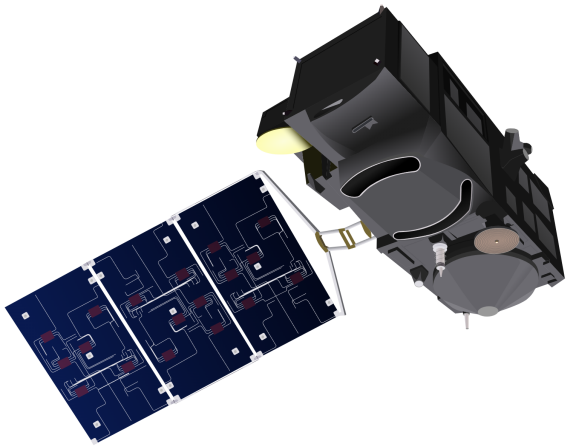
University of Stuttgart
Institute of Geodesy

Dealing with challenges of altimetry-based surface water height derivation over boreal catchments: case study of Mackenzie river



Bachelorarbeit im Studiengang
Geodäsie & Geoinformatik
an der Universität Stuttgart

Jiaxin Liu

Stuttgart, February 2022

Supervisor: M.Sc. Sajedeh Behnia

Examiner: Prof. Dr.-Ing. Nico Sneeuw

Erklärung der Urheberschaft

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Ort, Datum Unterschrift

Abstract

The Earth is a watery place, which fills the oceans, rivers, and lakes. Approximately 71% of the Earth's surface is water-covered. Measurements of surface water level in oceans, lakes, rivers and coastal waters are important for a variety of reasons. In the short term, this can, for example, help to alert to dangerous water level so that actions can be taken in advance, while in the long term, monitoring water levels can provide even greater insight into patterns of water dispersal in the area, and measure, for example, the effects of global warming.

Satellite altimetry, which was originally designed for oceanography in the 1970s, has revolutionized our knowledge of the marine gravity field, of the dynamics of the oceans and even of land hydrology. It is a space measurement technique that uses artificial satellites to measure the altitude from the satellite to the Earth's surface. Due to its high resolution, global coverage and short revisit time, it is playing an increasingly important role in measurements of water level. For some years, this technology has also been used to retrieve water levels from rivers, lakes, and any inland water body as well. However, compared with the wide seas, measurements of inland water bodies involve many challenges.

In this paper, we will take the Mackenzie River in northern Canada as a research object and process the data through the Matlab-based program Atlbundle+ to study the potential of inland altimetry, the problems it faces and the possible ways to find the solution. In general, there are two perspectives that will be considered: Firstly, how we can accurately remove outliers in the measurements, and secondly, how to improve the altimetry-driven water level time series by improving the retracking methods. Finally, based on the results and the procedure, a systematic analysis of the inland altimetry can be carried out.

Keywords: Inland altimetry, AtlBundle+, Time Series, Water Level, Outlier Identification, ASTERDEM/SWORD, LEIPI

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Altimetry

Satellite Altimetry, which was originally designed for oceanography in the 1970s, is a space measurement technique that uses artificial satellites as carriers to measure the altitude from the satellite to the Earth's surface using radar, laser, and other ranging techniques to obtain the topography of the Earth's surface. The concept was first introduced at the Conference on Solid Earth and Ocean Physics held at Williamstown in 1969 by American geodesist William M. Kaula and was regarded as a key technique for studying solid Earth and ocean dynamics.

Satellite altimetry can be used to measure sea surface topography and altitude with high resolution. It can cover the globe, rather than being limited to a particular region. Since the first missions in the mid-1970s, the accuracy of the single sea surface height measurements has improved by a factor of 100 and is today approaching the 1 cm level. In addition, it also provides more information than the surface observations did. By measuring sea surface topography, altimeters provide information about the Earth's gravity field, the shape and structure of the seafloor, the integrated heat and salt content of the ocean, and geothermal ocean currents. Much progress has been made in developing operational ocean applications, and altimeter data are now routinely utilized in near real-time to help predict El Niño and monitor coastal circulation (Cazenave, 2019).

1.2 Radar Altimetry

1.2.1 Basic principle

Satellite altimetry can be realized by different techniques. In this thesis, all measurements are based on radar altimetry. So the next focus is the principle and process of radar altimetry.

The principle of radar altimetry is simple and easy to understand, but it is quite difficult to implement and achieve high accuracy in reality. The basic idea is that the satellite sends a radar pulse to the Earth and measures the two-way travel time between the satellite antenna and the ocean surface at the nadir of the spacecraft. The time here refers to the time required for the signal to make the round trip from the satellite to the ocean surface and back to the satellite. Since the signal travels at the speed of light, the first requirement is that the measurement of the round-trip time must be very precise, and of course, the position of the satellite must also be accurately determined, which is accomplished with radio ranging, laser tracking, and the

Global Positioning System (GPS). Because the slight variations in the Earth's gravitation can cause changes in the satellite orbit, this factor must also be considered. In order to determine the variability of the ocean surface height field, the measurements must be made with reference to the geoid, a hypothetical shape of the Earth's gravity field that shapes the average (mean) sea level, which is also a very important part of the many corrections that will be mentioned later (Acker et al., 2003).

1.2.2 Basic formula

After introducing the basic principle, the calculation process of radar altimetry will be derived. Satellite altimeter transmits signals at high frequencies (over 1,700 pulses per second) and measures the distance from the satellite to the sea level denoted $\tilde{R}$. This range is measured from a time taken for incident radiation of a signal to reflect back to the satellite altimeter, which enables determining of the sea surface height (Grgić and Bašić, 2021). Without refraction accounted for, based on the speed of light in vacuum c and the round-trip travel time t, the range $\tilde{R}$ is defined as:

$$\tilde{R} = \frac{ct}{2} \tag{1.1}$$

When electromagnetic waves travel through the atmosphere, they are decelerated by water vapor or ionization and are also affected by some other factors. After applying the corrections, the corrected range R can be present as:

$$R = \tilde{R} - \sum_j \Delta R_j = \tilde{R} - (\Delta R_{\text{tropoD}} + \Delta R_{\text{tropoW}} + \Delta R_{\text{iono}} + \Delta R_{\text{SSB}} + ...) \tag{1.2}$$

where $\Delta R_j, j = 1, ...$ is the sum of the corrections, which encompasses the ionospheric correction, wet and dry tropospheric correction, sea state bias (for great lakes), tidal loading correction, Earth tide correction, pole tide correction and the other corrections.

Next is the altitude of the satellite, which depends upon a number of constraints. Factors such as inclination, atmospheric drag, gravity acting on the satellite all affect the altitude of the satellite. The satellite can be tracked in several ways so as to measure its altitude with the highest possible accuracy and thus determine its precise orbit, accurate to within 1 or 2 cm (Rosmorduc.V et al., 2018). The satellite altitude h_{sat} is determined through different techniques, including Doppler shift, Global Navigation Satellite System (GNSS), and laser tracking. The ellipsoidal surface height h (see Fig. 1.1) is then the difference between the satellite's altitude relative to the reference ellipsoid and the corrected measured range:

$$h = h_{\text{sat}} - R = h_{\text{sat}} - \tilde{R} + (\Delta R_{\text{tropoD}} + \Delta R_{\text{tropoW}} + \Delta R_{\text{iono}} + \Delta R_{\text{SSB}} + ...) \tag{1.3}$$

Normally, for rivers and oceans, we use the orthometric height H more often:

$$H = h_{\text{sat}} - R = h_{\text{sat}} - \tilde{R} + \sum_j \Delta R_j - N \tag{1.4}$$

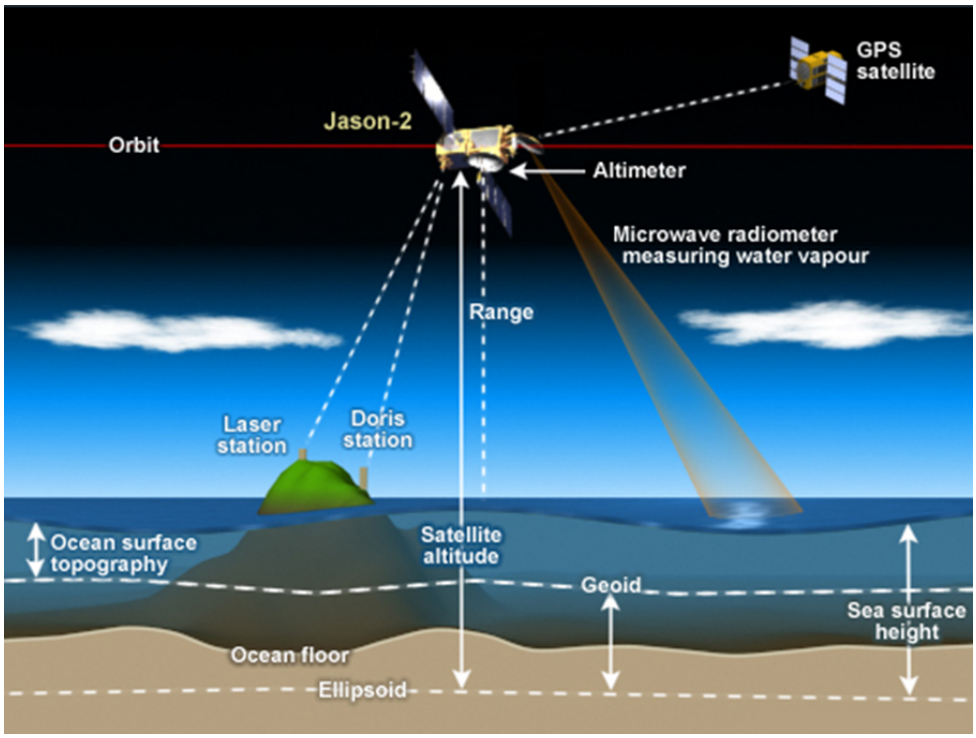


Figure 1.1: *How satellite radar altimetry works (STAR, 2020)*

1.3 Relevant missions

Accurate satellite altimetry missions provide us with new ways and ideas for observing the Earth and ocean. The high-precision altimeter measurement enables us to systematically observe and record the sea level, which provides a basis for subsequent research. Until the advent of Seasat (1978), altimeters were presented only as concepts, but the advent of Seasat was the first demonstration of scientific results. And then, since the application of Geosat (1986), these altimeter missions have been providing important data to the international user community (Rosmorduc.V et al., 2018).

Figure 1.2 and Table 1.1 show the construction of the satellite radar altimetry constellation in the past, present and future. Throughout the development of these platforms, incremental increases in precision have reduced the errors associated with range and orbit measurements from 1 m/500 cm, to 2 cm/3 cm in Sentinel-3.

The satellites used as data sources in this paper are Jason series (Jason-1, 2, 3), Envisat and Saral, Envisat XT, Sentinel-3A and 3B. All these satellites provide altimetry data which is the basis for subsequent time series, water level etc. Next, some of these satellites will be briefly introduced in order to have a general understanding.

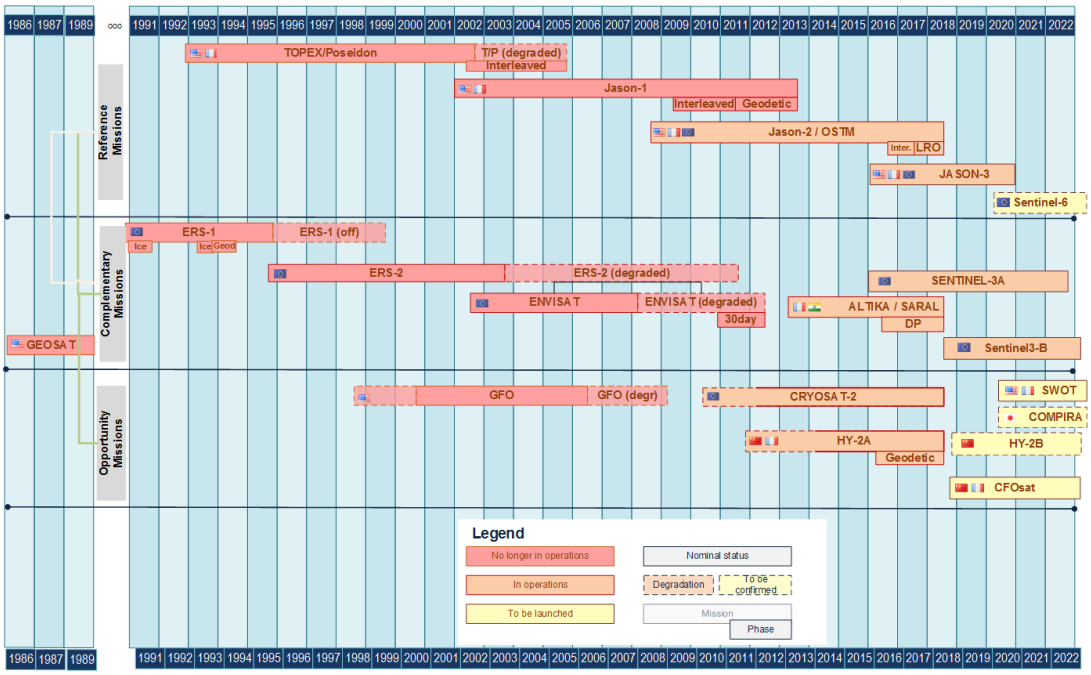


Figure 1.2: *Timeline of modern radar altimetry missions (DUACS, n.d.)*

Satellite Mission	Mission Period	Inclination (deg)	Revisit Time (Days)	Along-Track Resolution (km)
Geosat	1985-1990	108	17	1.7
ERS 1 and 2	1990-2011	98.5	35	1.7/3.4
TOPEX/Poseidon	1992-2006	66	10	2.2
GFO	1998-2008	108	17	1.7
Jason 1, 2, and 3	2001-present	66	10	2.2
Envisat	2002-2012	98.55	30-55	1.7
CryoSat-2	2010-present	92	369	0.25/1.6
SARAL	2013-present	98.55	35	1.4
Sentinel-3	2016-present	98.6	27	0.3
SWOT	Planned on 2022	77.6	21	0.1

Table 1.1: *List of satellite altimetry data (Gao et al., 2019)*

1.3.1 Jason Series

Jason is the name of a joint CNES/NASA oceanography mission series, which is used to monitor global ocean circulation, explore the tie between the oceans and atmosphere, improve global climate predictions, and can also monitor events such as El Niño conditions and ocean eddies (Kramer, 2019).

Jason-1, the first satellite of the Jason series, was the first follow-on mission to the TOPEX/-Poseidon mission, which measured ocean surface topography from 1992 through 2005. It was launched on December 7, 2001 an American Delta II rocket from Vandenburg Air Force Base in California and was flown on a French spacecraft with a radar altimeter on board. It was

decommissioned in July 2013 after the failure of its last remaining transmitter left it unable to send data to Earth.

Jason-2 or OSTM (Ocean Surface Topography Mission) is a follow-on to the Jason 1 mission. It carries on Jason-1's predecessor's mission, such as monitoring global ocean circulation (Krebs, 2021). After 11 years of operation, it was finally deactivated on 1 October 2019.

Jason-3 is then the fourth satellite in the framework of the cooperation between CNES, NASA, Eumetsat and NOAA. It is proposed to take over and also continue the missions of Topex, Jason-1 and 2. The satellite was launched by a Falcon-9 rocket in January 2016 and is still in operation.

1.3.2 Sentinel

Measurements with high accuracy of the sea surface topography, sea/land surface temperature, and sea/land surface color are used to support ocean forecasting systems, environmental monitoring, and climate monitoring. These measurements are exactly the missions of Sentinel-3. It consists of three satellites (SENTINEL-3A, 3B and 3C) and is jointly operated by ESA and EUMETSAT to deliver operational ocean and land observation services.

The main process is in the following Table 1.2:

Date	process
16 February 2016	Successful Launch of S3A from Plesetsk, Russia.
18 February 2016	LEOP phase concluded successfully
26 February 2016	Platform In-Orbit Verification completed
04 March 2016	Payload In-Orbit Verification completed
27 April 2018	Successful Launch of Sentinel-3B from Plesetsk, Russia.

Table 1.2: *Major milestones of Sentinel-3*

Sentinel-3 is 3 axis stabilized satellite flying in a frozen sun-synchronous orbit. It has a near-polar, sun-synchronous repeat orbit with the ratio 385/27 because of the requirements of the topography mission.

1.3.3 Envisat

Envisat, which was ESA's successor to ERS, is the largest free-flying and the most complex satellite ever built in Europe (Dubock et al., 2001). It was launched on 1 March 2002 on an Ariane-5 rocket from Europe's spaceport in French Guiana. But just a few weeks after celebrating its 10th year in orbit, ENVISAT's communications were suddenly interrupted. After rigorous attempts to re-establish contact and investigation of the failure, contact with the satellite was still not made and the missions had to end. This is a very big loss because it carries more advanced imaging radar, radar altimeter and radiometer instruments. In addition, there are new instruments, including medium-resolution spectrometer which is sensitive to land features and ocean colors and also atmospheric sensors monitoring trace gases.

1.3.4 SARAL

SARAL is a cooperative altimetry technology mission of ISRO (Indian Space Research Organization) and CNES (Space Agency of France) for oceanographic studies. CNES provides the payload module consisting of the AltiKa altimeter, DORIS, LRA, and Argos-3 DCS (Data Collection System), and the payload data reception and processing functions, while ISRO is responsible for the platform, launch, and operations of the spacecraft. The overall objectives are to realize precise, repetitive global measurements of sea surface height, significant wave heights and wind speed. The SARAL mission is complementary to the Jason-2 mission. The combination of two altimetry missions in orbit has a considerable impact on the reconstruction of SSH, reducing the mean mapping error by a factor of 4 (ESA, 2022). Beyond the expected contributions and SARAL's very good integration into the objectives of the constellation of altimetric satellites, more important is its application to Ka-Band. It uses a single frequency in Ka-band (35.75 GHZ) when other satellites use Ku-Band or C-Band. It is the first oceanographic altimeter using such a high frequency. The advantages of the Ka-band are in short the reduction of ionosphere effects, the smaller footprint, the better horizontal and vertical resolution.

1.4 Problem Statement

1.4.1 Brief Introduction

In general, satellite altimetry is playing an increasingly important role in measurements of water level. Over the years, this technology is no longer limited to the ocean but has also been used to retrieve water levels from rivers, lakes and inland water bodies. Of course, measurements of inland water level face many challenges compared to the wide ocean. Compared with the relatively simple sea surface environment, the more complex environment of inland rivers such as complex channels and tributaries, changing terrain, excessive human intervention and many other factors lead to distortion of the telemetry signal when reflected from the surface. Difficulties that are closely related to this paper are for example, how to correctly determine the measurement objects or extract measurements that really come from the water surface, how to accurately identify outliers and assess the altimetry results without in-situ data. Based on these, how to improve the results is also the focus.

Our research object in this paper is a river in Canada called Mackenzie River. By assessing altimeter's performance in the Mackenzie River (see Fig. 1.3), we explore how to deal with these challenges of altimetry-based surface water height derivation.

1.4.2 Workflow

The whole work is mainly divided into three parts:

1.4.2.1 Data preparation and analysis

First, the measurement data of the altimeter was extracted through the Matlab program as a basis. Altimeter data along the Mackenzie river was collected and analysed. Problems were identified through specific case studies.

1.4.2.2 Improvements

Second, how to find suitable improvements were the key points. In this paper, by using external datasets (see Section 4.1-4.2), or improving the so called "retracking "method (see Section 2.4 and 4.3), we tried to improve the initial results and then through case analysis to assess the effects and defects of these approaches.

1.4.2.3 Final Analysis and Conclusions

Finally, after several improvements the whole process is compared, analyzed and concluded.

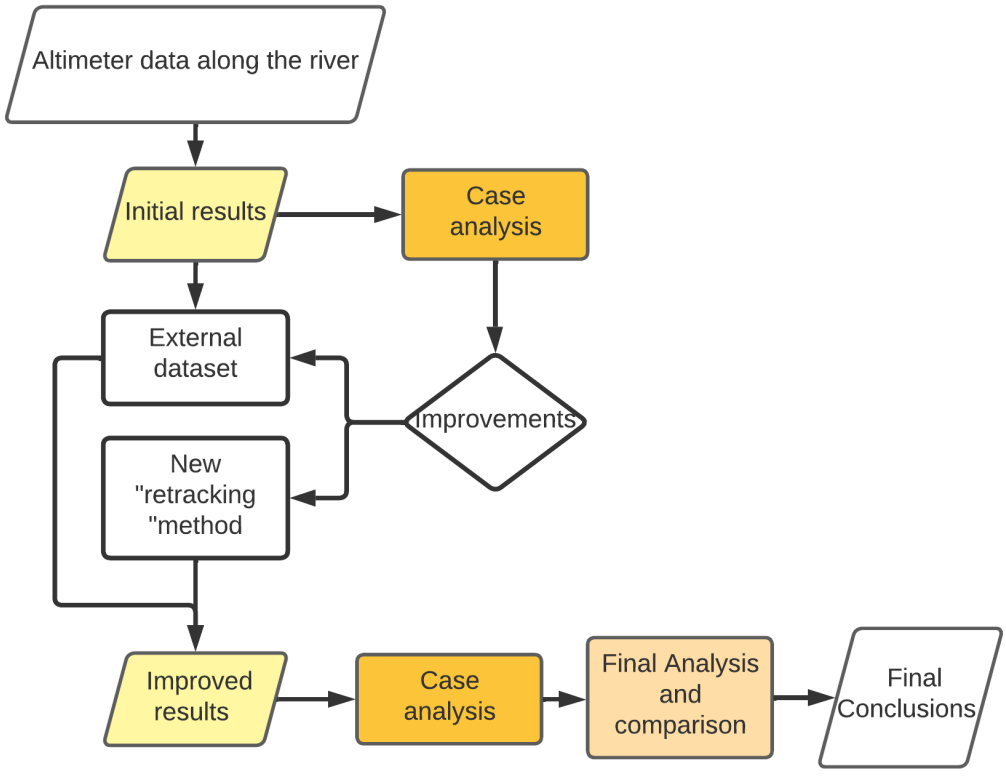


Figure 1.3: *Brief Workflow: deal with the challenges of altimetry-based surface water height derivation*

Chapter 2

Altimetry Waveform

2.1 Waveform

As we mentioned above, the altimeter measures the distance from the satellite to the sea level. The fundamental principle of any kind of altimeter is based on the shape and timing information of the returned pulse. Figure 2.1 shows the interaction of a pulse with the scattering surface and the procedure of constructing the returned waveform.

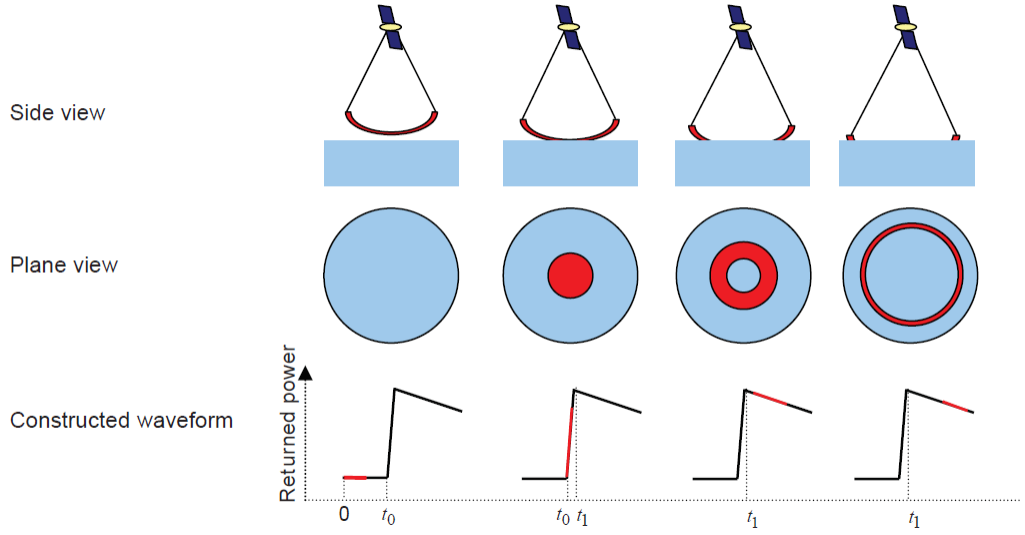


Figure 2.1: *Interaction of a pulse and scattering surface, and the procedure of constructing the returned waveform (Tourian, 2013)*

Initially ($0<t_0<t_0$), a pulse of electromagnetic energy (nadir-pointing radar) is transmitted from the antenna of satellite altimeter, propagating in a spherical wavefront (Guo et al., 2009).

At $t = t_0$, the wavefront meets the nearest wave crest directly under the satellite. A point is illuminated at this moment, and the signal starts to return to the altimeter.

From t_0 onwards, the returned signal rises up, the footprint being a disc linearly spreading with time, which makes the corresponding return signal increase up to the maximum at $t = t_1$ (Rosmorduc.V et al., 2018).

Thereafter ($t > t_1$), the backscattered energy begins to decline due to the limited antenna beamwidth and the fewer proper reflected facets, and the footprint turns to a ring with increasing radius but constant area.

Satellite records the averaged returned waveform, which is a time series of the mean returned powers. The mean returned waveform is the result of the convolution of three terms, i. e., the average impulse response of the smooth sea surface, the sea surface elevation distribution and the system's point target response (Hayne, 1980).

This time series consists mainly of three parts (see Fig. 2.2):

- The thermal noise: The altimeter generates thermal noise power before the scattering surface returns the signal for the first time. The returned waveform will therefore have a constant power level (Tourian, 2013).

- The leading edge: The leading edge of a pulse echo corresponds to the first interaction with the sea surface. Information about the significant wave height (SWH) and of course, the range between the satellite altimeter and the mean sea surface are closely related to the properties of the leading edge. In addition, smooth surfaces have steep leading edges and, conversely, rough surfaces have shallow leading edges.

- The trailing edge: The trailing edge of the pulse is also related to the scattering properties of the surface. It represents the returned power from the scattering surface outside the footprint and can ideally be approximated by a straight line.

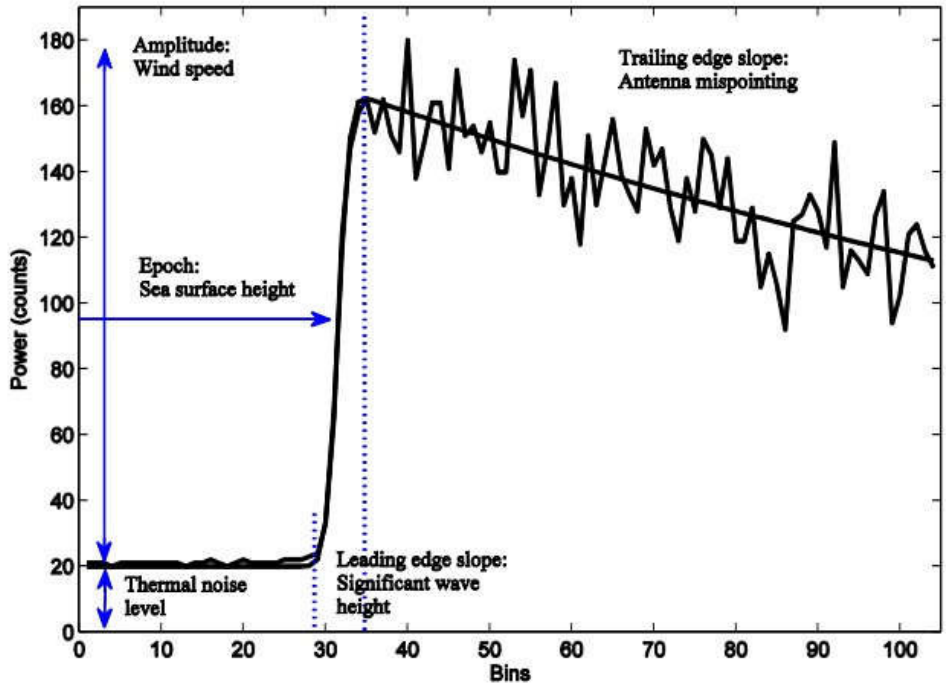


Figure 2.2: *Schematic altimeter waveform with the geophysical parameters that correspond to different parts of the waveform over homogenous ocean surface waveform (Abdullah, 2018)*

2.2 On board tracking system

In general, with appropriate corrections for atmospheric refraction and other factors, the range to the nadir can be determined from the two-way travel time, that is, time taken by the radar pulse to travel from the satellite antenna to the surface and back to the satellite receiver (Fu and Cazenave, 2001; Frappart et al., 2006). The returned power is proportional to the illumination area, so at the midpoint of the leading edge the normalized power is half or middle of the highest normalized power, and the travel time can also be determined by determining the time that the midpoint of the pulse needs to return from the water surface at nadir. Therefore it is essential to find this half-power point on the leading edge if we want to get the travel time t at the leading edge midpoint.

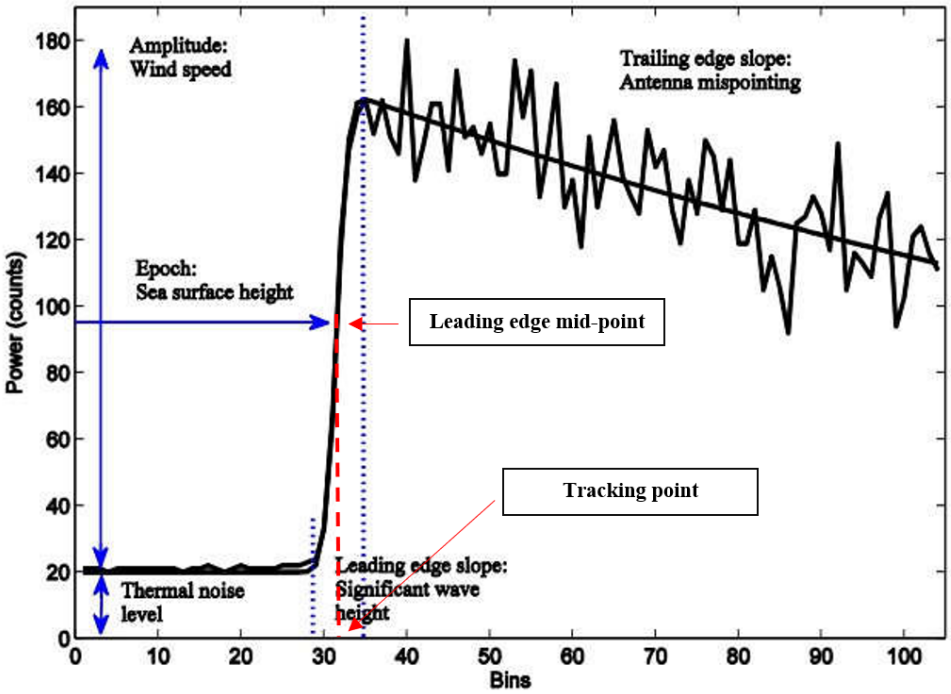


Figure 2.3: *Characteristics of an ocean waveform*

Figure 2.3 shows the same figure as Figure 2.2 but with some extra labels. To effectively extract the two-way time from Leading edge midpoint, SSH and other information from the waveform, one important function of the altimeter is to keep the leading edge of the waveforms well centered in range and power in the tracking window (Biancamaria et al., 2016), the so-called tracking point. These are all implemented by the Adaptive Tracking Unit (ATU) in on-board tracker. On-board tracker continuously adjusts the range window to keep the leading edge of the waveform at the center of the range window. Before the satellite launch, the tracking points are already accurately set. For example, the tracking point of TOPEX/Poseidon and ENVISAT are in the range bins of 24.5 and 32.5, respectively. Then the on-board tracker finds the half-power point on the leading edge of the waveform, determine its offset to the tracking point and realigns subsequent waveforms (Tourian, 2013).

2.3 The Brown-Hayne Theoretical Ocean Model

Over an ocean surface, the radar altimeter echo waveform has a characteristic shape that can be described analytically. The interaction of the radar signal with the sea surface as a function of time delay can be described as the convolution of three terms (Brown, 1977):

$$W(t) = P_{\mathrm{FS}}(t) * P_{\mathrm{PT}}(t) * q(t) \tag{2.1}$$

where t is the time delay, P_{FS} is the average flat surface impulse response, P_{PT} is the radar point target response, and $q(t)$ is the sea surface elevation probability density function of the specular point for weak, non-linear ocean waves. The radar point target response is expressed as a sinc function and then can be usually approached by a Gaussian function in order to perform the convolution of the three terms:

$$P_{PT}(t) = |\frac{\sin \pi Bt}{\pi Bt}|^2 \approx \exp \frac{-t^2}{2\sigma_p^2} \tag{2.2}$$

where σ_p is the width of the radar point target response function and can be approximated as time resolution r_t multiplied by a scale:

$$\sigma_p = \frac{1}{2\sqrt{2 \ln 2} r_t} \approx 0.425 r_t \tag{2.3}$$

The formulation of the theoretical shape of an echo over an ocean surface can be expressed after some derivation as (Brown, 1977):

$$P(t) = T_n + a_\xi P_u * (\frac{1 + \mathrm{erf}(u)}{2}) \exp(-v) \tag{2.4}$$

and refined by Hayne (1980) to include a fixed skewness parameter, λs. The formula will not be explained too much here because it is not the focus of this thesis, but it is important to understand that the Brown-Hayne Theoretical Ocean Model reveals the relationship between the waveform and the parameters implied in the waveform.

2.4 Retracking

The tracking window of a satellite is open within a specific time interval so that the transmitted pulses and received signals are correctly correlated with each other. Over the open ocean, most waveforms are well described by the Brown mathematical model. However, over non-ocean surfaces, i.e. ice, river, land, ice-sea and land-ice interfaces, the situation is different. In these cases, the waveform shape changes due to the influence of terrain and other factors. For example, raised land can also give returns within the altimeter's range window. In addition, the terrain relief can also interfere with the satellite's ability to track the water bodies, meaning that the altimeter is unable to get the echo waveform at the nominal tracking position. Therefore, there is an offset between the leading edge mid-point of the waveform and the pre-designed, fixed tracking gate (Tourian, 2013) (see Fig. 2.4).

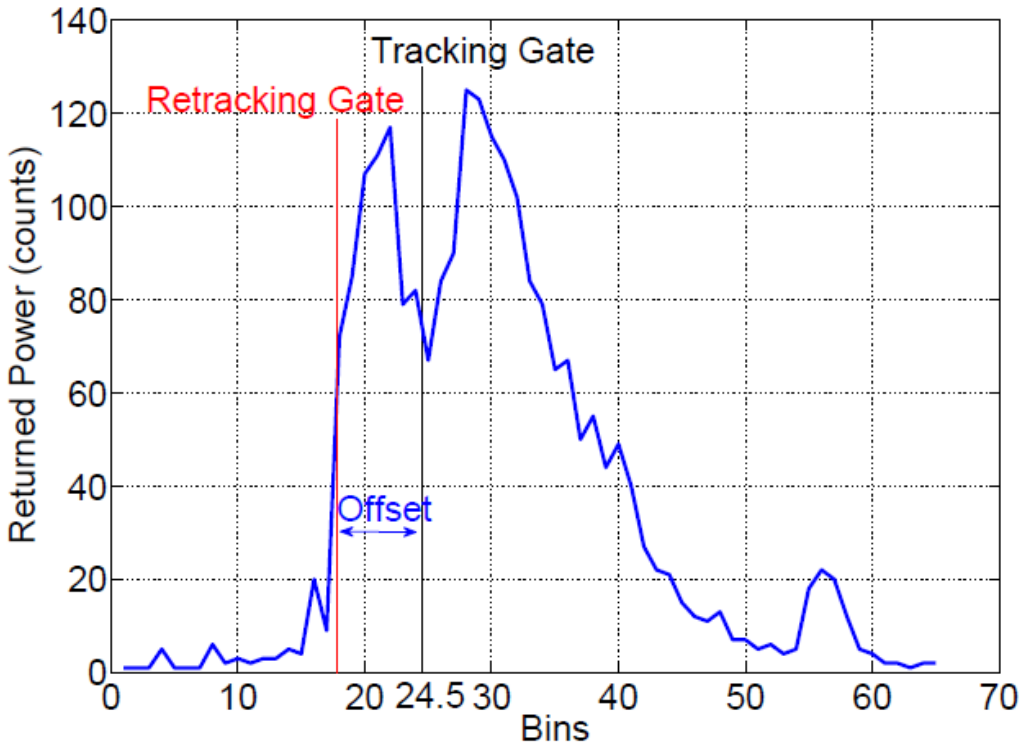


Figure 2.4: *The concept of waveform retracking (Tourian, 2013)*

In order to compute the offset and correct the range measurements, altimeters downlink the waveforms to Earth, and the final retrieval of geophysical parameters from the waveforms is performed on the ground. This post-processing is called "retracking". There are two specific categories of retracking methods after years of development, namely empirical retrackers based on empirical observations and practical experience (such as Offset Centre of Gravity Retracker (OCOG), Threshold Retracker, The β-parameter Retracker) and physically-based retrackers derived from theoretical knowledge of microwave scattering at nadir (such as The Brown-Hayne Theoretical Ocean Model, The NOCS Non-linear Ocean Retracker) (Gommenginger et al., 2011).

In this section, some waveform retracking methods related to this thesis will be introduced.

2.4.1 The Off Center Of Gravity (OCOG) retracker

The Off Center Of Gravity (OCOG) retracking algorithm is a purely statistical approach that does not consider the functional model developed by Wingham and Rapley (1986) to achieve robust retracking. The algorithm estimates the center of gravity (COG) of a rectangular box. Amplitude A is the twice the height of the center of gravity and the length of the box is determined by it. The retracking gate (see Fig. 2.5) is then determined by the width W of the box. In general, the rectangular box is defined by the parameterization of the waveform using three parameters: the center of gravity (COG) of the box; the width, W; and the amplitude, A, of the box, and can be expressed as:

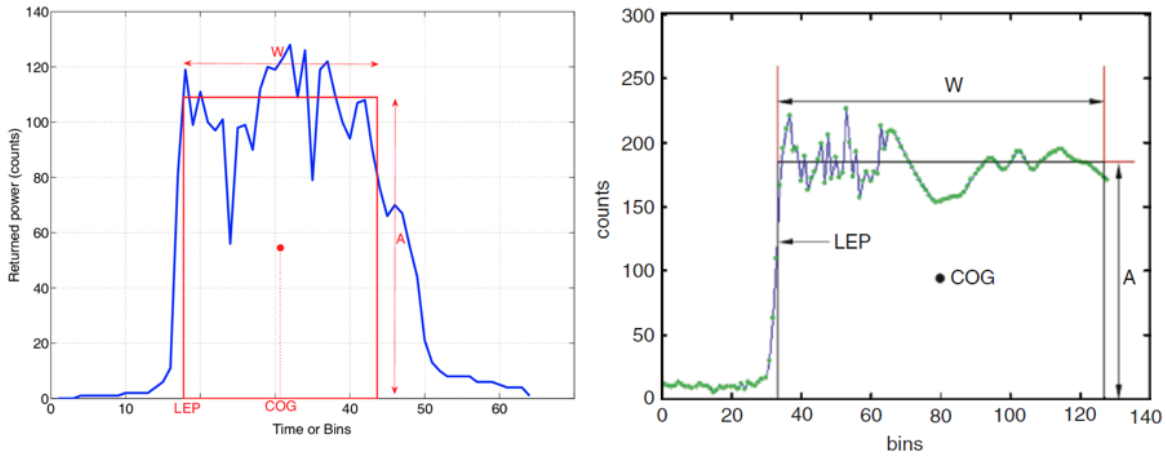


Figure 2.5: *Schematic description of the OCOG retracker, where W is the width of the box, A is the length of the box that represents the amplitude of waveform and COG is the center of gravity of the box*

$$A = \sqrt{\frac{\sum_i^N P_i^4(t)}{\sum_i^N P_i^2(t)}} \tag{2.5}$$

$$W = \frac{(\sum_i^N P_i^2(t))^2}{\sum_i^N P_i^4(t)} \tag{2.6}$$

$$COG = \frac{\sum_i^N iP_i^2(t)}{\sum_i^N P_i^2(t)} \tag{2.7}$$

$$N = N_w - n_2, \qquad i = 1 + n_1 \tag{2.8}$$

where P_i is the waveform sample value at the i^{th} bin, N_w is the total number of samples in the waveform, n_1 and n_2 are the numbers of bins affected by aliasing at the beginning and end of the waveform. The mid-point Leading Edge Position (LEP) is then given by:

$$LEP = COG - \frac{W}{2} \tag{2.9}$$

2.4.2 The Ice-1 retracker

The OCOG algorithm is a statistical algorithm that is simple to implement but has nothing to do with any physical properties of the reflective surface, and it is the algorithm behind the Ice-1 retracker.

Ice-1 retracker is a retracker based on the OCOG retracking method and the threshold retracking scheme.

In the threshold retracking, the position on the leading edge is sought by locating the first range bin to exceed a percentage of the maximum waveform amplitude (see Fig. 2.6).

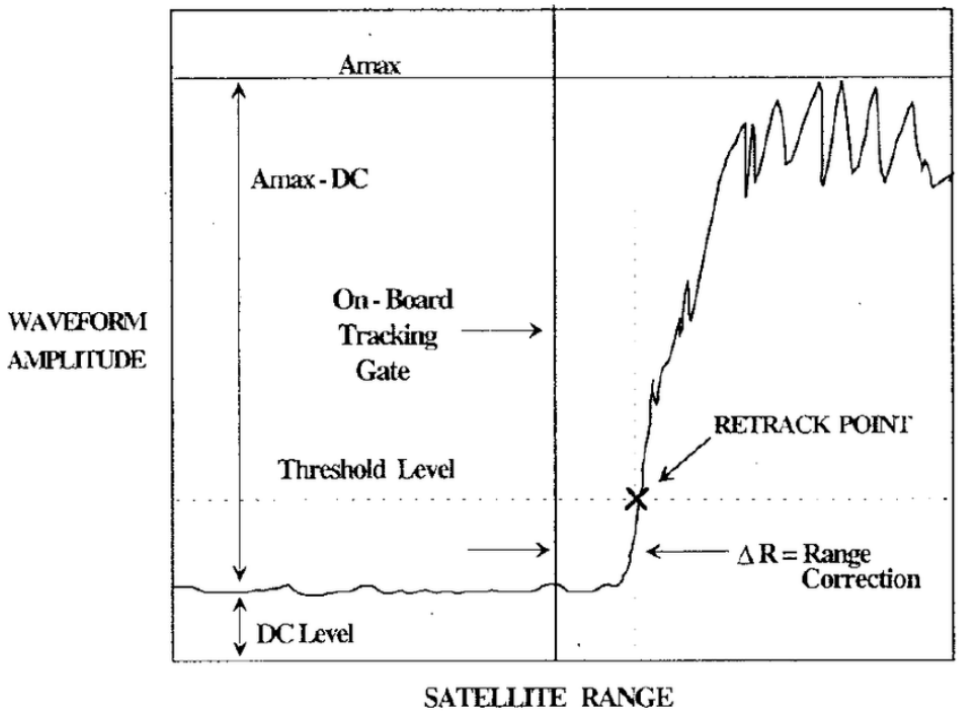


Figure 2.6: *Threshold retracker with a typical waveform from ice-sheet altimeter measurements. It is to mention that the threshold is referenced to a defined percentage of the maximum waveform amplitude above the DC level in front of the leading edge (Davis, 1997)*

The thermal noise P_N or DC level can be computed by averaging the value of power in the first five gates, and the amplitude is the maximum of these powers. According to the suggestion, the threshold should be set to 50% for surface-scattering dominated waveforms, and 10% or 20% for volume-scattering surface (Davis, 1997).

In general, the threshold retracking algorithm/Ice-1 retracker estimates the thermal noise, P_N, and the Retracked gate, T_L, as:

$$A = \max P(i) \tag{2.10}$$

$$P_N = \frac{1}{5}\sum_{i=1}^{5} P_i \tag{2.11}$$

$$T_L = P_N + T(A - P_N) \tag{2.12}$$

where T is the threshold value (e.g., 30% for Ice-1retracker).

2.4.3 The Ice-2 retracker

The Ice-2 retracker is based in the Brown model. The basic principle of the ice-2 retracker is to detect waveform edge, fit an error function to the leading edge and then an exponential decrease to the trailing edge. According to a Least Square Error (LSE) minimisation, the ice-2 retracker fit the measured waveform with a return power model:

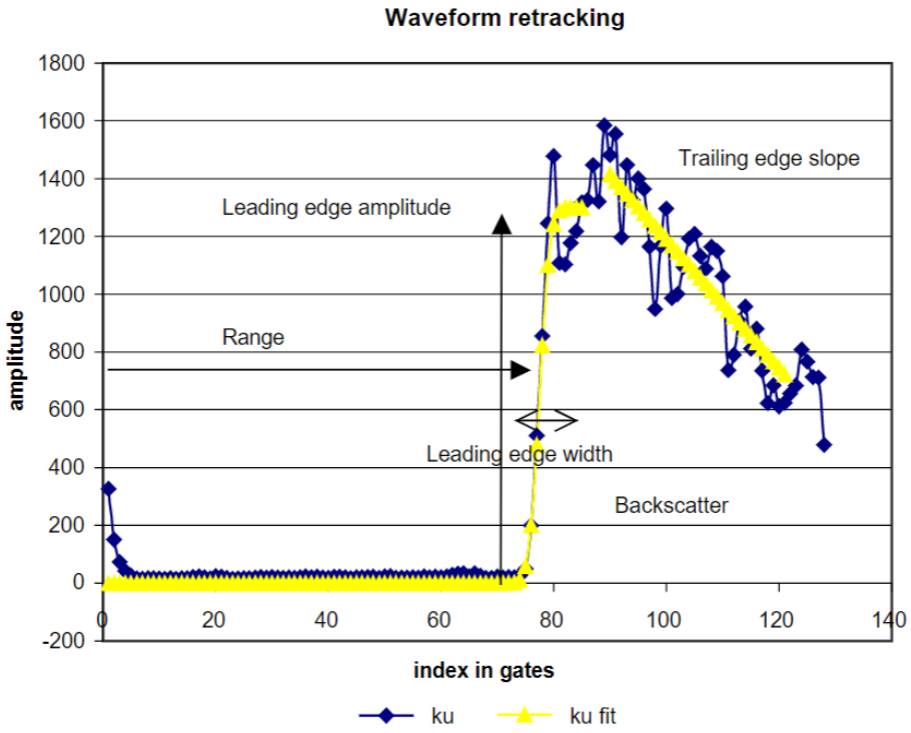


Figure 2.7: *Schematic diagram of the Ice-2 retracking method. The blue line represents the altimetric waveform and the yellow line represents the fitting function from Ice-2 retracking (Legresy et al., 2005)*

Ice-2 retracker has the following main output: the Leading edge amplitude (LeBs), the Leading edge width (LeW), the trailing edge slope (TE) and the Backscatter coefficient (Bs) corresponding to the waveform integration (Blarel et al., 2015).

Chapter 3

Case study

3.1 Study Area

Mackenzie River, as shown in the Figure 3.1, is one of the superlatives in Canada. It is also the longest river in Canada with the largest watershed, covering almost one-fifth of the country, and also the largest northwardflowing river in NorthAmerica (Yang et al., 2015). Sir Alexander Mackenzie, a Canadian explorer, discovered the river in 1789, and the river was named after him. Because of its location far from human activities and its sparse population, the Mackenzie River basin can be considered a relatively pristine area. From a geographic point of view, Mackenzie River is located in northern Canada and covers the Northwest Territories, Alberta, British Columbia, and the Yukon Territory and collects a vast system of rivers that flow from Great Slave Lake. To be precise, the river starts at the Great Slave Lake in the Northwest and flows northward into the Beaufort Sea in the Arctic Ocean, where it forms a large delta at its mouth (Millot et al., 2003).

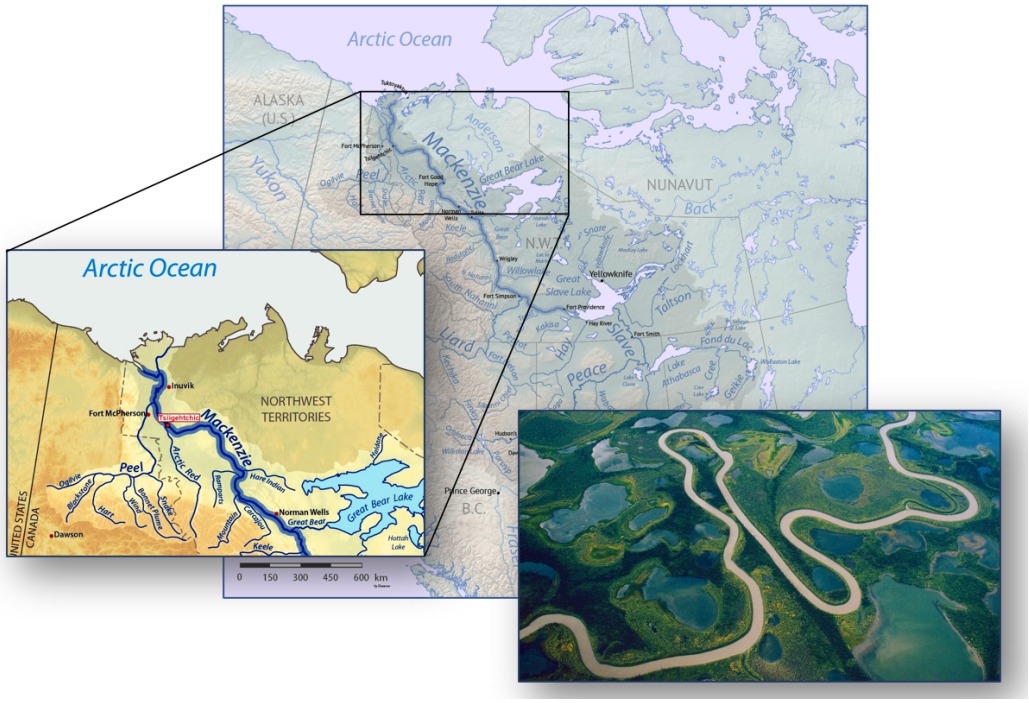


Figure 3.1: *Map of the Mackenzie River watershed*

Mackenzie River drains an area of 1.8 km^2 and delivers more than 325 km^3 of fresh water each year, accounting for roughly 11% of the total river flow into the Arctic Ocean. It is unfrozen from about June until October each year, and the rest of the year, it is frozen, with the ice typically breaking up by early to mid-May in the south and late May-early June in the north (McColl, 2014).

By assessing altimeter's performance in the Mackenzie River, we explore how to deal with the challenges of altimetry-based surface water height derivation. In fact, the altimetry data processing methods that have been developed for the ocean are generally not suitable for inland waters. As mentioned, compared with the relatively simple sea surface environment, inland waters are more complicated. As a typical representative of the boreal catchment, the Mackenzie River is strongly affected by the seasons, and the freezing of the water surface is one of the crucial features that distinguish it from other water bodies and is bound to have an impact on the altimetry results.

3.2 Virtual stations

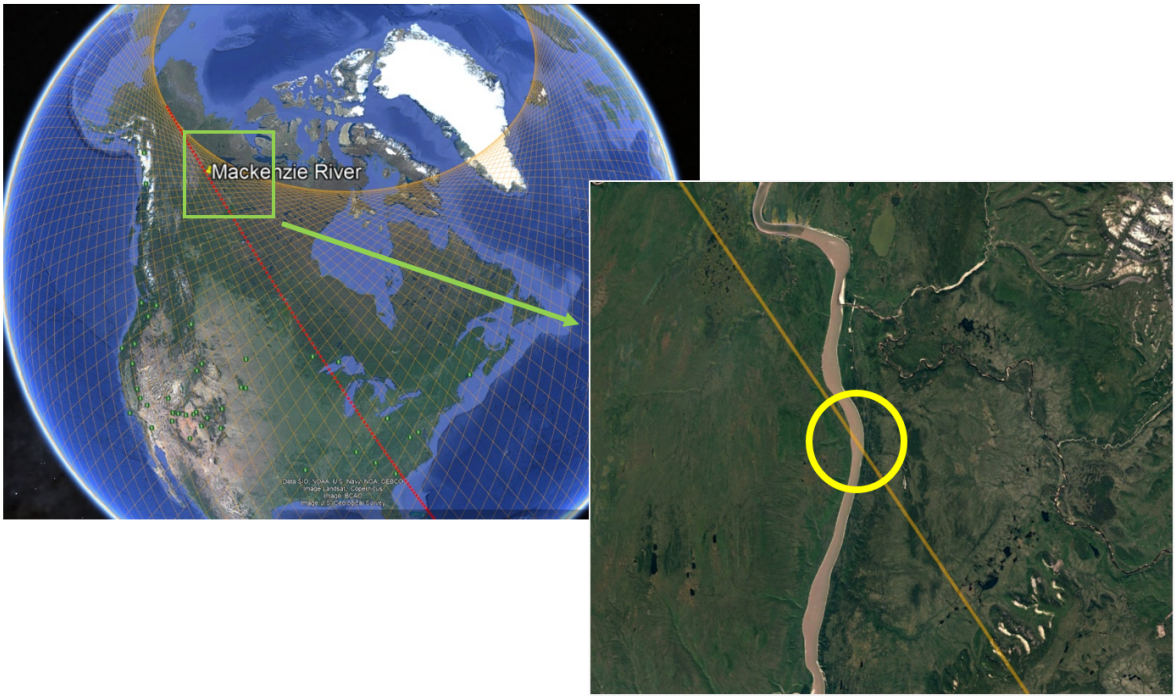


Figure 3.2: *Example of a virtual station located at Mackenzie River, Canda*

In order to assess the performance of the altimeter, the measurements of the altimeter need to be collected first. The data collection can be achieved by setting up a so-called "virtual station (VS)". "Virtual" means that the station is not an actual station built on the ground but a virtual one that realizes data collection through codes and programs. The software we use here is called "AltBundle+", which is a series of programs based on Matlab written by S. Behnia, T.Xiao, and M.J . Tourian from the University of Stuttgart.

AltBundle+ is an object-oriented, modular software for processing inland altimetry data. Currently, AltBundle+ is capable of converting Level-2 altimetry data of missions Jason-1, Jason-2, Jason-3, Envisat, Saral/AltiKa, Sentinel-3A, Sentinel-3B, CryoSat-2, and ICESat-2 into 'GIS-standard' objects and provide research- and product outputs. All the results of this paper are based on Altundle+.

In the case of specifying the mission, track number, and the location of the VS, users can create a corresponding VS (see Fig. 3.2), and set different parameters (search range, shapefile, Water Occurrence Frequency, etc.) to finally get the water level time series, along track view of altimetric heights, radargram stacks, etc:

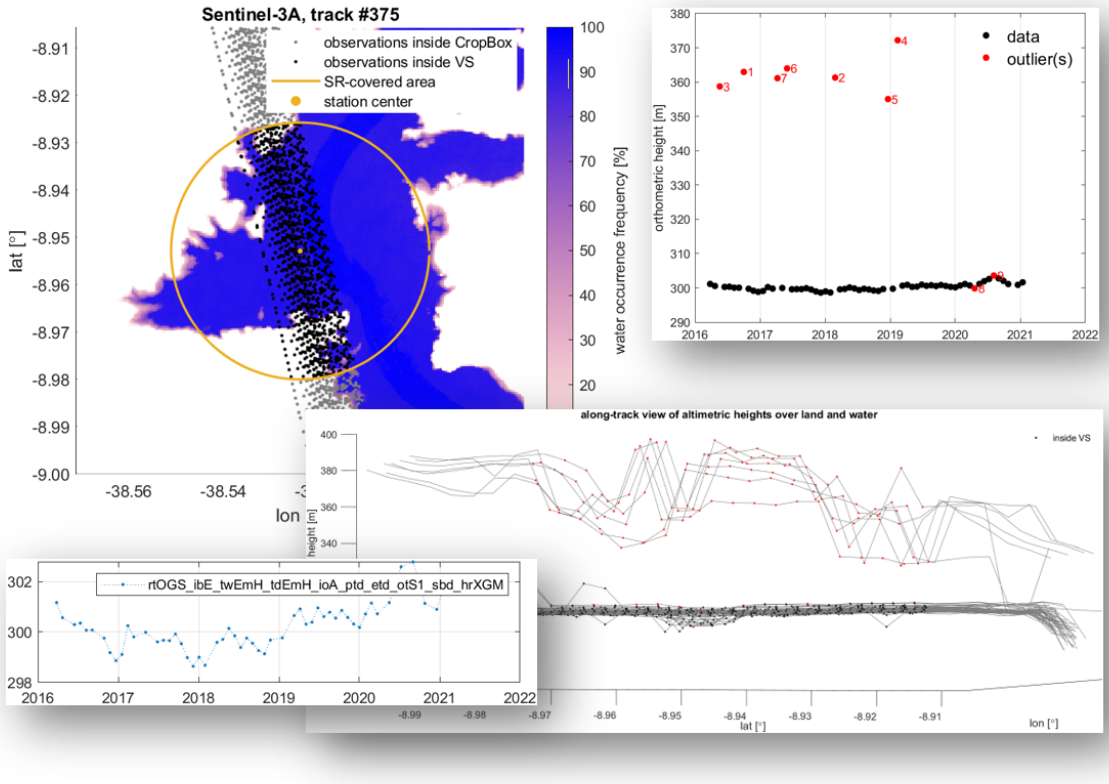


Figure 3.3: *Output figures from AltBundle+: water Level Times series, along track view, times series with Outlier Identification, etc*

3.2.1 Set up virtual station

To comprehensively asess the performance of different altimeters in the Mackenzie River, as many VSs as possible should be set up at various locations along the river. The basic principle is to ensure an optimal location for each VS to fully utilize the altimeter data. With the help of the river shapefile and the satellite ground track data files, the intersection of the river and the tracks passing the river can be easily calculated. A nominal VS takes the midpoint of this line as the location and half of the length as the search radius (see Fig. 3.4). Then AltBundle+ can define a search area based on shapefile and the search range/radius. The water level time series

(TS) can be obtained by only taking the average or median of the water-related measurements from satellite altimeters inside the search area.

Particular attention should be paid here that there is a specific deviation between the shapefile of the river and the watermask (the blue area in the background, which will be introduced in detail in Section 3.2.3), which can also be noticed in the figure. To avoid data missing due to this problem, we can add a buffer to the shapefile or search range to get a wider search area.

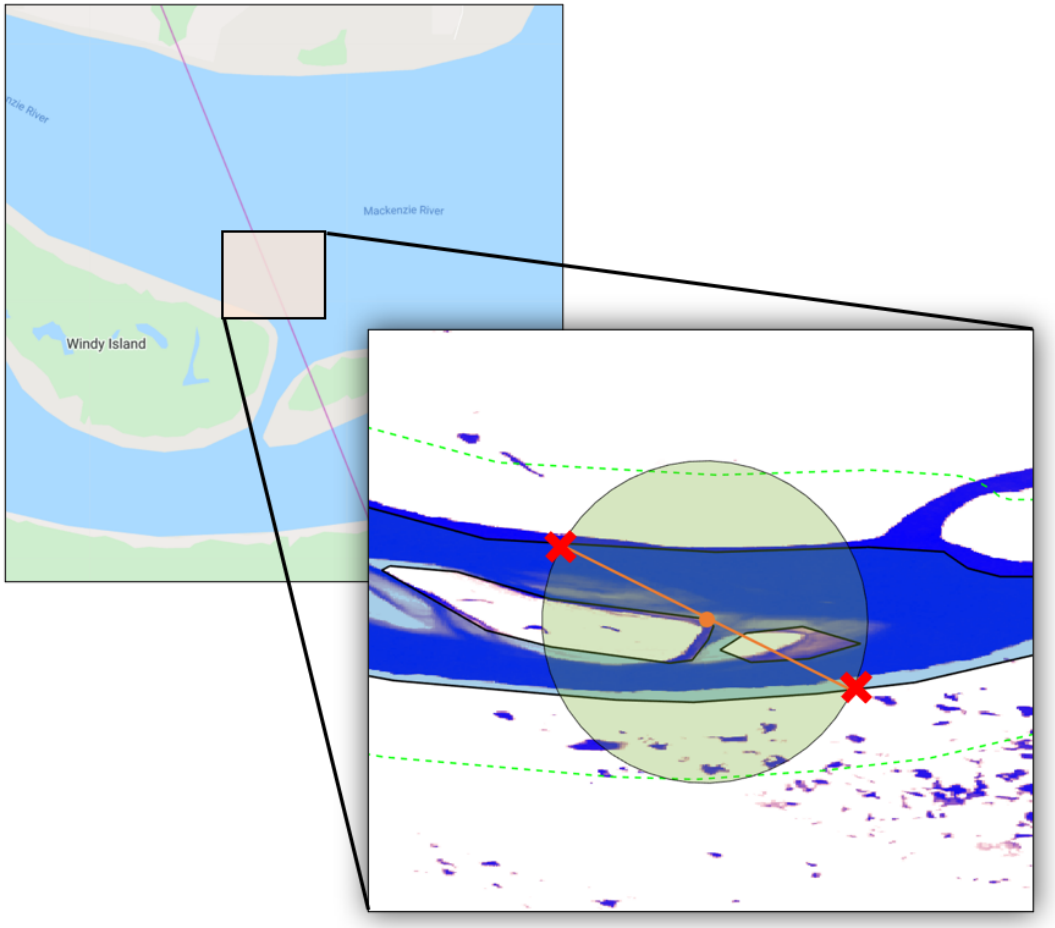


Figure 3.4: *Virtual station with search range -> Orange dot: the location of the VS; Red cross: intersection; Light green circle: search range*

As shown in Figure 3.5, in the end, a total of 265 VSs were defined along the Mackenzie River. The five previously mentioned satellites are distinguished by different colors and each search range, added buffer is also displayed. These densely distributed VSs along the river will provide comprehensive altimetry data for analysis.

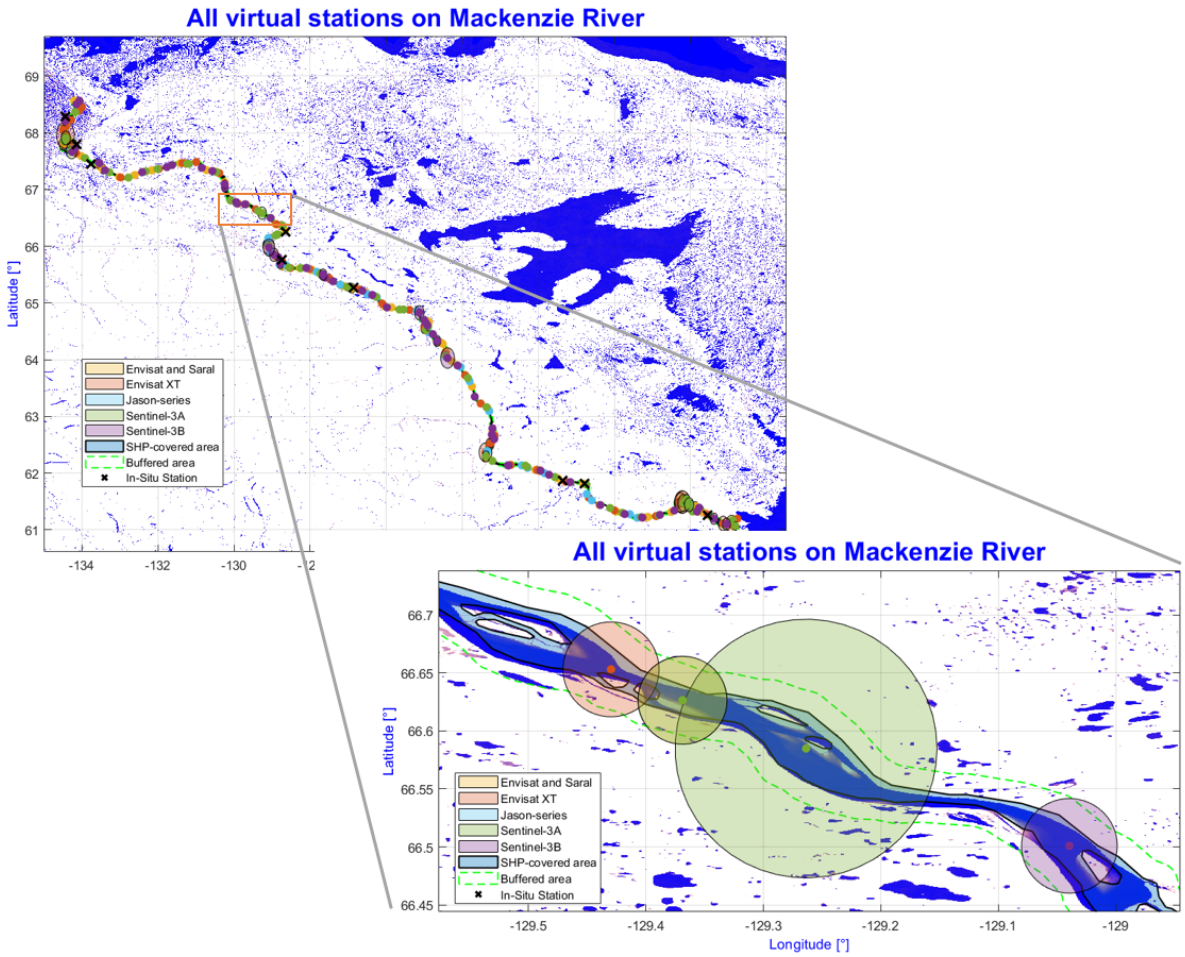


Figure 3.5: *All virtual stations along the Mackenzie River*

3.2.2 Set up parameters

In order to run AltBundle+, the following basic parameters need to be set:

Station	name of the VS
lat, lon	location of the VS
SR	search range
shp	shapefile
VS-, SR-, SHPflag	1: on / 0: off
SHPbuffer	shapefile buffer
Track, SAT	satellite information

Table 3.1: *List of basic parameters*

In addition to these basic parameters, some other important parameters that can strongly affect the results will be further explained.

3.2.3 Water Occurrence Frequency (WOF)

Global Surface Water Explorer (GSWL) is a virtual time machine that maps the location and temporal distribution of water surfaces at the global scale and provides statistics on their extent and change to support better informed water-management decision-making. AltBundle+ extracts the water occurrence map according to the location of VS and the extent of Crop Box based on GSWL dataset (raster layer). Then it filters out the measurements whose WOF is smaller than a set threshold (see Fig. 3.6).

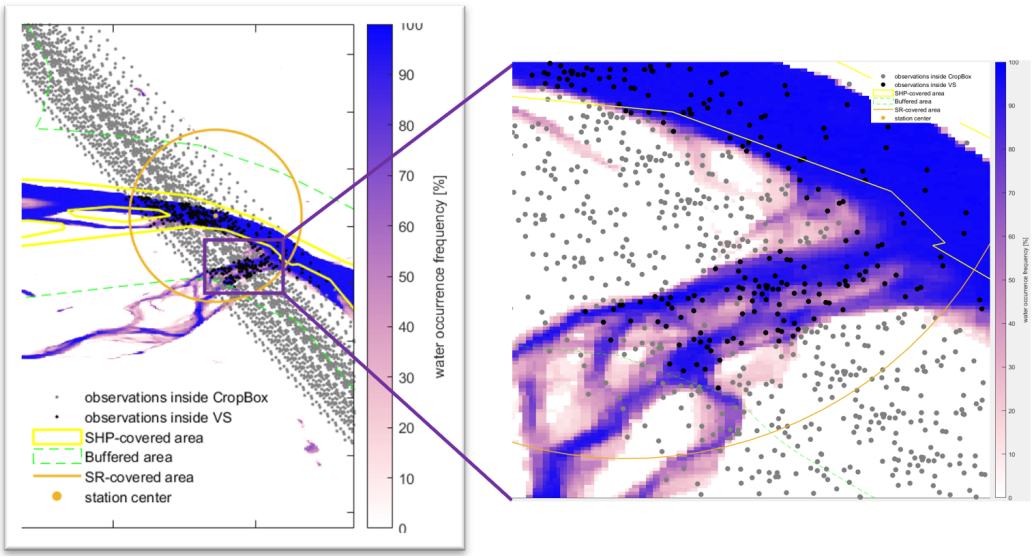


Figure 3.6: *Water Occurrence Frequency = 50%, each point is a measurement where the grey points are filtered*

Therefore, a suitable choice is crucial. Too strict a threshold will lead to a sharp decrease in the amount of data, while too low results in too much noise remaining (We set it to 50%).

3.2.4 Outlier Identification

Outlier Identification (OI) is a critical part of altimeter data processing, which identifies and removes outliers from the data and is the basis for subsequent processing. The inevitable emergence of outliers in altimetry water level time series is due to many problems, including failure in retracking, loss of track, inaccurate atmospheric and geophysical corrections, etc. However, a number of factors hamper the identification of such outliers. For instance, on the one hand, there may be no sufficient measures of uncertainty to help with the screening. On the other hand, short time series, low signal to noise ratio, irregular sampling, etc., can also negatively affect such data-driven approaches.

In AltBundle+, for any input time series, the sampling will be first regularized, and the gaps in the signal will be filled with SSA-based estimation (see Fig. 3.7). Singular Spectrum Analysis (SSA) is a time series analysis technique to decompose a signal based on Singular Value Decomposition (SVD). It is a data-driven method that does not require stationarity conditions and can help us fill the gaps easily.

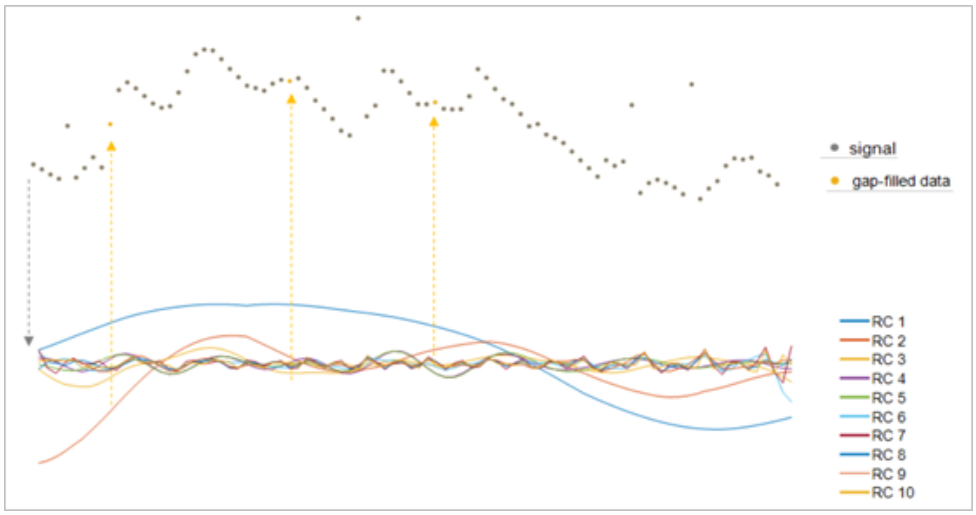


Figure 3.7: *An Example of SSA-based estimation*

Then AltBundle+ smooths the signal using a Savitzky-Golay (polynomial) smoothing filter, which is a data smoothing method based on local least squares polynomial approximation (see Fig. 3.8) with a given window size and polynomial order.

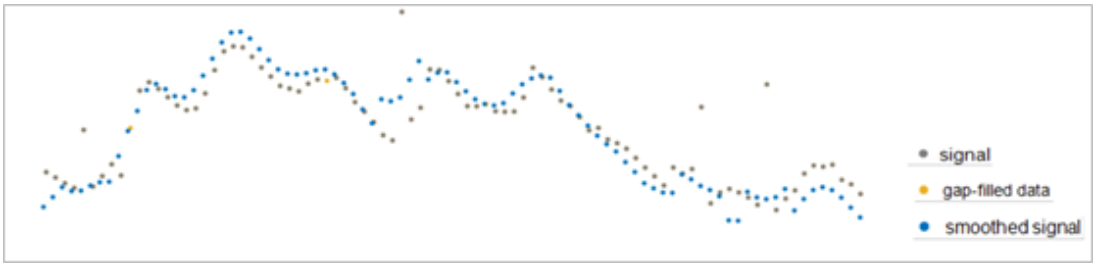


Figure 3.8: *Data smoothing using a Savitzky-Golay smoothing filter*

Based on the obtained smoothed signal $\overline{H}(t)$ and the initial signal $H(t)$, the water level residual can be calculated by:

$$r(t) = H(t) - \overline{H}(t) \tag{3.1}$$

Then a confidence level α (for example 99%) can be set up. With a two-tailed hypothesis test (normal distribution) a point will be rejected if:

$$\left| \frac{r(t)}{\sigma} \right| \geq k_{\frac{\alpha}{2}} \tag{3.2}$$

where σ is the standard deviation of the water level residuals. Outliers will be iteratively deleted. At each iteration, the algorithm gives all rejected outliers a second chance for a possible retention:

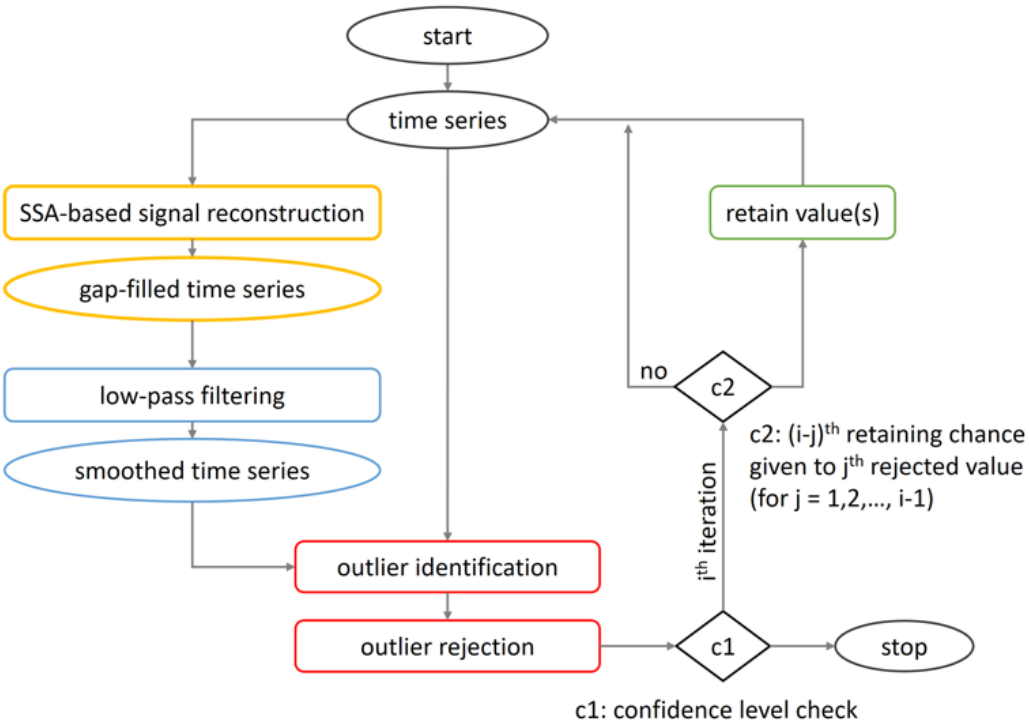


Figure 3.9: *Flowchart of the Outlier Identification*

3.3 Initial result

The initial result after running AltBundle+ is shown in the figure:

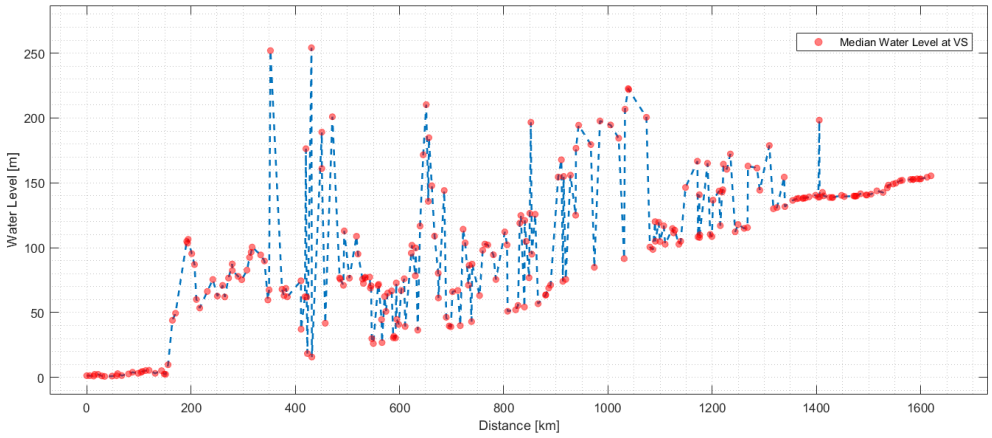


Figure 3.10: *Altimetric water level along the Mackenzie River. Starting point: VS with the highest latitude (Distance = 0 km) / End point: VS with the lowest latitude (Distance ≈ 1600 km)*

Figure 3.10 shows the median altimetric water level at each virtual station along the Mackenzie River. All VSs are connected to each other along the river, so ideally, the water level should show a gradual increase or decrease trend (depending on the defined direction). Obviously,

this is a bad result full of bad data everywhere. However, as mentioned many times before, due to the complexity of inland altimetry it is not surprising to get such a result.

Next round, the focus is on identifying problems through case studies and finding solutions to improve the results.

3.4 Case analysis

Before the case analysis, one principle must be clarified: the correctness of the altimetry results can never be completely determined in any case. The lack of in-situ data makes it impossible to do the validation. However, it would also make sense if the results could be improved to seem more plausible than before. At least it can show that altimeter has development potential in this field.

3.4.1 "Plausible" water level

Reviewing the water level shown in Figure 3.10, there is no doubt that the water levels at distances from about 0-170 km and after 144 km show a relatively flat upward trend, which seems to be plausible.

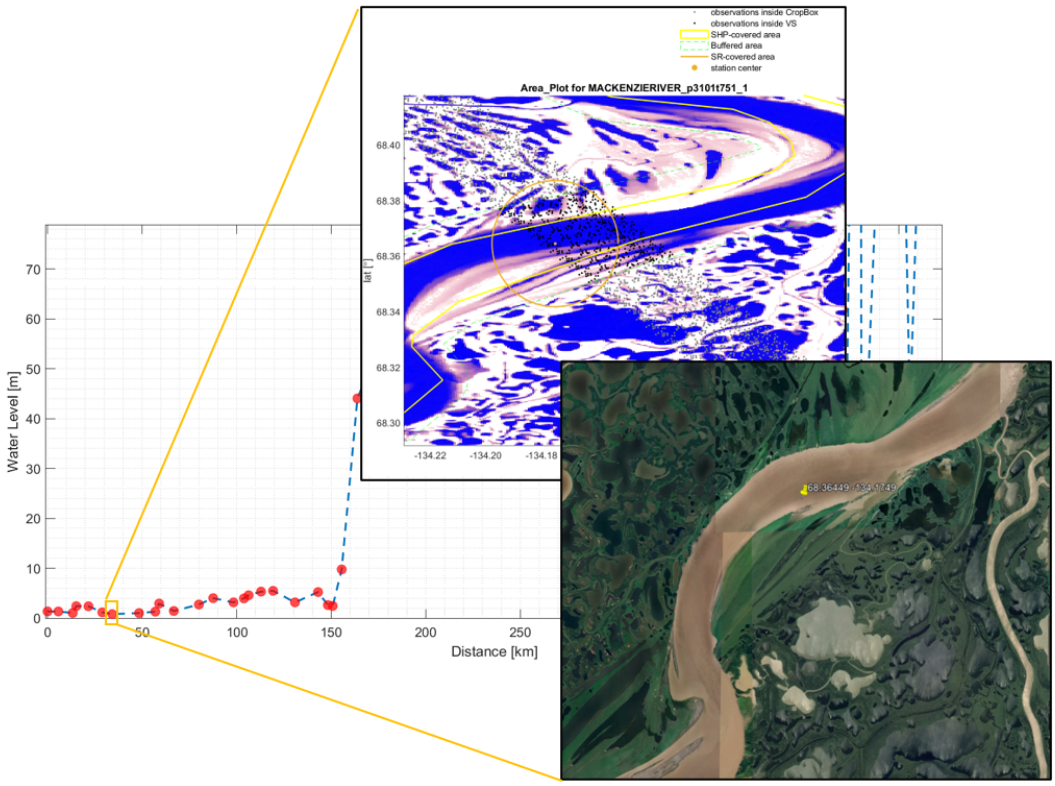


Figure 3.11: *MACKENZIERIVER_p3101t751_1 related information*

Taking one of the VS as an example, Figure 3.11 shows the geographic information of this VS (MACKENZIERIVER_p3101t751_1), where "3101" is the Satellite ID of Sentinel-3A for identification in AltBundle+, "751" is the track number, and "1" means that it is the first part of this track across the river. "p3101t751_1" means that it is the first part of the Sentinel-3A's track 751 over the Mackenzie River, and this VS is located at 68.36 N, 134.17 W in the northern section of the river.

The poorly visible circle in the figure is exactly the search radius we introduced before, and of course, with a buffer added. Each poorly visible point in this figure represents one measurement, where the black points are data filtered by the search radius, shapefile, and WOF, and are considered by AltBundle+ to be from the water surface. The orbital cycle of Sentinel is 27 days (14+7/27 orbits per day, 385 orbits per cycle), which means it passes over the same geographical point every 27 days, and this repetition is why all the measurements are arranged in parallel. The data measured each time the satellite passes the VS will be collected, and the relevant information will also be retained to distinguish each cycle.

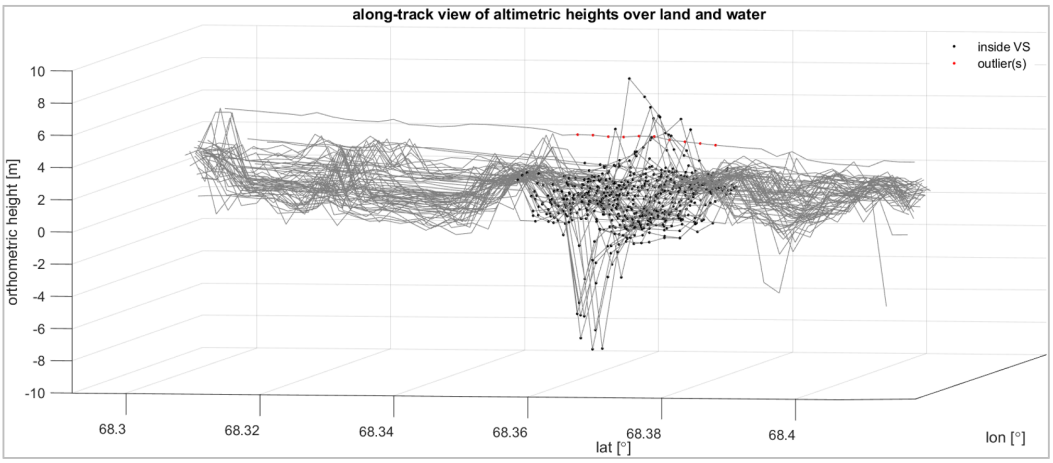


Figure 3.12: *Along-track view of altimetric heights over land and water at MACKENZIERIVER_p3101t751_1*

For each VS, AltBundle+ saves all information in a mat file and generates multiple figures for analysis. Figure 3.12 shows the along track view of the altimetric heights over land and water at this VS, which gives us an intuitive display of the results with latitude and longitude in XY-plane and orthometric height in Z-axis. The black points are the measurements over the river, while the gray points are the measurements over land, and the red points are outliers. Points on the same cycle are connected to each other, which makes the comparison between measurements of each cycle also possible. Since there is no in-situ data, the only thing that can be done at present is to for example, use the height of google earth as a reference. The height at this VS in google earth is about 2 m which also matches the altimeter result here.

However, an approximate height is not enough to assess the performance of the altimeter. More reliable information from the water level time is series (see Fig. 3.13) is needed. What AltBundle+ does is to take the median or average of all data in each cycle, providing a distribution of heights over time.

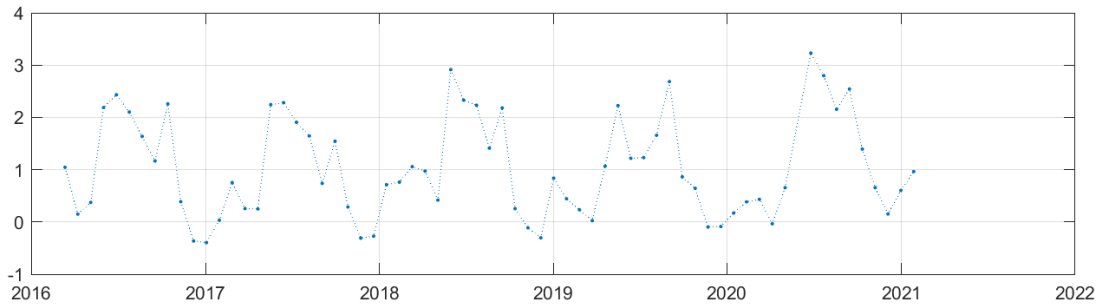


Figure 3.13: *Water level time series without outliers at MACKENZIERIVER_p3101t751_1*

The water level in this case shows an apparent periodic change, which is also in line with a normal water level time series. The water level in each year's first and fourth quarters is significantly lower than that in the second and third quarters. It is reasonable to speculate that such changes in the water level may be caused by the influence of the seasons. The middle of the year may be the flood season of the river, and in winter, due to the high latitudes, river freezing is inevitable, which will definitely affect the height of the river.

In general, the conclusion that can be drawn is that the results obtained here are in line with a "water level time series". Although it is impossible to determine its correctness, it is at least reasonable and plausible.

3.4.2 "Implausible" water level

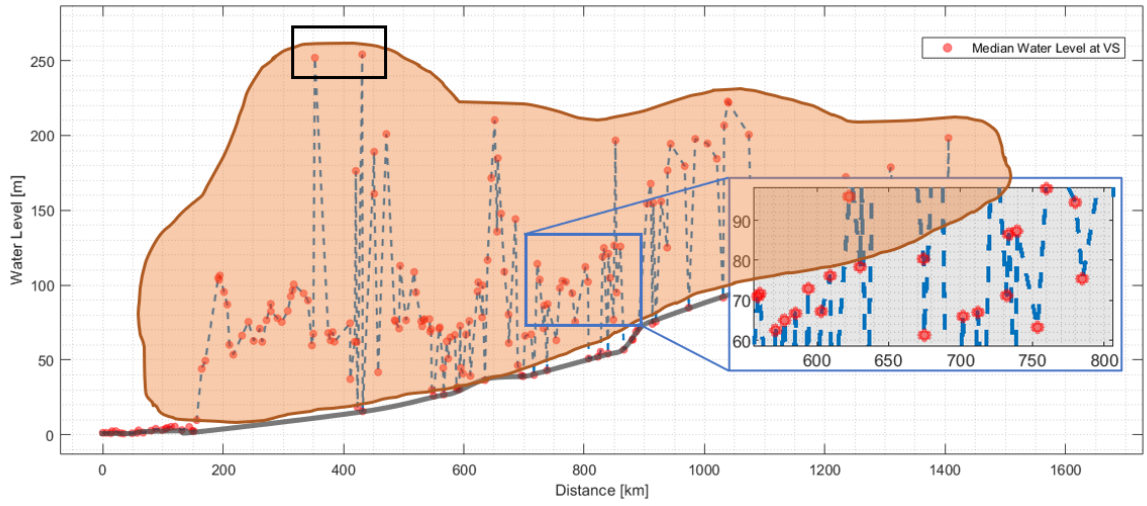


Figure 3.14: *Altimetric water level along the river (Orange area: implausible results)*

Figure 3.14 is the water level along the river we showed before but with a few more annotations. The grey line along the data shows the general trend in the height of the river. These data along the grey line seems to be plausible. However, it can be seen that most of the data is located in the orange area, which is obviously implausible.

The results in the black box are the cases we want to analyze first. Even in implausible data, these two results, marked in the black box, are extremely obvious. Checking the problems en-

countered here is of great help in finding improvement solutions later. Since the two situations are similar, common analysis and comparison are helpful.

Figure 3.15 shows the along track view of these two VSs. Most of the data are clustered above 200 m, while only a tiny part is gathered at about 10-20 m. AltBundle+ considers data around 10-20 m to be outliers and removes them, which leads to the high water level in the results. However, google earth shows a height of about 20 m, while the hills nearby are more than 200 m high. Therefore, a possible reason could be that most of the measurements are not the height of the water surface but the nearby hills.

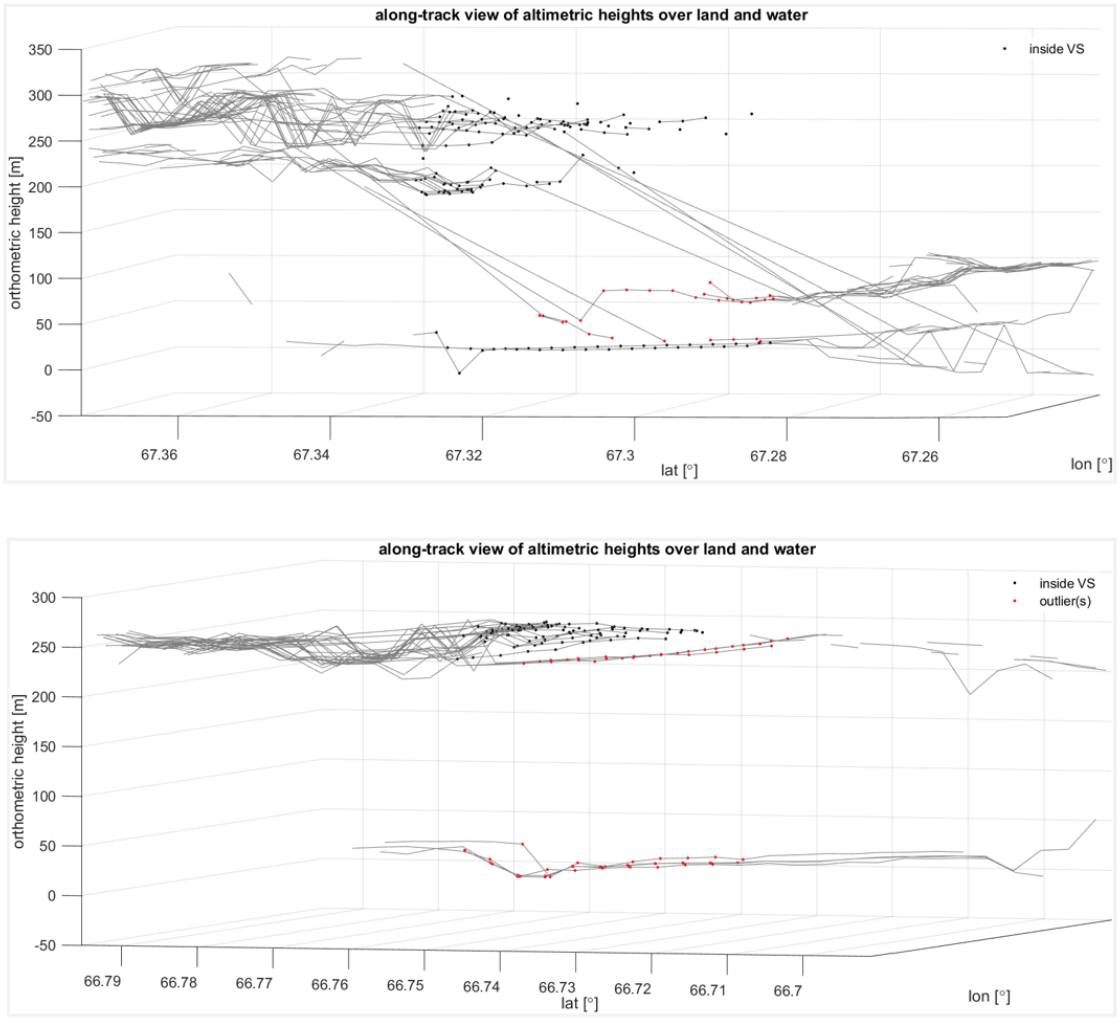


Figure 3.15: *Along-track view of altimetric heights over land and water at MACKENZIERIVER_p3101t484 (above) and MACKENZIERIVER_p3201t712 (below)*

Figure 3.16 shows the water level time series at these two VSs:

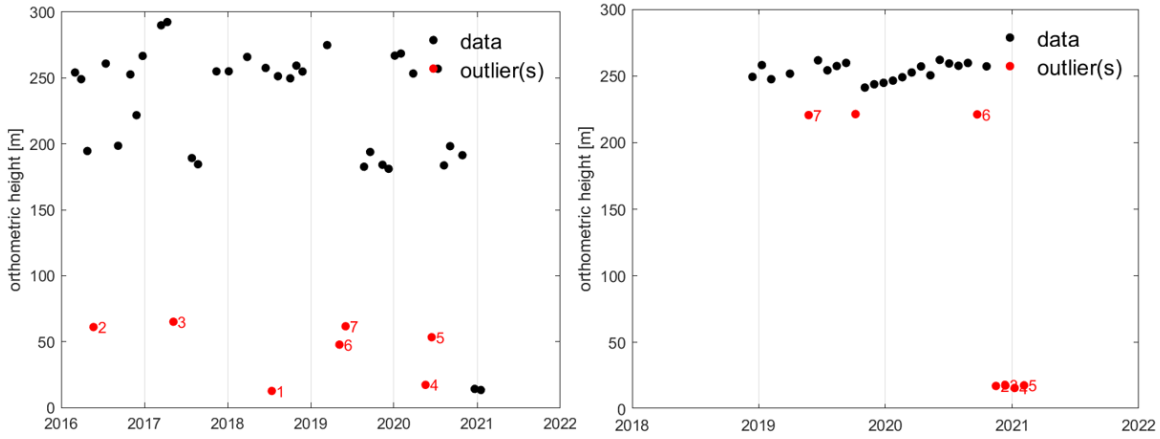


Figure 3.16: *Water level time series with outliers at MACKENZIERIVER_p3101t484 (left) and MACKENZIERIVER_p3201t712 (right)*

Neither of these two time series is consistent with the characteristics of the water level time series, especially in the second case. Almost all the results are clustered around a value, which implicitly indicates that the height of the measured object may not be affected by the time, the surrounding hills are therefore "suspects". What's more tricky is that even if there are some seemingly plausible measurements, Outlier Identification removes them. It is not surprising to get such a result. Outlier Identification used by AltBundle+ is completely a data-driven algorithm. The premise of correctly identifying outliers is that there is no huge error in the data itself or that the wrong data is not dominant. Data quality is unpredictable, but results may be improved if a relatively reliable reference height could be added to assist Outlier Identification. In Chapter 4, we will introduce related improvement methods.

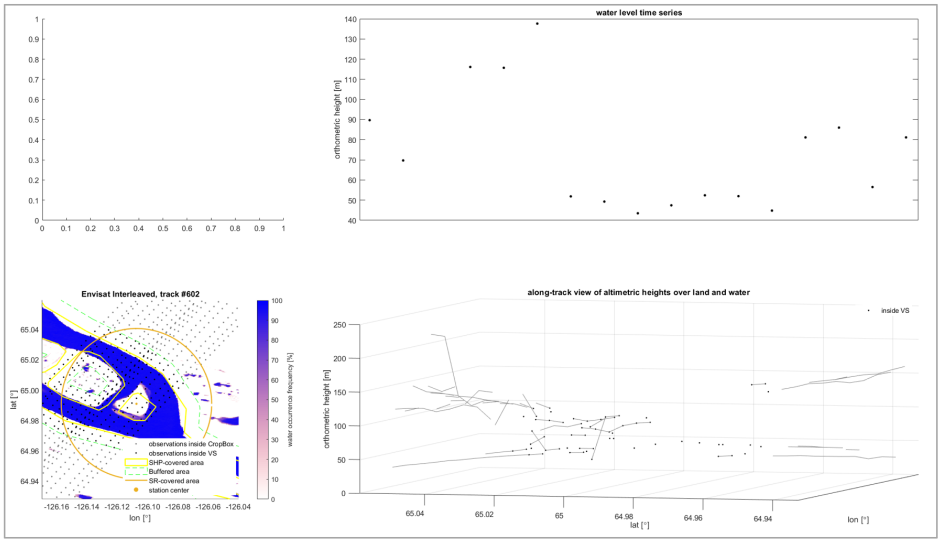


Figure 3.17: *Figures generated by AltBundle+ at MACKENZIERIVER_p1391t602 (Water Level time series + geographic information + along track view)*

Now, let us see some other cases, for example data in blue box (see Fig. 3.14). What needs to be clear is that both the data in the black box and the blue box belong to the implausible data one the one hand, but on the other hand, they do have potential for improvement. A specific analysis can inspire us to find ways to improve results. Two VSs located in the blue box will be used as cases for analysis.

Figure 3.17 shows some figures generated by AltBundle+ at one of these VSs. None of the data were identified as outliers, mainly because the amount of data is too small and the data is too scattered. It's hard to explain the reason here. The error of retracking, wrong object, or other problems, all of these can lead to problems. In the reprocessing section (Chapter 4), such errors will be discussed in particular. More important now is just an understanding.

Taking a look at another VS, something similar happened. Although there is one outlier identified (see Fig. 3.18), the results are obviously not ideal.

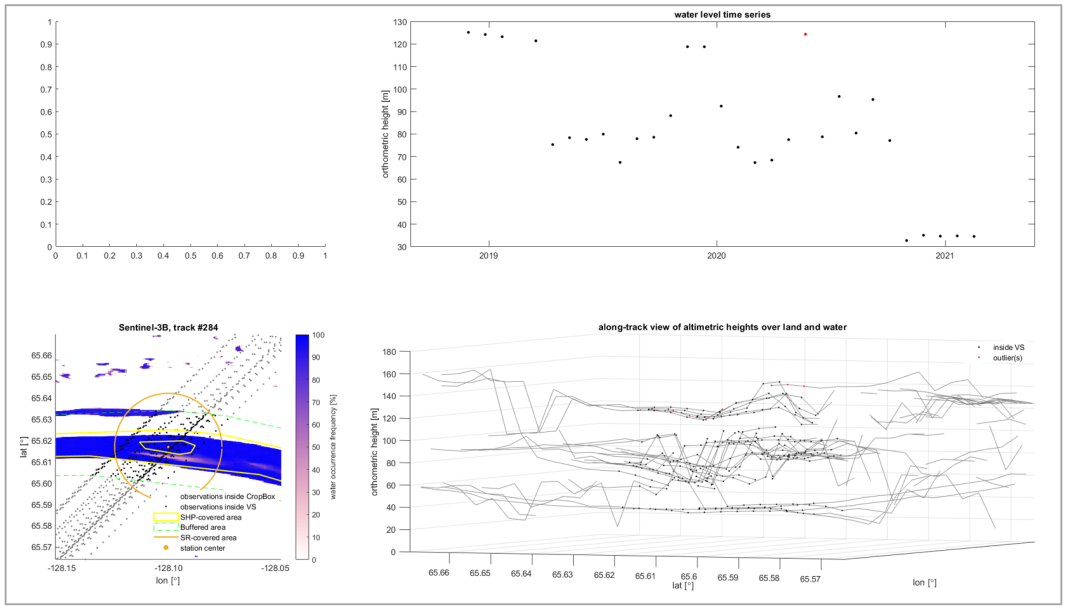


Figure 3.18: *Figures generated by AltBundle+ at MACKENZIERIVER_p3201t284 (Water Level time series + geographic information + along track view)*

The data is clearly divided into three "layers". Combined with the location of the VS (see Small figure in the lower left corner of Figure 3.18), it is noteworthy that the small island located in the middle of the river divides the river into two branches. The complex terrain may be the reason behind this phenomenon.

After analyzing the initial results, some solutions for improving the results will be elaborated in the next chapter.

Chapter 4

Altimetry processing

4.1 Digital Elevation Model (DEM)

The correct identification of outliers is crucial. However, no reference data is used to assist the identification. So finding a reliable reference is the first idea to improve the results. Here, the first concept that comes to mind is the Digital Elevation Model.

Our planet is full of peaks, valleys, natural habitats, and human-made objects. Digital Elevation Model (DEM) is a digital cartographic dataset in 3D coordinates to model the terrain of the Earth. The terrain elevations are sampled at regularly spaced horizontal intervals through techniques such as photogrammetry, LiDAR, IfSAR, land surveying und saved as a raster GIS layer (Gandhi and Sarkar, 2016).

Although DEM cannot be directly used as a real theoretical value, it can be used as a reference value to improve the accuracy of the Outlier Identification.

4.1.1 NASADEM

NASADEM is a state-of-the-art global DEM dataset (see Fig. 4.1), which provides global elevation data at 30 m spacing. The data products were derived from a combination of Shuttle Radar Topography Mission (SRTM) processing improvements, elevation control, void-filling and merging with data unavailable at the time of the original SRTM production. SRTM is a collaboration between the United States National Aeronautics and Space Administration (NASA) and the National Geospatial-Intelligence Agency (NGA). In addition, German and Italian space agencies are also involved. SRTM generated a near-global DEM of the Earth using radar interferometry and also relied on Ice, Cloud, and Land Elevation Satellite (ICESat) Geoscience Laser Altimeter System (GLAS) ground control points of its lidar shots to improve the geolocation accuracy (Buckley, 2020).

NASADEM is indeed a good option and it's already built into AltBundle+. The data are distributed in 1° latitude by 1° longitude tiles and consist of all land between 60°N and 56°S latitude. However, Mackenzie River is located above 60°N. Even though NASADEM has been used in AltBundle+, it is not applicable to our current research, it is necessary to find an alternative.

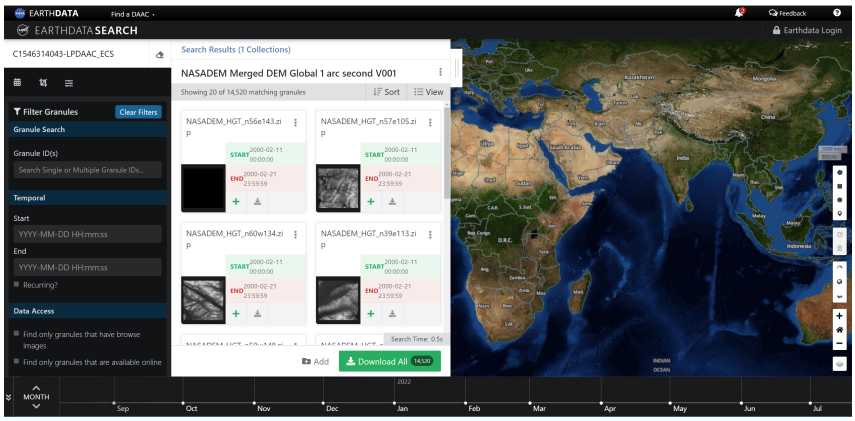


Figure 4.1: *EARTH DATA: NASADEM related data can be downloaded*

4.1.2 ASTER GDEM

An alternative dataset to replace NASADEM is ASTER GDEM (see Fig. 4.2). It is a global Digital Elevation Model published by the Ministry of Economy, Trade, and Industry (METI) of Japan and NASA. They jointly announced the release of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 3 (GDEM 003), and the ASTER Water Body Dataset (ASTWBD) on August 5, 2019 (NASA and Teams, 2019).

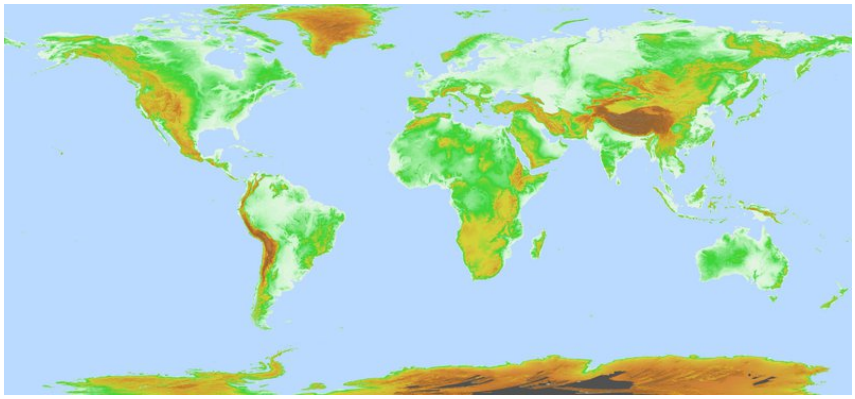


Figure 4.2: *ASTER Global DEM (NASA and Teams, 2019)*

ASTER obtains high-resolution (15-90 m^2/pixel) images of the Earth in 14 different wavelengths of the electromagnetic spectrum, ranging from visible to thermal infrared light to create the DEM dataset. The ASTER GDEM covers land surfaces between 83°N and 83°S and comprises 22,912 $1° \times 1°$ tiles. It is distributed in GeoTIFF file format with Geographic latitude/longitude coordinates referenced to the WGS84/EGM96 geoid at 1" spacing. Research shows that accuracies for this global product are 20 m at 95% confidence for vertical data (NASA and Teams, 2019).

Although ASTER GDEM is the best choice that can be found, its performance at high latitudes is still unknown. Using ASTER GDEM in AltBundle+ is also an approach to test its accuracy.

4.1.3 Add DEM in AltBundle+

As mentioned above, the DEM data is distributed with geodetic latitude/longitude coordinates. The entire data is a grid, and each node has a corresponding elevation. The latitude and longitude of each measurement target from the satellite altimeter are also stored in AltBundle+. By cubic interpolation, which means the height at a query point is based on a cubic interpolation of the height at neighboring grid points in each respective dimension, a height can be estimated for each measurement coming from the water surface. The reference height of this area can be determined by taking the average.

In addition, an interval can be defined by setting up a hard threshold in AltBundle+ and combining it with the reference height. The data larger or smaller than this interval will be regarded as outliers. This is a very straightforward algorithm, but it can effectively identify some outliers. Next, we will analyze the new results obtained and assess the reliability of ASTER GDEM.

4.1.4 Results after using ASTER GDEM

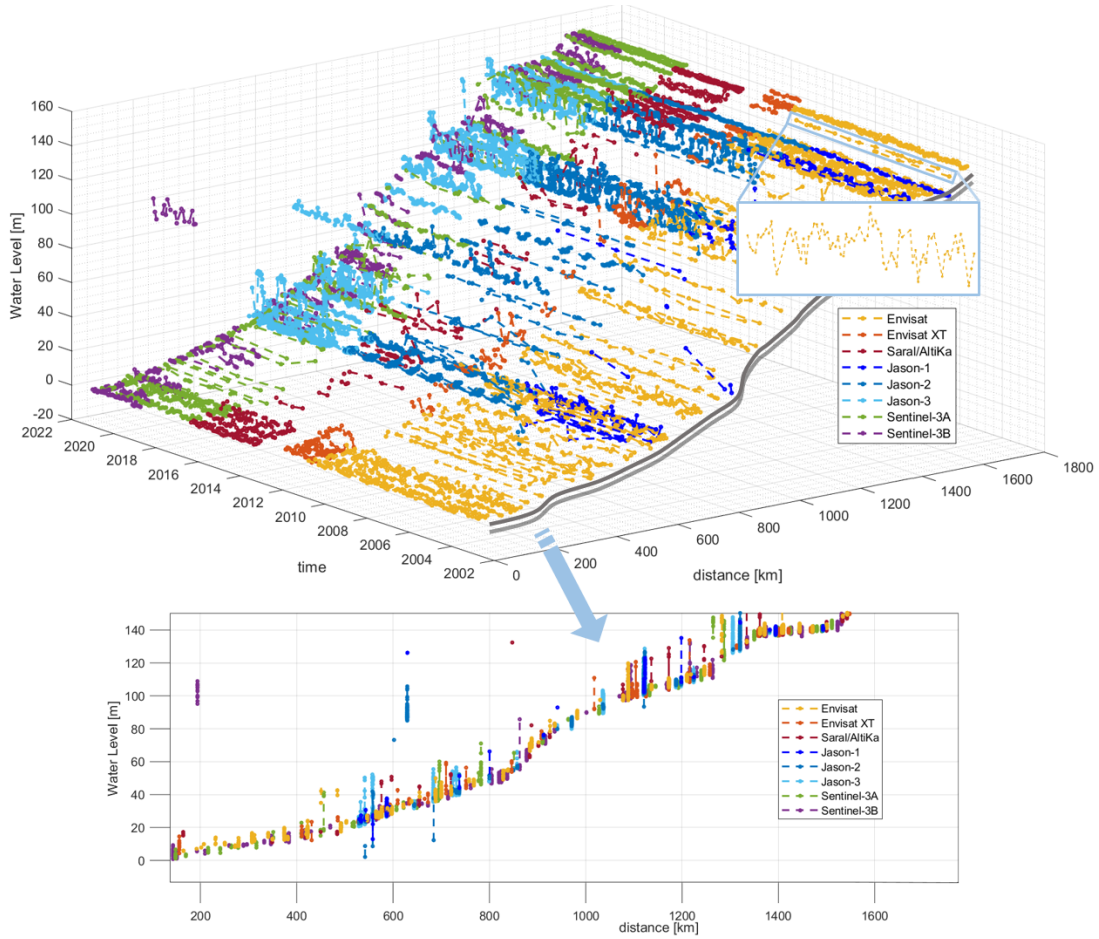


Figure 4.3: *Altimetric water level time series along the river (after using ASTER GDEM) Top Figure: 3D water level time series; Bottom Figure: cross-section*

The results obtained by running AltBundle+ with ASTER GDEM added are first as a 3D figure shown in Figure 4.3. Compared to 2D figure showing only the median or average height, the 3D figure with distance on the X-axis, time on the Y-axis and height on the Z-axis also retain all the time series along the river.

Although time series are full of fluctuations due to various factors, the entire height in the cross-section shows a clear trend. Moreover, compared to the initial results obtained before, the results are indeed greatly improved as expected.

Leaving aside the discussion of improvement, more notable are some significant problems that these TSs reveal. For those VSs with multiple satellites passing by, there are significant differences in the TSs obtained by different satellites, which does not seem normal since these satellites are clearly measuring the same area. For example, the height of the light blue TS representing Jason-3 always seems to be much higher than the height of the blue TS representing Jason-1. This difference may be caused by the retracking method, although they all belong to the Jason series but use different retracking methods.

So, there is no doubt that the so-called median altimetric water level along the river, obtained by taking the average or median, is affected by these differences. For those VSs with multiple satellites passing by, if the water level displayed in the 2D figure is bad, we should not blindly judge that the data is wrong because we cannot rule out that there is a correct TS but affected by the wrong one. Therefore, it is always necessary to study the TS behind the value.

On the other hand, as mentioned in Chapter 3, the periodic fluctuation of the TS also reflects the possible changes in water level. For example, seasonal changes will bring periodic effects on the water level, the freezing of river surface in winter is a problem that cannot be ignored, and the effects of flood and dry seasons are also big problems, all of which may be reflected in TS and affect the final result influence.

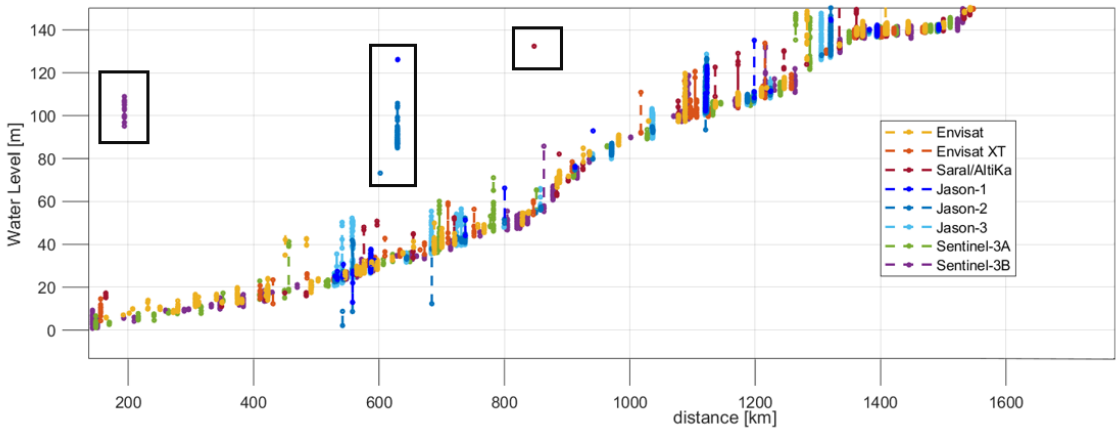


Figure 4.4: *Cross-section: altimetric water level along the river after using ASTER GDEM (black boxes: bad results)*

In addition to the 3D figure showing the properties of TS, the corresponding cross-section in Figure 4.4 shows the trend of the whole results, and the comparison of the median water level along the river before and after using ASTER GDEM are also shown in Figure 4.5. The results did improve, however, the still-existing bad data is also evident.

According to the procedure of AltBundle+, excessively large data will be filtered based on the set reference height and threshold. The only explanation that these outliers still exist is that the reference height may be set incorrectly and cause the outlier to not be correctly identified.

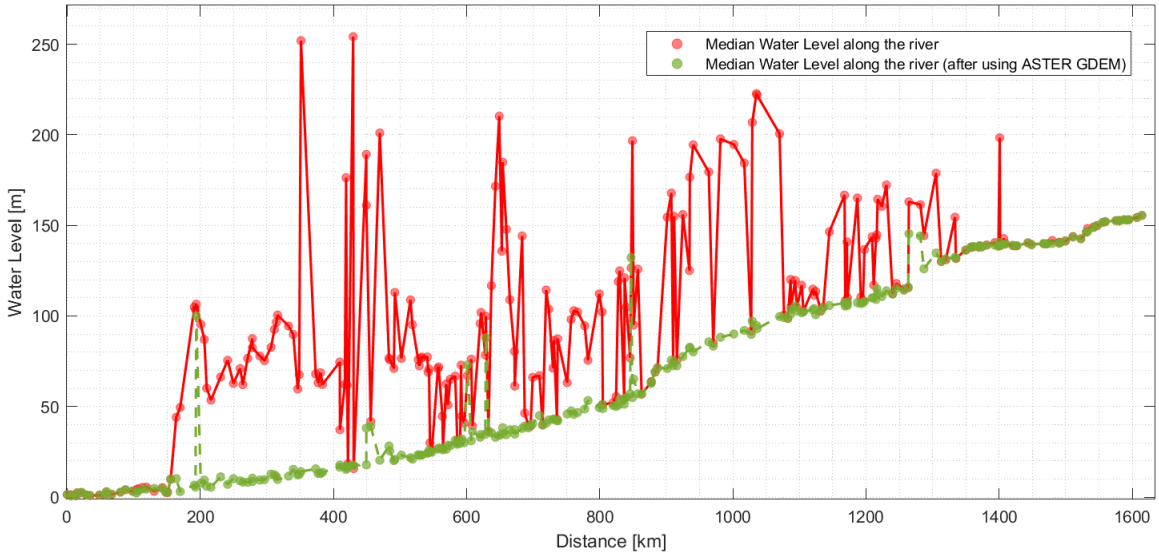


Figure 4.5: *Altimetric water level along the river (after using ASTER GDEM und before using it)*

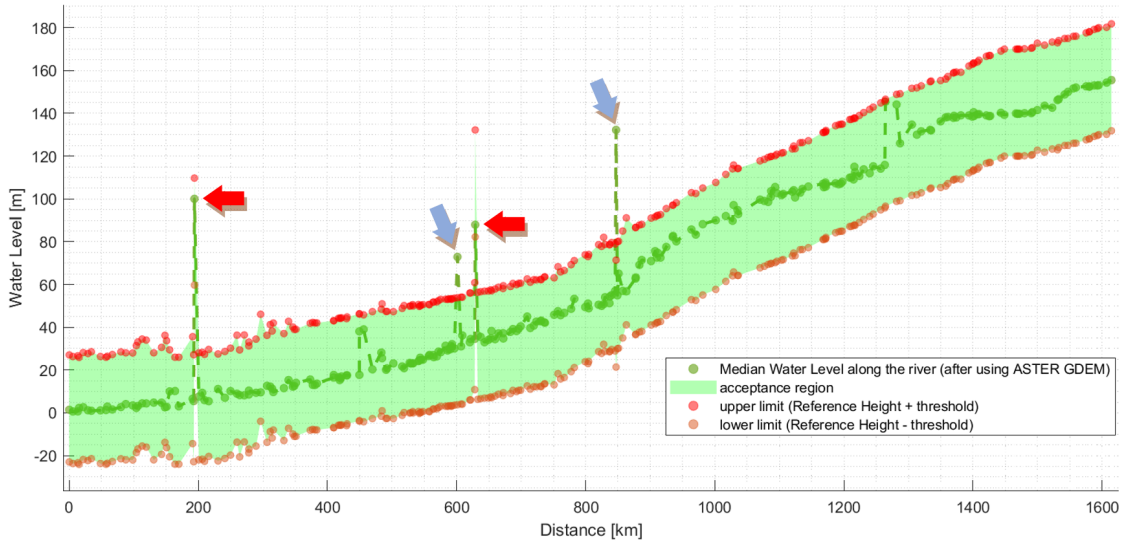


Figure 4.6: *Altimetric water level along the river (after using ASTER GDEM) and the corresponding acceptance region*

Figure 4.6 shows the acceptance region bounded from the reference height and threshold (25 m). There are many reasons for choosing such a large threshold. Small thresholds can make results look better, but prevent a precise assessment of altimeter's performance. A larger thresh-

old will keep a lot of outliers on the one hand, but on the other hand, it can show the performance of the altimeter here as much as possible.

The data pointed to by the red and blue arrows in the figure is exactly these implausible data. But the first thing to note is that the problem pointed to by the blue arrow here is not caused by ASTER GDEM, but by the algorithm of Outlier Identification in AltBundle+. Looking at the upper and lower thresholds, it can be found that these two data are not in the "acceptance region" but are still accepted. The reason is that Outlier Identification in AltBundle+ will only be executed when the length of time series is greater than 5. In both cases, it was due to the low amount of data that the data should have been rejected was retained, which eventually led to the problem.

The data pointed by the red arrow is really caused by the error of the ASTER GDEM. Implausible data is not identified due to the wrong acceptance region bounded by the inappropriate reference height.

Now we will take the data pointed by the first red arrow (around 200 km) as an example to see why this occurred.

Figure 4.7 gives us a visualization of the results obtained in this VS:

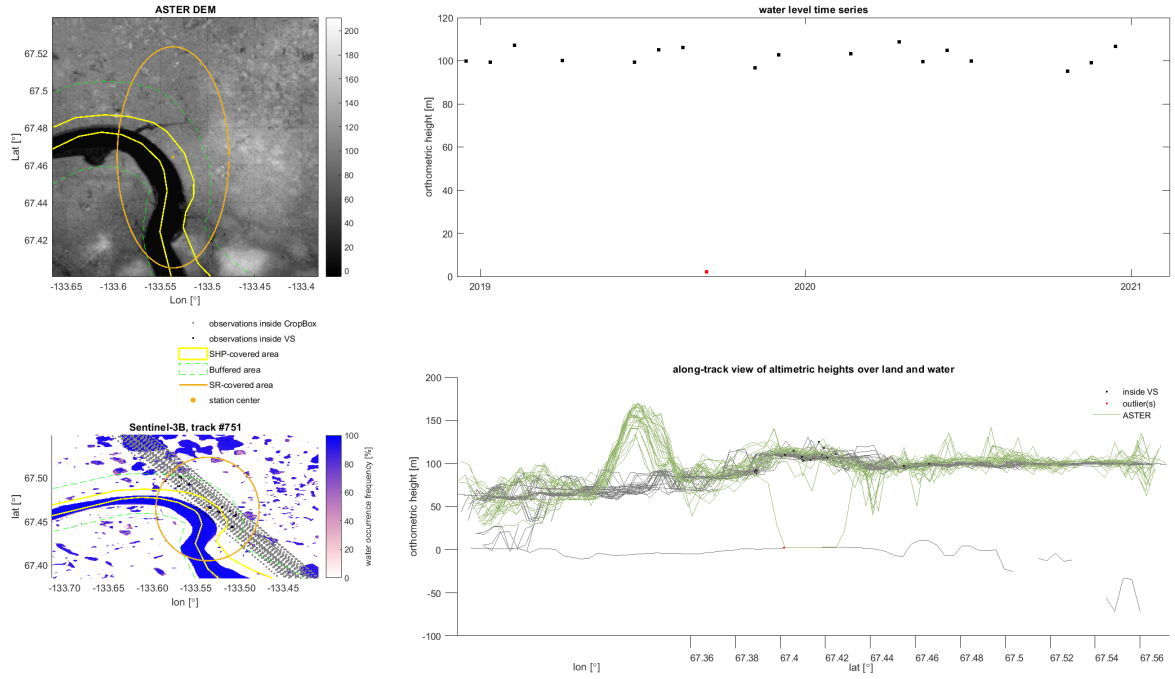


Figure 4.7: *Figures generated by AltBundle+ at MR_p3201t751_1 (Water Level time series + geographic information + along-track view)*

Now extract the along-track view separately:

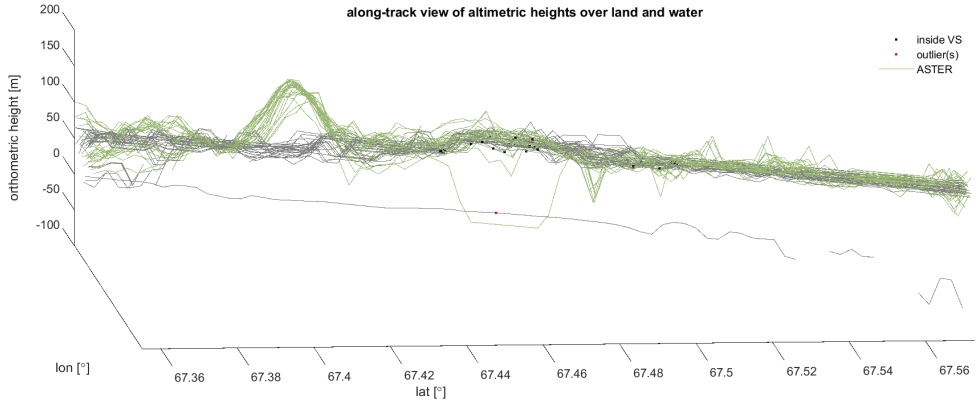


Figure 4.8: *Along-track-view of altimetric height over land and water at MR_p3201t751_1*

Figure 4.8 shows the along-track view of altimetric height over land and water in this VS and the height based on ASTER GDEM represented by the green line. The reference height is about 60 m because most of the heights obtained by interpolation based on ASTER GDEM are clustered at 50-100 m. However, the height from google earth is only about 20 m, which is totally inconsistent with the result obtained by ASTER GDEM. Observing the median water level of the surrounding VSs, 20 m is a more reasonable result. Therefore, there should be something wrong with the reference height.

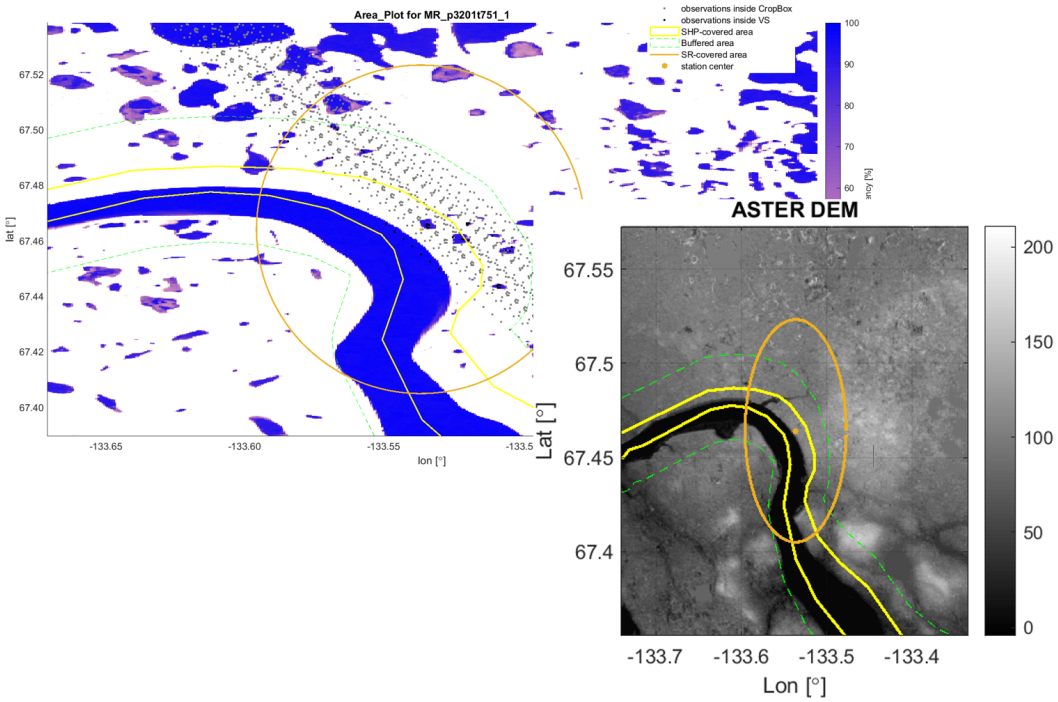


Figure 4.9: *Area plot and ASTER GDEM information at MR_p3201t751_1*

The problem with the DEM seems to be found in Figure 4.9. The yellow dotted line in the figure is the outline of the Mackenzie River extracted from the shapefile, the green line is a

slightly larger search range defined based on this outline, and the blue area is the Mackenzie river extracted from the GSWL. It is worth noting that the yellow outline and the blue area do not completely overlap. There is a certain difference between the data in the shapefile and reality or other data sets such as ASTER DEM and GSWL.

The height we found from the ASTER GDEM is not the water level but the height of the surrounding land, which is also confirmed by querying google earth. The problem is that DEM is just a dataset in 3D coordinates to model the terrain without judging whether a data corresponds to water surface or land, etc. All judgments are done in AltBundle+ based on shapefile, GSWL, etc. The difference between these datasets is the cause of the problem.

The analysis of this one case is enough to reveal the unreliability of DEMs for this application. Since DEM is not a dataset developed specifically for waters, there is no way to judge whether the data obtained is really from the water surface. To be more precise, the problem is not really the DEM itself but our options in deriving the desired height from a DEM. While there are rooms for a smarter usage of a DEM as a reference height, there is no easy way to derive the right height in every scenario automatically. Moreover, in the event of an error, the obtained height may completely deviate from the true water height and have a huge impact on Outlier Identification.

If a more suitable dataset can be found to make up for this defect, the results may be further optimized.

4.2 The SWOT A Priori River Database (SWORD)

4.2.1 Introduction

To expand observations of river water surface elevation (WSE) and slope, the upcoming Surface Water and Ocean Topography (SWOT) satellite mission will be launched in 2022. A couple of a priori datasets are provided to help with the SWOT data analysis after launch. One of these products is the river vector product. For practical interpretation and application of SWOT measurements, a global a priori database with defined SWOT reaches and nodes a priori are required so that SWOT data can be assigned to them. The SWOT A Priori River Database (SWORD) was built to support the development of RiverObs. Multiple global rivers and satellite related datasets are combined to define the nodes and reach that will constitute SWOT river vector data products (Altenau et al., 2021).

SWORD provides fixed node locations with high-resolution (200 m), reach boundaries (10 m) in vector and netCDF formats with attached hydrologic variables (WSE, width, slope, etc.). Table 4.1 provides a summary of data sets and the attributes they contribute to the final product.

The data that is useful for us is WSE, which is similar to DEM and can be used to assist us in identifying the outliers of altimeter.

Dataset	Attribute Contribution
Global River Widths from Landsat (GRWL)	Provides river centerline locations at 30 m resolution
MERIT Hydro	Provides river surface elevation and flow accumulation at 3" resolution
HydroBASINS	Provides Pfafstetter nested basin codes
Global River Obstruction Database (GROD)	Provides global locations of river obstructions along the GRWL river network
Global Delta Maps	Provides the spatial extent of 48 of the world's largest deltas

Table 4.1: *Summary of data sets used in the development of SWORD (Altenau et al., 2021)*

4.2.2 Add SWORD in AltBundle+

The extraction of SWORD data is relatively complex, not like DEM. The core is to determine the correct reach and node. The centerline of the river is stored in the SWORD, which describes the midpoint along the length of a body of flowing water, and each centerline also has its corresponding reach ID and Node ID. AltBundle+ finds the centerline closest to VS and also the reach and node. After a series of operations, AltBundle+ finally extracts the WSE of the Node closest to VS as a reference value. In fact, we expect more from SWORD than the reference value extracted from DEM. After all, SWORD is specially developed for rivers.

4.2.3 Results after using SWORD

Figure 4.10 shows the acceptance region bounded from the reference height based on SWORD and the threshold. Although the implausible data pointed by the blue arrow caused by the algorithm still exists, the data pointed by the red arrow in Figure 4.6 has returned to more plausible.

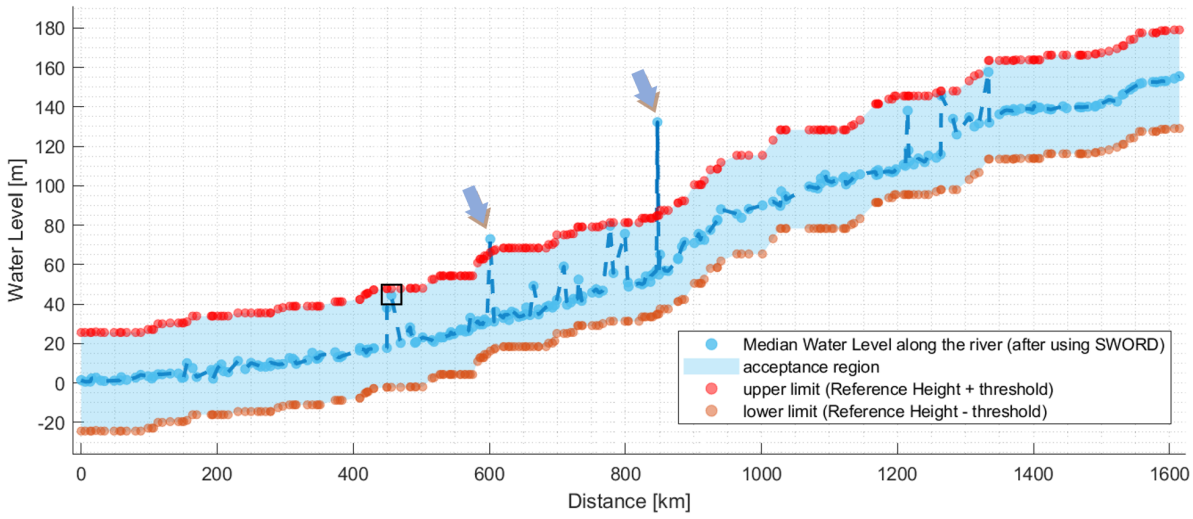


Figure 4.10: *Altimetric water level along the river (after using SWORD) and the corresponding acceptance region*

Comparing the reference heights based on ASTER GDEM and SWORD (see Fig. 4.11), the reference height obtained based on SWORD is generally more reasonable. The occurrence of excessive errors can be effectively avoided because of the characteristics of SWORD based on nodes and reach.

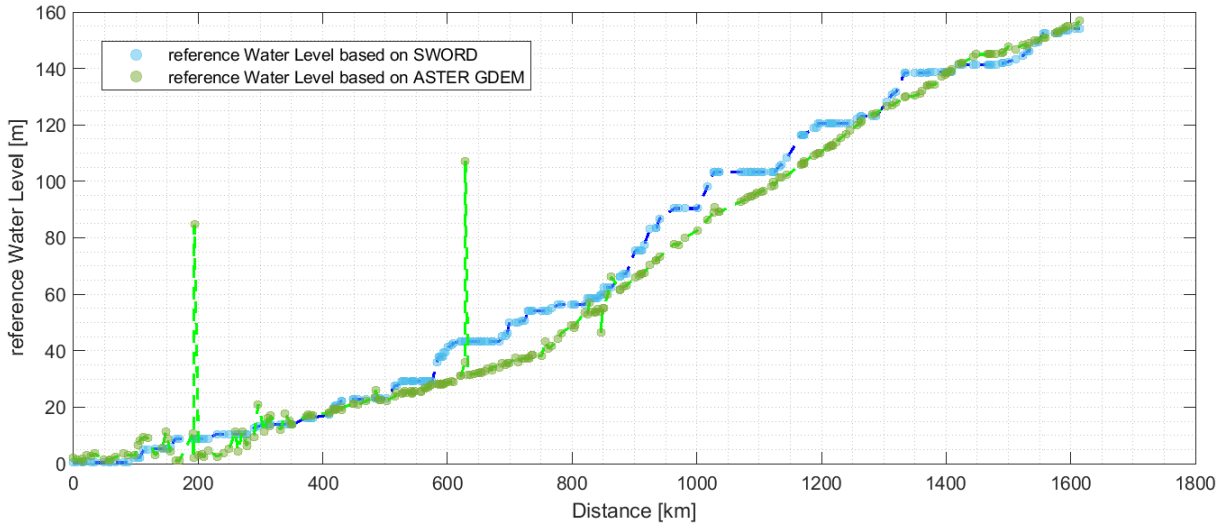


Figure 4.11: *Reference water level based on SWORD and ASTER GDEM*

After solving the excessively large outlier, there are still many potential problems worthy of study for the data in the acceptance region. For example, the data marked by the black box in Figures 4.10 represent many cases that are in the acceptance region but are relatively bad compared with the adjacent VSs. For this VS, when the surrounding median water levels are only about 15 meters, the median water level by this VS does reach 40 meters.

Figure 4.12 shows the along-track view in this VS, where the blue plane represents the reference height based on SWORD. It is worth noting that the reference height obtained by SWORD is also about 20m, so the problem here is not due to the wrong reference height.

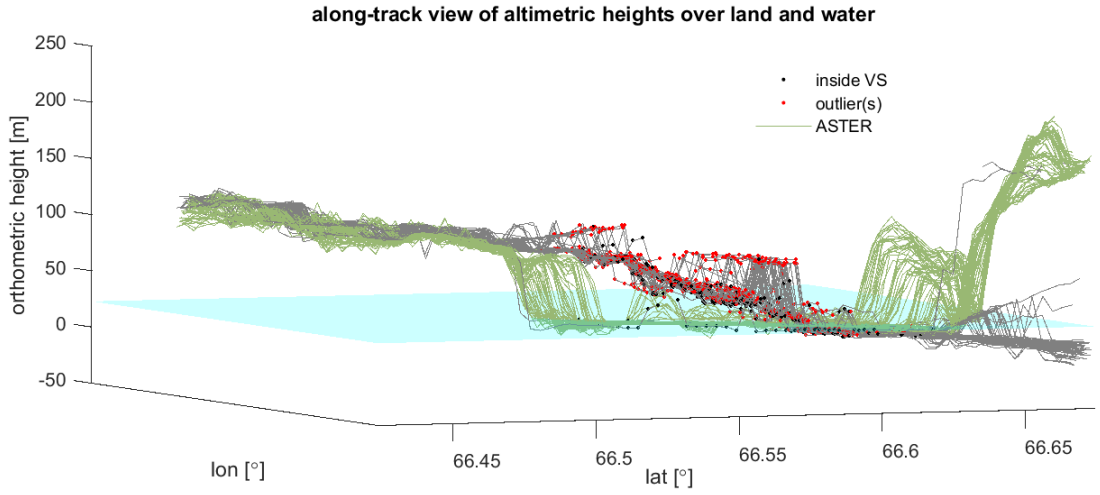


Figure 4.12: *Along track view of altimetric height over land and water at MR_p3101t323_3*

By studying the water level time series (see Fig. 4.13), the problem can be found. As mentioned, a threshold of 25 m was set to bound the acceptable region. However, since the area limited by 25 m is not quite small, in this case, a large number of data located at 40 m, which we consider implausible, are retained after Outlier Identification.

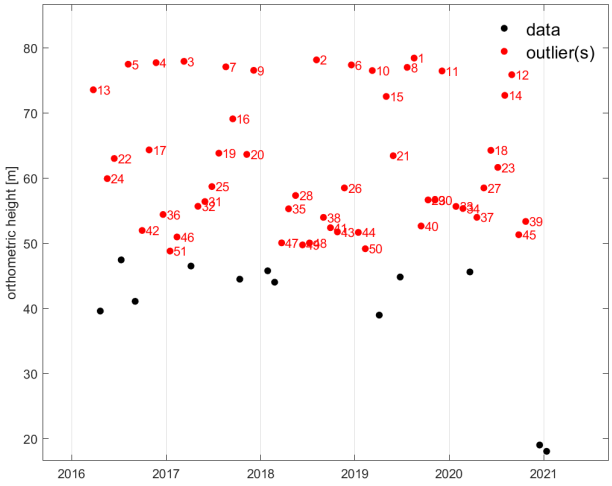


Figure 4.13: *Water level time series with outliers at MR_p3101t323_3*

The normal logic is that we can reduce the threshold to make the Outlier Identification "stricter", but in fact this does not make sense. The reference height we get is only an estimated value, not the real value. Simply setting a smaller threshold will give too much weight to the reference height and could be a huge challenge to the correctness of the reference value.

What we want is a reliable water level rather than a height closer to the reference. More importantly, with a small threshold the results of the altimetry may not be truly reflected. For this case, a threshold of 5 m or even 1 m can allow most of the data around 40 m to be filtered and only keep the two reasonable data in the lower right corner. However, although the results seem reasonable, the altimeter actually performs very badly here.

So far, there is nothing more to improve the accuracy of the Outlier Identification. However, as mentioned before, an important process in satellite altimetry is retracking. If the retracking is not accurate enough, the height obtained will be imprecise. In other words, if the retracking algorithm can be optimized, the results may be improved. Next, we will introduce a new retracking algorithm.

4.3 Leading Edge Identification with Prior Information (LEIPI)

4.3.1 Introduction

In situations where best efforts have been made to identify outliers, a new retracking algorithm called Leading Edge Identification with Prior Information (LEIPI) was used to improve the accuracy of the retracking.

Most of the existing retracking algorithms post-process the altimetry waveforms, find the midpoint of the leading edge, determine its offset to the predefined tracking point and correct

the water level estimates. Due to the complexity of non-ocean surfaces, new algorithms have been developed over non-ocean surfaces since 1977, e.g. the ocean retracker, the 5β parameter retracker, the OCOG method, and the sea ice retracker etc. However, most of the existing methods take a case independent and forward approach in finding the retracking gate, and none of them can handle all waveform complexities. These complex factors will systematically affect the inland water level time series. More accurate results should be obtained if reliable prior data is available, i.e. if we can analyze the prior time series, infer the types of contamination during the radar acquisition, and confirm how they were translated into waveform properties.

4.3.2 Work process

Leading Edge Identification with Prior Information (LEIPI) is a new approach developed @ Institute of Geodesy, University of Stuttgart to retracking inland altimetry waveforms that relies on prior information. This prior information is a model of the time series, which is obtained via the Outlier Identification analysis. Figure 4.14 shows an example to explain the process:

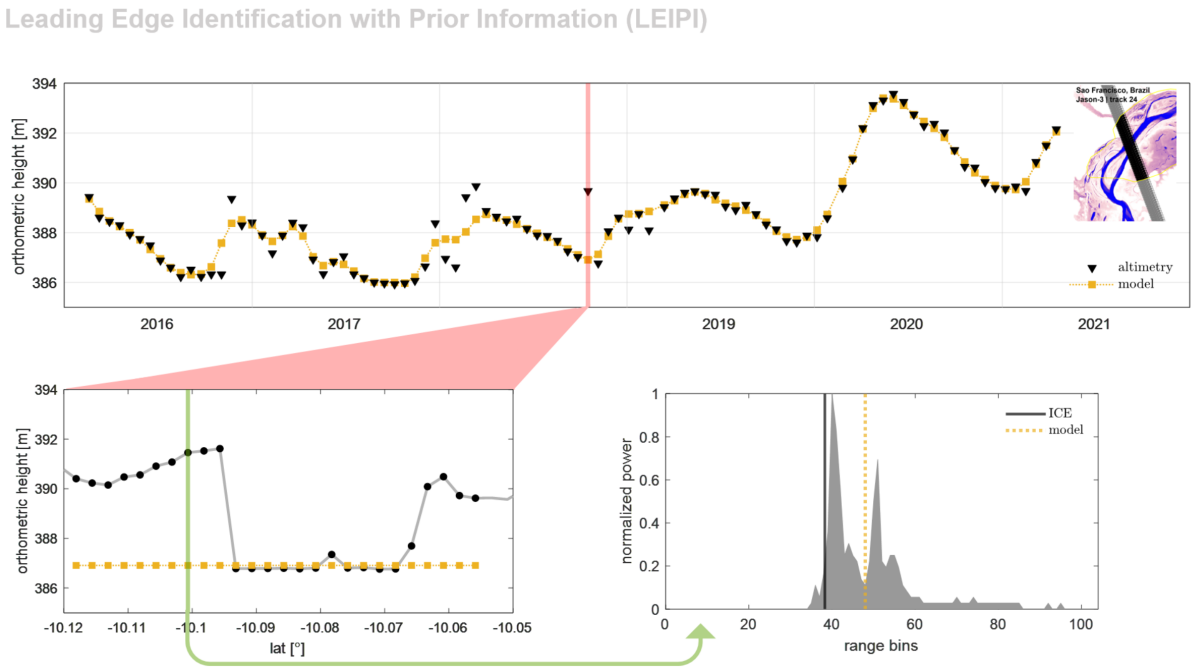


Figure 4.14: *Water level time series and the model at San Francisco, Brazil*

The yellow line is the mentioned model. The deviation from the model value for every single outlier is calculated and translated into radar bin units (single outlier: the black dots that the green straight line passes through). The retracking should be reconsidered if the deviation falls into the covered range spectrum of the waveform. The figure in the lower right corner shows the difference between the default retracker gate and the one obtained by LEIPI.

The retracking principle of LEIPI is essentially a Bayesian estimation. Figure 4.15 roughly shows the basic principles. First, a normal prior distribution is represented by a searching area bounded by the model value, and the model value should also be reliable enough to bound the searching area where the leading edge can be located.

In addition to the prior distribution, the likelihood function is defined as the derivative of the waveform. Then the posterior probability can be calculated. LEIPI nominates simply the tracking gate with Maximum A Posterior probability as the best retracking gate.

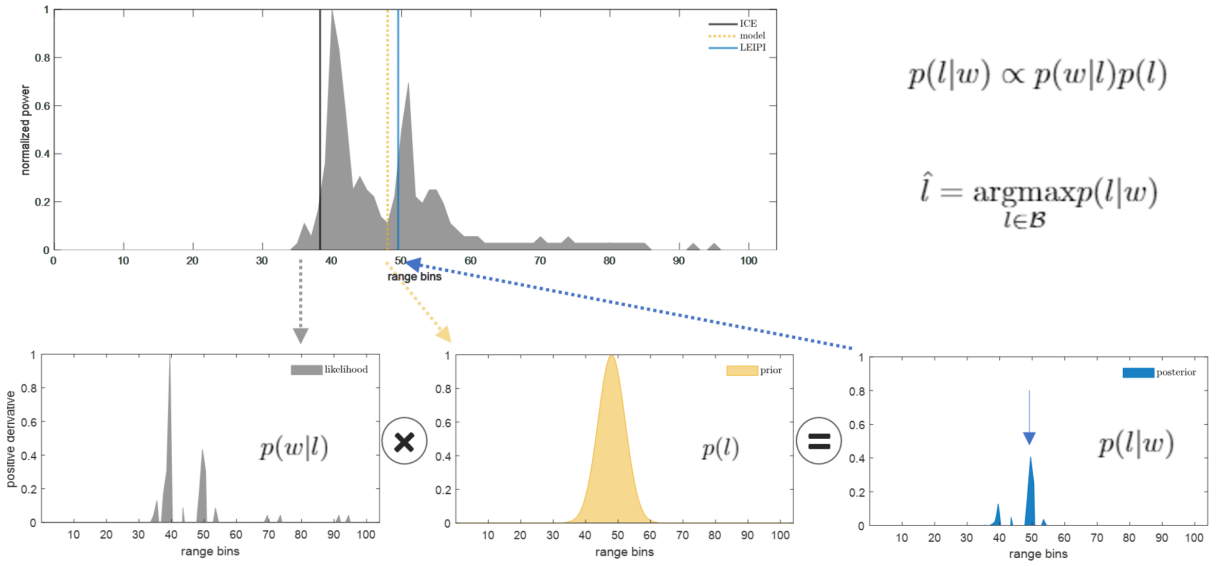


Figure 4.15: *LEIPI: a Bayesian approach*

Finally, as shown in Figure 4.16, the blue triangle depicts the result corrected by LEIPI. Validation with In-situ data confirms that LEIPI does get a good retracking result in general.

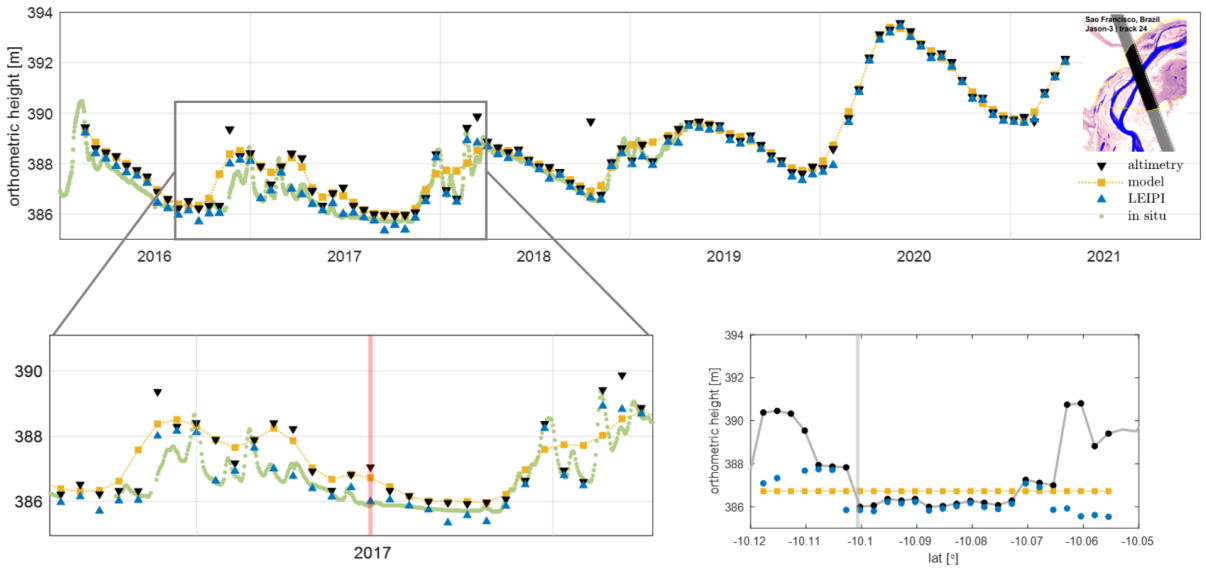


Figure 4.16: *Water level time series after using LEIPI*

4.3.3 Results and Visualization

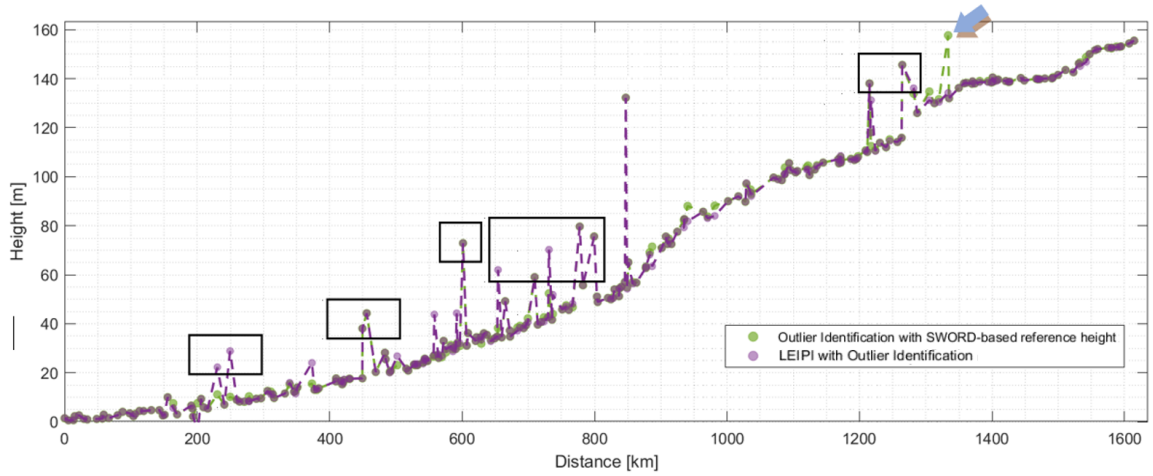


Figure 4.17: *Altimetric water level along the river in two scenarios: 1. after OI, 2. after LEIPI+OI*

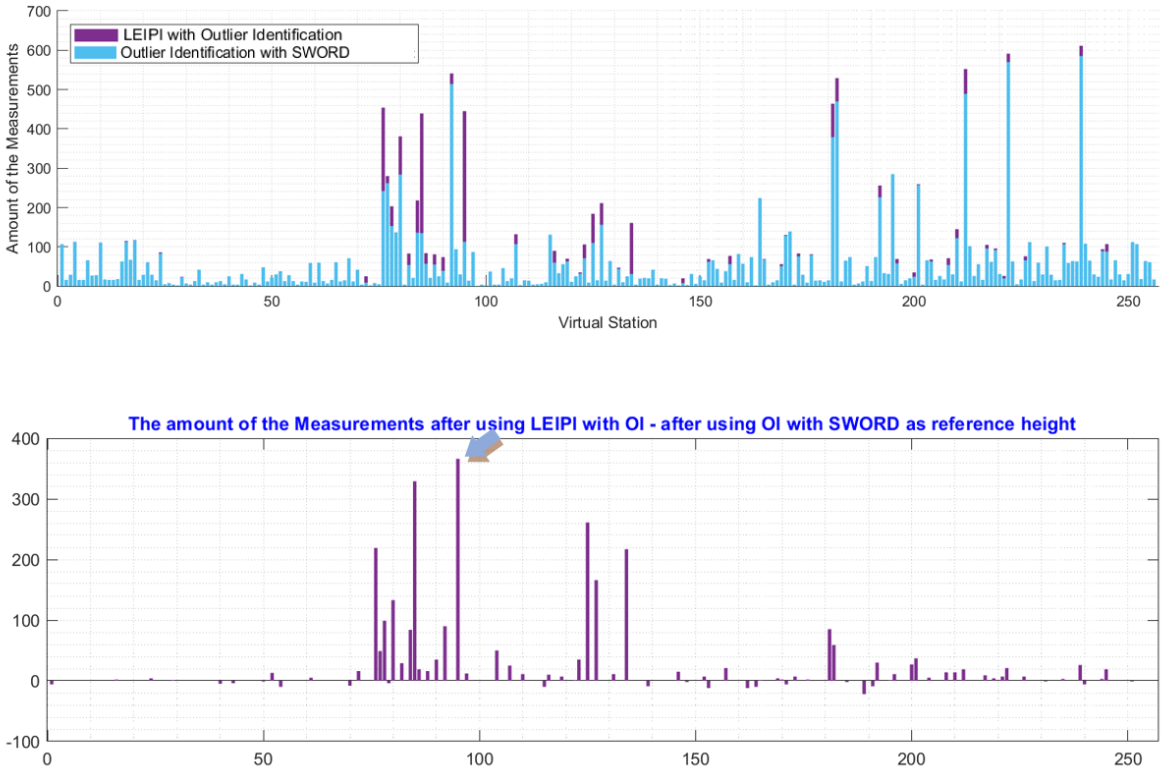


Figure 4.18: *Difference between available data after using LEIPI+OI and after using OI with SWORD-based reference height*

The results after using LEIPI+OI and the previous results after using Outlier Identification with SWORD-based reference height are shown in the Figure 4.17 as a comparison.

What needs to be mentioned here is that since the performance of SWORD is better than that of ASTER GDEM, subsequent comparisons will only use the results obtained based on Outlier Identification with using SWORD. At first glance, compared to only using Outlier Identification, there are obvious improvements in some places after using LEIPI+OI, but at the same time, some data also seems to be getting worse. The data pointed to by the blue arrow serves as a good case to show the role of LEIPI here, but on the other hand, the data in the black box seems to be less than "ideal".

Figure 4.18 clearly reveals the effect of LEIPI by showing the amount of data available for each VS after using LEIPI or SWORD. After using LEIPI, the data volume of almost all VSs has been improved because most of the heights obtained by uisng LEIPI are in the acceptable region, which is exactly what we expected.

Next, taking the VS pointed by the blue arrow in Figure 4.18 as an example, which represents the huge increase in the amount of data after using LEIPI. The working process of LEIPI will be explored through visualization. Figure 4.19 shows the time series at this VS after only using Outlier Identification and using LEIPI+OI, respectively. In fact, this is a very interesting case. On the one hand, the amount of data is significantly improved after using LEIPI, and it also shows periodic changes, which is also in line with a real water level time series, but on the other hand, there are fluctuations of tens of meters between the data, which seems unreasonable.

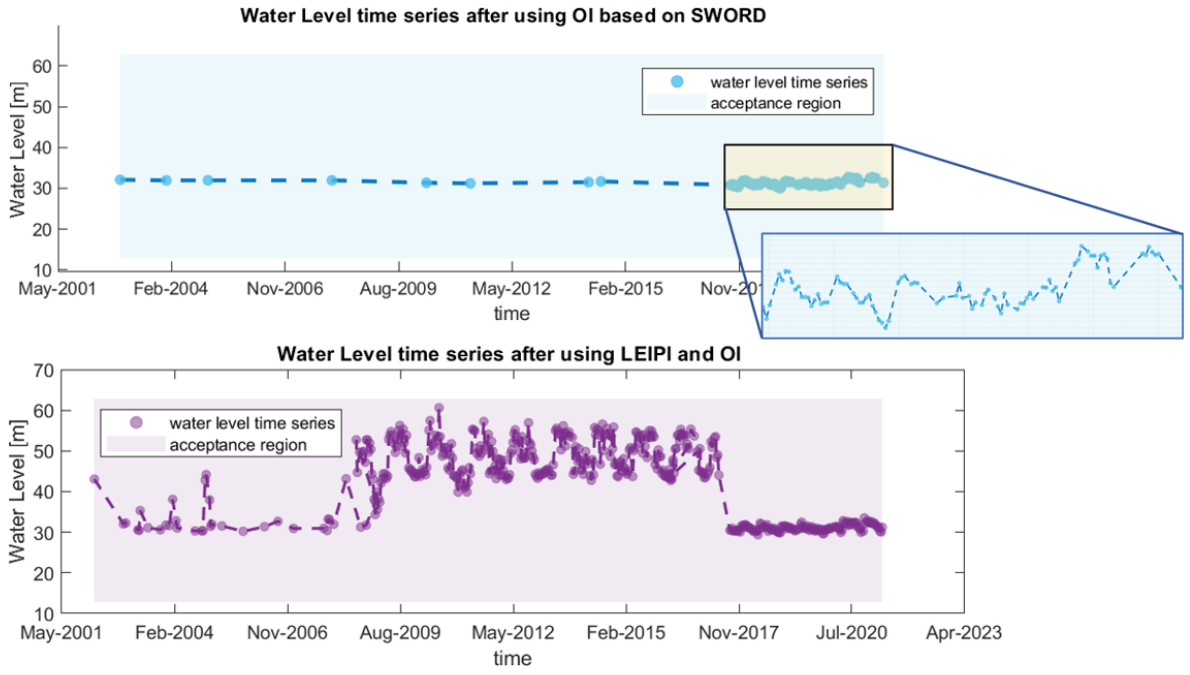


Figure 4.19: *Water level time series at MR_p2000t228_3: Top figure: using OI based on SWORD; Bottom Figure: using LEIPI+OI*

Since the track number here is 2000, the final result is actually composed of the measurement results of Jason 1, 2, 3. Taking the result of Jason-2 (around 2008 to 2017) as an example, Figure

4.20 shows the results of after only using Outlier Identification, and we can find that the data was filled with a large number of outliers. If combined with Figure 4.19, it is not difficult to draw the conclusion that these outliers become acceptable results after using LEIPI.

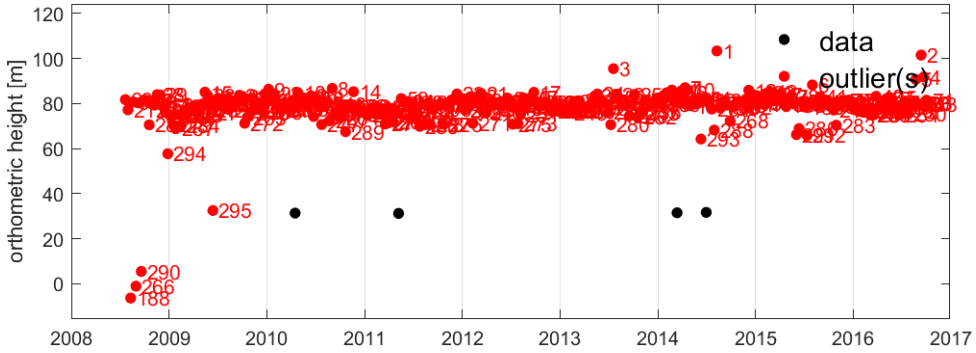


Figure 4.20: *Water level time series with OI at MR_p2000t228_3 (only using OI based on SWORD)*

In addition, if we check the approximate height of this water body in google earth, the obtained height of about 40 m also reveals the reasonableness of the results from Jason-2 here.

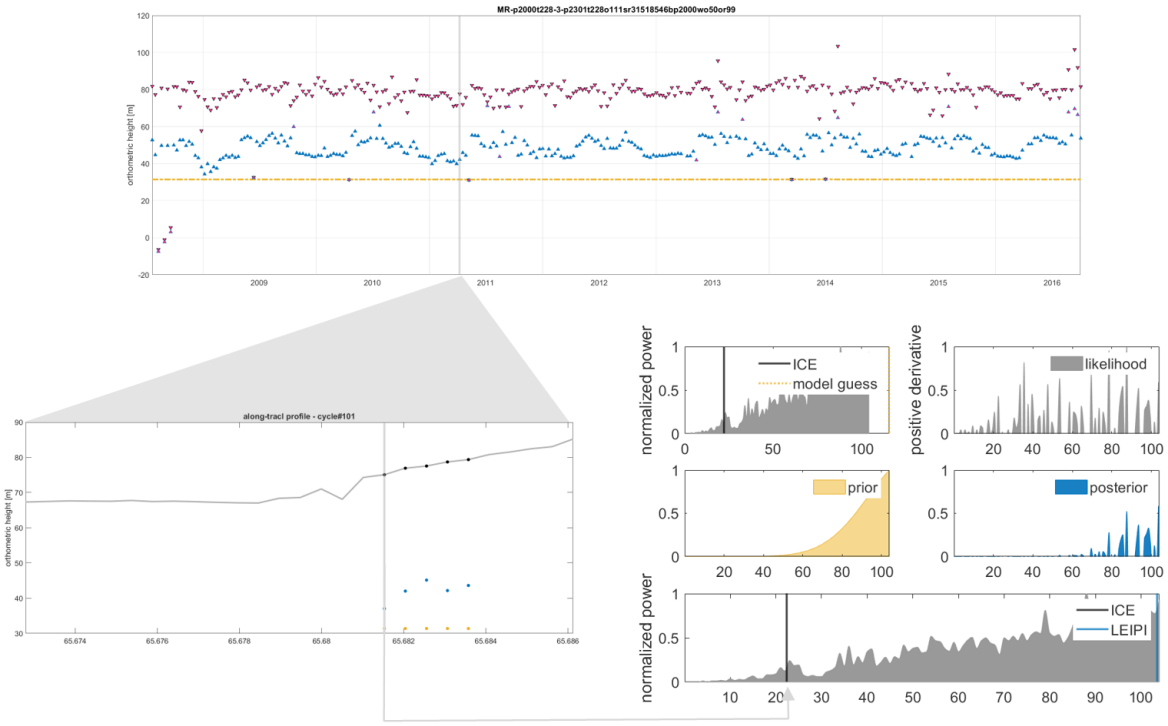


Figure 4.21: *Visualization of LEIPI at VS MR_p2000t228_3*

Now, if we only focus on the process of LEIPI (see Fig. 4.21), at least the whole procedure is also logical. The red dots representing outliers have been significantly improved after using LEIPI. After combining with the prior information, the corrected results is somewhat more like a real time series. But the exact reason for the huge difference in the water level time series cannot be determined.

Returning now to water level along the river shown in Figure 4.17, as mentioned, the data indicated by the blue arrow is what seems to be the most successful or typical example. Checking the TS of this VS in Figure 4.22, we can find that although the amount of data obtained by LEIPI is still very small, the result is more like a normal TS and also closer to the reference height.

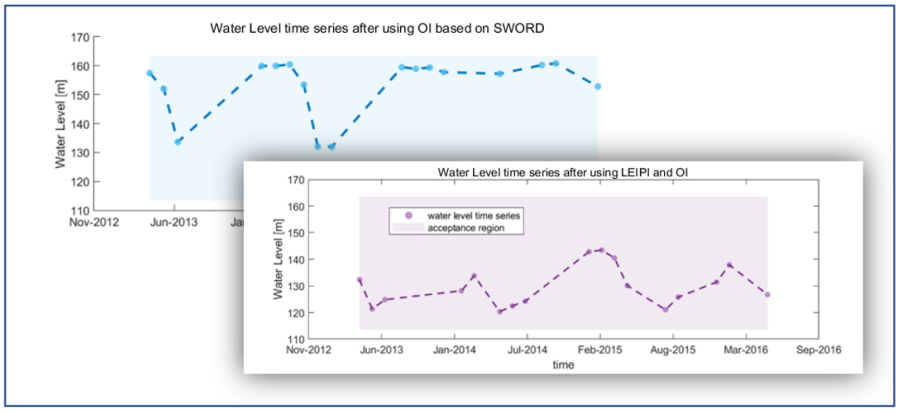


Figure 4.22: *Water level time series after using OI with SWORD-based reference height and after using LEIPI+OI*

Finally, let's focus on these corrected but worse results in the black boxes in Figure 4.17. Figure 4.23, 4.24 shows the TSs and the visualization from one of the VSs in this area:

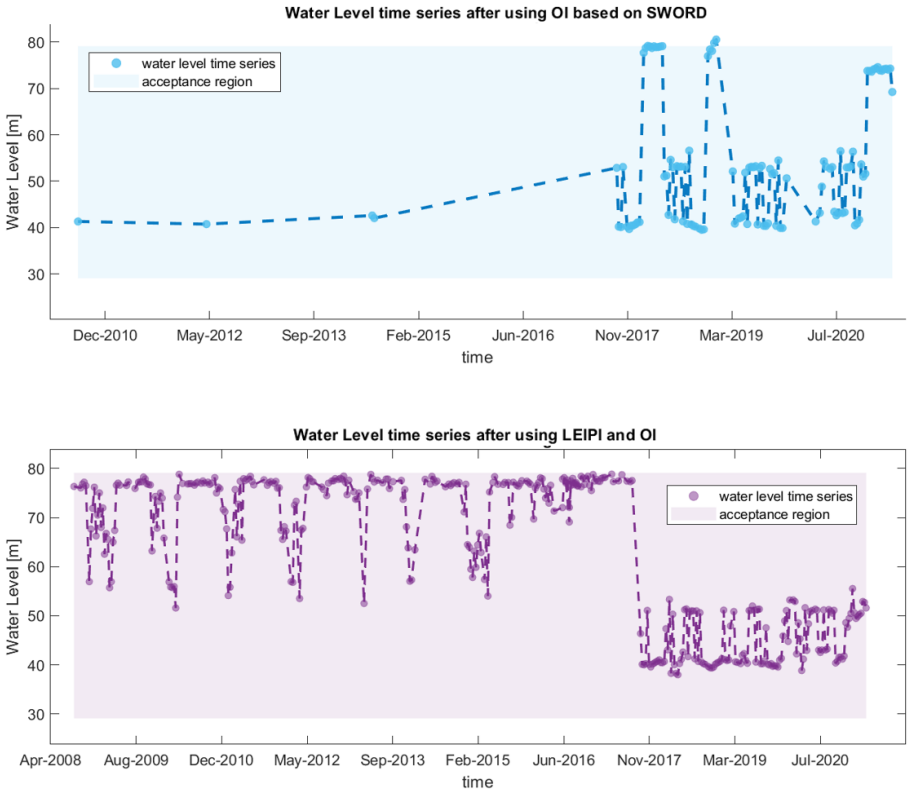


Figure 4.23: *Water level time series after using OI with SWORD-based reference height and after using LEIPI+OI*

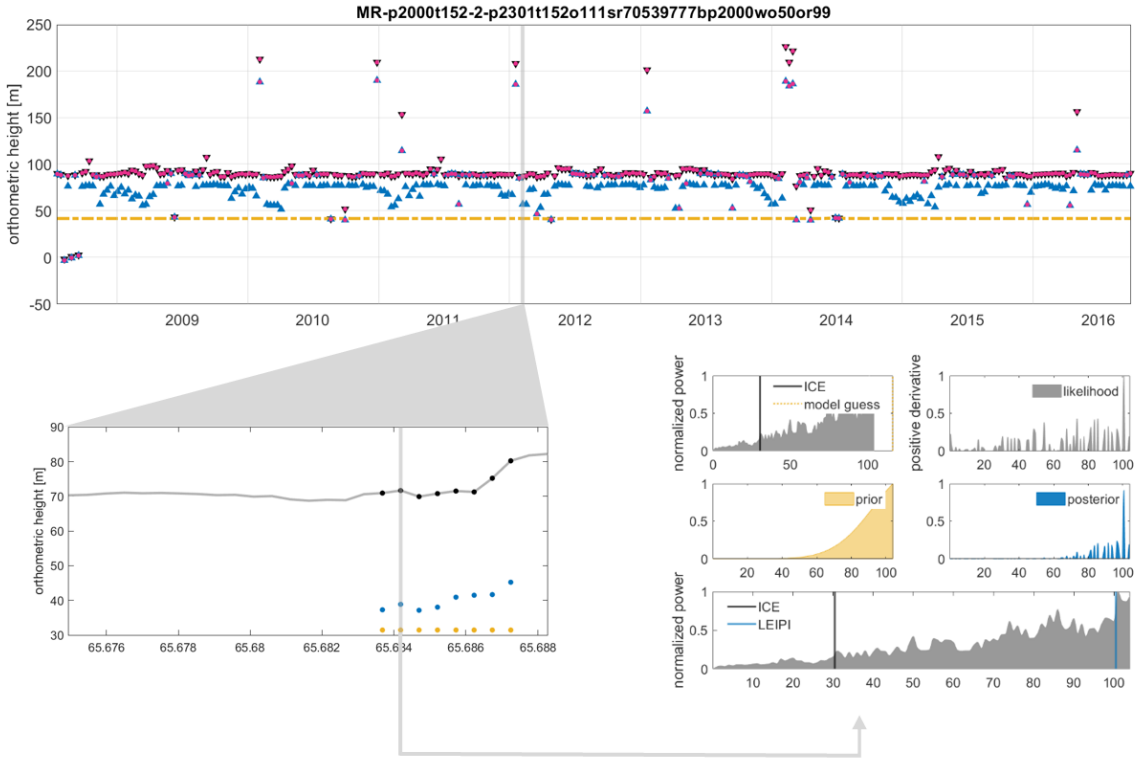


Figure 4.24: *The visualization of LEIPI*

A similar thing happened again, the results obtained by may look like normal time series, but the results obtained by different satellites are slightly different, although they share the same orbit. The next chapter will present the final results and summarize this series of questions.

Chapter 5

Results and conclusions

5.1 Final Results

So far, we have tried different methods to optimize the results, and the results have indeed improved. However, there seems to be another possibility for further improvement. The general process is to smooth the results obtained by using LEIPI+OI to obtain a model, update the reference in Outlier Identification to this model, and then reprocess the data with LEIPI+OI again. The basic idea is that the Outlier Identification based on the reference height obtained from the LEIPI+OI processed results itself can better reflect the true state of the data without ignoring reality.

Here the filter we use to smooth the whole result is again the Savitzky-Golay filter, which is a digital filter that can be applied to a set of points for the purpose of smoothing the data. On the basis of using the Savitzky-Golay filter, the data will be iterated continuously until the standard deviation of the difference between the data value and the smoothed data is small enough.

The results have no need for further optimization after four iterations:

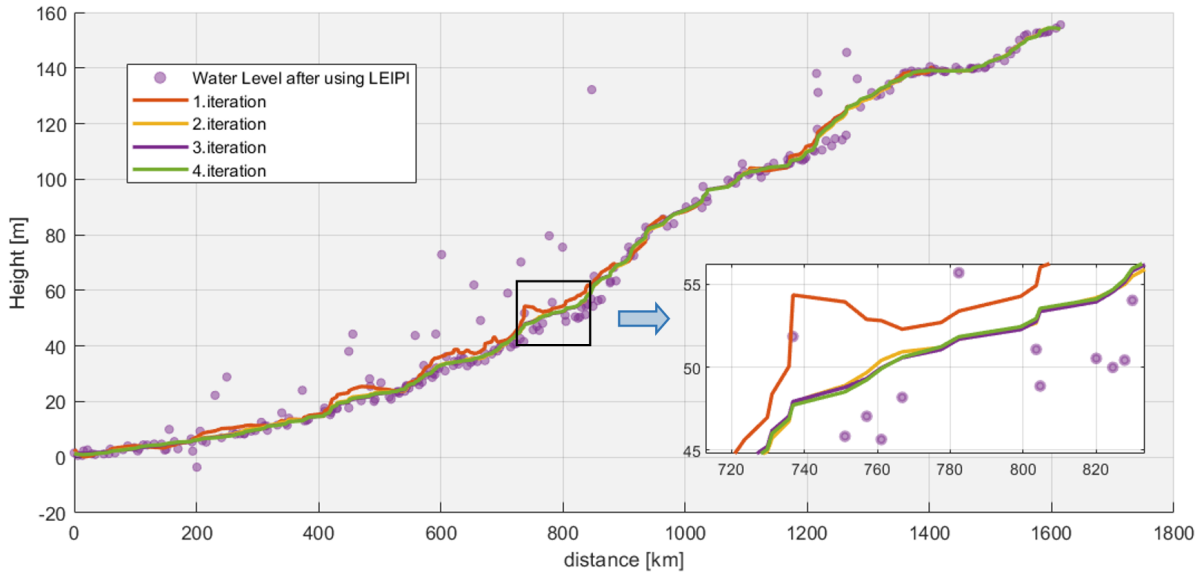


Figure 5.1: *Data smoothing based on Savitzky-Golay filter*

Then AltBundle+ reprocesses the data with these results as the new reference height. It should be noted here that although the final reference looks more accurate than the one we set before, the previously set threshold of 25 m remains unchanged because we have emphasized many times that the whole process is to study the specific performance of the inland altimeter and not to get a result that looks good. Although a small threshold may help us get a good result, it will potentially hide the problem of the altimeter's results.

The result is shown in the Figure 5.2, which is also the final result of the whole process. The previously mentioned data for which Outlier Identification was not performed due to too little data is removed and is not considered in the final result.

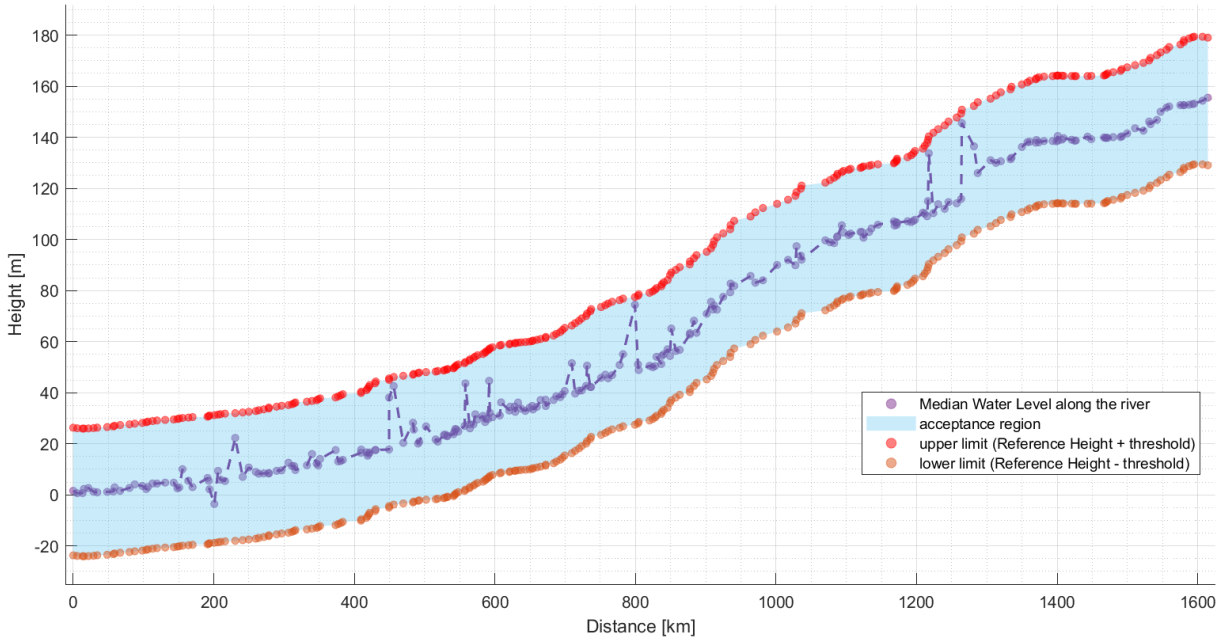


Figure 5.2: *Final results: median altimetric water level along the river and the corresponding acceptance region*

Compared with the results of using LEIPI+OI for the first time (see Fig. 5.3), except for some local optimizations, there is no overall huge change this time. As shown in the black box, some results were further optimized. The overall result looks more reasonable and more in line with a real water level along the river. In addition, if compared with the results using Outlier Identification based on SWORD, it must be noted that some bad data disappeared indeed, but at the same time, some results got worse.

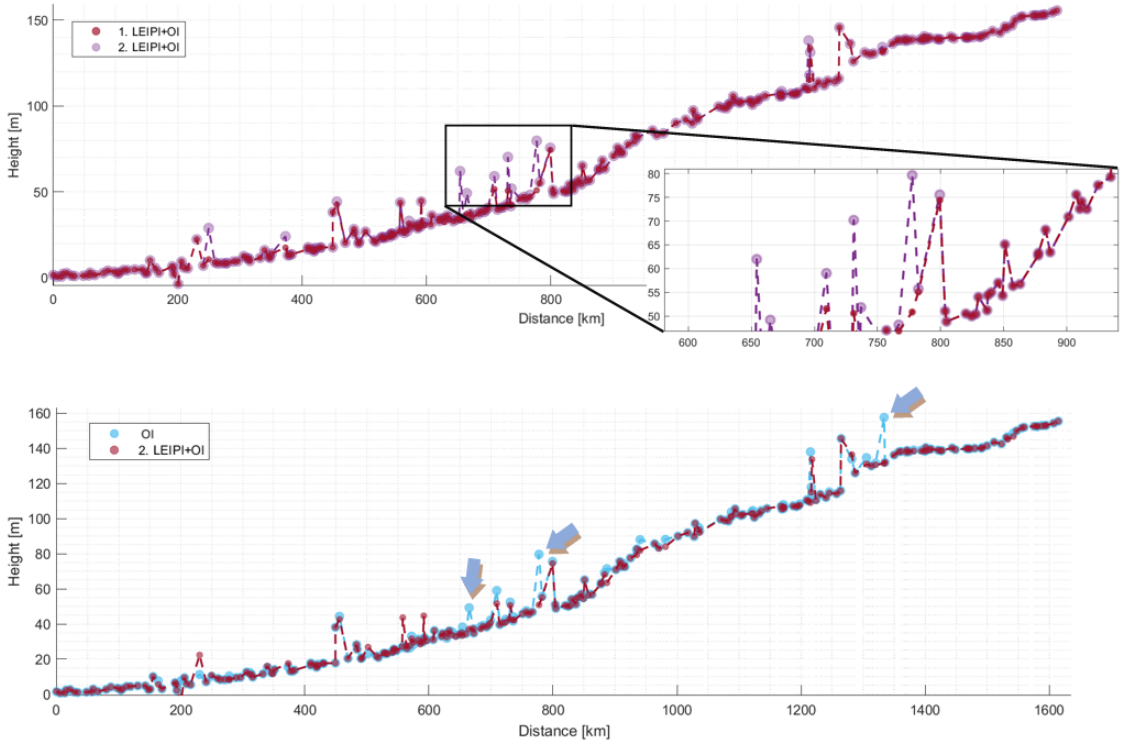


Figure 5.3: *Comparison in different scenarios -> Top Figure: Comparison between using LEIPI + OI (1.time and 2.time) Bottom Figure: Comparison between using LEIPI + OI (2.time) and OI based on SWORD*

5.2 Summary

In general, as mentioned at the outset, the focus of this work is on algorithmically improving the results of existing altimetry data and exploring the usability of inland altimetry. The improvement of the data is divided into two directions, optimizing the Outlier Identification and the retracking methods.

5.2.1 Optimizing the Outlier Identification

ASTER GDEM and SWORD datasets were applied to aid in the identification of outliers. Overall this is a very simple but efficient way to improve the Outlier Identification. Figure 5.4, 5.5 show the amount of the data without using Outlier Identification, using Outlier Identification without reference height, and using Outlier Identification with ASTER or SWORD-based reference height, respectively. The trick here is to output the data in the order of "without Outlier Identification" - "with Outlier Identification" - "with Outlier Identification based on ASTER/SWORD", because in most cases, the amount of data in these three scenarios decreases in turn, such an output method can show the changes in the amount of data in three scenarios as much as possible.

We already know from the previous analysis that the distribution of the data itself will greatly affect the results of Outlier Identification when the reference height is not used. It can also be seen in the Figure that in many cases, even when Outlier Identification is used, the amount of data barely changes. Then compared with only using Outlier Identification, when the reference height is used as an auxiliary, the amount of data decreases sharply because a clear "condition" is introduced.

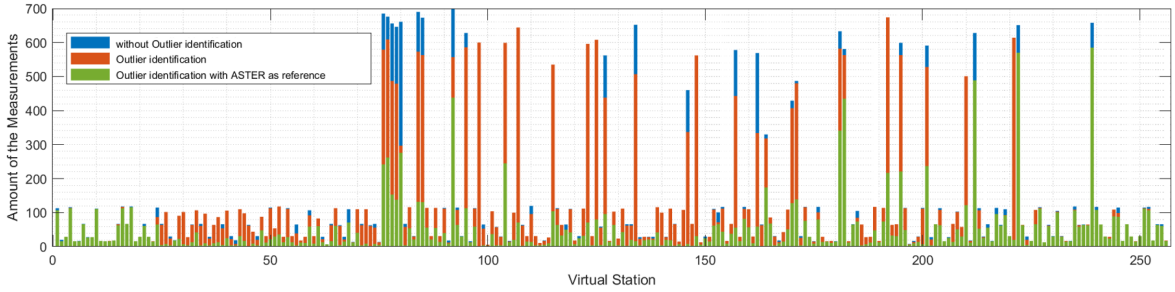


Figure 5.4: *The amount of data when OI is not used, and when only OI is used, and when OI based on ASTERGDEM is used*

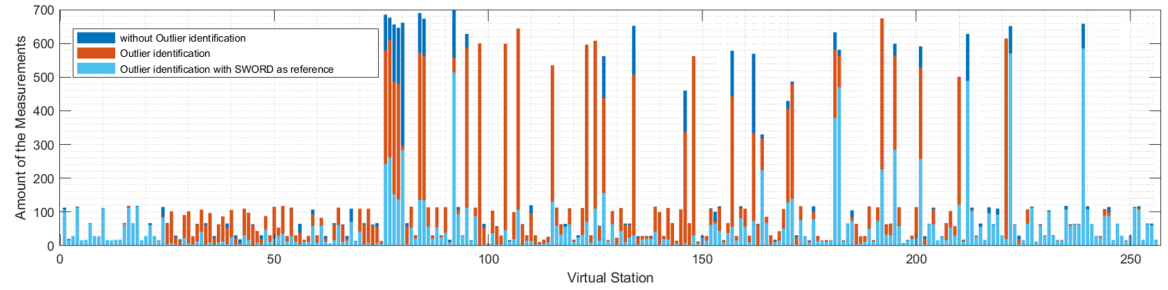


Figure 5.5: *The amount of data when OI is not used, and when only OI is used, and when OI based on SWORD is used*

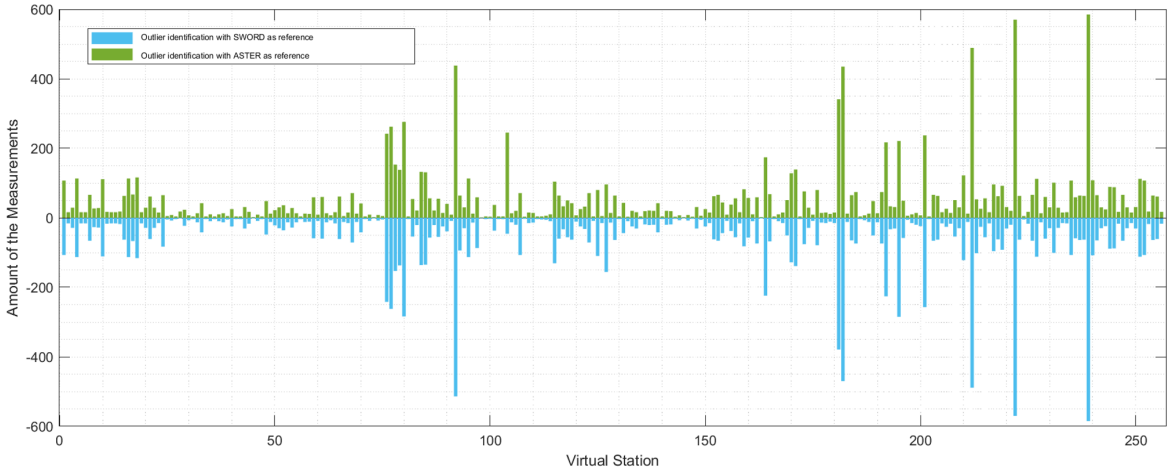


Figure 5.6: *The amount of data when OI based on SWORD or ASTER GDEM is used (turn all results from SWORD into negative numbers for comparison)*

However, if the amount the data after using SWORD and ASTER are compared (see Fig. 5.6), the difference between the two is not significant.

Although a reference can greatly improve the results, due to the gap between the reference and the true height, a small threshold is not desirable, which also makes further optimization impossible. With the help of the reference height, it is possible to limit the results to a range of 30, 20, or even 10 meters from the true value, but it is impossible for Outlier Identification to go further and it also does not make sense.

5.2.2 Optimizing the retracking methods

The second idea of optimizing the result is to optimize the data itself, that is, try a new retracking method to correct the outlier to a plausible value.

The new approach to retracking inland altimetry waveforms that relies on prior information named LEIPI was used.

As shown in Figure 5.7, the visible purple bars represent an increase in the amount of data remaining after using LEIPI + OI compared to only using Outlier Identification with SWORD-based reference height. This shows that after LEIPI's correction, many data have been restored to a reasonable range, and LEIPI has indeed played a role.

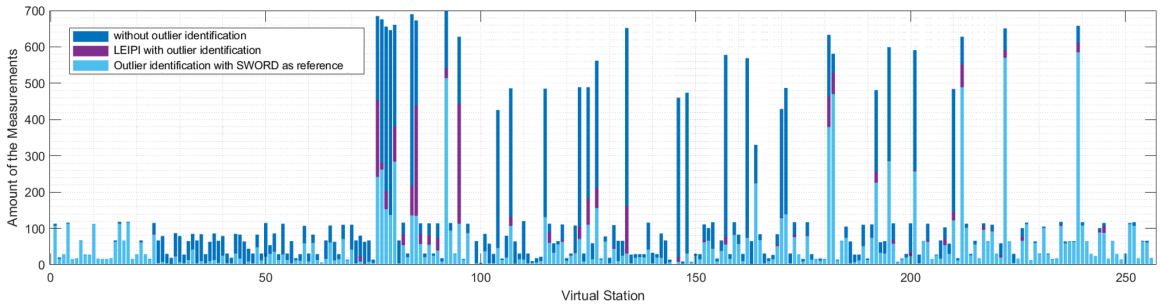


Figure 5.7: *The amount of data when OI is not used, when OI based on SWORD is used and when LEIPI with OI is used*

But on the other hand, recalling the previous case study on LEIPI, there are many cases that we retrieve something which looks like every viable time series, but compared with other data obtained at the same time, there are quite a few differences. Then the comparison with only using Outlier Identification also shows that the use of LEIPI needs to be cautious.

5.2.3 Global River Ice Dataset

As the research object of this paper, we always expect that there should be some internal relationship between the characteristics of the Mackenzie River at high latitudes and the performance of the altimeter. A brand new attempt is to make use of the global ice database to analyze the altimetry results.

The global river ice dataset used here is a global multitemporal river ice extent dataset produced by Yang et al. (2020) based on 407,880 satellite images from 34 years of observations from the Landsat 5-8 mission (1984-2018). It is a simple yet highly predictive, empirical model of river ice extent to reveal patterns of change in global river ice cover and to apply the model to future climate projections. The basic idea is to extract snow/ice conditions from the Landsat image quality band on the Google Earth Engine platform based on the Fmask algorithm used by the US Geological Survey, which labels each pixel as clear, water, cloud, cloud shadow or snow/ice. However, to reduce the volume of data, pixel-level snow/ice conditions were aggregated into the percentage of total river length covered by ice, or river ice extent, for each Landsat image. The dataset consists of one CSV file (Yang et al., 2019) containing estimated river ice (length) fraction from 1984 to 2018 based on images from USGS Landsat satellite missions, where useful to us are:

- date: the date on which the image was captured
- river_ice_fraction: the length fraction of ice cover (range: [0, 1])
  - 0: no ice covered / 1: all lengths in this area are ice-covered
- PATH: the path of the Landsat image (under WRS-2)
- ROW: the row of the Landsat image (under WRS-2)

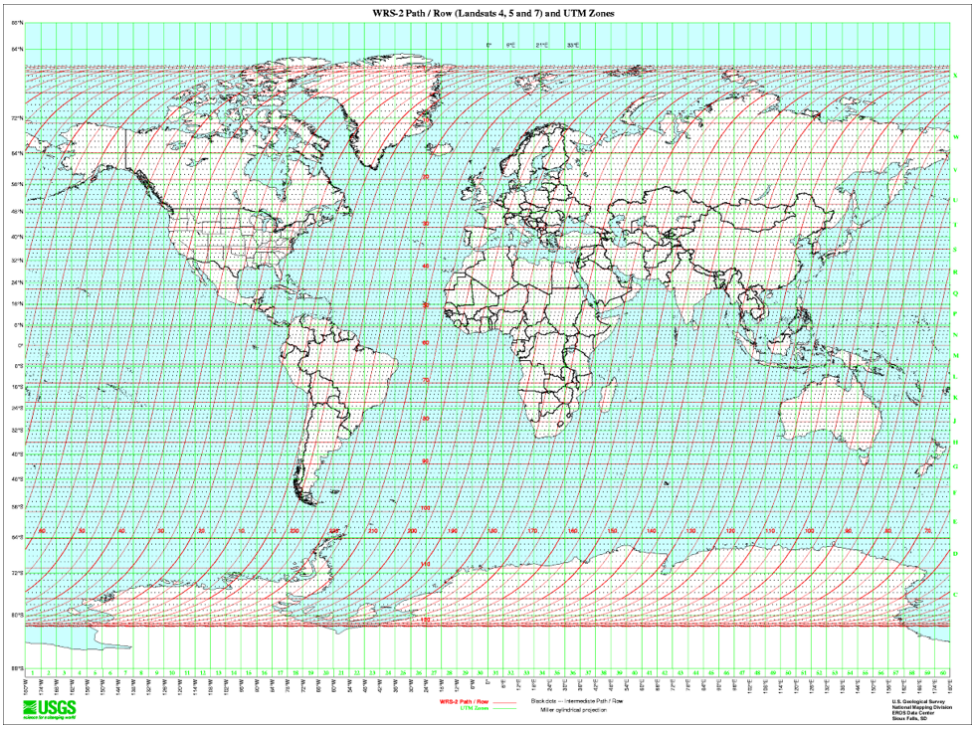


Figure 5.8: *World Reference System-2 (WRS-2) - Day/Descending*

where WRS-2 (see Fig. 5.8) is the global notation system for Landsat data. It enables users to query satellite imagery over any portion of the world by specifying the nominal scene center specified by PATH and ROW numbers.

In our work, we combined the ice fraction with the altimetric water level along the river and tried to find the link between the ice fraction and the results. The basic idea is to determine the PATH and ROW of the VS with known latitude and longtitude, and then find the corresponding Landsat image and ice fractions.

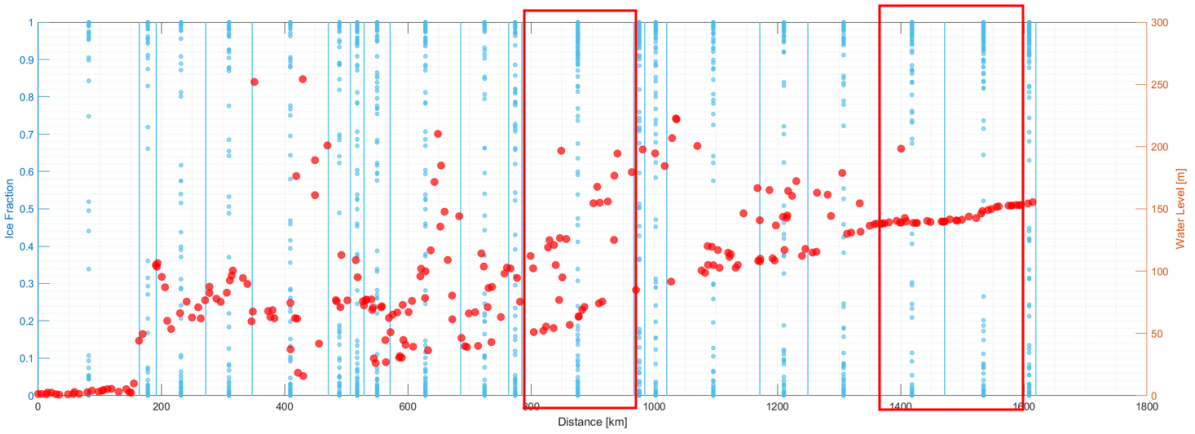


Figure 5.9: *Initial altimetric water level along the river (red dot) and ice fractions in each Landsat image (blue dot: ice Fraction, blue line: Landsat image covered area)*

The entire Makenzie River is segmented by approximately 20 Landsat images, each of which contains the ice fractions from 1984-2018. All available ice fractions in each image are output as background to just show the aggregation of ice fractions without the time involved(see Fig. 5.9). Our expectation is that implausible data generally owns larger ice fractions over the entire time period, which will prove that the altimetry results are indeed strongly correlated with ice conditions. However, that is not what we see. Almost all of these 20 areas have all kinds of ice fractions without any characteristics. For example, the two sets of data in the red boxes both have all kinds of ice fractions and are mostly around 0 or 1. However, one set of data shows normal trends in water levels while the other set is full of anomalies.

It is not easy to draw clear conclusions from the ice fractions. The main problem can be that, on the one hand, only counting the ice fraction in each Landsat image leads to very low resolution or too rough data. On the other hand, in fact, the Landsat images have overlapping areas. How to assign a correct ice fraction to the points in the overlapping area is a problem.

In any case, this is a meaningful attempt, which represents the idea of combining realistic factors with altimetry results. A better result may be achieved with more research in the future.

5.2.4 Validation

Here is a little addition to the validation. In fact, if in situ data is available, the output time series can be validated against in situ data in Altundle+. In situ data and altimetric time series belonging to closest VSs to these in situ stations will be compared. Appropriate statistics, which include pre and post-outlier detection statistics for single TS, as well as pre and post-bias correction statistics for the merged time series, can be derived and help analyze the results of the altimeter. The black cross in Figure 3.5 represents the location of all the about 10 in situ stations. However, less than 10 in situ stations can not support such a scenario with 265 VSs.

Very few VSs that can be done the validation cannot represent the whole. Therefore, we are more of an analysis of the whole water level along the river.

5.3 Conclusions

Inland altimetry based on ocean altimetry faces many challenges. Multiple sources of uncertainty, insufficient measures of uncertainty, and nature of inland water level time series all influence the measurement results, which can be fully verified by the initial results that we achieved in Chapter 3.

Our approach is to use LEIPI with Outlier Identification based on the reference height. On the one hand, more data can be achieved, and outliers can also be identified more accurately. But on the other hand, the reliable result is not always achieved. Analysis throughout chapter 4.3 reveals that LEIPI is not always doing the right thing sometimes, and we get results that are not our targets. Of course, this is not the problem with LEIPI. It is mainly the problem about, for example, the frozen area in this region and multiple water bodies. LEIPI is a fully data-driven algorithm, and there is no way to guarantee the correctness of the LEIPI for every case study. Moreover, sometimes the waveforms are not available or are not usable, and that is why we do not see improvements for Envisat, for instance. One should carefully analyze the achieved results before taking them as reliable.

In the future some possibilities for a better inland altimetry can be:

- using other available missions like CryoSat, ICESat, and Saral/AltiKa on a drifting orbit. The slow drift of the satellite in longitude makes satellites cross rivers more frequently than if they were in repeating orbits and can be used to measure water level.
- studying auxiliary data related to geographic information, such as ice coverage maps. Potentially, it can help analyze the data and the influencing factors behind it.
- Advances in technology. With SAR altimetry missions, with the application of OLTC, or with the launch of Sentinel-6, which allows Fully-Focused SAR analysis, the improvement can be achieved.

Bibliography

Abdullah, N. (2018), The Optimal Retracked Sea Level from SARAL/AltiKa Satellite Altimetry over the Southeast Asia, PhD thesis.

Acker, J., Williams, R., Chiu, L., Ardanuy, P., Miller, S., Schueler, C., Vachon, P. W. and Manore, M. (2003), Remote sensing from satellites, *in* R. A. Meyers, ed., 'Encyclopedia of Physical Science and Technology (Third Edition)', third edition edn, Academic Press, New York, pp. 161–202.
URL: *https://www.sciencedirect.com/science/article/pii/B0122274105009388*

Altenau, E. H., Pavelsky, T. M., Durand, M. T., Yang, X., Frasson, R. P. d. M. and Bendezu, L. (2021), 'The surface water and ocean topography (SWOT) mission river database (SWORD): A global river network for satellite data products', *Water Resources Research* **57**(7), e2021WR030054. e2021WR030054 2021WR030054.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021WR030054*

Biancamaria, S., Frappart, F., Leleu, A.-S., Marieu, V., Blumstein, D., Desjonquères, J.-D., Boy, F., Sottolichio, A. and Valle-Levinson, A. (2016), 'Satellite radar altimetry water elevations performance over a 200 m wide river: Evaluation over the Garonne river', *Advances in Space Research* **59**.

Blarel, F., Frappart, F., Legrésy, B., Blumstein, D., Rémy, F., Fatras, C., Mougin, E., Papa, F., Prigent, C., Niño, F., Borderies, P., Biancamaria, S. and Calmant, S. (2015), Altimetry backscattering signatures at Ku and S bands over land and ice sheets, p. 963727.

Brown, G. (1977), 'The average impulse response of a rough surface and its applications', *IEEE Transactions on Antennas and Propagation* **25**(1), 67–74.

Buckley, S. (2020), 'NASADEM: Creating a new NASA digital elevation model and associated products'.
URL: *https://earthdata.nasa.gov/esds/competitive-programs/measures/nasadem*

Cazenave, A. (2019), Satellite altimetry, *in* J. K. Cochran, H. J. Bokuniewicz and P. L. Yager, eds, 'Encyclopedia of Ocean Sciences (Third Edition)', third edition edn, Academic Press, Oxford, pp. 397–401.
URL: *https://www.sciencedirect.com/science/article/pii/B9780124095489116240*

Davis, C. (1997), 'A robust threshold retracking algorithm for measuring ice-sheet surface elevation change from satellite radar altimeters', *IEEE Transactions on Geoscience and Remote Sensing* **35**(4), 974–979.

DUACS (n.d.), 'Altimetry data for sea level studies and applications'.
URL: *"https://duacs.cls.fr/history-and-context/"*

Dubock, P., Spoto, F., Simpson, J., Spencer, D., Schutte, E. and Sontag, H. (2001), 'The Envisat satellite and its integration', *ESA bulletin* **106**, 26–45.

ESA (2022), 'Saral: Satellite with ARgos and ALtiKa'.
URL: *https://earth.esa.int/web/eoportal/satellite-missions/s/saral*

Frappart, F., Calmant, S., Cauhopé, M., Seyler, F. and Cazenave, A. (2006), 'Preliminary results of ENVISAT RA-2-derived water levels validation over the Amazon basin', *Remote Sensing of Environment* **100**(2), 252–264.
URL: *https://www.sciencedirect.com/science/article/pii/S0034425705003585*

Fu, L.-L. and Cazenave, A. (2001), Preface, *in* L.-L. Fu and A. Cazenave, eds, 'Satellite Altimetry and Earth Sciences', Vol. 69 of *International Geophysics*, Academic Press, pp. xi–xii.
URL: *https://www.sciencedirect.com/science/article/pii/S0074614201801455*

Gandhi, S. and Sarkar, B. (2016), Chapter 3 - reconnaissance and prospecting, *in* S. Gandhi and B. Sarkar, eds, 'Essentials of Mineral Exploration and Evaluation', Elsevier, pp. 53–79.
URL: *https://www.sciencedirect.com/science/article/pii/B9780128053294000107*

Gao, Q., Makhoul varona, E., Escorihuela, M. J., Zribi, M., Quintana Seguí, P., García, P. and Roca, M. (2019), 'Analysis of retrackers? performances and water level retrieval over the ebro river basin using sentinel-3', *Remote Sensing* **11**, 718.

Gommenginger, C., Thibaut, P., Fenoglio, L., Quartly, G., Deng, X., Gomez-Enri, J., Challenor, P. and Gao, Y. (2011), *Coastal Altimetry*, pp. 61–101.

Grgić, M. and Bašić, T. (2021), Radar satellite altimetry in geodesy - theory, applications and recent developments, *in* B. Erol and S. Erol, eds, 'Geodetic Sciences', IntechOpen, Rijeka, chapter 5.
URL: *https://doi.org/10.5772/intechopen.97349*

Guo, J., Chang, X., Gao, Y., Sun, J. and Hwang, C. (2009), 'Lake level variations monitored with satellite altimetry waveform retracking', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2**(2), 80–86.

Hayne, G. (1980), 'Radar altimeter mean return waveforms from near-normal-incidence ocean surface scattering', *IEEE Transactions on Antennas and Propagation* **28**(5), 687–692.

Kramer, H. (2019), *Observation of the Earth and Its Environment: Survey of Missions and Sensors*, Springer Berlin Heidelberg.
URL: *https://books.google.de/books?id=T2W-DwAAQBAJ*

Krebs, G. D. (2021), 'Jason 2, 3 (OSTM)'.
URL: *"https://space.skyrocket.de/doc _sdat/jason-2.htm"*

Legresy, B., Papa, F., Remy, F., Vinay, G., Bosch, M. and Zanife, O.-Z. (2005), 'Envisat radar altimeter measurements over continental surfaces and ice caps using the ice-2 retracking algorithm', *Remote Sensing of Environment* **95**, 155–163.

McColl, R. (2014), *Encyclopedia of World Geography*, number Bd. 1 *in* 'Facts on File Library of World Geography', Facts On File, Incorporated.
URL: *https://books.google.de/books?id=DJgnebGbAB8C*

Millot, R., Gaillardet, J., Dupré, B. and Allègre, C. J. (2003), 'Northern latitude chemical weathering rates: clues from the Mackenzie River Basin, Canada', *Geochimica et Cosmochimica Acta* **67**(7), 1305–1329.
URL: *https://www.sciencedirect.com/science/article/pii/S0016703702012073*

NASA and Teams, M. S. (2019), 'NASA and METI Release ASTER Global DEM Version 3 by NASA and METI Science Teams'.
URL: *https://lpdaac.usgs.gov/news/nasa-and-meti-release-aster-global-dem-version-3/*

Rosmorduc.V, Benveniste.J, Dinardo.S, Lauret.O, Maheu.C, Milagro.M, Picot.N, Escola.R, Garcia-Mondejar.A, Schrama.E, Ambrozio.A, Terra-Homem.M and Restano.M (2018), 'Radar altimetry tutorial'.
URL: *http://www.altimetry.info/radar-altimetry-tutorial/how-altimetry-works/basic-principle/*

STAR (2020), 'Altimetric bathymetry'.
URL: *"https://www.star.nesdis.noaa.gov/socd/lsa/AltBathy/"*

Tourian, M. J. (2013), *Application of spaceborne geodetic sensors for hydrology*.

Wingham, D. and Rapley, C. (1986), New techniques in satellite altimeter tracking systems.

Yang, D., Shi, X. and Marsh, P. (2015), 'Variability and extreme of mackenzie river daily discharge during 1973-2011', *Quaternary International* **380-381**, 159–168. Larger Asian Rivers 8: Impacts from human activities and climate change.
URL: *https://www.sciencedirect.com/science/article/pii/S1040618214006715*

Yang, X., Pavelsky, T. M. and Allen, G. H. (2019), 'Global River ICE Dataset - validation dataset'.
URL: *https://doi.org/10.5281/zenodo.3372754*

Yang, X., Pavelsky, T. M. and Allen, G. H. (2020), 'The past and future of global river ice', *Nature* **577**, 69–73.