

Article

# Treating the Hooking Effect in Satellite Altimetry Data: A Case Study along the Mekong River and Its Tributaries

Eva Boergens \*, Denise Dettmering, Christian Schwatke and Florian Seitz

Received: 05 November 2015; Accepted: 18 January 2016; Published: 23 January 2016

Academic Editors: Magaly Koch and Prasad Thenkabail

Deutsches Geodätisches Forschungsinstitut, Technische Universität München, Arcisstraße 21, 80333 Munich, Germany; denise.dettmering@tum.de (D.D.); christian.schwatke@tum.de (C.S.); florian.seitz@tum.de (F.S.)

\* Correspondence: eva.boergens@tum.de; Tel.: +49-89-23031-1263

**Abstract:** This study investigates the potential of satellite altimetry for water level time series estimation of smaller inland waters where only very few measurements above the water surface are available. A new method was developed using off-nadir measurements to estimate the parabola generated by the hooking effect. For this purpose, a new waveform retracker was used as well as an adopted version of the RANdom SAmple Consensus (RANSAC) algorithm. The method is applied to compute time series of the water levels height of the Mekong River and some of its tributaries from Envisat high-frequency data. Reliable time series can be obtained from river crossings with widths of less than 500 m and without direct nadir measurements over the water. The expected annual variations are clearly depicted and the time series well agree with available *in situ* gauging data. The mean RMS value is 1.22 m between the resulting time series and *in situ* data, the best result is 0.34 m, the worst 2.26 m, and 80% of the time series have an RMS below 1.5 m.

**Keywords:** satellite altimetry; inland water; Envisat; RANSAC; Mekong basin; water level time series

## 1. Introduction

A prerequisite for an accurate description of the global water cycle, is the assessment of water storage of inland water bodies. In particular, long-term changes in water storage are important for the risk assessment of floods and drought and for water availability. However, the number of *in situ* gauges along inland waters has been declining worldwide since the 1980s. The number of available gauges at the Global Runoff Data Center (GRDC) decreased from 7300 in 1978 to 1000 today [1]. Especially in Asia and Africa, almost no *in situ* data are available.

More recently, satellite altimetry has begun to close this data gap for inland waters. Altimetry data are especially advantageous for inland waters because they are obtainable even in remote regions. Satellite altimetry was originally designed for sea surface observations, but with the development of new data processing strategies, altimetry measurements may now be used over inland water bodies such as lakes, reservoirs, rivers, and wetlands. For example, using Geosat data, Morris and Gill [2] obtained time series of the levels of the Great Lakes in North America. With the launch of the ERS-1 mission in 1991 and the Topex/Poseidon mission in 1992, more inland waters became observable (e.g., [3–6]). The ERS-1 mission was succeeded by ERS-2, Envisat, and SARAL/AltiKa in the same orbit; similarly, the Topex/Poseidon mission was succeeded by Jason-1 and Jason-2. Currently, we can access more than 20 years of continuous altimetry time series.

However, altimetry data collected over inland water bodies are more difficult to process than open ocean data. The over-land footprint size of Envisat's Ku-band radar echo is between 2 and 16 km [7], which limits the use of standard Geophysical Data Record (GDR) products for inland

water applications. The waveforms returned from complex surfaces differ from the Brown-like ocean waveforms on which the standard processing algorithms are based [8]. For reliable results over smaller water bodies, we require specific retracking algorithms, such as  $\beta$ -5 Retracker, Threshold Retracker, or Improved Threshold Retracker [8].

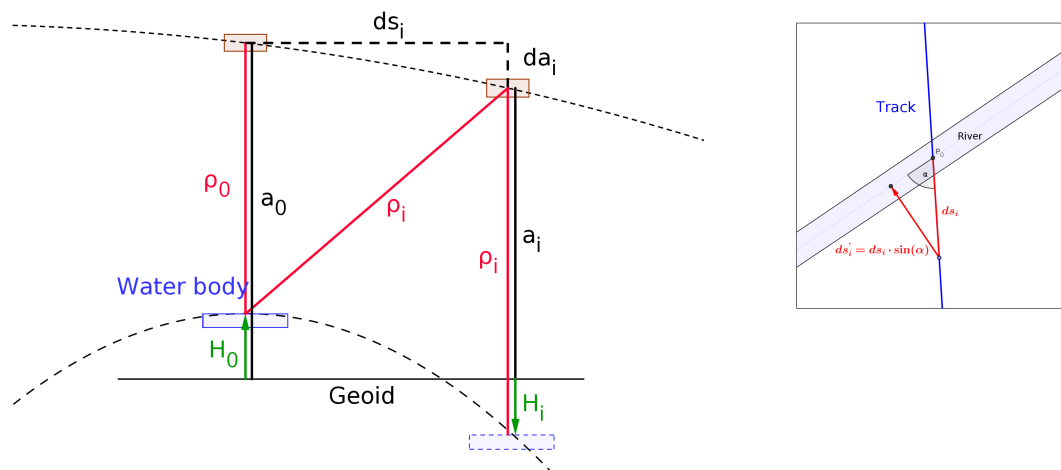
Reflecting the importance of inland water altimetry, four inland water altimetry databases have now been established: Hydroweb, developed by Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS) [9]; the River and Lake database provided by the European Space Agency (ESA) [10]; the Global Reservoir and Lake Monitor (GRLM), maintained by the Foreign Agricultural Service of the United States Department of Agriculture (USDA) [11]; and the Database for Hydrological Time Series over Inland Water (DAHITI), developed by the Deutsches Geodätisches Forschungsinstitut der Technischen Universität München (DGFI-TUM) [12].

To date, most of the small-river studies have focused on the Amazon basin (e.g., [13–15]) or on other rivers surrounded by flat topography (e.g., [16,17]). Flat topography is desirable, as it does not shadow the water body in the altimetric footprint, reducing the risk of the altimeter losing lock.

Off-nadir measurements are also problematic over inland water bodies. Because water reflects more strongly than the surrounding land, the altimeter instrument measures the slant distance to the water surface even when not in nadir. Such off-nadir measurements lead to a parabolic shape of the track-height profile, called the hooking effect (Figure 1). Frappart *et al.* [13] were the first ones who described this effect for inland waters. They corrected the effect using a migration method which integrated the energy over such a parabolic feature. Calmant *et al.* [18] addressed the hooking effect as one major issue which affects the altimetry of inland water bodies. In Santos da Silva *et al.* [15] the geometry behind the hooking effect is explained in more detail and the effect is corrected for some of the presented time series over the Amazon basin. Later works, such as Santos da Silva *et al.* [19] or Frappart *et al.* [20], corrected the hooking effect as well. Maillard *et al.* [17] developed a new method to correct the hooking effect by comparing the theoretical shape of the hooking parabola with the measured data profile. All of these studies, including the study presented in this paper, were conducted with conventional radar altimetry. SAR altimetry data, like Cryosat-2 data, is less prone to the hooking effect and does not require its correction.

In this study, we introduce a new method that derives reliable water-level time series over smaller rivers using measurements affected by the hooking effect. This method can even be used when no nadir data over the river is available at all but only parts of the hooking parabola. In contrast to the above mentioned studies by Frappart *et al.* [13], Santos da Silva *et al.* [15], or Maillard *et al.* [17], the presented method is almost fully automatic and enables a fast and reliable processing of almost all river crossings, even for very small rivers. We developed a special retracker in order to ensure that all off-nadir measurements can be used and the parabola is well represented in the data set. To identify off-nadir measurements and thus estimate the parabola, we employ the RANSAC algorithm [21]. RANSAC is commonly used in computer vision, for example, in image registration (e.g., [22–24]) and feature extractions from LiDAR data (e.g., [25,26]).

The method is trialed in a case study, namely the derivation of water-level time series along the Mekong River in South-East Asia and some of its tributaries. In terms of discharge and length, the Mekong River ranks among the 10 largest rivers in the world. Using satellite altimetry, Frappart *et al.* [27] have already investigated the lowest part of the Mekong basin south of Krong Stung Treng, where the river width exceeds 2 km. Birkinshaw *et al.* [28] studied the entire lower Mekong basin (the same region is investigated in the present study) using Envisat and ERS-2 data. They focused on the area south of Vientiane and one station further in the North, where the river is less than 500 m wide.



**Figure 1.** Off-nadir measurement of the slant distance to the water body ( $\rho_i$ ) generates a parabolic profile of the heights above the geoid. The vertex of this parabola is the height  $H_0$  at the nadir position. On the right side the situation of non-perpendicular intersection is shown.

## 2. Study Area

The Mekong River is one of the longest rivers in the world (approximately 4500 km). From the Tibetan Plateau, it flows through six countries in South East Asia: China, Myanmar, Laos, Thailand, Cambodia, and Vietnam. The present paper focuses on the lower Mekong basin, downstream of the South of the mountainous regions in China, and the triangle of Myanmar, Laos, and Thailand. From there, it flows 2600 km south and outflows into the South China Sea in Vietnam. This area is inhabited by roughly 60 million people who are provided with water from the Mekong River. At the beginning of the lower basin, the Mekong River is less than 500 m wide, but widens to more than 2 km at the beginning of the delta. Most of the tributaries and the main stream north of Vientiane are less than 1 km wide. North of Vientiane, the Mekong River not only narrows to below 500 m but is also surrounded by mountainous and hilly topography. Figure 2 displays the lower Mekong region. During the summer months (mid-May to early September), the climate in the lower Mekong basin is dominated by the yearly south-eastern monsoon [29], which can raise the water level of the Mekong to more than 10 m [28].

## 3. Data

### 3.1. Altimetry Data

This study uses the high-frequency altimetry data of the Envisat/RA-2, ERS-2 and SARAL/AltiKa satellite missions, focusing on Envisat data. Envisat had a 35 day-repeat orbit with an inclination of  $98.55^\circ$  and height of approximately 770 km. Envisat's orbit is a continuation of the ERS-1 and ERS-2 orbits. The mission started in March 2002 and was shifted to a new orbit in October 2010. Only data prior to October 2010 is used in this study, cycle 006 to 094. The altimeter instrument measures ranges between the satellite and the surface. The heights  $H$  are obtained by subtracting the observed ranges  $\rho$  from the satellite height  $H_{sat}$ , and are corrected for the atmospheric delay caused by the ionosphere (*iono*) and the dry and wet troposphere (*dry* and *wet*, respectively), the crustal motions caused by pole tides (*pt*) and solid earth tides (*et*), a geoid correction (*geoid*) and the radial bias between different altimeter missions (*oerr*). Thus, the physical heights are calculated as

$$H = H_{sat} - \rho - (iono + dry + wet + pt + et + geoid + oerr). \quad (1)$$

Table 1 summarizes the corrections applied to the data. For a proper comparison among altimeter missions, we adopted correction models that were available over the entire time period of the altimetry data. In addition to these corrections, the ranges  $\rho$  were retracked over the inland water bodies to obtain more reliable heights. The retracking process will be detailed in Section 5.1.

**Table 1.** Corrections applied to the altimeter measurements.

Correction	Model/Source	Reference
ionosphere	NOAA Ionosphere Climatology 2009 (NIC09)	Scharroo and Smith [30]
dry troposphere	ECMWF (2.5° × 2.0°) for Vienna Mapping Functions 1	Boehm <i>et al.</i> [31]
wet troposphere	ECMWF (2.5° × 2.0°) for Vienna Mapping Functions 1	Boehm <i>et al.</i> [31]
polar tides	IERS Conventions 2003	McCarthy and Petit [32]
earth tides	IERS Conventions 2003	McCarthy and Petit [32]
geoid	EIGEN-6C3stat	Förste <i>et al.</i> [33]
oerr	MMXO14	Bosch <i>et al.</i> [34]

Envisat's RA-2 instrument measures high-resolution data in the Ku- and S-bands with a frequency of 18 Hz, corresponding to a distance of approximately 365 m between two consecutive measurements. This study employs the Sensor Geophysical Data Record (SGDR) v2.1 data. In the study region, the bandwidth of the RA-2 instrument switches among 320 MHz, 80 MHz, and 20 MHz depending on the roughness of the topography [35]. Over rougher terrain the bandwidth is lowered which increases the tracking window of the altimeter. This prevents the altimeter to lose lock over rapidly changing heights. According to the smaller bandwidth the size of each bin is increasing and thereby the height resolution is decreasing. The different bandwidth are also called measuring modes.

The Envisat satellite mission was preceded by ERS-2 (1995–2011) and succeeded by SARAL/AltiKa (2013 onward) in the same orbit. ERS-2 was measuring in two different bandwidth like Envisat. SARAL/AltiKa does not change measuring bandwidth. For ERS-2 we used the cycles 000 till 085 (REAPER) and for SARAL/AltiKa the cycles 001 till 020 (GDR-T patch 2).

### 3.2. In-Situ Gauging Data

The time series derived from the altimetry data is validated by comparisons with *in situ* gauging data. The *Mekong River Commission* (<http://ffw.mrcmekong.org/>) maintains several gauging stations along the lower Mekong River and its tributaries. Among the 27 available time series of gauging-station data, we selected those satisfying the following criteria: First, because this study focuses on smaller rivers, we selected only the gauging stations upstream of Vientiane and along the tributaries, where the river width is about 1 km or less. The only exception is the Mukdahan station, where specific topographic influences can be studied. Second, we excluded gauging stations with large data gaps or unexplained jumps in their time series. Third, we filtered out all gauging stations with no usable close-by altimetry data. The nine remaining stations were located 15–80 km distant from the investigated virtual stations, *i.e.*, the crossing points of the altimeter tracks and the rivers. The locations of the gauging and virtual stations are shown in Figure 2.



**Figure 2.** Study area of the lower Mekong River, indicating the gauging stations used in the validation (name of the nearest settlement and ID issued by the Mekong River Commission), the virtual stations, and the Envisat altimetry tracks.

The temporal resolution of the time series of the gauging stations is one day. Gauging data along the main stream cover the entire time frame of the Envisat measurements (until 2010), while those along the tributaries terminate between 2005 and 2007.

#### 4. Hooking Effect

The hooking effect is an off-nadir distortion of altimetry measurements that occurs at the water-land transition. Because the water surface reflects usually more strongly than the surrounding land surface, the altimeter measures the distance to the water surface even when not vertically positioned over the water. The measured off-nadir distances form a parabolic shape in the along-track altimetric height profiles [15]. This can be explained from the relationship between the off-nadir measurements and the heights (Figure 1). The measured range  $\rho_i$  at a distance  $ds_i$  from the nadir over the water body is given by

$$\rho_i = \sqrt{ds_i^2 + (\rho_0 - da_i)^2}. \tag{2}$$

Here,  $da_i = \frac{\partial a}{\partial s} ds_i$  is the height difference of the satellite between the nadir point and the actual measurement points,  $\frac{\partial a}{\partial s}$  is the change in satellite altitude along the track, and  $\rho_0$  is the measured nadir range over the water body. Given that the water-level height is  $H_i = a_i - \rho_i$ ,  $H_i$  can be expressed as a function of  $ds_i$  with

$$H_i = \underbrace{a_0 - da_i}_{a_i} - \rho_0 \sqrt{\left(1 - \frac{da_i}{\rho_0}\right)^2 + \frac{ds_i^2}{\rho_0^2}}. \tag{3}$$



Neglecting the higher-order polynomials, Equation (3) can be rewritten as

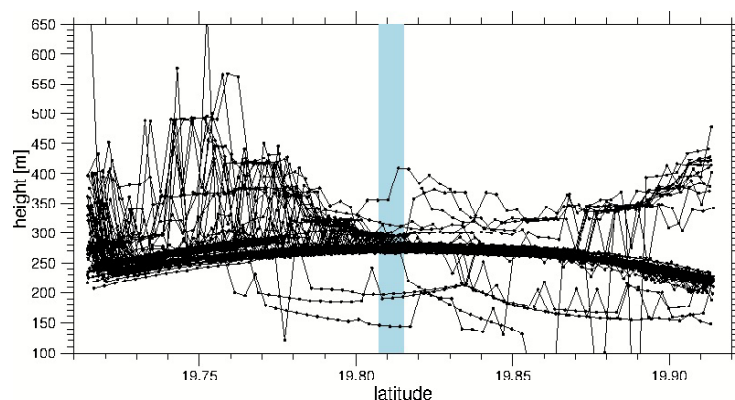
$$H_i = H_0 - ds_i^2 \left( \frac{1}{2\rho_0} \right) \left( 1 + \left( \frac{\partial a}{\partial s} \right)^2 \right). \quad (4)$$

Equation (4) shows that  $H_i$  follows a parabolic profile with a vertex at  $H_0$ , *i.e.*, the water height. The parabola is distorted by the term  $\left( \frac{\partial a}{\partial s} \right)^2$ , which changes over the water body region. The changes of height over short distances are small (approximately  $0.005 \frac{m}{m}$ ). For example, over 5 km from a water body away, the height would be distorted by approximately 8 cm, below the accuracy of the altimetry-derived inland water heights (a few decimeters). Therefore, the height distortion can be neglected.

Equation (4) assumes that the along-track distance equals the horizontal distance between the satellite and the water body, *i.e.*, the altimetry track is perpendicular to the water body. In practice, however, the track intersects the river at an angle  $\alpha$ . The strongest water return then originates not from the intersection of the altimetry track and the river, but from a closer upstream or downstream position. Assuming a straight river in proximity, the distance between the footprint center and the nearest water body is  $ds'_i = ds_i \cdot \sin(\alpha)$  (see Figure 1 right side). Therefore, the hooking parabola is dilated but not distorted.

However, depending on the intersection angle, the altimeter can measure river sections of length 10 km or more. Over a stretch of 10 km, the slope of the river cannot usually be neglected, and the measurements must be corrected accordingly.

Figure 3 shows an altimetry profile over the upper Mekong River, near Luang Prabang (Figure 2). The parabolic shape of the measurements is clearly visible. In this example the hooking effect extends to 10 km from the water body and the error in the parabolic heights is up to 100 m. Not all of the measurements are influenced by the off-nadir returns. The topography of the river and its surroundings can influence the measurements such that no hooking parabola appears in the data. Winding rivers and regions surrounded by mountains that shadow the river are especially prone to this effect. As the exact crossing point of the altimeter track and the river slightly changes (up to 1 km), the hooking parabola may be apparent in some measurements, while others show a completely different characteristic.



**Figure 3.** Profile of measured Envisat data over the Mekong River near Luang Prabang (see Figure 2). The figure shows all available passes at this location. The data were retracked beforehand using the Multi-Subwaveform Retracker (MSR) (see Section 5.1). The blue region indicates the actual location of the river.

Most land-water transitions generate a hooking effect in the altimetry data. However, the hooking effect can usually be ignored in water-level estimates of larger water bodies where enough nadir measurements allow for a reliable estimate of the water level. For smaller rivers or lakes, (<1 km) where fewer altimetric measurements over the water body are available, the altimeter instrument can

lose lock due to the rapidly changing topography. In these cases, the hooking parabola can yield more reliable water levels than those obtained solely from measurements over the water body. Especially in the Envisat data, no measurements may exist where the river width is less than 350 m which is the distance between two consecutive measurements of Envisat. If the altimeter loses lock and misses some water surface data, information can be retrieved from distant parts of the hooking parabola. When analyzing the parabola, the water body does not need to be precisely located; indeed, its position might change over time.

The whole hooking effect and its treatment is based on the assumption that water causes the strongest return of all objects in the footprint. However, other objects may cause a strong reflection as well, such as river banks, surrounding rice fields, or urban buildings. If these objects are the strongest reflector over the whole pass they might even form a parabola on their own.

In the Mekong region investigated in this work rice fields have the highest importance of these reflecting objects. However, only a few time series examined in this work are surrounded by rice fields as we focus on the more northern part of the river in hilly and mountainous areas. For some crossings of the altimeter over the river no parabolic shape is visible in the heights which can be caused by water bodies surrounding the river such as rice fields. In these cases, correction the hooking parabola is not applicable.

## 5. Method

In our approach, the hooking parabola is derived from all measurements during a single pass around the river crossing. Our method can be divided into three main steps. First, we process the data using a newly developed retracker, the Multi-Subwaveform Retracker (MSR), which improves the data along the hooking parabola. The second (and most important) step detects the useful measurements that are influenced by the hooking effect. For this purpose, we adopt the RANSAC algorithm. The final step combines the measurements extracted by the RANSAC algorithm into a model of the hooking parabola, from which the water-level is obtained.

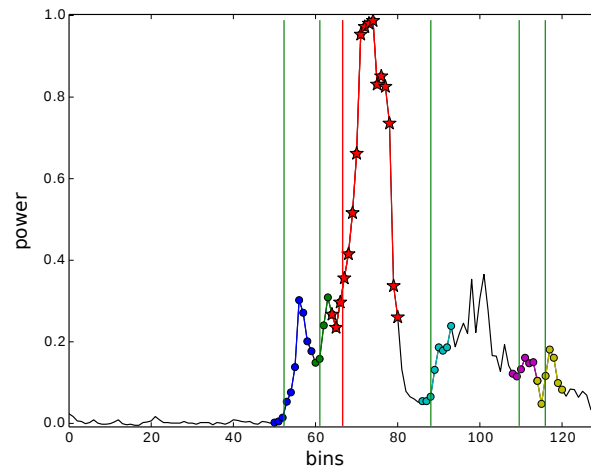
### 5.1. Multi-Subwaveform Retracker MSR

Our novel retracking method, MSR, was developed to improve the data quality of water-level time series, with special focus on the hooking effect. The method determines the sub-waveforms in Improved Threshold Retracker (ITR) data [8]. As an initial step, we consider here all ITR sub-waveforms, although only the first one is used for determining the water height. Figure 4 shows a typical waveform of a land-water transition, with the identified sub-waveforms marked by different colors. Over the water body itself, the first return and (and therefor the first sub-waveform) is caused by the reflection of the water. As the satellite retreats from the water, the first return signal most likely originates from land reflectance, and later sub-waveforms can be attributed to water returns. Thus, the first sub-waveform in the ITR waveform delivers good results over the water body, but yields no information on the hooking parabola. We assume that in each waveform, one of the sub-waveforms originates from water reflectance, but we cannot know which one. We can only assume that the water return yields the most intense sub-waveform. To find this sub-waveform, we define a weight for each subwaveform.

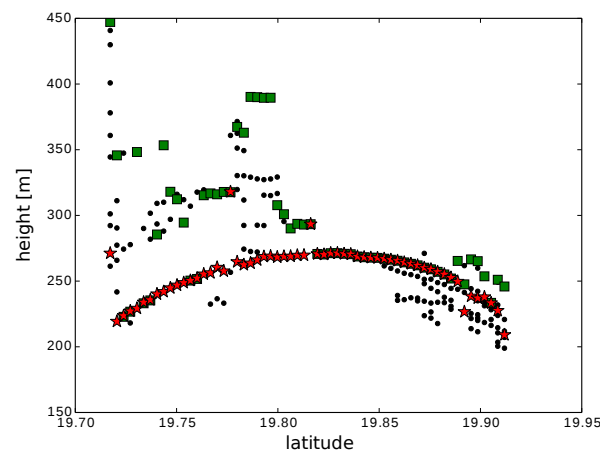
$$w_j = \sum_i P_i \quad i = 1 \dots \text{length}(\text{subwaveform}) \quad (5)$$

where  $w_j$  is the weighting factor of the  $j$ th sub-waveform and  $P_i$  denotes the return power at gate  $i$  of the  $j$ th sub-waveform. The number of sub-waveforms  $j$  and the number of gates  $i$  for each sub-waveform are determined in the MSR. The sub-waveform with the largest weight is assumed as the water-return sub-waveform and is therefor used in the height determination. This factor  $w_j$  defines the dominant sub-waveform as the one with the most power and highest amplitude. The weighting factor also defines the accuracy of the heights estimated in further calculations.

Figure 5 shows the MSR results for one crossing. The hooking parabola is clearly visible in the bundle of heights derived from all sub-waveforms (black dots), and the best sub-waveform (red stars) is usually caused by water reflection. The ITR results (green squares) turn out to be less suitable for extracting the hooking parabola.



**Figure 4.** Typical waveform near the water-land transition. The identified sub-waveforms are marked with different colors. The most intense sub-waveform (red stars) is assumed to come from the water surface. The vertical lines indicate the leading edge of each sub-waveform.



**Figure 5.** Altimeter profiles obtained by two retracker during one pass over the Mekong River. The heights obtained from all subwaveforms of the MSR are marked in black dots, the heights of the best waveforms in the MSR and Improved Threshold Retracker (ITR) data are marked as red stars and green squares, respectively.

The analysis of the heights retrieved from all subwaveforms affirms the assumption that the most prominent return is caused by the river. Other objects in the footprint might cause a strong reflection as well (river banks, surrounding rice fields, or urban buildings). However, heights caused by these targets would probably not lie on the hooking parabola, save they are such strong reflectors that they form a parabola on their own. Figure 5 shows clearly that most of the heights derived from the most prominent sub-waveform from a parabolic shape.

Nonetheless, it is possible to use all heights by sub-waveforms in the following algorithm and not only the best. By this, one can be certain that the water return is included. The effects of including all heights and not only the best in the algorithm will be examined in Section 6.

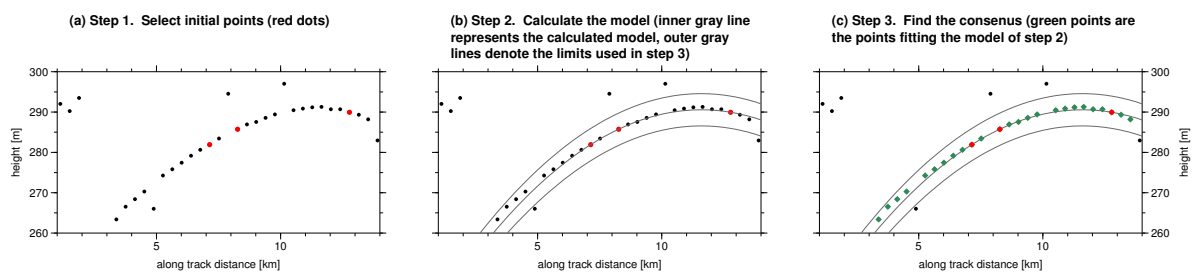


The retracking process is not impaired by the changing mode of the altimeter mentioned in Section 3.1. The nominal tracking gate and bin size are changing with each mode, but both are known for all bandwidths. Therefore, the retracking method can be used regardless of the actual measuring mode.

### 5.2. RANSAC Algorithm for Hooking Effect Estimation

The off-nadir altimetry returns from the water surface along the hooking parabola are identified by the RANSAC algorithm. This non-deterministic algorithm estimates the model parameters from data with outliers [21]. By virtue of its iterative approach, RANSAC more robustly handles data with outliers than conventional estimation algorithms. The iterative process begins with the best-fitting model within a given uncertainty range. The success of the algorithm depends on the accuracies of the input data. Moreover, it is helpful, but not necessary, to know the approximate number of outliers in the data.

The RANSAC algorithm proceeds through three steps (an example is illustrated in Figure 6):



**Figure 6.** The main steps of the RANDOM SAMPLE CONSENSUS (RANSAC) algorithm.

1. Select the initial values: A sufficient number of points to unambiguously define the model are randomly picked from all data points (e.g., 3 for a parabola and 2 for a line; see Figure 6a).
2. Calculate the a-priori model: This step uses the randomly chosen points from step 1 (See Figure 6b).
3. Find the consensus set:
  - (a) The consensus set contains all data points that fit the model within a specified limit, which is determined by the accuracy of the points. Given the uncertainty in the data, if many data points fit the model the randomly picked starting points have probably homed-in on the correct model (See Figure 6c).
  - (b) Recalculate the model using all points in the consensus set, and determine and save the new consensus set.

These three steps are repeated  $N$  times. The number of repetitions depends on the proportion of outliers and the targeted level of confidence.  $N$  can be estimated by

$$N = \frac{\log(1 - prob)}{\log(1 - (1 - \epsilon)^s)} \quad (6)$$

Here  $\epsilon$  is the relative number of outliers (as a percentage),  $prob$  denotes the confidence level (e.g., 99%), and  $s$  is the number of points required in the model determination. The confidence level defines the probability of finding the correct model. Although the number of outliers needs not be exactly known, an overly optimistic assessment reduces the number of repetitions, impeding the estimation of the correct model. If the number of outliers exceeds 60% (as in the present study),  $N$  can be 1000 or even larger.

Finally, the algorithm selects the “best” model. In the original RANSAC algorithm, the best model acquires the largest consensus set. We modified the RANSAC algorithm to account for the goodness of

fit of the points to the model. Instead of maximizing the number of points in the consensus set, we seek the model that minimizes the following error function

$$C = \sum_i v(\text{dist}(p_i, \text{model})) \quad (7)$$

with

$$v(\text{dist}(p_i, \text{model})) = \begin{cases} \text{dist}(p_i, \text{model}) & : \text{dist}(p_i, \text{model}) < \text{limit} \\ \text{constant} & : \text{dist}(p_i, \text{model}) > \text{limit}. \end{cases} \quad (8)$$

In these expressions,  $\text{dist}(p_i, \text{model})$  denotes the difference between the model and the  $i^{\text{th}}$  data point, and  $\text{limit}$  defines the threshold below which points are considered to fit the model (cf., Step 3a). The constant can be freely chosen but must be larger than the limit. In this study, we set  $\text{constant} = 2 * \text{limit}$ . Different values for the  $\text{constant}$  were tested on their performance in the algorithm, but  $\text{constant} = 2 * \text{limit}$  yielded overall the best results. The proportion of outliers determines whether a model is accepted. If the best model has a smaller consensus set than expected from the outliers, the model is discarded, and no model is returned. Again, it is preferable to overestimate rather than underestimate the outlier proportion.

In our study, the RANSAC algorithm finds the points belonging to the hooking parabola, not the final parameters of the parabola. Two parabola fittings are required, one each for the northern and southern banks of the water. The two half-parabolas are estimated separately because their shapes are influenced by the topography in the footprint (among other factors), which might differ on both sides of the river. Moreover, two half-parabolas might not share a common vertex. To estimate these parabolas, the algorithm requires only points from the northern or southern half of the data (with an additional 10% margin). For wider rivers with more than 500 m width, the algorithm also identifies points along a horizontal line over the river.

Sometimes, parabolic height profiles can result from the topography around the water body. To avoid such profiles, we impose three restrictions on Step 2 of the RANSAC algorithm. First, the hooking parabola is known to open toward the downside. Second, the parabola opening can be roughly estimated from Equation (4). Third, the water height is roughly known ( $\pm 25$  m) from a topography model. Therefore, models that predict water heights outside of this window are discarded. Under these restrictions, the RANSAC algorithm may not find any model and the height determination may fail in that epoch.

We also found that, in general, models with vertexes close to the middle of the water body yielded superior results.

### 5.3. Final Parameter Estimation

The RANSAC algorithm outputs up to three sets of points, one set for each of the two parabolas and one set for the line (optional). These models are not necessarily consistent, as they define different water levels. Therefore, to determine which of the models fit together, we use restricted parameter estimation with hypothesis testing [36] (Chapters 3.2 and 4.2). The three models were pairwise-tested if they predicted the same height with a level of confidence of 95%.

In the first step, the points of the two test models are used in a parameter estimation of the water-level height. The second step forces the parameters of both models to predict the same height of the water body. The subsequent hypothesis test evaluates whether the same-height restriction appreciably distorts the models. To this end, it compares the standard deviations of the unrestricted and restricted parameters.

Provided that the models fit together, their points are used in a parameter estimation to determine the water-level height. If none of the models fit together, the model with the smallest standard deviation is selected. The height of the vertex of the estimated hooking parabola is adopted as the final height when composing the time series.

#### 5.4. Post-Processing of the Time Series

##### 5.4.1. Slope Correction

The slope of the river influences the data processing in two ways. First, it influences the altimeter measurements at non-perpendicular intersections as described in Section 4. This influence is corrected in the retracked altimeter heights prior to the main algorithm. The correction is minor and exerts no significant influence on the results.

The second influence is caused by the orbital variation. The crossing location might change as much as 1 km, leading to variable biases in the time series. If the slope is known, the water levels can be corrected to the mean point of the river crossing. The slopes of the main Mekong River are reported in Gupta and Liew [37], but those of the tributaries are unknown. However, this correction improves the root mean square (RMS) (relative to data from a neighboring gauging station) by only a few centimeters, and is therefore a secondary correction. Hence, it is applied only to measurements along the main river.

##### 5.4.2. Outlier Detection

The time series resulting from the final parameter estimation might contain some outliers. Residual outliers are detected with prior information; namely, we expect that an annual signal dominates the time series. To this end, we fit an annual sinusoidal signal to the time series, and detect all points whose residual is outside the 95% quantile of all residuals. However, points outside of this quantile are not necessarily outliers. For instance, the peak of a flood, drought, or some other unusual event would be detected as an outlier. If the neighboring measurements show a similarly large residual, we attribute these points to extreme events in the time series, rather than discarding them as artefacts.

For the absolute standard deviation  $\sigma_{h_i}$  it is necessary to know the variance factor  $\sigma_0$ :

$$\sigma_{h_i}^2 = \sigma_0^2 \cdot \frac{1}{w_{h_i}}. \quad (9)$$

We assume, that  $\sigma_0$  is constant for all heights in one time series and for all time series. To derive  $\sigma_0$  we compare altimeter time series with close by *in situ* gauging data. The RMS is considered to be identical with the standard deviation of the whole time series, which is

$$(RMS)^2 = \sigma_T^2 = \frac{\sum_i^N \sigma_{h_i}^2}{N} = \frac{\sigma_0^2 \cdot \sum_i^N \frac{1}{w_{h_i}}}{N} \quad \text{with } N \text{ number of points in time series.} \quad (10)$$

In a strict sense the accuracies of the corrections and orbit have to be included here on the left hand side. However, as they are assumed constant over both time and the Mekong River region and small compared to the RMS ( $\approx 5$  cm) their contribution will be absorbed in the variance factor  $\sigma_0$ .

## 6. Results, Validation and Discussion

This section presents the water-level time series of the Mekong River and its tributaries, and validates them in comparisons with gauging data. For this purpose, we ensured that all virtual stations are close to an existing gauging station. The validation stations have different topographic surroundings and river widths, and thus, represent the diversity of real inland waters.

For the following results we had to set parameters in the RANSAC methods (see Section 5.2). Where Envisat data were measured at the highest resolution with a bandwidth of 320 MHz, we set *limit* = 1 and the outlier proportion to 70%. All virtual stations south of Chiang Khan (latitude 18° N) were measured in this bandwidth. At more northern stations, the bandwidth altered during a single pass and also among passes, increasing the proportion of outliers to 80%. In these cases, we set

$limit = 1$  if most of the measurements remained within the 320 MHz bandwidth; otherwise, we set  $limit = 1.5$ .

### 6.1. Results and Validation of the Water-Level Time Series Derived by the Hooking Approach

In this section, we derive the water-level time series from altimetry data using our proposed method, hereafter referred to as the hooking approach. Table 2 compares the results of the virtual stations with those of their nearest gauging stations. To compare the time series derived from our method with that of the gauging station data, we compute the RMS, the coefficient of determination ( $R^2$ ), and the number of points in the time series and all available Envisat cycles. For both the RMS and  $R^2$  coefficient we remove the long term mean from both the altimetry and gauge time series as the gauge data is not height referenced. The long term mean is only calculated with time points which are in both time series. An interpolation is not needed because of the daily resolution of the gauge data. The hooking approach outputs two values: the first from the time series without outlier detection and the second with outlier detection (see Section 5.4.2).

The hooking approach yields water-level time series for over three-quarters of the available cycles. The exception is the Ban Mixai Station, where the topography is atypical (see Section 6.2). We also identified three sources of missing cycles. First, even when Envisat collects measurements over a complete cycle, data might be missing over a specific region. Accordingly, data from this cycle is missing in all altimetric time series. Second, our algorithm might be unable to fit a parabola or a line to the available data; in this case, it would return no height measurements. Finally, even when the height data are well-fitted to a parabola or line, they might be discarded because they fall outside the specified height window.

The RMS of the time series derived by the hooking approach ranges from 0.34 m to 2.26 m, with a mean of 1.22 m. The RMS is below 1.5 m in 80% of the cases. The  $R^2$  lies between 0.55 and 0.97 with a mean of 0.83 (improving to 0.91 for the main river only). The RMS exceeds 2 m in only one time series (Luang Prabang 1), where the river topography is especially prone to seasonal effects (see Section 6.2). Excluding outliers (at the expense of reducing the number of data points) improves the derived time series in 8 out of 14 cases. In three cases, outlier detection does not change the results; in two cases, it slightly deteriorates the results; and in one case, the RMS decreases but also the correlation.

The quality ( $R^2$ ) of the gauging-station data is generally poorer along the tributaries of the Mekong River than along the main river. At only one station (Voeun Sai 1),  $R^2$  exceeds 0.8 and the RMS is very low (0.34 m). However, the low  $R^2$  at other stations does not necessarily imply defects in the hooking approach; rather, it depends largely on the quality of the gauging data. Despite our choice of gauging data (see Section 3.2), both the quality and length of the *in situ* data are inferior. As the time series shortens, the determination of  $R^2$  becomes more unstable and more strongly influenced by single outliers. In addition, the amplitudes of the annual variations are smaller along the tributaries (6–11 m) than along the main river (10–18 m). Assuming the same uncertainty in all water levels, smaller signal amplitudes will yield smaller  $R^2$ , although the magnitude of the absolute RMS difference is independent of amplitude. Along the tributaries, where the RMS is less sensitive to the signal amplitude, the RMS values are comparable (within the same order of magnitude) to those of the main river.

We also conducted this examinations with using the heights derived by all sub-waveforms and not only the best one (see Section 5.1). With all heights we can be certain to always include the height of the water return which should improve the results. However, we found that in general the results are neither improving nor deteriorating but the computational time is increasing. With an increasing number of points used in the RANSAC algorithm the portion of outliers is increasing as well. With this the number of necessary iterations of the RANSAC algorithm increases accordingly (see Equation (6)). This lead us to the conclusion that it is not worth using all heights instead of the best. However, it is possible that at other locations at other rivers using all heights prove to be useful.

**Table 2.** Comparison of results obtained by the hooking approach and the median approach. The MCR code is the code of the gauging station assigned by the Mekong River Commission and the station name is the nearest settlement. The stations above the double lines reside along the main Mekong River; those below the double lines reside along the tributaries. Listed are the pass numbers of the intersecting Envisat track, the location of the intersection, the distance between the gauge and the intersection in kilometer, the intersection length of the water body measured by the altimeter in meter, and the approximate amplitude measured at the gauge in meters. The quality of the results is indicated by three measures: the root mean square error (RMS) in meter between the altimetry time series and the gauging data, the squared correlation coefficient ( $R^2$ ), and the number of epochs in the time series (compared to the number of all available altimeter epochs).

MCR Code	Station Name	Dist.	River Name	Pass	Lon	Lat	Intersect. Length	max. Amplitude	Hooking Approach						Median Approach						
									without Outlier Detetction			with Outlier Detetction			RMS	$R^2$	# Epochs	RMS	$R^2$	# Epochs	# Avail. Epochs
									RMS	$R^2$	# Epochs	RMS	$R^2$	# Epochs							
010501	Chiang Saen	30	Mekong River	294	100.339	20.390	350	10	2.25	0.61	59	1.83	0.84	55	6.04	0.29	50	80			
011201	Luang Prabang 1	24	Mekong River	651	101.949	20.027	250	15	2.51	0.94	81	2.26	0.97	77	3.64	0.71	77	81			
	Luang Prabang 2	16	Mekong River	651	102.000	19.814	500	15	1.23	0.88	76	1.20	0.91	73	6.96	0.26	76	79			
011903	Chiang Khan 1	60	Mekong River	193	101.612	18.424	240	13	0.87	0.94	72	0.86	0.94	72	3.58	0.48	71	80			
	Chiang Khan 2	5	Mekong River	193	101.730	17.919	2860	13	1.28	0.86	65	1.08	0.89	62	10.23	0.00	52	80			
	Chiang Khan 3	35	Mekong River	666	101.943	18.084	340	13	1.46	0.91	70	1.48	0.90	67	1.96	0.75	73	80			
011901	Vientiane	19	Mekong River	651	102.436	17.980	1800	11	1.63	0.76	71	1.22	0.86	69	6.30	0.03	82	82			
013402	Mukdahan 1	39	Mekong River	21	104.984	16.283	3220	12	1.35	0.78	71	0.97	0.89	67	4.61	0.25	79	83			
	Mukdahan 2	60	Mekong River	952	105.068	16.109	1000	12	0.51	0.97	79	0.50	0.97	77	5.47	0.16	84	86			
120101	Ban Mixai	18	Nam Khan	666	102.3240	19.6856	90	4.50	1.79	0.58	46	1.68	0.70	43	3.90	0.32	67	81			
350101	Ban Keng Done	42	Xe Bangfai River	479	105.6986	16.3180	180	14	1.44	0.78	74	1.40	0.55	68	6.32	0.25	80	85			
440102	Vooun Sai 1	18	Tonle San River	322	106.7130	13.8421	460	7	0.97	0.79	73	0.34	0.88	69	3.44	0.39	82	84			
	Vooun Sai 2	16	Tonle San River	937	106.9437	14.0426	320	7	0.98	0.61	63	0.89	0.59	61	3.11	0.30	81	85			
430102	Siempang	31	Tonle Kong River	479	106.2653	13.8467	430	10	1.49	0.72	69	1.49	0.72	69	2.29	0.44	84	85			



## 6.2. Effects Influencing the Accuracy of the Water-Level Time Series

The topography of the Mekong River strongly influences the consistency of the altimetry time series and the gauging data. The gauging data are always maintained at settlement sites, where the river valley is usually widened. In contrast, altimetry measurements can be collected upstream or downstream of settlements, where the river is narrower. Consequently, the amplitudes are higher in the altimetry water-level time series than in the gauge time series. The amplitude difference is revealed in the RMS values but not in the determination coefficients ( $R^2$ ). This effect becomes obvious at the virtual stations of Chiang Saen and Luang Prabang 1.

The time series of the two virtual Luang Prabang stations are compared with that of the gauge station in Figure 7 (for visual clarity, the means of all time series are removed). The virtual station Luang Prabang 1 lies *ca.* 20 km upstream of the gauging station and Luang Prabang 2 *ca.* 5 km downstream. The amplitude of the time series at Luang Prabang 1 is clearly exaggerated during periods of low water, possibly because of the changing river width. The river width at different water levels can be measured from Landsat 7 panchromatic images (data available from <http://earthexplorer.usgs.gov/>), with a pixel size of 15 m [38]. For this purpose, we used two 2006 images: one from the dry season in March and the other from the flood season in August. The Landsat images of the three locations at both dates are shown in Figure 8. The measured river width is also shown. The Mekong River clearly widens at Luang Prabang 2 and at the gauging station, from 400 m and 240 m during the dry season to 520 m and 450 m, respectively, during the flood season. The wet season covers many sand banks that are visible during the dry season. On the other hand, the river at Luang Prabang 1 is little expanded during the flood season (from 240 to 300 m); instead, the water rises in the valley. This effect is clearly visible in the amplitudes of the three time series.

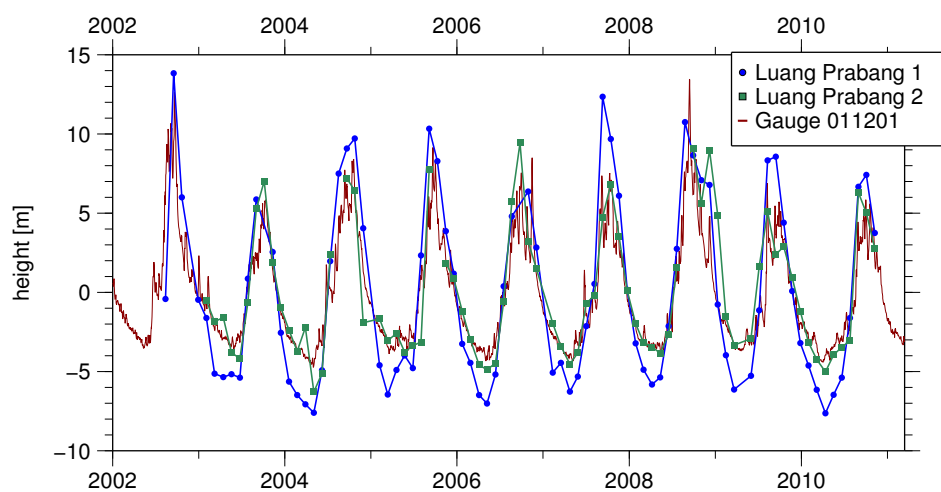
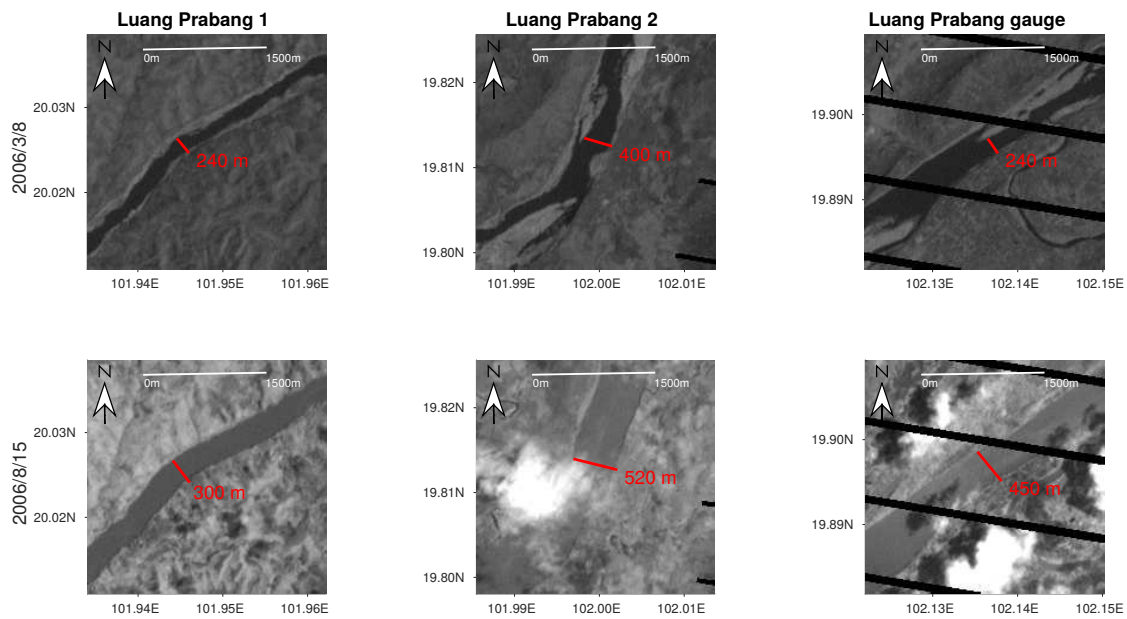


Figure 7. Water-level time series at the stations near Luang Prabang.

The change of river width can be used to estimate the river discharge at the three sites. This parameter should be similar for all three stations at a given time epoch. If assuming a trapezoidal shape of the river intersection, which is clearly a strong simplification, one is able to estimate the discharge variation from the change of intersection area. Furthermore, we assume an identical flow velocity at all locations. For the altimetry height differences we took the measured values closest in time to the Landsat images. Heights from the gauge are taken at the same epoch. For Luang Prabang 1 we obtain a height difference of 9 m and a change of the intersection area of 2430 m<sup>2</sup>; for Luang Prabang 2 we find a height difference of 5 m, and area change of 2300 m<sup>2</sup>; and for the gauge the height difference is 7 m which leads to an area change of 2450 m<sup>2</sup>. Considering the assumption of the simplified river shape and the accuracy of width measurement which of least one pixel on each side of the river and a pixel size of 15 m, no significant difference can be determined.



**Figure 8.** Landsat 7 images at the Luang Prabang sites during the dry season (**top**) and the wet season (**bottom**).

The influence of different topography even between close by stations shows the importance of co-location of the observations. Unfortunately, co-locating the gauge data and altimetry data is not available in most cases which should be kept in mind when gauge data and altimetry data are compared.

Besides the topography surrounding the station, the time series is affected by the angle of intersection between the altimeter track and the river. The two virtual stations at Mukdahan are closely spaced and surrounded by comparable landscape, but yield significantly different time series. These two virtual stations differ only by the angle at which the altimetry track intersects the river. Track 21 (Mukdahan 1 further upstream) intersects very obliquely and passes along the river for several kilometers, whereas track 952 (Mukdahan 2 further downstream) intersects nearly perpendicular to the river. Equivalently, the intersection length is three times longer at Mukdahan 1 than at Mukdahan 2. Reflecting this difference, the time series of Mukdahan 2 has an RMS against the gauge only half the size then the Mukdahan 1 series, and contains more data points. Because of the oblique intersecting angle and long intersection length, the altimetry data at Mukdahan 1 are influenced by the slope in the water surfaces, which can be compensated as described in Section 5.4.1. However, to properly correct this influence one has to know the exact slope, which points are influenced by the slope, and which points are cross track off-nadir measurements from the river. Gupta and Liew [37] provide river slopes for the Mekong. However, Mukdahan is located at the transition of two river units. For the upstream unit a slope of  $0.06 \frac{m}{km}$  is given, and for the downstream unit a slope of  $0.2 \frac{m}{km}$ . Using the altimetry data set, we calculate slopes between  $0.09 \frac{m}{km}$  and  $0.11 \frac{m}{km}$ , which seems to be realistic and in accordance to Gupta and Liew [37]. The time series of Mukdahan 1 and 2 are compared with that derived from the gauge data in Figure 9. Although neither of the virtual time series properly reproduces the maxima, Mukdahan 2 generates better data overall and a more accurate time series at the minima than Mukdahan 1.

The results are influenced not only by acute intersection angle between the altimetry track and the water body but also by the shape of the river at the intersection. This is exemplified at Chiang Khan 2, where the intersection length seems to not fit the expectations. At Chiang Khan 2, the river makes an s-bend which expands its apparent width under the altimetric track, although the river is only around 300 m wide at this point. The meandering of the river under the track also degrades the quality of the

results, as the measured water-level height can no longer be regarded as constant. Consequently, the time series deteriorates and displays more data gaps.

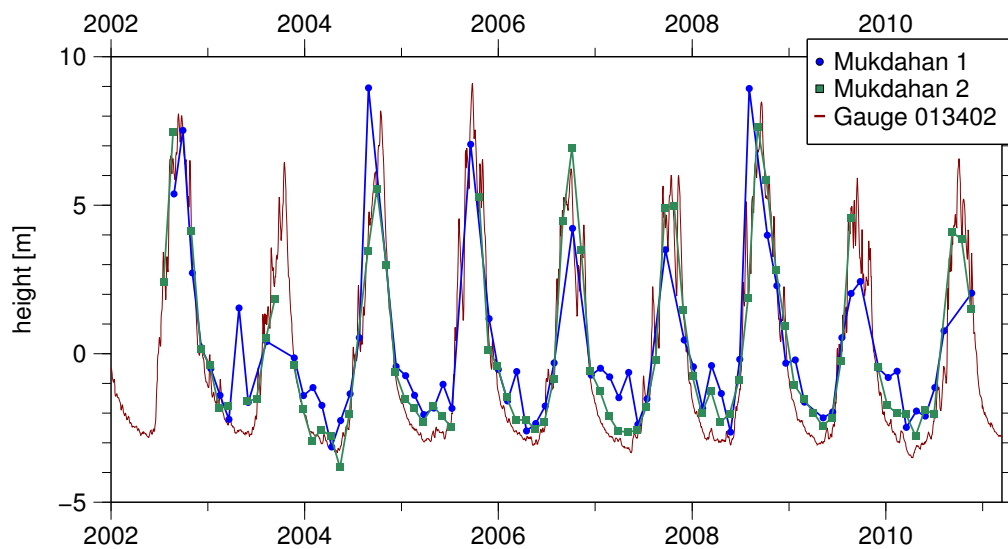
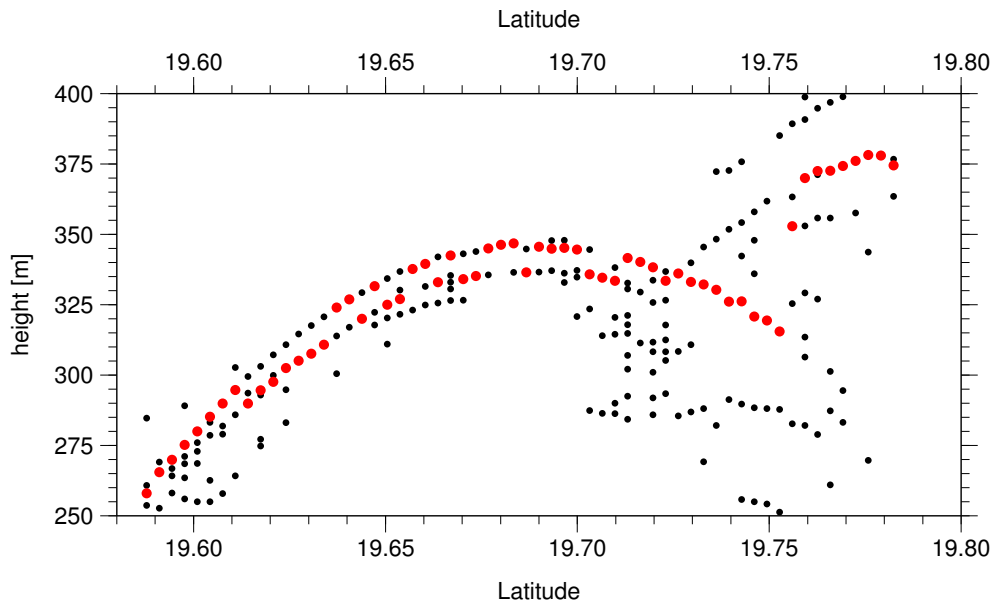


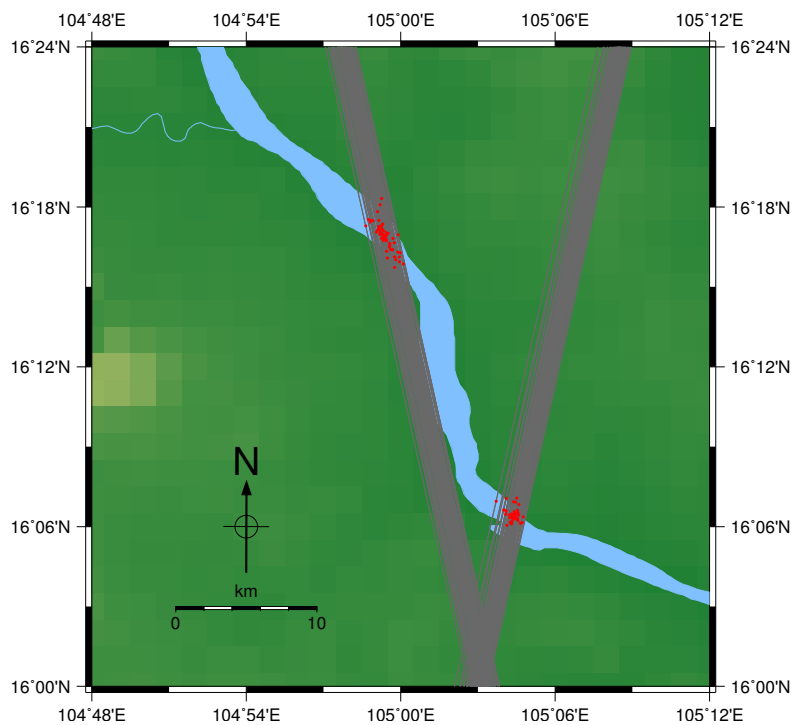
Figure 9. Water-level time series at the stations near Mukdahan.

As discussed above, the differences in the time series derived from the virtual stations and the gauging data depend on the valley width. However, at Ban Mixai station, the valley introduces an additional problem. In this stretch, the river flows through a narrow valley with high and very steep banks (indeed, the Ban Mixai site has the steepest banks among the sites investigated in this study). These steep banks reflect the radar signal multiple times, causing offsets in some of the height measurements. The water levels acquired through one cycle at Ban Mixai are plotted in Figure 10. We show both the results of all sub-waveforms from the MSR and the best one. Clearly, there are two shifted parabolas and partially also a third one. Two similar peaks are visible in the waveforms: the first from the original water reflection and the second from the multiway reflection. The applied retracker detects the dominant peak in either of these parabolas. Because the algorithm cannot distinguish between water returns and multiple reflections, it indiscriminately extracts measurements from either the upper or lower parabola. The two parabolas can be offset by up to 15 m, three times higher than the amplitude of the seasonal signal at Ban Mixai (5 m). Thus, the multireflections preclude a reliable time series at this location as many cycles cannot be used.

In Section 4 we discussed about possible reflections from objects other than the river. In the Mekong region the biggest concern are neighboring paddy fields which might show a strong reflectance as well. However, only a few virtual stations south of Vientiane are surrounded by rice field. In order to investigate these stations for possible non-river influences we can also use the hooking approach. Since it is not only possible to extract the height of the vertex of the hooking parabola but also its along-track position we are able to check whether the vertexes are in a close proximity to the river or not. A vertex further away of the river is probably caused by a reflection not originating from the river. In Figure 11 we show as an example all tracks and the position of the vertex for the two stations near Mukdahan which are surrounded by paddy fields. For both virtual stations all vertexes are in close vicinity to the river. Moreover, the water leverage of rice fields is season-dependent since they are only flooded during the wet season. Therefore, the reflection from the paddy fields would only degrade the measurements during these months, but we do not see any season dependency of the time series quality. Additionally, the consistency of the two parabolas we estimate for each target (see Section 5.3) can be used as indicator, that only desired reflections are included: if both parabolas fit together in height they originate most probably from the same reflector, *i.e.*, the river.



**Figure 10.** Along-track heights at Ban Mixai, Envisat cycle 16, retracked with MSR, showing all heights of all sub-waveforms in black and the best one in red.



**Figure 11.** The position of the vertices along the track for the two virtual stations near Mukdahan.

### 6.3. Comparison with Other Altimetry Products

This subsection compares the results of the hooking approach with those of an established median approach (e.g., [17]). The median approach uses the median of all MSR retracked altimeter measurements within a 3 km radius around the river center. Similar to our hooking approach, it also discards all heights outside a specified height window. The median approach is the easiest one for deriving the water-level time series of inland waters, requiring only the river location as additional information.

The hooking approach always yields superior results to the median approach for all virtual stations. The results of the median approach should improve with increasing intersection length; on the contrary, both RMS and  $R^2$  values reduce with increasing intersection length, whereas the results of the hooking approach are independent of intersection length. The median approach yields a meaningful time series only at the Chiang Khan 3 station. At all other stations, the resulting time series are of insufficient quality for further analysis.

Birkinshaw *et al.* [28] published altimetry time series at Luang Prabang 2 and Vientiane and at downstream stations not considered in this study. In their study they use ERS-2 and Envisat data between 1995 and 2003 and 2002 and 2008 respectively. The altimetry data is retracked according to the shape of each waveform (ocean-like, flat patch and two quasi-specular). They adopted two outlier detection methods in their time series determination. Comparing the altimetric and gauging-station time series at Luang Prabang 2, they reported RMS values of 1.99 m and 1.24 m with 22 points and 16 points in the time series, respectively. The hooking approach yields an RMS value of 1.23 m and 1.20 m without and with outlier detection and the time series contains 76 and 73 data points of 79 available epochs respectively (see Table 2). Over the same period, the Vientiane station yielded RMS values of 0.41 m and 0.44 m with 29 points and 33 points in the time series, respectively. For the Envisat data between 2002 and 2008 there should be approximately 60 repeats, but due to the retracking scheme and the measurement quality only 25 repeats for Luang Prabang and 37 for Vientiane are available after retracking. Our results for the same station show RMS values of 1.63 m and 1.22 m with 70 and 67 data points of 82 available epochs in the time series. Although the time series obtained by Birkinshaw *et al.* [28] and our hooking approach are comparable in quality measured with the RMS, the former is constructed from fewer data points.

We could not compare our results with data from the GLRM [11] or Hydroweb [9] databases, because no common virtual stations are available. In the ESA River and Lake database [5] two virtual stations near Mukdahan are included. For Mukdahan 1 the ESA time series has an RMS value of 2.42 m and  $R^2$  of 0.76 compared to 0.97 m and 0.89 in our results; the Mukdahan 2 ESA time series has an RMS value of 0.43 m and  $R^2$  of 0.99 compared to 0.50 m and 0.97 in our results.

#### 6.4. Application of the Hooking Approach to Other Missions

The Envisat satellite mission was preceded by ERS-2 and succeeded by SARAL/AltiKa. Therefore, we should consider whether the hooking approach is applicable to data from these satellites, and consequently, whether the time series can be expanded. To this end, we tested the hooking approach on a previously used virtual station, Chiang Khan 1. Although the hooking approach yielded reasonable time series from ERS-2 and SARAL/AltiKa data, the results were limited by inherent defects in the datasets. The ERS-2 data were discontinuous due to lost lock of the altimeter and changing bandwidth of the measurements, which were not properly documented. However, compared with the gauging data, the 1995–2003 time series derived by the hooking approach exhibited an RMS and  $R^2$  of 1.33 m and 0.85, respectively. Given the aforementioned limitations, the quality of these results is comparable to that of the Envisat analysis.

Data gaps also limited the processing of the SARAL/AltiKa data, but in this case, the gaps were caused by weather influence. SARAL/AltiKa measurements are more sensitive to atmospheric water content [39] than the other two missions, and are more degraded during the monsoon season. Nonetheless, the time series of SARAL/AltiKa derived by the hooking approach yielded good statistics (RMS = 0.76 m;  $R^2$  = 0.93 versus the gauging station data). The SARAL/AltiKa measurements should yield better results than Envisat, especially over smaller inland waters, because of the higher repetition rate (40 MHz) and smaller footprint. In fact, the time series derived from SARAL/AltiKa and Envisat were comparable in quality because the SARAL/AltiKa data are degraded by atmospheric water content.

According to these results, the hooking approach is applicable to ERS-2 and SARAL/AltiKa data, as well as to Envisat data. To collate the time series derived from the three altimetry missions, we must



tackle the range biases between the missions. One part of these biases is the radial range error *oerr* between the altimetry mission. This is determined over the open ocean and interpolated over the continents [34]. The correction of this bias is already included in the general corrections mentioned in Section 3.1. The other part mostly depends on the retracking algorithm and is not constant for even one retracker. Solutions to the later one are currently being debated. At Chiang Khan, the between-mission bias can be reduced by establishing the long-term mean. This approach enables us to merge the three time series into a single time series of nearly 20 years (see Figure 12). However, if the observed water body has no constant long-term mean, correcting the range biases becomes a complex problem, and is not attempted here.

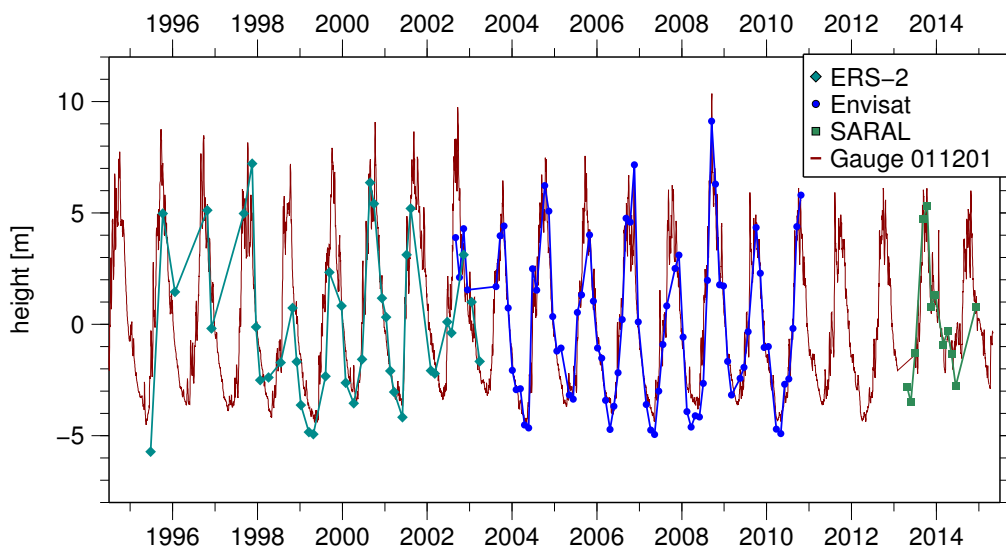


Figure 12. Twenty years multi-mission time series at the station Chiang Khan 1.

## 7. Conclusions

In this study, we demonstrated a method that derives reliable time series from off-nadir measurements acquired in satellite altimetry over rivers narrower than 500 m. These off-nadir measurements form a parabola in their height profiles whose vertex is the water level. Without further a-priori information, our method identifies the measurements affected by the hooking effect. From these measurements, it derives one water level per epoch. The method of deriving water levels is even applicable if no nadir measurement over the river is available as long as parts of the hooking parabola can be identified in the data.

The off-nadir measurements are detected by the RANSAC algorithm. Because the only additional information is the estimated location of the river, the operation of the algorithm is almost fully automatic. To validate our method, we derived water-level time series along the Mekong River and some of its tributaries, and compared them to the time series acquired by *in situ* gauging stations. The RMS of our time series ranged between 0.34 m and 2.26 m, with a mean of 1.22 m, and the  $R^2$  value ranged between 0.97 and 0.55, with a mean of 0.83 (improving to 0.91 when the tributaries were excluded).

The topography exerted a major influence on the altimetry data in the Mekong region. For example, the amplitude of the time series is increased in narrow valleys. Another influence was the intersection angle between the altimetry track and the river; as the intersection became more oblique, the data quality declined.

Comparing the results of our hooking approach with those of the median approach, we find that the hooking approach clearly outperforms the median approach in the study area. Using the hooking approach, we can measure water-level time series of rivers in the Mekong region even when no nadir

altimetry data are available. The approach is also well-suited to very small rivers (width below 500 m). Over wider rivers, there are usually sufficient altimetric measurements for conventional approaches such as used in DAHITI or other altimetry databases. Due to the automatization of the algorithm we are able to extract a larger number of water level time series in a short time. The hooking approach seems equally applicable to smaller river systems around the globe.

Here, we demonstrated the presence and correction of the hooking effect in data from three satellites: Envisat, ERS-2, and SARAL/AltiKa. After removing the range biases between consecutive missions, we derived a water-level time series over the Mekong River that spans 20 years.

In the next stage of our research, we will combine our hooking approach with the DAHITI database operated by the DGFI-TUM [12] in order to be able to process altimetry data for large and small rivers within one common software. Toward this goal, we are planning to correct the measured altimetric heights of the hooking parabola, providing more data points at the height of the water surface. This correction should yield reliable water-level time series for all kinds of inland water bodies, lakes, reservoirs, and rivers, including very small bodies (width <500 m).

**Acknowledgments:** The altimeter observations and geophysical corrections are taken from OpenADB (<http://openadb.dgfi.tum.de>). The altimeter missions are operated and maintained by ESA (ERS-2, Envisat) and ISRO/CNES (SARAL/AltiKa) and the original data sets are disseminated by AVISO, ESA, and PODAAC. This work was supported by the German Research Foundation (DFG) through fund SE 1916/4-1 and by the TUM International Graduate School of Science and Engineering (IGSSE). The open access was supported by the TUM Open Access Publishing Fund. We also thank the five anonymous reviewers whose constructive remarks helped to improve this paper.

**Author Contributions:** Eva Boergens developed the method, conducted the data analysis and wrote the majority of the paper. Denise Dettmering and Christian Schwatke helped with discussion of the method and results and writing the manuscript. Florian Seitz supervised the research and contributed to manuscript writing and organization.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Global Runoff Data Center. *Long-Term Mean Monthly Discharges and Annual Characteristics of GRDC Stations*; Technical Report; Federal Institute of Hydrology: Koblenz, Germany, 2013.
2. Morris, C.S.; Gill, S.K. Variation of Great Lakes water levels derived from Geosat altimetry. *Water Resour. Res.* **1994**, *30*, 1009–1017.
3. Birkett, C. The contribution of TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes. *J. Geophys. Res.: Oceans* **1995**, *100*, 25179–25204.
4. Birkett, C.M. Contribution of the TOPEX NASA radar altimeter to the global monitoring of large rivers and wetlands. *Water Resour. Res.* **1998**, *34*, 1223–1239.
5. Berry, P. Two decades of inland water monitoring using satellite radar altimetry. In Proceedings of the Symposium on 15 Years of Progress in Radar Altimetry, Venice, Italy, 13–18 March 2006; Volume 15.
6. De Oliveira Campos, I.; Mercier, F.; Maheu, C.; Cochonneau, G.; Kosuth, P.; Blitzkow, D.; Cazenave, A. Temporal variations of river basin waters from Topex/[Poseidon satellite altimetry. Application to the Amazon basin. *Comptes Rendus L'Académie Sci.-Ser. IIA-Earth Planet. Sci.* **2001**, *333*, 633–643.
7. Chelton, D.B.; Ries, J.C.; Haines, B.J.; Fu, L.L.; Callahan, P.S. Satellite altimetry. *Int. Geophys.* **2001**, *69*, i–ii.
8. Gommenginger, C.; Thibaut, P.; Fenoglio-Marc, L.; Quartly, G.; Deng, X.; Gómez-Enri, J.; Challenor, P.; Gao, Y. Retracking altimeter waveforms near the coasts. In *Coastal Altimetry*; Benveniste, J., Cipollini, P., Kostianoy, A.G., Vignudelli, S., Eds.; Springer: Berlin, Germany, 2011; pp. 61–101.
9. Crétaux, J.F.; Jelinski, W.; Calmant, S.; Kouraev, A.; Vuglinski, V.; Bergé-Nguyen, M.; Gennero, M.C.; Nino, F.; Del Rio, R.A.; Cazenave, A.; *et al.* SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. *Adv. Space Res.* **2011**, *47*, 1497–1507.
10. Berry, P.; Bracke, H.; Jasper, A. Retracking ERS-1 altimeter waveforms over land for topographic height determination: An expert systems approach. *ESA SP* **1997**, *1*, 403–408.
11. Birkett, C.; Reynolds, C.; Beckley, B.; Doorn, B. From research to operations: The USDA global reservoir and lake monitor. In *Coastal Altimetry*; Springer: Berlin, Germany, 2011; pp. 19–50.

12. Schwatke, C.; Dettmering, D.; Bosch, W.; Seitz, F. DAHITI – an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 4345–4364.
13. Frappart, F.; Calmant, S.; Cauhopé, M.; Seyler, F.; Cazenave, A. Preliminary results of ENVISAT RA-2-derived water levels validation over the Amazon basin. *Remote Sens. Environ.* **2006**, *100*, 252–264.
14. Zhang, M.; Lee, H.; Shum, C.; Alsdorf, D.; Schwartz, F.; Tseng, K.H.; Yi, Y.; Kuo, C.Y.; Tseng, H.Z.; Braun, A.; et al. Application of retracked satellite altimetry for inland hydrologic studies. *Int. J. Remote Sens.* **2010**, *31*, 3913–3929.
15. Santos da Silva, J.; Calmant, S.; Seyler, F.; Rotunno Filho, O.C.; Cochonneau, G.; Mansur, W.J. Water levels in the Amazon basin derived from the ERS-2 and ENVISAT radar altimetry missions. *Remote Sens. Environ.* **2010**, *114*, 2160–2181.
16. Kuo, C.Y.; Kao, H.C. Retracked Jason-2 altimetry over small water bodies: Case study of Bajhang River, Taiwan. *Mar. Geod.* **2011**, *34*, 382–392.
17. Maillard, P.; Bercher, N.; 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 Sens. Environ.* **2015**, *156*, 226–241.
18. Calmant, S.; Seyler, F.; Cretaux, J.F. Monitoring continental surface waters by satellite altimetry. *Surv. Geophys.* **2008**, *29*, 247–269.
19. Santos da Silva, J.; Seyler, F.; Calmant, S.; Rotunno Filho, O.C.; Roux, E.; Araújo, A.A.M.; Guyot, J.L. Water level dynamics of Amazon wetlands at the watershed scale by satellite altimetry. *Int. J. Remote Sens.* **2012**, *33*, 3323–3353.
20. Frappart, F.; Papa, F.; Marieu, V.; Malbeteau, Y.; Jordy, F.; Calmant, S.; Durand, F.; Bala, S. Preliminary assessment of SARAL/AltiKa observations over the Ganges-Brahmaputra and Irrawaddy Rivers. *Mar. Geod.* **2015**, *38*, 568–580.
21. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395.
22. Chen, C.S.; Hung, Y.P.; Cheng, J.B. RANSAC-based DARCES: A new approach to fast automatic registration of partially overlapping range images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 1229–1234.
23. Kim, T.; Im, Y.J. Automatic satellite image registration by combination of matching and random sample consensus. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1111–1117.
24. Dung, L.R.; Huang, C.M.; Wu, Y.Y. Implementation of RANSAC algorithm for feature-based image registration. *J. Comput. Commun.* **2013**, *1*, 46–50.
25. Tarsha-Kurdi, F.; Landes, T.; Grussenmeyer, P. Hough-transform and extended ransac algorithms for automatic detection of 3d building roof planes from LiDAR data. In Proceedings of the ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, Espoo, Finland, 12–14 September 2007; Volume 36, pp. 407–412.
26. Yan, J.; Jiang, W.; Shan, J. Quality analysis on RANSAC-based roof facets extraction from airborne LiDAR data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *1*, 367–372.
27. Frappart, F.; Do Minh, K.; L’Hermitte, J.; Cazenave, A.; Ramillien, G.; Le Toan, T.; Mognard-Campbell, N. Water volume change in the lower Mekong from satellite altimetry and imagery data. *Geophys. J. Int.* **2006**, *167*, 570–584.
28. Birkinshaw, S.; O’Donnell, G.; Moore, P.; Kilsby, C.; Fowler, H.; Berry, P. Using satellite altimetry data to augment flow estimation techniques on the Mekong River. *Hydrol. Process.* **2010**, *24*, 3811–3825.
29. Mekong River Commission, Ed. *Overview of the Hydrology of the Mekong Basin*; Mekong River Commission: Vientiane, Lao PDR, 2005.
30. Scharroo, R.; Smith, W.H. A global positioning system-based climatology for the total electron content in the ionosphere. *J. Geophys. Res.: Space Phys.* **2010**, *115*, doi:10.1029/2009JA014719.
31. Boehm, J.; Kouba, J.; Schuh, H. Forecast Vienna Mapping Functions 1 for real-time analysis of space geodetic observations. *J. Geod.* **2009**, *83*, 397–401.
32. McCarthy, D.D.; Petit, G. *IERS Conventions (2003)*; Technical Report; Verlag des Bundesamtes für Kartographie und Geodäsie: Frankfurt am Main, Germany, 2004.

33. Förste, C.; Bruinsma, S.; Shako, R.; Abrikosov, O.; Flechtner, F.; Marty, J.C.; Lemoine, J.M.; Dahle, C.; Neumeyer, H.; Barthelmes, F.; *et al.* A new release of EIGEN-6: The latest combined global gravity field model including LAGEOS, GRACE and GOCE data from the collaboration of GFZ Potsdam and GRGS Toulouse. *Geophys. Res. Abstr.* **2012**, *14*, 2821.
34. Bosch, W.; Dettmering, D.; Schwatke, C. Multi-mission cross-calibration of satellite altimeters: Constructing a long-term data record for global and regional sea level change studies. *Remote Sens.* **2014**, *6*, 2255–2281.
35. Benveniste, J.; Baker, S.; Bombaci, O.; Zeli, C.; Venditti, P.; Zanife, O.; Soussi, B.; Dumont, J.; Stum, J.; Milagro-Perez, M. *ENVISAT RA-2/MWR Product Handbook, Issue 2.2*; Technical Report; PO-TN-ESR-RA-0050; European Space Agency: Frascati, Italy, 2007.
36. Koch, K.R. *Parameter Estimation and Hypothesis Testing in Linear Models*; Springer Science & Business Media: Berlin, Germany, 1999.
37. Gupta, A.; Liew, S.C. The Mekong from satellite imagery: A quick look at a large river. *Geomorphology* **2007**, *85*, 259–274.
38. NASA. *Landsat 7 Science Data Users Handbook*; Technical Report; National Aeronautics and Space Administration: Washington DC, USA, 2011.
39. Schwatke, C.; Dettmering, D.; Börgens, E.; Bosch, W. Potential of SARAL/AltiKa for inland water applications. *Mar. Geod.* **2015**, *38*, 626–643.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).