final time series.)

P4833, line 16-17: Have the results from the other databases also been corrected for possible outliers? E.g. the unsmoothed GRLM results which, judging from Figure 7, are used here might still contain large outliers in some cycles, which will have a significant effect on the computed RMS. (see also general comments).

We do not modify the external time series before the validation process and also do not use the in-situ
data for a outlier detection of the DAHITI time series (as done in Ricko et al) since we want to document the quality of the time series without further modification. We agree that the RMS will decrease after outlier removal but each altimeter database has the possibility to add own outlier rejection before releasing the time series.

P4833, line 22: Figure 8 deals with a different lake so it does not show the "same time series" like Figure 7.

- Phrasing will be updated

P4833, line 24: The GRLM time series for Lake Athabasca is no longer available on the GRLM website.

The time series is still available. The following archive which is available via the GRLM website http://www.pecad.fas.usda.gov/lakes/images/envisat_lakes.txt.zip contains Lake Athabasca in file lake0348.N.1.4.txt

P4833, line 24-25: Figure 8 highlights the year 2004 (like Figure 7) and not 2010. Again, what makes this year special?

There is no certain reason for the selection of the year. But, it must be a year in which all investigated time series and in-situ data are available. The reason for zooming into a single year is to show the temporal resolution of the different time series.

P4834, line 1: "very good" The results for the Topex/Poseidon era seem to be more noisy compared to LEGOS.

Maybe it looks more noisy because there was only one mission used compared to the period starting in 2002. Furthermore, the time series are rarely affected by ice coverage compared to the LEGOS time series.

P4834, line 2: Since "outliers" should have been removed during the processing, I'd suggest to maybe rephrase this.

- We will rephrase the sentence

P4834, line 2-7: Generally, the same concerns regarding mentioning the number of outliers and also applying the same outlier removal to the time series of the other groups.

We do not want to modify time series of other groups as mentioned above.

P4834, line 9: Suggestion: Maybe remove the first sentence.

- The sentence will be removed

P4834, line 13: "only one mission" → Figure 9 includes results based on Envisat and SARAL. These are two different missions.

Both mission are not flying at the same time.

- Therefore we will update the following sentence: "All altimeter time series reach a temporal resolution of about 1~month since they are based purely on missions with 35-day temporal resolution (Envisat and SARAL)."

P4834, line 14-15: Similar to P4819, line 3-5, Envisat stopped working in April 2012 and SARAL was launched in March 2013, so there is a gap in the coverage.

This point was already commented in P4819, line 3-5:

P4834, line 9-24: What pre-processing criteria (mentioned in Sect. 3.1, e.g. SD thresholds) have been used? Why is the year 2010 highlighted?

There is no certain reason why we selected 2010. There requirement was to select a year in which in-situ and all other altimeter time series are available.

- We add information about the applied outlier criteria in Table 3 and 4 and especially for the 3 selected examples
- Figure 7,8,9 will be updated to be more understandable for the user
- We will explain why we select certain highlighted year

P4835, line 7: “used points (No)” →Are the number of removed outliers included? (Same for Table 4).

The number of used points is defined by the number of days for which a water level height was computed and used for the validation. This is not the complete number of days of the whole time series.

- We will explain in more detail what is meant by used points

P4835, line 9-10: The smallest and largest RMS are not marked in color in Table 3.

- Table 3 and Table 4 the smallest RMS will be highlighted in bold and the largest RMS in italic. A highlighting in red and green is technically not possible in this journal.

P4835, line 15-18: Wind effects or other local influences might have an impact on the gauge station measurements. Again this raises the question whether stations closer to the track show better agreement.

This was already answered before in P4833, line 8-9

P4835, line 19-27: In case the other products still contain outliers, the comparisons might need to be improved before a general better performance of DAHITI can be validated.

This was already answered before in P4833, line 16-17

P4836, line 1-4 (Table 4): I think it would be helpful if the authors would include information on the utilized satellite data in each row (Maybe, e.g.: Station Name (J2) …).

- The used altimeter satellites will be added to Table 3 and Table 4

P4836, line 15-16: Is there a reference for that statement? Especially for rivers, the quality of the reprocessed water heights is quite depending on the surrounding terrain (in this case mainly rain forest?), etc. I think it would be nice to rephrase this statement to make it more clear.

- We will rephrase this statement in more detail

— Conclusions

P4837, line 15: Related to the previous point, the authors could elaborate (either here or in the previous section) what is meant by “surrounding conditions”.

- We will rephrase this statement in more detail

— References

- Well presented, except for a few missed “{}” (see “technical comments”).

- {} will be removed from references

TECHNICAL COMMENTS

I will only list a few things here, since the manuscript will probably be reworked anyway.

— Abstract

P 4814, line 2: “However, since some years” → “However, for some years”.

- Phrasing will be updated

P 4814, line 16: “from available other altimeter ...” → “from other available altimeter ...”

- Phrasing will be updated

— Introduction

P 4814, line 26: “However, in the last years … “ → “However, over the last years ...”

- Phrasing will be updated

P4815, line 27: “... fitting curves on the resulting ...” →“... fitting curves to the resulting ...” Altimeter data and preprocessing

- Phrasing will be updated

P4820, line 14: Missing words. → “Finally, each single altimeter measurement is corrected for its ...”

- Phrasing will be updated

— Kalman filter approach

P4822, line 5: “data base” → “database”

- Phrasing will be updated

P4822, line 14: “Due to lacking absolute ...” → “Due to the lack of absolute ...”

- Phrasing will be updated

P4823, line 19-21: Just a suggestion: “... input heights by combining of the […] at different locations (Kalman, 1960).”

- Phrasing will be updated

P4824, line 15: “At the beginning an initialization ...” → “In the beginning, an initialization ...”

- Phrasing will be updated

P4824, line 18: “contant” → “constant”

- Phrasing will be updated

P4824, line 18-19: Rephrase: “That means that our system each time is updated if a new ...” → “That means that our system is updated each time a new ...”

- Phrasing will be updated

P4826, line 15: “actual” → “current”

- Phrasing will be updated

— Results and validation

P4830, line 21: To be consistent with the other times it is mentioned, write “Lago Argentino”

- All Spanish lake and river names were translate into English

P4831, line 5-6: Maybe merge the two sentences to one.

- Both sentences will be merged into one

P4831, line 7: Suggestion: replace “wet” with “humid”

- Phrasing will be updated

P4831, line 8: Suggestion: replace “show” with “reach”

- Phrasing will be updated

P4831, line 10: Suggestion: replace “diversification” with “variety”

- Phrasing will be updated

P4832, line 10-14: Suggestion: These 3 sentences contain the word “differences” five times. Maybe rephrasing some of it will make it more easy to read.

- Phrasing will be updated to “These results are based on various altimeter missions and diverse approaches were performed to compute the water level time series. As a consequence, these external time series cover different time periods with temporal resolutions between 10 and 35 days. This has to be kept in mind when comparing the time series of the four databases.”

P4832, line 18: “disparate” → “divers”

- Phrasing will be updated

P4833, line 21: “reaching” → “providing”

- Phrasing will be updated

P4834, line 19-20: “... when taking the […] into account” → “... when taking into account the ...”

- Phrasing will be updated

— Conclusions
-
— References

P4839, line 31: The names “ERS” and “ENVISAT” are in curly brackets

- Phrasing will be updated

P4840, line 33: The name “SARAL” is in curly brackets.

- Phrasing will be updated

P4841, line 15: “Shum, C.” → “Shum, C. K.”

- Phrasing will be updated

J.F. Crétaux Referee #1

Received and published: 29 Jun 2015

I will not comment here the methodology and will not suggest a full review of the paper, but I am surprised by the results given in table 3.

A previous study was done in Ricko M., C.M. Birkett, J.A. Carton, and J-F. Cretaux, Intercomparison and validation of continental water level products derived from satellite radar altimetry, J. of Applied Rem. Sensing, Volume 6, Art No: 061710, DOI:10.1117/1.JRS.6.061710, 2012.

it was an article with the aim of comparing the water level products for lakes of 3 database that are also used in the present article. The methodology of comparison between both article looks to be the same (calculation of RMS of differences between satellite product and in situ level, and calculation of the R2 coefficient). In the present article the product from the DAHITI database is added to the general comparison, and there are between both articles, seven common lakes. The fact I wish to point out and wish the authors give explanation, and at least refer in their article to ricko et al., 2012, is that for the 3 other database (Hydroweb, GRLM and ESA/DMU) the accuracy and the correlation calculated in the present article are much worse than what was present in the Ricko et al;, 2012 article.

I made a summary table of both comparisons: those made in Schwatke et al, those made in Ricko et al. See below.

Lake name	**Schwatke et al., 2015** **RMS (cm) / R2**	**Ricko et al., 2012** **RMS (cm) / R2**
Superior	5-6 / 0.94-0.95	6 / 0.97
Michigan	7-12 / 0.82-0.95	11 / 0.98
Ontario	6-7 / 0.94-0.95	6 / 0.98
Erie	9-18 / 0.69-0.92	10 / 0.95
Huron	6-11 / 0.92-0.98	8 / 0.99
Athabasca	33.7 / 0.79	28 / 0.91
Woods	43-44 / 0.58-0.63	27 / 0.81

Table 1. Hydroweb comparison from Schwatke et al analysis and published in Ricko et al., 2012

Lake name	**Schwatke et al., 2015** **RMS (cm) / R2**	**Ricko et al., 2012** **RMS (cm) / R2**
Superior	11-12 / 0.62-0.75	5 / 0.97
Michigan	8-10 / 0.73-0.95	8 / 0.98
Ontario	11 / 0.85	6 / 0.98
Erie	14-20 / 0.61-79	6 / 0.97
Huron	7-12 / 0.90-0.96	6 / 0.99
Athabasca	55.7 / 0.27	Not calculated
Woods	Not calculated	19 / 0.86

Table 2. GRLM comparison from Schwatke et al analysis and published in Ricko et al., 2012

Lake name	Schwatke et al., 2015 RMS (cm) / R2	Ricko et al., 2012 RMS (cm) / R2
Superior	8-9 / 0.75-0.82	5 / 0.95
Michigan	5-7 / 0.69-0.78	7 / 0.93
Ontario	5 / 0.96-97	7 / 0.96
Erie	13-17 / 0.50-0.74	10 / 0.86
Huron	6-9 / 0.80-89	7 / 0.93
Athabasca	80.5 / 0.30	28 / 0.85
Woods	36 / 0.40-41	24 / 0.81

Table 3. ESA/DMU comparison from Schwatke et al (in revision) analysis and published in Ricko et al., 2012

Lake name	Schwatke et al., 2015 RMS (cm) / R2
Superior	4-6 / 0.85-0.96
Michigan	5-7 / 0.82-0.95
Ontario	4-5 / 0.94-0.98
Erie	5-13 / 0.78-0.96
Huron	4-9 / 0.92-0.98
Athabasca	17 / 0.88
Woods	15-16 / 0.75-79

Table 3. DAHITI comparison from schwatke et al.,

It is worth to note that almost all the time the results are degraded from the accuracy assessment done by Ricko et al., 2012 to the present study. I'm for example very surprised by the high degradation observed for the GRLM solution of the great lakes with accuracy quite always higher than 10 cm while it has been assessed to be between 5 and 8 cm from Ricko et al., 2012 and if so keeps at the same order than what is obtained in DAHITI. The difference for lake Woods of Legos is also significantly degraded (28 cm to 43-44 in this article) and for the lake Athabasca ofthe DMU solution (28 cm from Ricko et al;, to 80.5 from this article) the difference is extremely high. for the correlation coefficient the degradation is very significant and for all of the cases. This needs explanation. I did for example the calculation for lake Erie with the Legos time series and the in situ and I found 9 cm and 0.97 (between 9 and 18 cm for Schwatke and 0.69 to 0.92, and 10 cm and 0.95 for Ricko et al., 2012).

Second point: I have downloaded the figure from hydroweb of the lake superior which is also reported in the figure 7. The legos solution (Hydroweb) on this figure looks very poor in term of number of valid measurements (maybe due to the fact that all solution sare put on the same plot) but figure 7 is not representative of the real distribution of the corresponding time series on Hydroweb, particularly at the end of the time series

The time series of Hydroweb in Figure 7 looks similar to your plot from the hydroweb web site.
We will improve the plots of Figure 7,8,9 to display that the different time series more clearly.

(see figure attached). It looks like the distribution of measurements is very irregular and sparse, which is not the case. please make this more realistic.

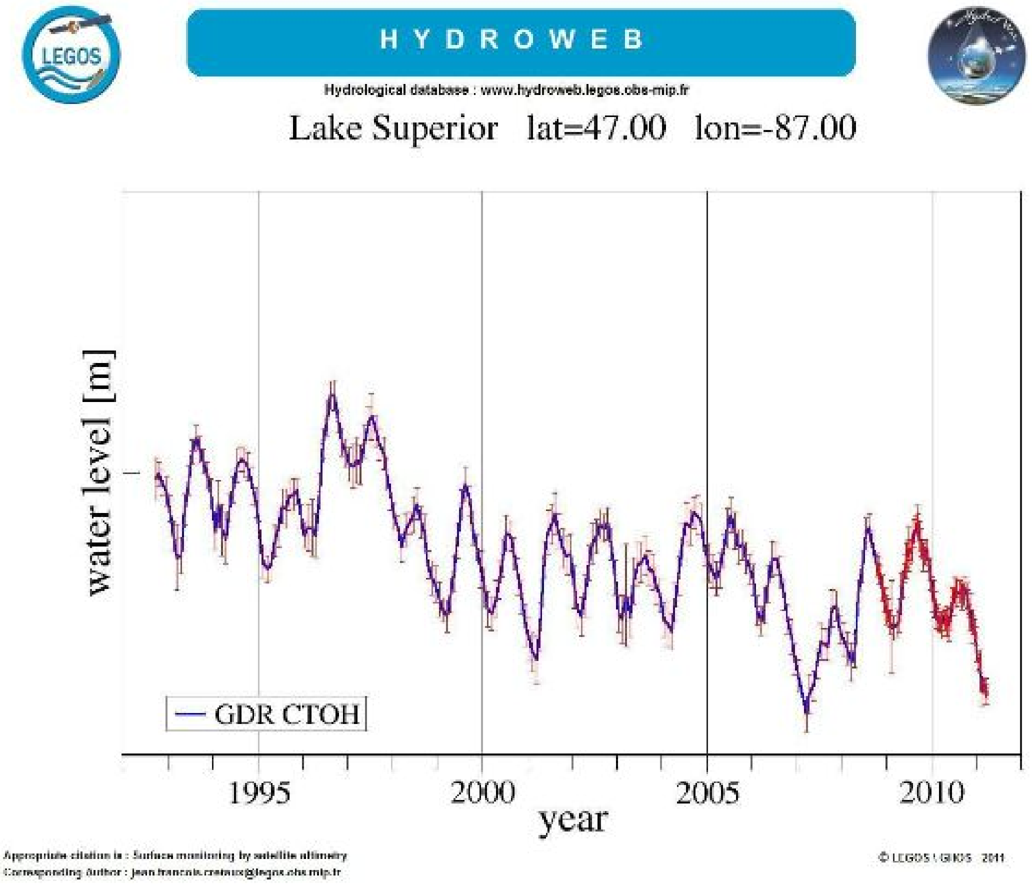


Fig. 1. Lake superior: image down loaed from Hydroweb web site.

The differences between the results from Ricko et al, 2012 and our paper can be explained by the different outlier rejections performed before the validation.

In Ricko et al. 2012, an outlier rejection and filtering of time series with respect to in-situ data was performed and is described in ‘3.2. Product Filtering’: “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.”

In contrast, we use the original time series from Hydroweb, GRLM and River & Lakes without any changes for the comparison with in-situ data. In our opinion, using the validation data set (in-situ) for outlier detection is sub-optimal. Moreover, when applying a outlier rejection by means of standard deviations a rejection level of 3sigma is recommended. We do not modify the external time series before the validation process and also do not use the in-situ data for a outlier detection of the DAHITI time series since we want to document the quality of the time series without further modification. This differences in the validation strategy cause our RMS values to be larger than those of Ricko.

In order to verify the DAHITI results and to check the dependence on time series length we re-downloaded the Hydroweb time series. The new comparisons (see table) show small improvements with respect to the old ones but the results are still much worse than the Ricko results meaning that there are still some outliers in the original Hydroweb time series.

	(Schwatke et al.) RMS [m] / R² (old data)	(Schwatke et al.) RMS [m] / R² (new data)	(Ricko et al.) RMS [m] / R / R²
Erie - Cleveland	0.094 / 0.91	0.090 / 0.92	0.10 / 0.95 / 0.90
Ontario - Oswego	0.067 / 0.94	0.066 / 0.95	0.06 / 0.98 / 0.96
Michigan - Milwaukee	0.100 / 0.95	0.088 / 0.96	0.11 / 0.98 / 0.96
Huron - Harbor Beach	0.075 / 0.97	0.077 / 0.97	0.08 / 0.99 / 0.98
Superior - Marquette	0.059 / 0.94	0.055 / 0.95	0.06 / 0.97 / 0.94
Woods - Warroad	0.333 / 0.60	0.323 / 0.51	0.27 / 0.81 / 0.65
Athabasca - Crackingstone	0.337 / 0.79	0.321 / 0.79	0.28 / 0.91 / 0.82

We will discuss the differences to Ricko results in the revised version of the paper and point out the different outlier rejection strategies.

Moreover, we will try to improve the paper figures 7/8/9 and to display the different time series more clearly. At the moment part of the Hydroweb time series is hidden behind other data sets – probably we will shift the time series to each other in order to allow to distinct between them.