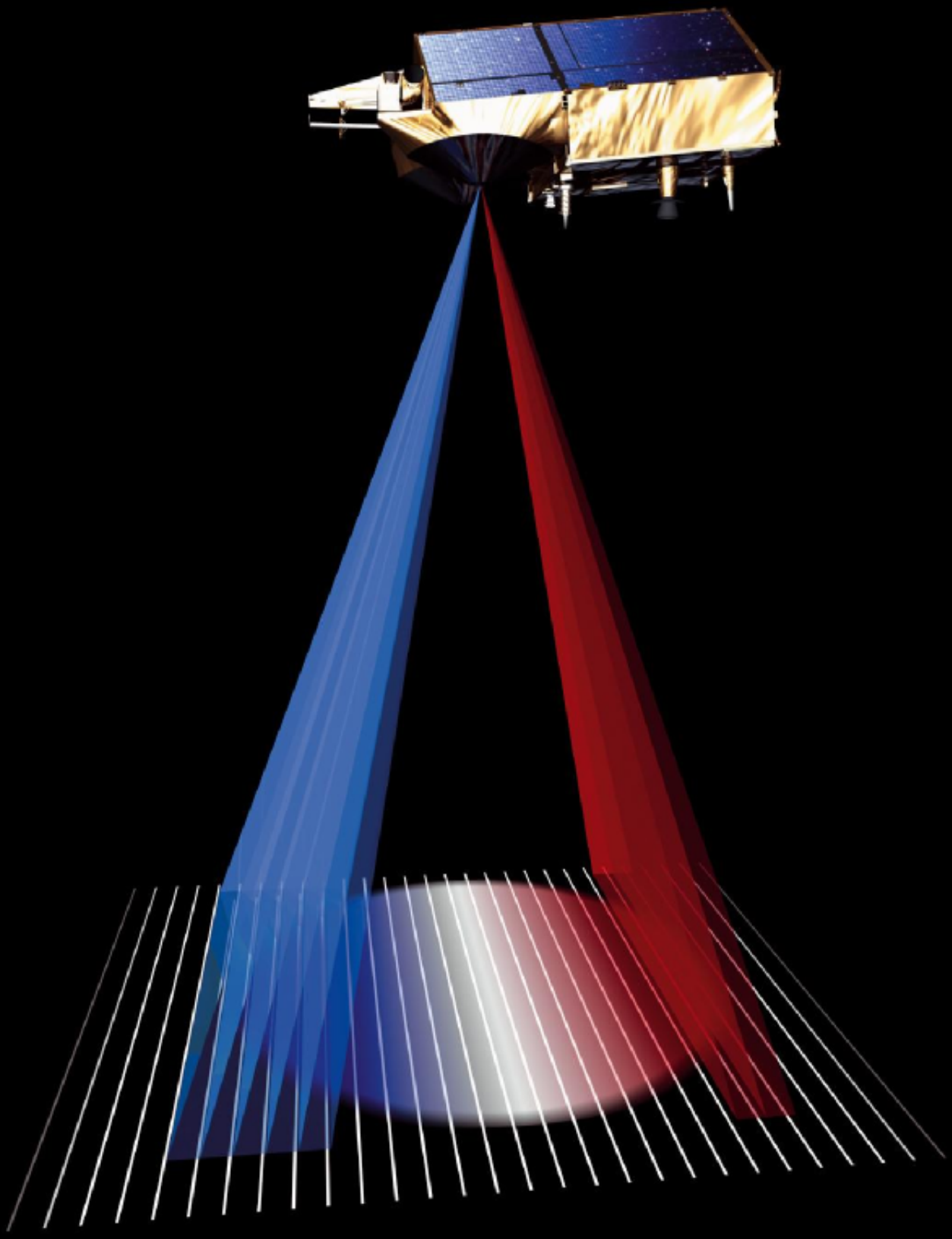
Implementation and application of a numerical Cryosat retracker to retrieve lake level estimates on the Tibetan plateau



Delft University of Technology

Geoscience and Remote Sensing

Optical and Laser Remote Sensing

Implementation and application of a numerical Cryosat retracker to retrieve lake level estimates on the Tibetan plateau

Author:
Marcel Kleinherenbrink

Supervisors:
Dr. Roderik Lindenbergh
Dr. Pavel Ditmar
-
Graduation professor:
Prof. Dr. Massimo Menenti

September 26, 2013

Abstract

The Tibetan plateau holds the largest storage of fresh water outside of the polar regions. More then one billion people in South and East Asia depend on rivers originating from the Tibetan plateau. Therefore it is important to understand the hydrology and monitor the state of hydrology in Tibet.

One of the important indicators for the state of the hydrology in Tibet are lake levels. There are over 900 lakes on the Tibetan plateau larger than 1 km^2. These lakes contain larger amounts of water. A decrease or increase of water in these lakes can be determined from lakes level changes.

Until the beginning of the 1990s the way the measure lake levels was with gauges. However, due to the harsh environment and the remoteness of Tibet it is difficult to install gauges at the majority of the lakes. Satellite altimetry is used to get a more complete image of lake level changes in Tibet. However, due to the large footprints of radar altimeters only a few lakes can be monitored. Laser altimeters have a smaller footprint, but only one was flown between 2003-2009.

Since April 2010 Cryosat orbits the Earth using a new type of altimeter: Synthetic aperture Radar Interferometric ALtimeter (SIRAL). It is able to operate in three different modes. Over Tibet it uses Synthetic Aperture Radar (SAR) to enhance its along-track resolution and two antennas to compensate for cross-track slope. Due to the SAR technique Cryosat is able to measure more lakes than a conventional altimeter. Due to its 369 day repeat orbit, it has a dense cross-track spacing of approximately 6.5 km over Tibet. Therefore, it passes over many lakes at the cost of temporal resolution.

To compute geophysical parameters from Cryosat measurements, a theoretical waveform is fitted into a received waveform, a processed called waveform retracking. However, these theoretical waveforms are based on open ocean responses. Over Tibetan lakes the waveforms are polluted by surrounding topography, and therefore the shape changes. In various cases more peaks are present in the waveforms near the shores of lake, causing the retracker to find invalid measurements. As a consequence the measurements are less accurate and less reliable.

In this thesis a new retracking procedure is applied, which is specifically designed to estimate lake levels. This retracker uses a simulated Cryosat waveform from a lake

surface to fit it in the received waveforms. It furthermore applies a procedure called multiple peak retracking, in order to mitigate the number of invalid measurements.

For lake level estimation two different cases are distinguished: open water and near shore measurements. Over open water a simulated waveform is correlated with the received waveforms. The estimated elevations over a track have a precision of approximately 0.07 m, which is better than the offical elevation products of Cryosat. A discrepancy of approximately 0.45 m was found with the official Cryosat data over open water. Near shore a weighting procedure is implemented to increase the weight of the important part of the waveform. This makes the retracker more robust, at the cost of accuracy. The estimated elevations over a polluted track have a precision of approximately 0.24 m, which is again better than the official elevation products of Cryosat. The discrepancy is 0.57 m with the official Cryosat data over open water.

Preface

In this thesis a first step is made into using Cryosat for lake level measurements. The results showed that the new technologies onboard Cryosat have much potential in the field of lake level monitoring. Because the new type of altimeter onboard Cryosat is also used in upcoming satellite mission like Jason-CS and Sentinel-3. Therefore I expect that in the near future a more complete image of lake level changes in Tibet will emerge. This can strongly contribute to the understanding of the hydrology of Tibet.

My MSc Space Applications at the faculty of Aerospace Engineering at the TUDelft comes to an end with this thesis. The research was performed over the course of more than a year at the section of Optical and Laser Remote Sensing, which is part of the Geoscience and Remote Sensing group at the faculty of Civil Engineering. During this research I primarily learned that if something is hard and nothing seems to work, it is better to work harder. If things are going well again, take your time and enjoy the results.

Ultimately, I want to express my gratitude towards everyone who contributed to this research. First of all, I would like to thank my supervisors Roderik Lindenbergh and Pavel Ditmar for their supervision. Not only did they check my thesis, but they also gave me useful insight and feedback to improve the results. Roderik primarily gave more insight into the hydrology and Pavel helped primarily with signal processing.

Furthermore, I would like to professor Massim Menenti for his role as graduation supervisor, but also for the opportunity for visiting China for my internship.

I would like to thank Marc Naeije and Riccardo Riva for joining the graduation committee as co-readers.

And finally, I would like to thank my roommates for the nice times during writing this thesis. I enjoyed the fruitful discussions about a large range of topics.

Contents

Abbreviations

ACU	Along-track Cross-track Up (system)
DFT	Discrete Fourier Transform
ESA	European Space Agency
ECEF	Earth Centered Earth Fixed (system)
GLAS	Geoscience Laser Altimeter System
FFT	Fast Fourier Transform
LEO	Low Earth Orbit
LPF	Low-Pass Filter
LRM	Low Resolution Mode
MODIS	MOderate-Resolution Imagining Spectrometer
NEU	North East Up (system)
POCA	Point Of Closest Approach
PLF	Pulse Limited Footprint
PRF	Pulse Repetition Frequency
RMS	Root Mean Square
SARInM	Synthetic Aperture Radar Interferometric Mode
SARM	Synthetic Aperture Radar Mode
SIRAL	Synthetic aperture, Interferometric Radar ALtimeter
SRC	Slant Range Correction

List of symbols

a	Pulse number in burst
A_m	Amplitude of the mixed signal
A_{st}	Amplitude of signal transmitted
A_{sr}	Amplitude of signal received
B	Bandwidth
c	Speed of light
C_θ	Circumference of the Earth at a particular latitude
$d_{a,res}$	Along-track resolution SARM
d_g	Horizontal distance between the target ground location and the position where the beam intersects the surface
D	Doppler effect
e	Euler number
E	Evaporation
f	Frequency
f_0	Maximum frequency
f_c	Center frequency
f_m	Frequency of the mixed signal
$f_{m,\text{Doppler}}$	Effect of Doppler on the frequency of the mixed signal
F_r	Frequency of the deramping chirp
F_s	Frequency of the received signal
G	Antenna gain
G_0	Peak antenna gain
GZ	Water storage in ground water
h	Elevation
i	Imaginary number
I	Intensity
$\bar{I}$	Average intensity
K	Weighting function
K_{ml}	Multilooked coherence
L_{km}	Length of the lake in kilometer
m_s	Sampled mixed signal
m_{imag}	Quadrature phase mixed signal

m_{real}	In-phase phase mixed signal
$m_{imag,LPF}$	Low-pass filtered quadrature phase mixed signal
$m_{real,LPF}$	Low-pass filtered in-phase mixed signal
$M_{s,bs}$	Beamsteered sampled mixed signal
M_{strip}	Signal per strip
N_L	Number of looks/number of waveforms in the stack
n	Sample number in echoe
n_s	Direction of scatterer
n_v	Normalized velocity vector
p_t	Probability density function
p_i	Power in unfiltered waveform
P_n	Power in filtered waveform
noc	Number of cycles
noc_m	Number of cycles of the mixed signal
P	Precipitation
Q	Frequency slope
Q_r	Runoff
r_{imag}	Quadrature phase deramping chirp
r_{real}	In-phase deramping chirp
R	Range to scatterer
R_{res}	Range resolution
RMS	Root Mean Square
SM	Soil moisture mass
SP	Snow pack and glacier mass
s_r	Signal received
s_t	Signal transmitted
t	Time
t_r	Time of the deramping chirp
t_{sr}	Slant range corrected time
t_{src}	Time slant range correction
t_s	Time of the received signal
t_{res}	Time resolution
t_{2w}	Two-way travel time
V	Satellite velocity
w	Single look waveform
W	Water storage in water bodies
W_{km}	Width of a lake in kilometers
w_{LRM}	Waveform in low resolution mode
w_n^-	Waveform n in the stack received at the right antenna
w_n^+	Waveform n in the stack received at the left antenna
w_s	Simulated waveform
w_s	Waveform averaged over a ground-track
w_t	Theoretical waveform

$W\bar{W}_{ml}^{(+,-)}$	Multilooked waveform
x_b	x-position of the burst in the ACU system
x_p	x-position of start-stop location in the ACU system
x_s	x-position of the scatterer in the ACU system
x_0	z-position of the orbit above a specific ground location in the ACU system
y_b	y-position of the burst in the ACU system
y_p	y-position of start-stop location in the ACU system
y_s	y-position of the scatterer in the ACU system
y_0	z-position of the orbit above a specific ground location in the ACU system
z_b	z-position of the burst in the ACU system
z_p	z-position of start-stop location in the ACU system
z_s	z-position of the scatterer in the ACU system
z_0	z-position of the orbit above a specific ground location in the ACU system
γ_1	Antenna elliptical constant along-track direction
γ_2	Antenna elliptical constant cross-track direction
$\Delta noc_{a,a}$	Difference in number cycles between between two pulse positions
Δt_a	Difference in time of arrival between the signal received at the center of the burst and at pulse position a
ΔR_a	Difference in range between the signal received at the center of the burst and at pulse position a
$\Delta R_{steering}$	Difference in range between the target location and the location where the beam intersects the surface
λ_c	Wavelength central frequency
λ_b	Latitude of burst
λ_g	Latitude of ground location
λ_s	Latitude of scatterer
ϕ_V	Flight angle
Φ_m	Phase of the mixed signal
Φ_{ml}	Multilooked Phase
ρ_w	Density of water
ρ_t	Transmitted pulse length in meters
τ	Sampled pulse length
τ_t	Transmitted pulse length
θ	Polar coordinate with respect to the antenna boresight
θ_b	Longitude of burst
θ_g	Longitude of ground location
θ_s	Longitude of scatterer
Θ_{Doppler}	Phase multiplier for Doppler correction
Θ_{SRC}	Phase multiplier for slant range correction
Θ_{steering}	Phase multiplier for beam steering

v	Azimuth coordinate with respect to the antenna boresight	

List of Figures

List of Tables

Chapter 1

Introduction

This chapter describes the importance of understanding Tibetan hydrology and the role of lake level measurements. Additionally, it describes why and how data from the Cryosat satellite can be used to measure lake levels. Furthermore, it outlines the objectives and research questions of this study.

In section 1.1 a global introduction into the hydrology of Tibet is given. Section 1.2 describes the application of satellite altimetry to lake level measurements. In section 1.3 suggests an alternative approach to find lake level estimates. Section 1.4 discusses the objectives and research questions. Finally, in section 1.5 the outline of the thesis and the methodology is described.

1.1 Tibetan lake levels

Lakes are an important source of fresh water on the Tibetan plateau. Their levels are related to changes in local variables as glacier runoff, precipitation and evaporation. This section will discuss the role of lake levels within hydrological research.

1.1.1 Importance of understanding Tibetan hydrology

The Tibetan plateau holds the largest storage of fresh water outside the polar regions [4] and is therefore called the Water Tower of Asia [5]. Except for thousands of glaciers, there are 900 lakes with a size over 1 km^2. These glaciers and lakes are connected to river systems like the Indus, Mekong, Yangtze, Yellow River and the Brahmaputra. More than one billion people depend on the water from these rivers for drinking, farming, etc. [4][6][7].

Due to increasing population and climate change, understanding the hydrology of the Tibetan plateau has become increasingly important. Research shows that the total water

storage on the plateau is decreasing [5]. Eventually, this could lead to an increased frequency of droughts, desertification [5] and possibly a shortage of water for metropoles like Islamabad, Bangkok, Shanghai and Beijing.

1.1.2 Role of lake levels

To show how lake levels are related to other hydrological processes a model is introducted based on the Hydrologiska Byrans Vattenbalansavdelning (HBV) model [8]. The model equates the precipitation P, the evaporation E, the runoff Q_r to the change in mass of the snow pack (and glaciers) SP, soil moisture SM, the groundwater zone GZ and water bodies W. The water balance in an area, i.e. the change in the total storage of water, is given by [8]:

$$P - E - Q_r = \frac{d}{dt}(SP + SM + GZ + W). \tag{1.1}$$

A water bodies is for example a lake of a river. It is not possible to directly measure the water storage change of an individual water body from satellite observations. Therefore, the water storage change $\frac{dW}{dt}$ is determined by relating it to the lake level change $\frac{dh}{dt}$, if it is assumed that the area A is constant, then the water storage change is related as:

$$\frac{dW}{dt} = \rho_w \frac{dh}{dt} A, \tag{1.2}$$

with ρ_w the density of water.

As the lake levels are related to the other variables in Eq. (1.1), they do not only help to estimate the mass balance, but they are also indicators for a change in the other variables. If a lake level rises over a period of time, it possibly indicates that the glaciers in the catchment of the lake are melting or that precipitation is increasing, which could be an effect of climate change.

1.2 Lake level measurements

In the previous section it is shown how lake levels are related to various hydrological variables and how they contribute to the overall water mass balance. This section will show how lake levels are measured and how satellite altimetry contributes to these measurements. In the final part of this section the application of Cryosat to lake level measurements is discussed.

1.2.1 Rationale for satellite altimetry measurements

At the Tibetan plateau the conditions are often harsh due to weather and elevation [9][10]. Many of the lakes are at altitudes higher than 4000 m, which makes them hard to reach. Therefore gauges are only installed at a few large lakes, like Nam Co and Qinghai [4].

An alternative solution for measuring lake levels is satellite altimetry. Altimeters estimate the elevation by measuring the two-way travel time of a transmitted pulse. Since the 1990s, inland water levels (lake levels and river levels) have been studied using satellite altimetry [11][12][6]. Satellite altimetry has proven to be a reliable solution to monitor lake levels, with accuracies ranging from centimeters up to several decimeters [4][6][13]. Two types of satellite altimetry measurements are available: laser altimetry and radar altimetry.

1.2.2 Laser Altimetry

IceSat is the only satellite with a laser altimeter onboard that has been flown around Earth so far. The Geoscience Laser Altimeter System (GLAS) onboard IceSat had centimeter to decimeter level measurement accuracy and a footprint of approximately 70 m [6]. The footprint spacing along-track was approximately 170 m. Due to its footprint size and along-track spacing it is suitable to measure many lakes on the Tibetan plateau.

A drawback is the limited temporal resolution, since IceSat was only operational three times per year for a period of 33 days. As a consequence, it is not possible to determine accurately seasonal variations in lake level. Another drawback is the limited density of ground-tracks. Nevertheless, 154 lake levels could be monitored in its operational period between 2003-2009 [6].

1.2.3 Radar Altimetry

The first studies of lake levels measurements with a radar altimeter were performed in 1993 with GeoSat [12]. Since then plenty of radar altimeter missions have been flown, like TOPEX/Poseidon, ERS-1, ERS-2, Envisat, Jason-1 and Jason-2. Compared to the laser altimeter of IceSat a much longer timeline of radar altimetry measurements is available. The accuracy of radar altimeters is at the same order as that of the GLAS instrument. The sampling interval is approximately 350 m in along-track direction.

Radar altimeters have however the disadvantage of a large footprint size, which ranges from 2 km for a flat surface up to 19 km for terrains with substantial topography. A theoretical waveform shape is fitted in the received waveform to estimate the elevation,

a procedure called waveform retracking. The received waveforms are in many cases polluted by the surrounding topography, due to the large footprint [14]. As a consequence, the fit is not optimal and the measurement accuracy deteriorates. The accuracies over smaller lakes were estimated to decrease from around 10 cm to 30-50 cm [4][15]. In some cases no lake level could be retrieved at all. Measurements over lakes with a size of at least 7 km in diameter are considered to be reliable with an altimeter like the RA-2 onboard Envisat [6].

Figure 1.1 demonstrates the problem of the fitting procedure. It becomes clear that fitting the theoretical shape into the received waveform over open water is likely to give accurate elevations. However, fitting the theoretical waveform into the topography polluted waveform decreases either the accuracy of the estimated waveform parameters or it finds the wrong peak in the waveform and a wrong elevation.

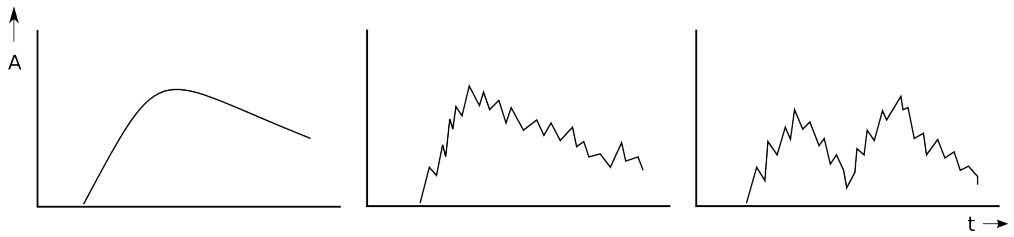


Figure 1.1: Left: theoretical representation of a radar altimeter waveform. Middle: received waveform over open water, or a large lake. Right: complex waveform received over a small lake surrounded by topography. Fitting the theoretical waveform into the received complex one is problematic and gives inaccurate elevations.

The quality of measurement is sometimes improved by using another type of retracker. For Envisat four different retracking algorithms are applied in the standard processing proceduring [16]. Other studies tried to increase the accuracy of land observation by first analysing the waveform shape followed by selecting one of eleven different retracking algorithms [14].

1.2.4 Cryosat

In April 2010 Cryosat was launched with a Synthetic aperture Interterferometric Radar ALtimeter (SIRAL) onboard. SIRAL is able to operate in three differen modes: Low-Resolution Mode (LRM), Synthetic Aperture Radar Mode (SARM) and Synthetic Aperture Radar Interferometric Mode (SARInM). It uses Synthetic Aperture Radar (SAR) in SARM and SARInM to improve the along-track resolution to approximately 300 m [1]. For every strip of 300 m along-track and approximately 15-20 km wide it produces a waveform. Therefore the waveform shape is different from conventional radar altimetry. In SARInM SIRAL uses a second receiving antenna to estimate the cross-track slope.

In a project called SAMOSA, researchers design an analytical retracker (SAMOSA-3) for SAR altimeters. In an analytical retracker, the description of the theoretical waveform is analytic, not the fitting procedure. Up to date no perfect analytical retracker is produced yet, therefore the current products are still processed using simplified (SAMOSA-1) or numerical retrackers (SAMOSA-2). The estimated parameters, like elevation, are therefore not optimal. Unfortunately, a theoretical description of the SAMOSA retrackers is not publicly available.

Once the SAMOSA-3 retracker is implemented, the elevation retrieval is expected to be more accurate than that from conventional radar altimeters [17]. Due to improved along-track resolution lake level waveforms will be less polluted by surrounding topography, especially in the case of a small lake size. Still the SAMOSA-3 retracker is designed for flat surface and ocean-like waveforms, so the computed lake elevations will not be optimal.

1.3 Alternative retracking

Based upon the problems with retracking radar altimeter waveforms discussed in the preceding paragraphs, a new approach is suggested to find lake level estimated on the Tibetan plateau. It involves the design of a numerical retracker tailored to retrack Cryosat waveforms over lakes in Tibet. The term numerical implies that the shape of the retracker waveform is constructed by simulating a Cryosat waveform from scatterers with a predefined spacing on the ground. Only the signal from the water surface will be simulated. By correlating the simulated waveform with the received Cryosat waveforms a lake levels elevation is computed.

By applying this approach two main problems should be mitigated. The first problem is that an analytical retracker adjusts the waveform parameters to the shape of the received waveform, while it is optimized for open water. As a consequence, estimated waveform parameters are affected by the signal of surrounding topography in the waveform. If an approximate state of the water surface is set in the numerical retracker the surrounding topography will have less effect on the estimated lake level. The second problem is that an analytical retracker in various cases find the wrong peak in the waveform. The estimated lake level is therefore actually not based on the signal from the lake but from another signal of the surrounding topography. Both peaks (a signal from the lake and from the surrounding topography) are however often visible in the waveform. With the numerical approach we will aim to find an increased number of valid lake level estimates.

1.4 Objectives and research questions

In order to maximize the number of lakes for which the water level is estimated accurately, a retracker is to be designed based upon numerical modelling a waveform from

the lake surface. The accuracy we are aiming for is at least of decimeter level. The retracker to be designed will take into account shape, size, orientation and position of a lake with respect to the satellite track.

This can be summarized in the following research objective:

Design a numerical retracker optimized for waveforms over Tibetan lakes.

With the main research question:

How to exploit the measurement principle of Cryosat in the most efficient way, so that a suitable numerical retracker is designed for lake level measurements over the Tibetan plateau?

To answer the main research question, first several subquestions are answered:

What are the measurement principles of Cryosat?
In this step a research is performed to find out how Cryosat operates. First, the transmitted and received signal are discussed. Then an investigation is done into the satellite processing procedures. After that the procedures of the ground processor are investigated. The ground processor should eventually compute a waveform from the satellite measurements and retrieve geosphysical parameters, like elevation.

How accurately can one model a Cryosat waveform?
To answer this question, the theory discussed in the previous step is applied. For several scenarios waveforms are computed and compared with received Cryosat waveforms. Furthermore, the question is answered how the waveform shape is affected by topography, water wave height and other effects.

How can a received waveform be retracked with a modelled waveform?
Once an accurate waveform is modelled it is used to retrack waveforms over lakes in several scenarios. At first a waveform is retracked in the middle of a large lake. After that the retracking algorithm is used to retrack waveforms at the edges of a lake, where the waveform is polluted by surrounding topography.

How accurately can lake levels be retrieved with a numerical retracker?
In this step the precision of the elevations retrieved from numerical waveform retracking is computed over several tracks. The results are compared to Cryosat level 2 data to verify the results. Furthermore, a lake level trend and seasonal effects will be computed over a lake and compared to trends and seasonal effects found in other studies. The final aim is to improve the retracking performance in areas where the waveform is polluted by surrounding topography.

1.5 Research Methodology and report outline

This thesis discusses the theory behind Cryosat measurements and processing particularly in SARInM over Tibet. It is shown how a Cryosat waveform can be simulated from a set of scatterers. And it shows how water levels are retrieved from waveforms over a lake. Finally, a comparison is made between the developed numerical retracker and Cryosat level 2 data.

Chapter 2 provides an overview of the Cryosat mission and its measurements. It gives analytical descriptions of the received echoes and the measurement data downlinked by the Cryosat. Furthermore, it provides the general theory of the processors that convert measurement data into waveforms and it shows how retracking is applied.

Chapter 3 gives a description of the developed algorithm to simulate Cryosat waveforms. Additionally, the shape of waveforms for several scenarios is discussed, which are computed from simulated water surfaces.

Chapter 4 describes how an numerical retracker, designed in this study is used to compute lake level estimations. It also shows the problems with a analytical retracker. In this chapter the error budget is addressed as well.

Chapter 5 shows the results of the developed retracking method, when applied to Nam Co. The results are verified with Cryosat level 2 elevation products to see if they give approximately the same elevation. Lake level trends are computed to see if they are comparable to other studies.

Chapter 6 provides the conclusions and answers to the research questions. Ultimately, recommendations are given for further research.

Chapter 2

The Cryosat Mission

This chapter gives an overview of the Cryosat mission and describes in detail the Cryosat measurement principles and data processing stages. It focusses on answering the subquestion: **What are the measurement principles of Cryosat?** By answering the subquestion the necessary background on Cryosats measuring principles are provided to simulate a Cryosat waveform, which is shown in chapter 3.

In section 2.1 Cryosat is introduced and the mission objectives and mission geometry are decribed. The rest of chapter is dedicated to the measurements principles and processing stages of Cryosat. To get a proper overview, figure 2.1 shows the measurements and processing stages describes in this chapter. Section 2.2 describes the signals transmitted by SIRAL and the received signals. Furthermore, it will be shown what measurement data Cryosat transmits to the ground stations. The data products of SIRAL and its processing stages are discussed in section 2.3. It will show how a waveform is created from the measurement data and how an elevation is computed using waveform retracking.

2.1 Mission Description

This section gives a brief overview of the Cryosat mission. The mission objectives are discussed in paragraph 2.1.1 and paragraph 2.1.2 describes the mission geometry.

2.1.1 Mission Objectives

Nowadays the most complete information about the cryosphere is delivered by Low Earth Orbit (LEO) satellites [1]. Examples of such LEO satellites are the gravimetry missions like GRACE, which measure the mass balance, and radar missions, which are able to measure glacier motion using interferometry. Altimetry missions are commonly used to

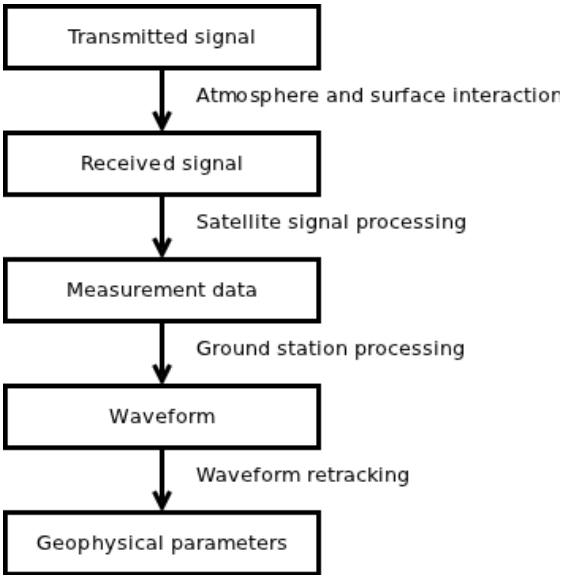


Figure 2.1: Overview of signals, measurements and processing stages discussed in this chapter.

measure elevation change of ice sheets in Antarctica and Greenland and are also used to provide an estimate of ice shelf (sea ice) thickness.

However, most altimeters are flown on missions that have an inclination around 81.5 degrees, which means that major ice masses like the Ross Ice Shelf are not covered [1]. Furthermore, due to their large footprints radar altimeters are only able to measure 5% of the largest flows in sea ice [1].

NASA solved these deficiencies by launching ICESat in 2002, with an inclination of 94 degrees, with a LIDAR onboard named GLAS, like described in chapter 1. ESA however, adapted the ideas described by [17] and [18] to modify the existing radar altimeters to improve its resolution [1].

In 1998 a group of researchers submitted a proposal to ESA for a mission that had the primary objective to measure sea-ice cover and ice sheets [19]. Out of 27 proposals Cryosat with its new type of radar altimeter was chosen. This radar altimeter was called the Synthetic Aperture, Interferometric Radar Altimeter (SIRAL), which operates in three different modes [19].

The choice of this type of instrument was made to meet scientific objectives which could not be met by any preceding satellite mission. The Cryosat mission has two primary goals [20]:

- Recovering regional and basin-scale trends in Arctic sea ice thickness and total mass.
- Estimating the contribution of shrinking of Antarctic and Greenland ice sheets to the

global sea level.

And the mission has two secondary goals [20]:

- Observation of the seasonal and interannual variability of Arctic and Antarctic sea ice thickness and mass.
- Observation of variations in thickness of ice caps and glaciers.

The launch of Cryosat was scheduled for October 2005, but the launcher failed on October 8th. Fortunately, ESA and governments realized the importance of an ice mission like Cryosat and soon after the failure Cryosat-2 was announced. The launch of Cryosat-2 took place on 8 April 2010 and on February 2011 the fist data was released to the public.

2.1.2 Mission Geometry

Since the Cryosat satellite mission must primarily make observations of the cryosphere, the polar areas have to be covered. Therefore the satellite is launched into a high inclination orbit. To meet the scientific objectives described in section 2.1.1 it is of particular importance that the dynamic glaciers at the edges of Antarctica and Greenland are measured on a regular basis. This is achieved by designing the orbit such that it has many crossovers at latitudes above 60 degrees where most glaciers are located [21], to yield a high temporal resolution. A purely polar orbit would cover the entire cryosphere, but only a few crossovers would be present. The design of the orbit is therefore mainly an optimization in the presence of two competing requirements: coverage and crossovers.

Orbit type	Near-circular, polar, LEO
Mean altitude	717.2 km
Repeat period	369 days
Inter-track spacing	7.5 km (equator)
Orbit control	5 km

Table 2.1: Cryosat orbit parameters

As a third requirement the ground-track spacing at the equator is considered [21]. It is required to have a dense enough spacing at the equator (a pseudo-continuous coverage) to meet the second secondary goal described in 2.1.1. Therefore a dense coverage over areas like Tibet is achieved. Ultimately, these requirements lead to the orbit parameters described in table 2.1 and Kepler elements described in table 2.2 [1][21]:

Semi-major axis	7095 km
Eccentricity	0.001407
Inclination	92.00068°
Right Ascension of Ascending Node	129.9971°
Argument of Perigee	115.6195°
True anomaly	283.8995°

Table 2.2: Initial Kepler elements of the Cryosat orbit

2.2 SIRAL

The SIRAL instrument is able to function in three different modes: Low Resolution Mode (LRM), Synthetic Aperture Radar Mode (SARM) and Synthetic Aperture Interferometric Mode (SARInM)[1]. The LRM is primarily activated above open waters and areas with minimal topography, like inner Antartica and Greenland. SIRAL operates in SARM above sea-ice. Due to the improved along-track resolution in this mode it is possible to detect more sea-ice flows. SARInM is applied over areas with substantial topography, like Tibet, the Alps and the coastal zones of Greenland and Antarctica. In SARInM Cryosat has an along-track resolution equal to that in SARM, but in SARInM the echoes are received at two antennas to determine the approximate cross-track slope[1]. A mode map is shown in figure 2.2.

In this section a detailed description of the SIRAL signal properties, signal processing performed onboard the satellite and the measurement principles of all modes is given. Paragraph 2.2.1 will mathematically describe the transmitted signals and received echoes of the SIRAL instrument. The transmitted pulse of SIRAL is frequency modulated, i.e. its frequency decreases over time, to increase the range resolution.

In paragraph 2.2.2 it will be shown how frequency modulation is exploited to increase the range resolution. For the exploitation it is necessary to mix the received signal with a replica and then apply a Low-Pass Filter (LPF). The mixing procedure and the LPF are also known as signal deramping, because after mixing the signal has a constant frequency.

In paragraphs 2.2.3,2.2.4 and 2.2.5 the mode specific operations onboard Cryosat are discussed. Not only the signal processing onboard the satellite for each mode is presented, but also the timing of pulses and bursts. Finally for all three modes a description is given of the measurement data that is telemetered to the ground station.

2.2.1 Signal Properties

This section describes the properties of the pulse transmitted by Cryosat. First it is shown why frequency modulation is necessary and how the transmitted signal can be

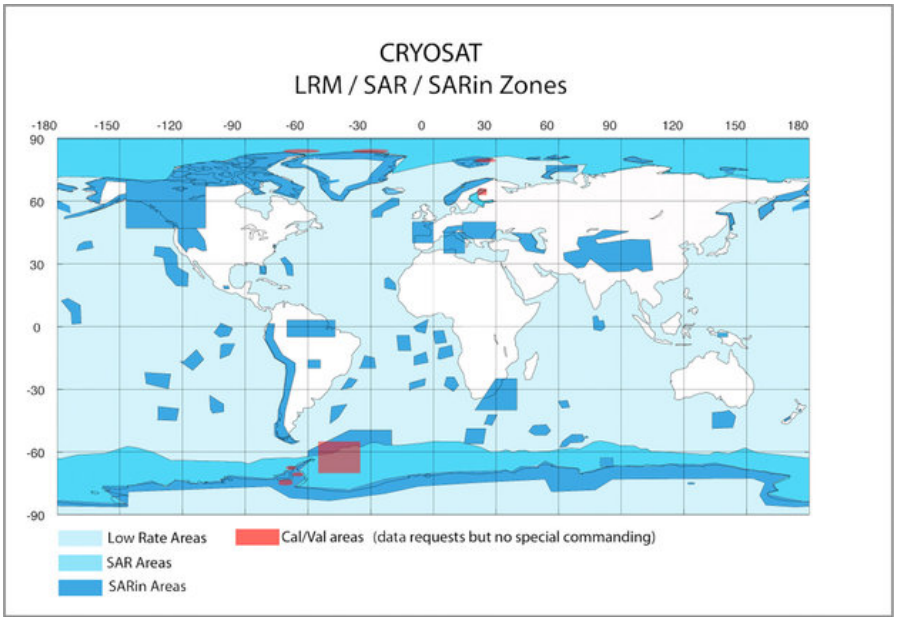


Figure 2.2: Initial mode map for Cryosat operating in LRM, SARM and SARInM.

described mathematically. Then an expression for the received signal is derived in terms of frequency, time of receival, frequency slope and amplitude.

In principle an altimeter measures the two-way travel time of a pulse transmitted towards the surface. The range R between the satellite and a scatterer is related to the speed of light c and the two-way travel time t_{2w} as:

$$R = \frac{ct_{2w}}{2}. \tag{2.1}$$

When the position and the attitude of Cryosat are known it is possible to approximate the elevation of a scatterer.

Using pulses with no frequency modulation it is possible to distinguish scatterers that are seperated by half the pulse length [22]. The transmitted pulse length in meters ρ_t can be computed from the pulse length in seconds τ_t using:

$$\rho_t = c\tau_t. \tag{2.2}$$

The pulse transmitted by SIRAL has a pulse length in seconds of 49 μs [1], which corresponds to pulse length in meters of 14700 m. Therefore scatterers have to be at least 7350 m apart in range to be distinguishable. As a consequence such a beam-limited altimeter would be suitable to find one range solution per pulse [1].

In order to distinguish between scatterers closer apart, either a shorter pulse should be transmitted, or a frequency modulated pulse is needed. A short pulse of a few nanoseconds requires high power to get above the noise level [23]. To limit the power requirements it is more convenient to transmit a pulse with a linear decreasing frequency, or a 'chirp'. By proper signal processing on these chirps it is possible to distinguish between scatterers that are seperated only nanoseconds, or decimeters, apart in range [22], which is described in section 2.2.2.

To describe the frequency variation within the pulse, the maximum frequency f_0 at the start of the 'chirp' is defined as a function of bandwidth B and the center frequency f_c:

$$f_0 = f_c - \frac{B}{2}. \tag{2.3}$$

The frequency $f(t)$ at a specific time during the pulse is a function of the frequency slope Q, the time t and the maximum frequency. Frequency slope is related to the bandwidth B of the signal with $Q = \frac{B}{\tau_t}$. The function for the frequency becomes:

$$f(t) = f_0 - Qt. \tag{2.4}$$

For brevity $f(t)$ will be written as f in the upcoming equations.

Since the frequency is the derivative of the number of cycles with respect to the time, an integration of the frequency yields the number of cycles $noc(t)$:

$$noc(t) = \int f dt = (f_0 - \frac{Q}{2}t)t + constant. \tag{2.5}$$

By assuming that $t = 0$ at the beginning of pulse transmission, the constant in equation 2.5 is set to zero. This reduces the equation for the number of cycles to:

$$noc(t) = (f_0 - \frac{Q}{2}t)t. \tag{2.6}$$

Therefore, the equation for the transmitted signal $s_t(t)$ with amplitude A_{st} can be written as:

$$\begin{aligned} s_t(t) &= A_{st}\sin(2\pi noc(t)) \\ &= A_{st}\sin(2\pi(f_0 - \frac{Q}{2}t)t)\ . \end{aligned} \tag{2.7}$$

An exaggerated example of such a frequency modulated signal is given in figure 2.3. It shows that the frequency is decreasing between $t = 0$ and $t = \tau_t$, the end of the pulse.

In reality the central frequency is much higher compared to the inverse pulse length and the decrease in frequency is hardly visible.

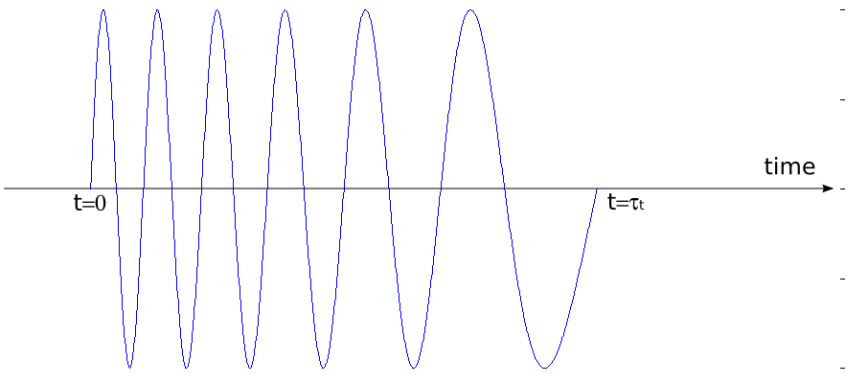


Figure 2.3: Exaggerated example of a 'chirped pulse', with τ_t the pulse length.

When this signal is received at an antenna of the satellite the frequency is affected by the Doppler effect. The satellite has a velocity V of approximately 7500 m/s, which causes a Doppler shift D_{shift} for all scatterers at an off-nadir angle (look angle) in along-track direction α not equal to zero. The Doppler effect D can be written as a function of the normalized velocity vector n_v, with a length of one, and the normalized direction vector towards the scatterer n_s:

$$\begin{aligned} D &= 1 + D_{shift} \\ 1 + \frac{2V\sin(\alpha)}{c} &= 1 + \frac{2Vn_v.n_s}{c}. \end{aligned} \tag{2.8}$$

The amplitude of the received signal depends among others on the antenna gain. Since SIRAL has two slightly elliptical antennas, the antenna gain $G(\theta, \upsilon)$, is different in cross-track and along-track direction. The antenna gain is a function of the polar θ and azimuthal υ directions with respect to the antenna boresight, where υ is zero degrees in the along-track direction and θ is 0 degrees in nadir direction if the satellite is pointing towards the Earth. The antenna gain is described by [1][24]:

$$G(\theta, \upsilon) = G_0 e^{-\theta^2\left(\frac{\cos^2(\upsilon)}{\gamma_1^2} + \frac{\sin^2(\upsilon)}{\gamma_1^2}\right)}, \tag{2.9}$$

in which G_0 is the peak gain of 42 dB and γ_1, γ_2 are antenna constants.

In order to compute the power received at the satellite, the radar equation is used. The received power P_{sr} depends on the transmitted power P_{st}, the wavelength of the signal

λ, the backscattering coefficient σ_0, atmospheric effects which are included in the path effect F, antenna gain and the range. The radar equation is written as:

$$P_{sr} = P_{st}\frac{G^2(\theta,\upsilon)\lambda^2\sigma_0 F^4}{(4\pi)^3 R^4}, \tag{2.10}$$

and can be rewritten to compute the amplitude received A_{sr} as:

$$\begin{aligned} A_{sr} &= \sqrt{P_{st}\frac{G^2(\theta,\upsilon)\lambda^2\sigma_0 F^4}{(4\pi)^3 R^4}} \\ &= A_{st}\sqrt{\frac{G^2(\theta,\upsilon)\lambda^2\sigma_0 F^4}{(4\pi)^3 R^4}}\ , \end{aligned} \tag{2.11}$$

because the amplitude is the square-root of the power.

Using the time at the beginning of signal receival t_s, the equation for the signal received $s_r(t)$ becomes:

$$\boxed{s_r(t) = A_{sr}\sin(2\pi D(f_0 - \frac{Q}{2}(t-t_s))(t-t_s)).} \tag{2.12}$$

2.2.2 Signal Deramping

Signal deramping consists of two steps: mixing the received signal with a replica signal, called a 'deramping chirp' and applying a Low Pass Filter (LPF) to remove high frequency signals. It will be shown that deramping of an echo creates a mixed echo consisting of signals with constant frequency, that depend on the time at which the signal from a scatterer arrives. Figure 2.4 shows the frequencies of signals arriving at different time stamps after deramping. It will be proven that the fundamental frequency (i.e. the lowest frequency in the mixed echo that can be retrieved) corresponds to the time resolution of 3.125 ns of SIRAL.

To avoid confusion between the terms 'echo' and 'signal': a signal is reflected from one particular scatterer, while an echo is a set of reflected signals received at the antenna. This section deals with an echo from one particular scatterer and therefore commonly uses the term signal, whereas in reality a set of signals is received.

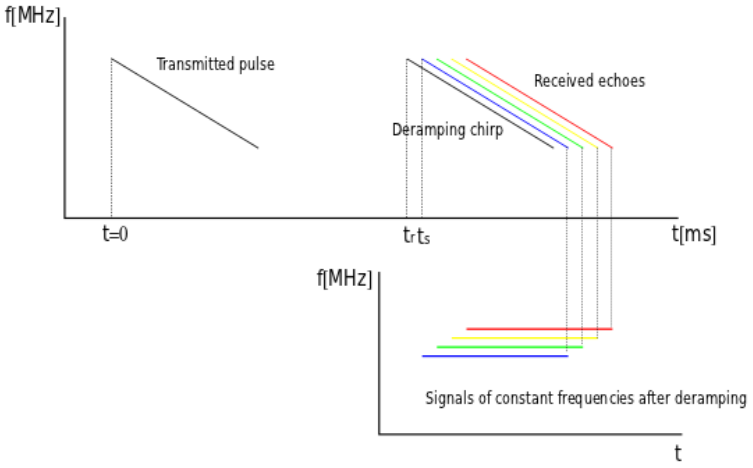


Figure 2.4: Incoming echoes are deramped, after which the echoes contain signals with constant frequency.

First it is necessary to rewrite Eq. (2.12) using Euler's formula, which results in:

$$s_r(t) = \frac{A_{sr}}{2i}[e^{i2\pi D(f_0 - \frac{Q}{2}(t-t_s))(t-t_s)} - e^{-i2\pi D(f_0 - \frac{Q}{2}(t-t_s))(t-t_s)}]. \tag{2.13}$$

SIRAL generates two deramping chirps just before the expected time of arrival of the signal from the Point Of Closest Approach (POCA), i.e. the first signal to arrive at the satellite antenna. The two deramping chirps are in-phase (cosine) and quadrature phase (sine), which means that there is a $\frac{\pi}{2}$ phase difference between the two chirps. In the equations these are referred to as 'real' and 'imag'. The equation for the deramping chirp can be written similarly to that for the received signal, with A_r the amplitude and t_r the time at which the deramping chirps are generated. This results in:

$$r_{real}(t) = \frac{A_r}{2}[e^{i2\pi(f_0 - \frac{Q}{2}(t-t_r))(t-t_r)} + e^{-i2\pi(f_0 - \frac{Q}{2}(t-t_r))(t-t_r)}] \tag{2.14}$$

and

$$r_{imag}(t) = \frac{A_r}{2i}[e^{i2\pi(f_0 - \frac{Q}{2}(t-t_r))(t-t_r)} - e^{-i2\pi(f_0 - \frac{Q}{2}(t-t_r))(t-t_r)}]. \tag{2.15}$$

A mixed signal is generated by multiplying the received signal with the deramping chirps. To write the mixed signal in a more convenient way, two additional parameters are introduced: the frequency F_s of the received signal without the Doppler effect and the frequency F_r of the deramping chirps, both of which are functions of time:

$$F_s = f_0 - \frac{Q}{2}(t - t_s) \tag{2.16}$$

$$F_r = f_0 - \frac{Q}{2}(t - t_r) \tag{2.17}$$

The mixed signal $m_{real}(t)$ based on the real deramping chirp can be written as:

$$\begin{aligned} m_{real}(t) &= s_r(t) \cdot r_{real}(t) \\ &= \frac{A_{sr}}{2i}[e^{i2\pi DF_s(t-t_s)} - e^{-i2\pi DF_s(t-t_s)}]. \\ &\quad \frac{A_r}{2}[e^{i2\pi F_r(t-t_r)} + e^{-i2\pi F_r(t-t_r)}] \\ &= \frac{A_m}{2i}[e^{i2\pi(DF_s(t-t_s)+F_r(t-t_r))} + e^{i2\pi(DF_s(t-t_s)-F_r(t-t_r))} - \\ &\quad e^{-i2\pi(DF_s(t-t_s)+F_r(t-t_r))} - e^{-i2\pi(DF_s(t-t_s)-F_r(t-t_r))}] \end{aligned} \tag{2.18}$$

with,

$$A_m = \frac{A_{sr}A_r}{2}. \tag{2.19}$$

After the mixing operation a LPF is applied to eliminate the high frequency signal. This allows the terms with $DF_s(t - t_s) + F_r(t - t_r)$ to be removed, which results in the following equation for the LPF signal:

$$m_{real,LPF}(t) = \frac{A_m}{2}[e^{i2\pi(DF_s(t-t_s)-F_r(t-t_r))} - e^{-i2\pi(DF_s(t-t_s)-F_r(t-t_r))}], \tag{2.20}$$

or:

$$m_{real,LPF}(t) = Im\{A_m e^{i2\pi(DF_s(t-t_s)-F_r(t-t_r))}\}. \tag{2.21}$$

In the equation above it is shown that the real signal, is the imaginary part of the complex signal. This is caused by the fact that the transmission of a sinusoidal signal is chosen over a cosinusoidal one, since this is technically more feasible. When the signal is created the the amplitude will not be immediately maximal, as sinusoidal signal starts at

zero this is not an issue. In parallel, mixing with the quadrature phase deramping chirp is performed. This results in a sinusiodal signal, written as the real part of a similar complex function for the mixed signal:

$$m_{imag,LPF}(t) = -Re\{A_m e^{i2\pi(DF_s(t-t_s)-F_r(t-t_r))}\}. \tag{2.22}$$

The number of cycles in the mixed signals, noc_m, is taken from the two preceding equation and can be described as:

$$noc_m = DF_s(t-t_s) - F_r(t-t_r). \tag{2.23}$$

Rewriting Eq. (2.23) using Eq. (2.16) and (2.17) results in:

$$\begin{aligned} noc_m &= (D-1)(f_0(t-t_s) - \frac{Q}{2}(t-t_s)^2) + \\ &\quad f_0(t-t_s) - \frac{Q}{2}(t-t_s)^2 - f_0(t-t_r) + \frac{Q}{2}(t-t_r)^2 \\ &= (D-1)(f_0(t-t_s) - \frac{Q}{2}(t-t_s)^2) \\ &\quad + f_0(t_r-t_s) + \frac{Q}{2}(t_r^2 - t_s^2) + Q(t_s-t_r)t\ . \end{aligned} \tag{2.24}$$

The maximum frequency f_0 is rewritten as a function of the center frequency as:

$$f_0 = f_c + \frac{B}{2} = f_c + \frac{Q\tau}{2}. \tag{2.25}$$

Using Eq. (2.25) the number of cycles is described with the following equation:

$$\begin{aligned} noc_m = (D-1)(f_c(t-t_s) + \frac{Q}{2}(\tau(t-t_s) - (t-t_s)^2)) + \\ f_c(t_r-t_s) + \frac{Q}{2}(t_r^2 - t_s^2 + \tau(t_r-t_s)) + Q(t_s-t_r)t\ . \end{aligned} \tag{2.26}$$

In order to verify if some terms in the equation above can be discarded, Eq. (2.8) for the Doppler effect and Eq. (2.4) for the frequency slope are used. Table 2.2.2 shows the values for the parameters used to compute the values of the terms in Eq. (2.26). The along-track offset of the scatterer is assumed to be maximally 7.5 km, since the diameter of the footprint of SIRAL is approximately 15 km. Using the values for along-track offset, satellite velocity and satellite altitude, the Doppler factor $D-1$ has a value of

$5.23 * 10^{-7}$. The frequency slope can be computed using the sampled pulse length and the bandwidth, which results in $Q \approx 7.14 \cdot 10^{12}$ Hz/s.

Parameter	Value
Maximum along-track offset of a scatterer	7.5 km
Satellite velocity(V)	7500 m/s
Satellite altitude	717 km
Sampled pulse length(τ)	44.8 μs
Bandwidth(B)	320 MHz
Time difference with arrival of signal($t - t_s$)	0-44.8 μs

Table 2.3: List of approximate parameter values to estimate the number of cycles.

Using the values for the Doppler factor and frequency slope the maximum value for $(D-1)\frac{Q}{2}(\tau(t-t_s)-(t-t_s)^2)$ becomes $9.4 \cdot 10^{-4}$ cycles and can be neglected. The term $(D-1)(f_c(t-t_s))$ causes a phase change up to several radians and cannot be neglected. This results in the following equation for the number of cycles:

$$noc_m \approx Q(t_s - t_r)t + \frac{Q}{2}(t_r^2 - t_s^2 + \tau(t_r - t_s)) + (D-1)(f_c(t-t_s)) + f_c(t_r - t_s) \quad (2.27)$$

Eq. (2.27) can be rewritten in the form:

$$noc_m \approx f_m t + f_{m,Doppler}(t - t_s) + \frac{\Phi_m}{2\pi} \quad (2.28)$$

with,

$$\boxed{f_m = Q(t_s - t_r)} \quad (2.29)$$

$$f_{m,Doppler} = (D-1)f_c. \quad (2.30)$$

$$\frac{\Phi_m}{2\pi} = f_c(t_r - t_s) + \frac{Q}{2}(t_r^2 - t_s^2 + \tau(t_r - t_s)). \quad (2.31)$$

By writing the equation for the number of cycles this way, it becomes clear that the time between the generation of the deramping chirp and the receival of the signal determines the phase of the signal. Furthermore the Doppler effect has an influence on the frequency of the signal. And the most important part is given in Eq. (2.29), which shows that the frequency of the signal depends upon the difference in arrival time of the signal and the generation of the deramping chirp.

These echoes are sampled over a period of 44.8 μs. The signals with the minimum frequency that can be retreived have a period of exactly 44.8 μs. Signals with this frequency are said to have a fundamental frequency f_F. After the application of a transform, discussed in section 2.2.3, frequencies which are an integer number of the fundamental frequency are transformed to range bins. The time interval between these range bins, or time resolution, is therefore related to the fundametal frequency. The fundamental frequency is computed using:

$$f_F = \frac{1}{\tau}. \tag{2.32}$$

With τ the sampled pulse length. By combining the equation for the fundamental frequency with Eq. (2.29), i.e. equating $f_m = f_F$, the time resolution t_{res} of SIRAL is (by setting $t_{res} = (t_s - t_r)$) computed as:

$$\begin{aligned} t_{res} &= \frac{1}{Q\tau} \\ &= \frac{1}{B} \\ &. \end{aligned} \tag{2.33}$$

As a result the resolution of SIRAL is only depending on the bandwidth. With a bandwidth of 320 MHz for the sampled signal, the time resolution of SIRAL is 3.125 ns. It is possible to modify the equation for the range, Eq. (2.1) to get the range resolution r_{res} as:

$$r_{res} = \frac{ct_{res}}{2} \tag{2.34}$$

This gives the range resolution of 46.9 cm.

Finally, the equation for both mixed signals, in-phase and quadrature phase can be rewritten as:

$$\boxed{m_{real,LPF}(t) = Im\{A_m e^{i(2\pi(f_m t + f_{m,Doppler}(t-t_s))+\Phi_m}\}} \tag{2.35}$$

$$\boxed{m_{imag,LPF}(t) = -Re\{A_m e^{i(2\pi(f_m t + f_{m,Doppler}(t-t_s))+\Phi_m}\},} \tag{2.36}$$

using Eq. (2.28).

To illustrate the result of deramping, figure 2.5 shows the real part of mixed signals arrived at 3.125 ns, 6.250 ns and 62.500 ns after the generation of the deramping chirp. The frequency of the signals increases if the signals are received later.

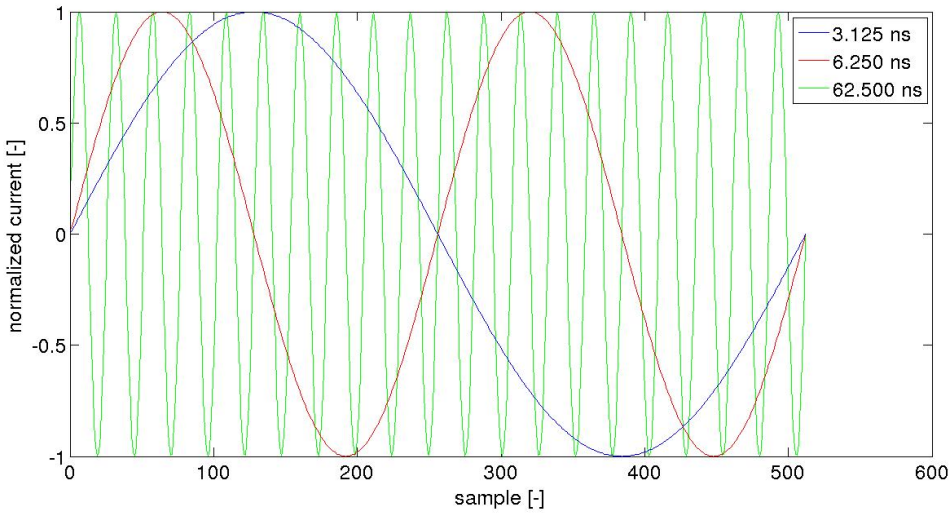


Figure 2.5: Plot of deramped signals received at various times after generation of the deramping chirp. The frequency of the deramped signals increases if the received signals arrive later.

2.2.3 Low Resolution Mode

In the LRM mode only the left of the two antennas is used to receive the echoes. SIRAL transmits pulses of 49 μs length and a bandwidth of 350 MHz at a Pulse Repetition Frequency (PRF) of 1971 Hz. Approximately 4.8 ms later the signal is received. In total ninety one waveforms are averaged over a 46.7 ms period as shown in figure 2.6.

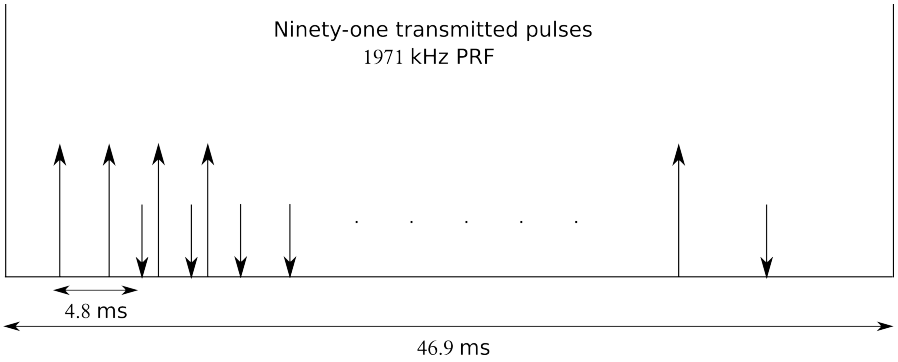


Figure 2.6: Low resolution mode echoes (arrow down) of transmitted pulses (arrow up) are received after 4.8 ms. Ninety-one echoes are accumulated per measurement. Adapted from [1].

Then the echoes are duplexed, which means that the echo is doubled and sent to the

two deramping chirps, one in-phase and one quadrature phase, described in the previous chapter. After deramping the echo is anti-alias filtered with a Hamming window, which is a specific type of LPF. This anti-aliasing is performed to remove high frequencies in the mixed echo. These high frequencies are caused by scatterers that fall outside of measurement window of 60 m in range. Ultimately 128 complex samples are formed from 44.8 μs of the in-phase (real) and quadrature (imaginary) phase mixed signals [1]. This reduces the bandwidth to 320 MHz.

In Eq. (2.29) it was shown that the LPF mixed signal is build up out of low frequencies depending on the travel time of the signal. To retrieve these frequencies a fast implementation of the Discrete Fourier Transform (DFT), the Fast Fourier Transform (FFT), is applied to the sampled signal [1]. This results in a waveform of 128 frequencies, which are proportional to the time increment between received signal and the generation of the replica. The DFT of the sampled mixed signal can be written as:

$$M_k = \sum_{n=0}^{N-1} m_s(t_n) e^{-i2\pi \frac{kn}{N}}, \tag{2.37}$$

in which N is the number of samples, k is the number of cycles in the sampling period and $m_s(t_n)$ is the sampled signal.

Ninety one consecutive waveforms (over 46.7 ms) are accumulated and averaged to reduce speckle. These are passed to the tracker, which provides closed-loop control for the timing of the deramping chirp and the receiver gain [1]. The closed-loop control provides an estimated elevation derived from the waveforms and the receiver gain is adjusted, so that saturation (the situation when the incoming power is higher than the maximum power that can be measured) does not occur. Furthermore, the waveforms are averaged, so one waveform per 91 measurements is produced, which is stored and finally transmitted to a ground station as measurement data. Due to its range resolution of 46.9 cm (Eq. (2.34)) and a waveform size of 128 bins the measurement window in LRM is approximately 60 m [1].

2.2.4 Synthetic Aperture Mode

In SARM, like in LRM, only one antenna is used to receive the echoes. In principle SARM is designed to increase the along-track resolution of SIRAL. Therefore pulses are transmitted not continuously, but in bursts.

The pulses are transmitted at a PRF of 18.182 kHz in bursts of 64 pulses, the length of the burst being 3.6 ms. The time between the onsets of two consecutive bursts is 11.7 ms. Like in LRM, the echoes are duplexed, Hamming window filtered and then deramped. The time between transmission of a pulse and the generation of the deramping chirp is held constant to avoid differential phase shift between echoes within the burst [1]. Over

46.9 ms (or 4 bursts) the echoes are FFT'd and acculumated, like in LRM to provide control for timing of the deramping chirp and receiver gain. However, the 128 real and imaginary samples of every pulse are transmitted to the ground station as measurement data [1]. This differs from LRM, when the waveforms, i.e. the data after applying the FFT are transmitted. A schematic representation of the burst and pulse timing is given in figure 2.7.

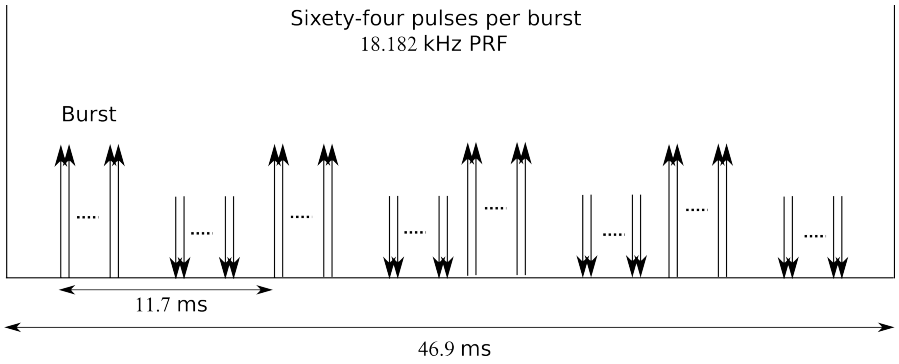


Figure 2.7: SARM pulses are transmitted in bursts of 64, with a burst repetition interval of 11.7 ms. Adapted from [1].

In the time interval between two pulse transmissions the satellite moves several decimeters along its orbit. This causes a phase change in the signal received at 64 successive along-track locations, since the range from the scatterer slightly varies. The corresponding difference in travel time Δt_a at a current along-track position with respect to the center of the burst is

$$\Delta t_a = \frac{2\Delta R_a}{c}, \tag{2.38}$$

which depends on the range difference ΔR_a and the speed of light.

The difference in range depends on the normalized velocity vector n_v and the direction to the scatterer n_s [1]:

$$\Delta R_a = (a - 65/2)\Delta I n_v \cdot n_s, \tag{2.39}$$

with an along-track sampling interval ΔI of

$$\Delta I = \frac{V}{PRF} \tag{2.40}$$

and a the position number in the burst, which ranges from 1 to 64. In order to clarify Eq. (2.39) the geometry of the pulses is shown in figure 2.8. The burst center is located

at position $a = 65/2$.

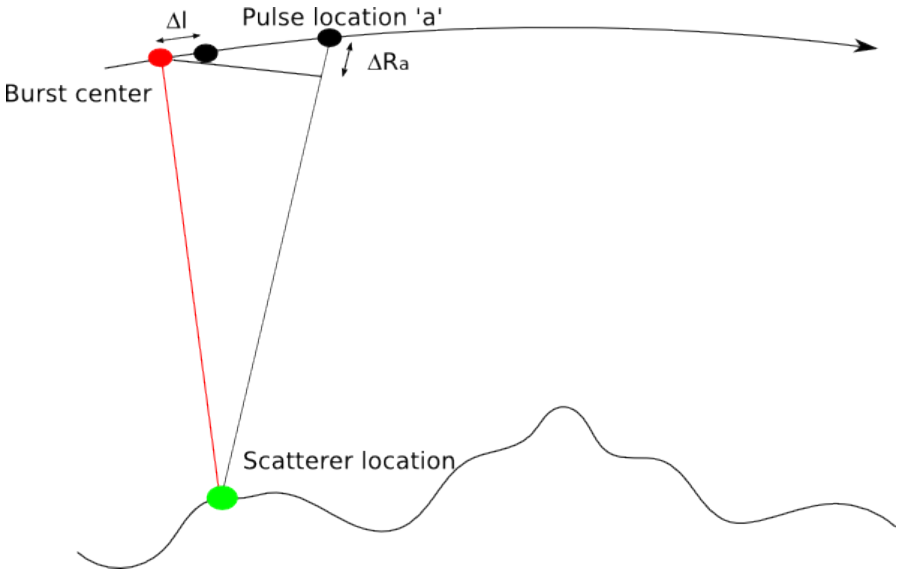


Figure 2.8: Along-track geometry of scatterer location and pulse locations.

As a result the echoes experience a slight phase shift as described by the second exponent in the following equation:

$$\boxed{m_{real,LPF}(t,a) = Im\{Me^{i(2\pi f_m t + \Phi_m + \Phi_{m,Doppler})} e^{i2\pi f_c \frac{2\Delta I}{c}(a-65/2)n_v.n_s}\}} \tag{2.41}$$

cf. Eq. (2.38) and Eq. (2.39) [1].

This slight phase shifts is used to distinguish scatterers in along-track strips (thus improving the along-track resolution). To that end an along-track FFT is applied, as will be described in section 2.3.2. The strip with minimum distance from nadir in along-track direction that can be distinguished has a 2π difference in phase, i.e. 1 cycle difference, between echoes 0 and 64. Echo 0 should be considered as the echo from the position of the satellite one sampling interval before the actual first echo.

The number of cycles difference in along-track between position 0 and position 64 is given by:

$$\Delta noc_a = f_c \frac{2\Delta I}{c}((64 - 65/2) - (0 - 65/2))n_v \cdot n_s. \tag{2.42}$$

Let us assume that the satellite flies along-track in a straight line. The position nadir of the satellite is set as the origin of the reference frame. The x-direction is along-track

direction of the satellite. The y-direction is the cross-track direction of the satellite. And the z-direction is the up-direction. Then, the normalized velocity vector becomes:

$$n_v = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and the normalized direction vector, which depends on the satellite altitude h, the along-track distance d_a and the cross-track distance d_c, is approximated by:

$$n_s = \begin{pmatrix} d_a/R \\ d_c/R \\ h/R \end{pmatrix} \approx \begin{pmatrix} d_a/h \\ d_c/h \\ h/h \end{pmatrix} = \begin{pmatrix} d_a/h \\ d_c/h \\ 1 \end{pmatrix}.$$

Using Eq. (2.42) and the normalized vectors, the along-track distance d_a can be computed with:

$$d_a = h\frac{\Delta noc_a c}{2\Delta I 64 f_c}. \tag{2.43}$$

By setting the number of cycles difference equal to one, the along-track resolution is computed. Assuming a satellite velocity of 7.5 km/s the along-track sampling interval is 0.41 m. The satellite altitude is approximately 717 km, the speed of light $3 \cdot 10^5$ km/s and the central frequency is 13.575 GHz. Using these values an along-track resolution of around 300 m is obtained. This is an approximation, which is valid for the look angles less than 0.76 deg of SIRAL. For larger look angles, the resolution becomes worse.

2.2.5 Synthetic Aperture Interferometric Mode

When SIRAL is operating in the SARInM, the PRF is equal to that in SARM (18.182 kHz). Again 64 pulses are transmitted per burst, but with a four times lower burst repetition frequency of 21.4 Hz, which yields a burst interval of 46.9 ms. The signals are received at both antennas, separated by 1.1676 meters cross-track [25], and the deramping chirps are generated exactly at the same time at both antennas [1]. An overview of the bursts and pulses is given in figure 2.9.

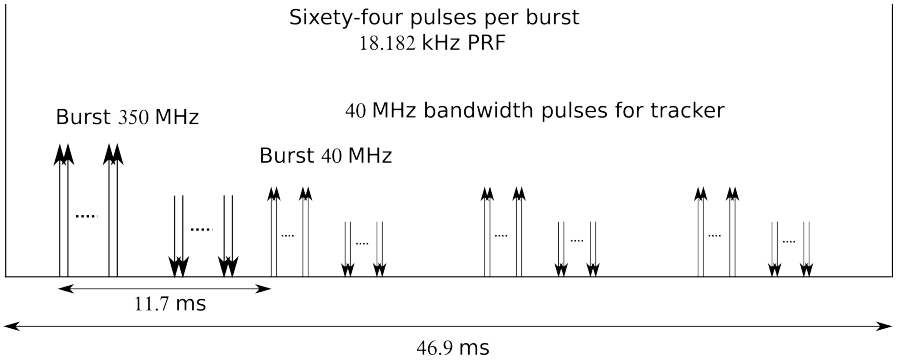


Figure 2.9: 350 MHz pulses are transmitted in bursts of 64, with a burst repetition interval of 46.7 ms. The 350 MHz bandwidth pulses are transmitted to obtain measurement data, whereas the three 40 MHz pulses are used for the tracker. Adapted from [1].

Since SARInM is active over areas with strong topography, the 60 m window of the LRM and SARM is not adequate anymore. Therefore the 44.8 μs long signal is discretized in 512 samples of in-phase and quadrature phase values. This increases the number of frequencies that can be retrieved with a DFT from 128 to 512 and therefore increases the measurement window to 512·46.9 cm, $\approx$ 240 m [1]. For both antennas 64 signals of 512 samples per burst are transmitted to the ground station as measurement data.

In contrast to LRM and SARM, the tracker that provides information for closed-loop control of timing of the replica and control of the receiver gain is not based on the measurement data. Instead another 40 MHz bandwidth signal is transmitted at a PRF of 18.182 kHz in bursts of 64, of which the echoes are passed through the FFT and accumulated over 46.7 ms. Due to the lower bandwidth the spectrum has decreased resolution, but spans a range window of 480 m to provide robust timing control for the replica above areas with significant topography [1].

The two antennas allow an estimation of the cross-track slope, by measuring a phase difference between the two antennas. The accuracy depends on the baseline, i.e. the cross-track distance between the antennas phase centers. If n_b is assumed to be normalized baseline vector and B the length of the baseline, which is approximately 1.1676 m, Eq. (2.41) can be extended to [1]:

$$\boxed{\begin{aligned} m_{real,LPF}(t,a,\pm) = Im\{ & M e^{i(2\pi f_m t + \Phi_m + \Phi_{m,Doppler})} \\ & e^{i2\pi f_c \frac{2\Delta I}{c}(a-65/2)n_v.n_s} e^{\pm i2\pi f_c \frac{B}{c} n_s.n_b}\} \, , \end{aligned}} \tag{2.44}$$

with the $\pm$ referring to the left or right antenna. The term $2\pi f_c \frac{B}{c} n_s.n_b/2$ shows the difference in phase between the antennas and the center of the satellite, because the distance to the antenna is slightly smaller or larger depending on whether the scatterer is right or left from nadir.

2.3 Data Products

For all three modes of SIRAL (LRM, SARM and SARInM) various processing products are available. An overview of the available data products is shown in table 2.3.1. As shown in the table level 0 data is nothing more than measurement data received from the satellite and is not available to the public.

The same operations are applied in all three modes to get level 1 data. It mainly concerns the datation and geolocation of data and converting the platform units to engineering units [2]. The processing stage from level 1 to level 1b data is different for all modes. In LRM only calibration of the product is performed, in SARM averaged echoes are created and in SARInM additional to SARM some additional parameters are computed concerning the two antennas. Ultimately level 2 data is the geophysical product, containing primarily elevations for LRM and SARInM and sea-ice thickness for SARM [2].

This section decribes the operations of the processing facility to get from level 0 to level 1b data. Additionally is describes the principles of processing the level 1b data to level 2 data, primarily for the SARInM since this mode is activated over Tibet.

2.3.1 Level 1 processing

The level 1 processing stage is described by the flowchart in figure 2.10. This processor is also called the Full Bit Rate (FBR) processor and is identical for all three modes. The level 1 FBR data is the lowest level product available to the public. Level 1 data will primarily be used by researchers interested in calibration of Cryosat data.

As the first step the level 1 processor computes the datation, which is the timing of the bursts and pulses, and geolocation. First precise orbit computations are required to enable more precise positioning of the burst locations [1]. A second input is the time of the transmission and receival of the pulses, to get the 'start-stop' locations. During the tranmission and receival of the pulses, the motion of the pulse and the satellite map a triangle, whose vertices are the start-stop locations and a scattering point on the ground. The received echo is indistinguishable from an echo received if the satellite were stationary half way the transmission and reception. Therefore its echo is treated as if it were measured from the half-way position. The center of the burst location is then the midpoint of the 64 'start-stop' locations, or halfway between the 32nd and 33rd 'start-stop' location [1], which is also shown in equation 2.44.

Product Type	User Distribution	Mode	Description
L0	No	LRM	Onboard average of 91 frequency domain waveforms.
		SARM	Bursts of 64 time domain complex waveforms.
		SARInM	Bursts of 64 time domain complex waveforms for two antennas.
L1	Yes	LRM SARM SARInM	Converted platform units geolocation and timing, no calibration, waveforms are a copy of L0.
L1b	Yes	LRM	Calibrated power echoes at 20Hz.
		SARM	Calibrated multilooked power echoes at 20Hz.
		SARInM	Calibrated multilooked power, interferometric phase difference and coherence waveworms at 20Hz.
L2	Yes	LRM SARInM	Surface elevation and across-track corrected echo location.
		SARM	Sea-ice thickness and free-board measurements.

Table 2.4: Overview of Cryosat-2 data products (Adapted from [2])

Then the saturation of the signal is checked and the baseline orientation is estimated using star-tracker data. This is followed by adding calibration data from the calibration modes CAL1, CAL2 and CAL4 to the level 1 data. The calibration data provides internal instrument corrections [1][25][26]. The corrections are stored, however not yet applied until the level 1b processing stage. After this the timing of the deramping chirp and the pulse transmission time are used to computed window delay, followed by the computation of receiver gain. Finally the records are created, in which the waveforms (LRM) or sampled echoes (SARM and SARInM) are copied, along with interpolated orbital state vectors and attitude information [27].

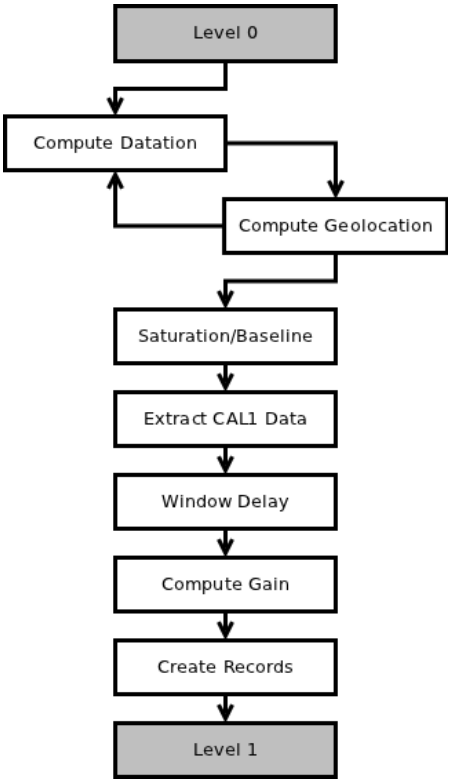


Figure 2.10: Flowchart Level 1 Processing (Adapted from [2]).

2.3.2 Level 1b processing

The second processing level is level 1b, which creates essentially an average echo for each location along the ground track [1]. For LRM the level1b processing mainly concerns applying the instrument calibration corrections CAL1, CAL2 and CAL4, since the averaging of 91 echoes is already applied onboard as described in paragraph 2.2.3. Additionally atmospheric delay compensation, like for the ionosphere and troposphere, and geophysical correction, like tides, are stored in data, but not applied yet. Finally, a correction for the elliptical antenna pattern is applied, which is described in [1]. This will not be discussed here, because over Tibet the SARInM is activated.

For SARM and SARInM averaged echoes have to be computed for strips of approximately 300 m width along-track, as described in paragraph 2.2.4. As input data for SARM there are 64 echoes of 128 complex samples per burst. In SARInM these echoes contain 512 complex samples.

The main goal of level 1b processing for SARM and SARInM is to create an averaged waveform per strip, called a multilooked waveform. This paragraph describes the procedure to get to a multilooked waveform. A flowchart is shown in figure 2.11. An

along-track FFT creates from the 64 echoes per burst, one signal per strip of approximately 300 m. Since the satellite does not move exactly 300 m over the ground-track these signals are unfocussed and therefore beamsteering is applied before the FFT. As shown in paragraph 2.2.1 the received signal is affected by the Doppler effect, for which is corrected after the along-track FFT. The bursts have different ranges to a particular strip and therefore a slant-range correction is applied to change them to a reference range. Then a range FFT is applied to get a single look waveform per 300 m strip. The single look waveforms from different waveforms are accumulated in a stack. Ultimately, the single look waveforms are multilooked to get one averaged waveform per stack.

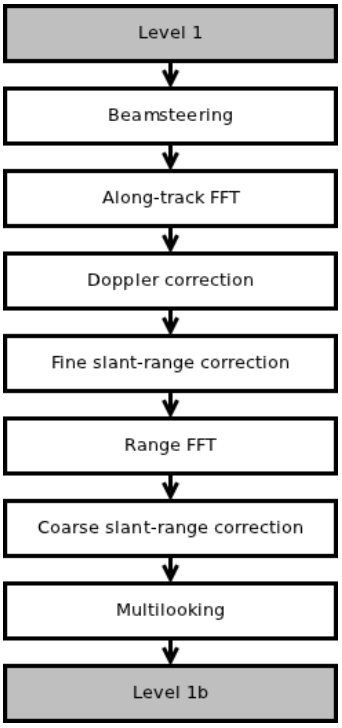


Figure 2.11: Flowchart Level 1b Processing (Adapted from [2]).

Figure 2.12 shows the geometry of what is the result of applying an along-track FFT over the 64 echoes within the burst. On top is the satellite trajectory with the burst locations. Would the along-track FFT not be applied, the footprint of Cryosat would be approximately 15-20 km along-track, in the figure this is from the first beam to the last beam. The along-track FFT creates in total 64 beams, which illuminate strips of approximately 300 m width on the ground.

In order to get a higher signal-to-noise ratio and to mitigate speckle, beams from different bursts are added. The ground locations are chosen such that they are exactly nadir the satellite 'start-stop' locations as computed in the level 1 processing stage. If an along-track FFT would be applied directly, the beams would not illuminate exactly the ground

location. To illustrate this, figure 2.13 shows beams from several bursts that illuminate one ground location, but only after approximation. To counteract this illumination problem, beamsteering is applied.

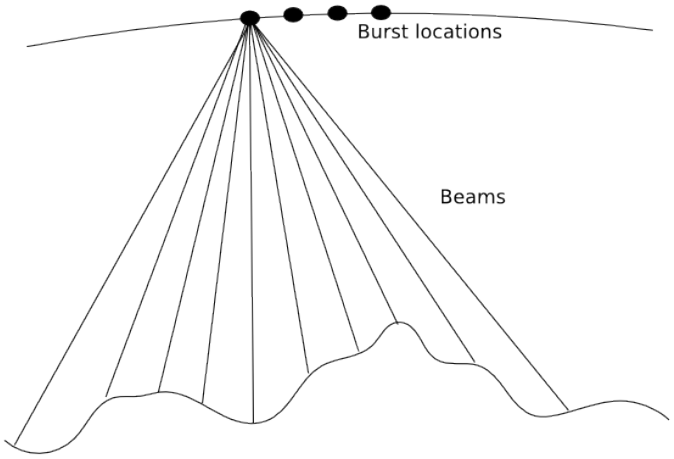


Figure 2.12: Sixty-four beams are formed from one burst (Adapted from [2]).

Beamsteering and along-track FFT

Beamsteering is performed in the form of a phase adjustment, which is shown in figure 2.14 [1]. The unfocussed beam points towards a point that is not exactly located at the ground location. This means that the center of the strip is located a certain distance Δd_a away from the ground location. By changing the phases of the signals received at the pulse locations or start-stop locations a, we can 'pretend' the ground location to be at the center of the strip. By changing the normalized direction vector n_s to the scatterer location to the one after beamsteering n_{bs}, a small difference in range $\Delta R_{bs}(a)$ is induced. The further along-track a scatterer is, the higher the phase change between the pulse locations in Eq. (2.39). In figure 2.14 it is shown that the ground location is further along-track than the center of the strip. Therefore, it is necessary to decrease the phase change, so we 'pretend' that the ground location is inside the center of the strip. This leads to the equation:

$$\begin{aligned} \Delta R_{bs}(a) &\approx \Delta I(a - 65/2) n_v \cdot (n_s - n_{bs}) \\ &\approx \Delta I(a - 65/2) n_v \cdot \Delta n_s \ , \end{aligned} \tag{2.45}$$

with

$$n_s - n_{bs} = \begin{pmatrix} -\Delta d_a \\ 0 \\ 0 \end{pmatrix}, \tag{2.46}$$

which is directly related to a phase change.

Since the echoes of SARM and SARInM are complex samples, the phases and amplitude can be computed. Therefore it is convenient to express the range adjustment in the form of a phase multiplier $\Phi_{bs}(a)$ for beam steering [17]. Using Eq. (2.38) the phase multiplier can be expressed as:

$$\boxed{\Phi_{bs}(a) = e^{i2\pi f_c \frac{2\Delta R_{bs}(a)}{c}},} \tag{2.47}$$

with the central frequency f_c and the speed of light c.

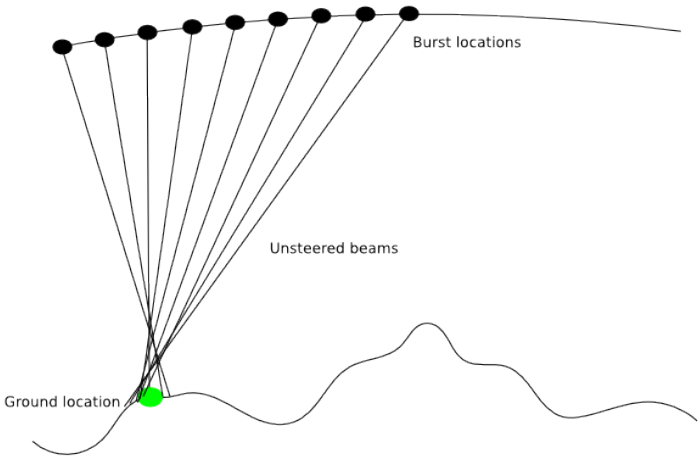


Figure 2.13: Beams from different burst locations do not exactly illuminate the same spot (Adapted from [2]).

For one ground location the phase multiplier is computed for every pulse in all bursts that illuminate the ground location. Then an along-track FFT is applied for all the bursts and for that particular ground location the resulting echoes are collected into a stack. For another ground location, the same procedure is applied again. The resulting effect of beamsteering is illustrated in figure 2.15. Beams formed from different bursts illuminate the same spot.

The signal after applying the FFT can be expressed in an equation. Based upon Eq. (2.44) let us assume that the beamsteered sampled mixed signal $M_{s,bs}$ before applying the along-track FFT can be expressed in several components as:

$$M_{s,bs}(a) = M_{burst} M_{antenna} M_{pulse} \Phi_{bs} \tag{2.48}$$

with

$$M_{burst} = e^{i(2\pi f_m t + \Phi_m + \Phi_{m,Doppler})} \tag{2.49}$$

$$M_{ant} = e^{\pm i 2\pi f_c \frac{B}{c} n_s . n_b} \tag{2.50}$$

$$M_{pulse}(a) = e^{i2\pi f_c \frac{2\Delta I}{c}(a-65/2) n_v . n_s} \tag{2.51}$$

and Φ_{bs} as described in Eq. (2.47).

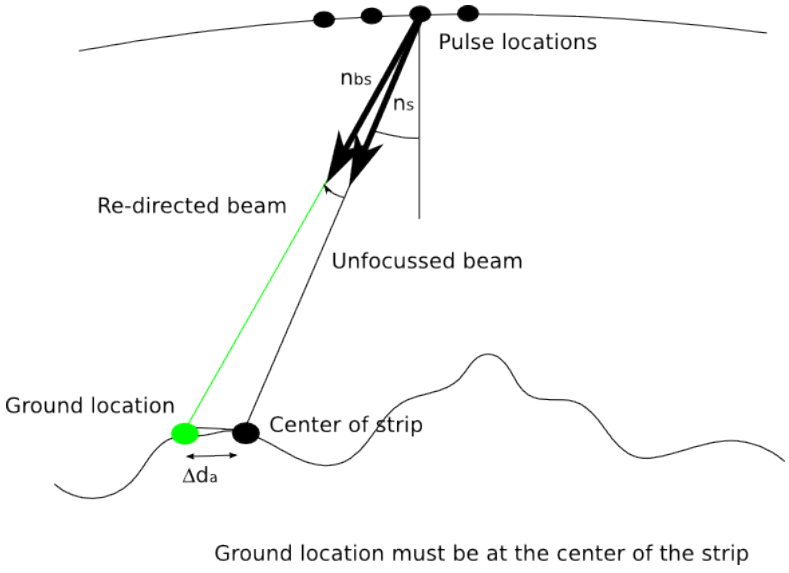


Figure 2.14: Geometry of the beamsteering correction (Adapted from [2]).

To get the signal M_{strip} per strip an along-track FFT is applied as:

$$
\begin{aligned}
M_{strip}(n) &= \mathcal{F}\{M_s(a)\} \\
&= \frac{1}{64}\sum_{a=1}^{64} M_s(a)\cdot e^{-i2\pi f_c\frac{2}{c}\Delta I(a-65/2)\alpha_n} \\
&= \frac{1}{64}\sum_{a=1}^{64} M_{burst}M_{ant}M_{pulse}\Phi_{bs}e^{-i2\pi f_c\frac{2}{c}\Delta I(a-65/2)\alpha_n} \\
&= \frac{1}{64}M_{burst}M_{ant}\sum_{a=1}^{64} M_{pulse}\Phi_{bs}e^{-i2\pi f_c\frac{2}{c}\Delta I(a-65/2)\alpha_n} \\
&= \frac{1}{64}M_{burst}M_{ant}\sum_{a=1}^{64} e^{i2\pi f_c(\frac{2\Delta I}{c}(a-65/2)n_v.n_s+\frac{2\Delta R_{bs}(a)}{c})}e^{-i2\pi f_c\frac{2}{c}\Delta I(a-65/2)\alpha_n} \\
&= \frac{1}{64}M_{burst}M_{ant}\sum_{a=1}^{64} e^{i2\pi f_c\frac{2}{c}(\Delta I(a-65/2)n_v.(n_s-\Delta n_s)}e^{-i2\pi f_c\frac{2}{c}\Delta I(a-65/2)\alpha_n}
\end{aligned}
\tag{2.52}
$$

with a the pulse position number in the burst and α_n the angle to the strip. Ultimately, this leads to the following expression for the signal of one scatterer in a strip [1]:

$$
M_{strip}(n) = \frac{1}{64}M_{burst}M_{ant}P(n_v\cdot(n_s-\Delta n_s)-\alpha_n), \tag{2.53}
$$

with

$$
\begin{aligned}
P(x) &= \frac{\sin(64f_c\frac{2}{c}\Delta Ix)}{64\sin(f_c\frac{2}{c}\Delta Ix)} \\
&\approx \frac{\sin(64f_c\frac{2}{c}\Delta Ix)}{64f_c\frac{2}{c}\Delta Ix}
\end{aligned}
\tag{2.54}
$$

The middle of the strips after beamsteering are located at $n_v\cdot\Delta n_s+\alpha_n$, with n assuming al values $\mathbb{Z}[-31,32]$. The resulting gain pattern in the along-track direction of the strip is shown in figure 2.16.

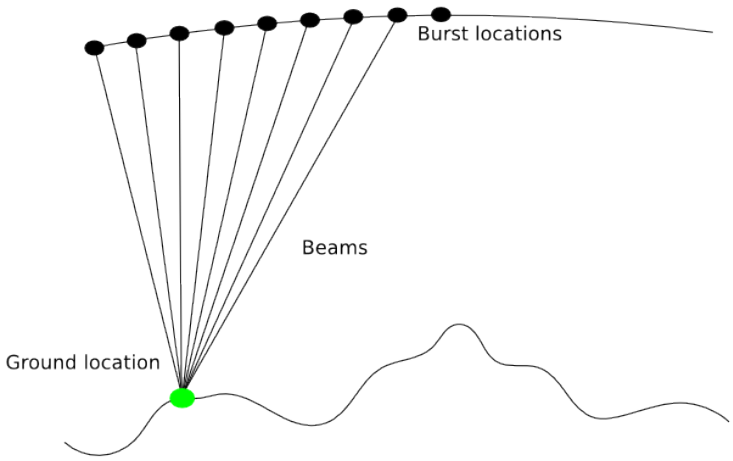


Figure 2.15: Beamsteering corrects the path of the beams to the ground location (Adapted from [2]).

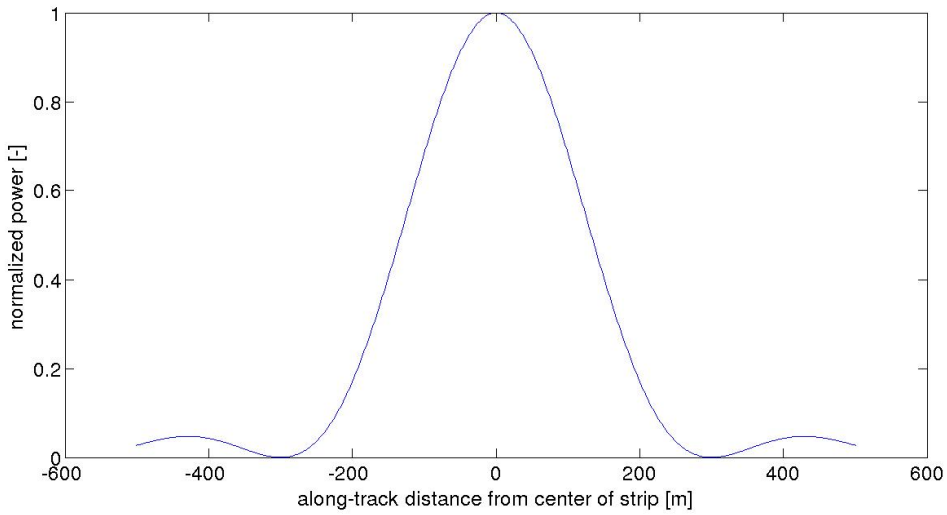


Figure 2.16: Power behaviour along-track in a strip (Adapted from [1]).

Doppler corrections

The implementation of the Doppler correction is not described in literature. Since the echoes contain low frequency signals that depend on the range, the spectrum of these frequencies forms a waveform. This spectrum can be retrieved using a range FFT,

which for LRM is already performed onboard as discussed in paragraph 2.2.3. However, before applying the range FFT first several corrections are applied, like the Doppler correction.

The Doppler effect is depending on the along-track position with respect to the satellite nadir point and is described by Eq. (2.8). As discussed in paragraph 2.2.2 the mixed signals are affected by Doppler. Therefore a different Doppler correction is applied for each individual beam in the stack [2].

Eq. (2.28) shows that the number of cycles, i.e the phase, of the signal depends on $t - t_s$. In which t_s is the arrival time of the signals from the scatterer at the receiver. The echo however contains signals from various scatterers, which all have a different arrival time.

In order to omit this problem, is will be shown that the effect of having a different arrival time can be neglected. First of all, Eq. (2.28) (the phase change) depends on the Doppler effect $D - 1$. The maximum value for $D - 1$ is computed in paragraph 2.2.2 as $5.23 * 10^{-7}$. Secondly, the phase change depends on the central frequency which is 13.575 GHz. Finally, it must be shown what the range of the values for t_s is. The resolution of SIRAL, or the binsize, computed in paragraph 2.2.2 is 3.125 ns. As discussed in paragraph 2.2.5 the number of complex samples of the echo in SARInM is 512. After applying a range FFT, this will lead to 512 range bins. The time difference between the last and the first bin is $512 \cdot 3.125 = 1600$ ns, so t_s ranges from t_r to t_r+1600 ns. The difference in value for $-(D-1)f_c \cdot (t - t_s)$ using the first and last bin is only 0.01 cycle and therefore t_s is set to t_r, the time at which the deramping chirp is generated. As a result, the correction for the Doppler effect can be written as:

$$-f_{m,Doppler} \cdot (t - t_s) \approx -(D-1)f_c \cdot (t - t_r) \tag{2.55}$$

using Eq. (2.28) and Eq. (2.30). Remember that the Doppler effect is related to the along-track look angle α as shown in Eq. (2.8).

Like for beamsteering it is again possible to turn this into a phase multiplier $\Phi_{Doppler}$ for the Doppler correction, using Eq. (2.38):

$$\boxed{\Phi_{Doppler} = e^{-i2\pi(D-1)f_c(t-t_r)}.} \tag{2.56}$$

This phase multiplier is used to compute Doppler corrected samples.

Slant range correction and range FFT

A second correction is applied before the range FFT is computed for the echoes. This is a slant range correction, of which the geometry is shown in figure 2.17. In the figure an along-track profile is sketched with a reference range, which is the range between the nadir point of the satellite and the antenna.

The aqcuisition of the echoes in the stack is at different burst locations. As a consequence the ranges vary, which makes it impossible to average them [2][24][1]. The slant range correction removes the additional range measured at the burst compared to the reference range.

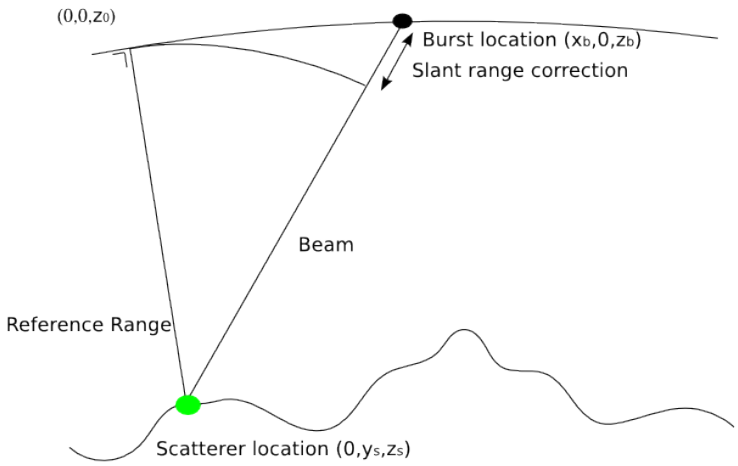


Figure 2.17: Geometry for slant range correction (Adapted from [2]).

The origing of the reference system used is the ground location. The position of the burst is (x_b, y_b, z_b) in which x_b is the along-track position, y_b the cross-track position and z_b the radial position. The satellite orbit is considered to be a straight line, and therefore y_b can be set to zero. The position of the point in the orbit directly above the ground location, or the strip is (x_0, y_0, z_0). The strip is approximately 305 m wide and is considered to be a line [24]. Since it is on the satellite track y_0 is again set to zero and because it is above the strip x_0 is set to zero as well. The term z_0 is the elevation above the surface estimated by the tracker onboard the satellite, see paragraphs 2.2.3 - 2.2.5 [1]. Furthermore, the position of the scatterer is (x_s, y_s, z_s), in which x_s is zero, because the scatterer is somewhere on the strip. Using these positions and Eq. (2.1) the two-way travel time t_s is written as:

$$t_s = \frac{2}{c}(x_b^2 + y_s^2 + (z_0 + (z_b - z_0) - z_s(y_s))^2)^{1/2}, \tag{2.57}$$

with c the speed of light[24]. Since the look angles of Cryosat are relatively small a Fresnel approximation is applied, which yields [24]:

$$t_s \approx \frac{1}{c}(2z_b + \frac{x_b^2}{z_0} + \frac{y_s^2}{z_0} - 2z_s(y_s)). \tag{2.58}$$

In order to compute the reference range a correction is applied to the slant range. The slant range correction Δt_{src} is computed with:

$$\begin{aligned} \Delta t_{src} &= -\frac{1}{c}(z_b - z_0 + \frac{x_b^2}{z_0}) \\ &\approx -\frac{1}{c}(z_b - z_0 + x_b \sin(\alpha)) \end{aligned} \tag{2.59}$$

where α is the along-track look angle. Using the slant range correction, the reference time t_{ref} is computed as:

$$t_{ref} = t_s - \Delta t_{src}. \tag{2.60}$$

Every slant range correction will be an integer number plus a fraction of a range bin. The slant range correction can therefore be applied in two steps, first a fine slant range correction before the range FFT and then an integer shift of bins after the FFT [2]. As shown in Eq. (2.29) the frequency of the mixed signal is depending on the range. In the fine slant range correction $t_{src,fine}$ it is therefore necessary to slightly decrease (the reference range is shortest) the frequency. In order to correct the 128 or 512 complex samples in the echoes the phase multiplier Φ_{src} for the fine correction is written as:

$$\boxed{\begin{aligned} \Phi_{src} &= e^{i2\pi(f_m)t} \\ &= e^{i2\pi(Qt_{src,fine})t} \\ &= e^{i2\pi(Q\mathrm{mod}[\frac{1}{c}(z_b - z_0 + x_b \sin(\alpha)), t_{res}])t} \end{aligned}} \tag{2.61}$$

in which mod is the modulus and t_{res} is the time resolution, or bin size of 3.125 ns.

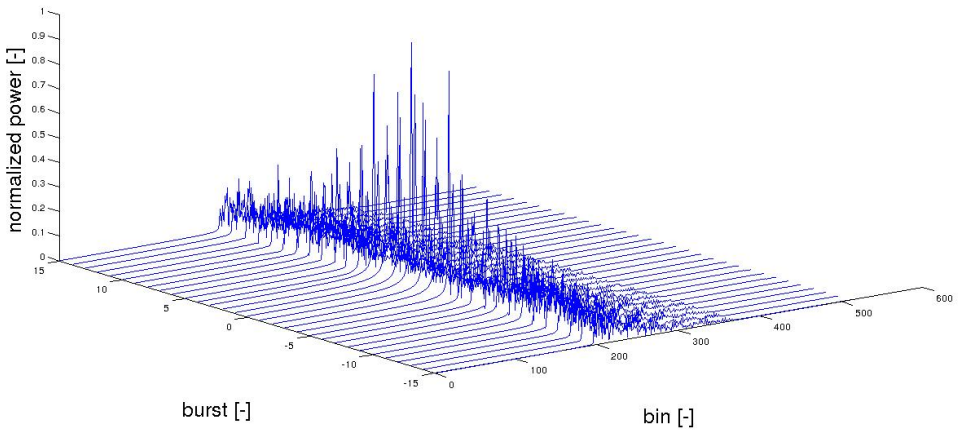


Figure 2.18: Stack of single look waveforms.

Then the range FFT is applied which results into single look waveforms. After this the range migration, or coarse slant range correction is applied to adjust their range to the reference range. Ultimately, the analytical description for a single look waveform w is:

$$w = \frac{1}{64} M_{ant} \mathcal{F}_{range} \{ M_{burst} \Phi_{src} \Phi_{Doppler} P(n_v \cdot (n_s - \Delta n_s) - \alpha_n) \} \tag{2.62}$$

where M_{burst} the signal received at the center of the burst at the center of the satellite, M_{ant} the change to this signal because it is received at an antenna left or right of the center, Φ_{src} the slant range correction, $\Phi_{Doppler}$ the Doppler correction, $P(n_v \cdot (n_s - \Delta n_s) - \alpha_n)$ the gain pattern of strip n and $\mathcal{F}_{range}$ the Fourier transform in the range direction. For every ground location the single look waveforms are accumulated in a stack. A result of such a stack is shown in figure 2.18. A detailed description of the shape of waveforms in given in chapter 3.

Multilooking

The final step in level 1b processing is the averaging of the waveforms. In the receiver chain of SARM only one antenna is used. The result is slightly different from the SARInM.

The cross product of a waveform is computed by multiplying the waveform with the conjugate of itself. In SARInM the cross-product between the left and right antenna is computed, i.e. the waveform of the left antenna is multiplied with the conjugate of the right one. If w_b^+ is waveform number b of the left antenna in the stack and $\bar{w}_b^+$ the conjugate, then the multilooked cross product of the waveform is computed as:

$$\boxed{W\bar{W}_{ml}^{(+,+)} = \frac{1}{N_L}\sum w_b^+ \bar{w}_b^+} \tag{2.63}$$

in SARInM as,

$$\boxed{W\bar{W}_{ml}^{(+,-)} = \frac{1}{N_L}\sum w_b^+ \bar{w}_b^-} \tag{2.64}$$

with N_L the number of looks or waveforms in the stack and w_b^- the waveform from the right antenna. Ultimately this leads to an averaged waveform at every ground location.

The level 1b processor computes two additional functions for the SARInM, which are the multilooked phase and the multilooked coherence. Both are used to compensate for cross-track slope. Details on this procedure are found in [1].

2.3.3 Level 2 processing

As described in the previous section, the level 1b data contains essentially one multilooked waveform for approximately every 300m in all three modes. The level 2 processor is mainly concerned with estimating an elevation from each multilooked waveform. First geophysical corrections are applied in all three different modes, but the retrieval of elevation is different for all modes [1]. The application and size of these correction are discussed in section 4.1. In this section only the level 2 processor of SARInM discussed, because the LRM and SARM modes are not activated over the Tibetan plateau. A detailed description of the retracking procedure for LRM is given in [16] and [28].

The goal in level 2 processing is the retrieval of geophysical parameters from the multilooked waveforms created by the level 1b processor, like elevation and Significant Wave Height (SWH), which is a measure for the height of ocean waves. This is done by waveform retracking, a procedure to estimate a theoretical function through the received waveform by means of an iterative least squares procedure. Therefore it is necessary to obtain a theoretical function describing the waveform with parameters related to the geophysical parameters that need to be estimated.

An analytical description of the SAMOSA level 2 retrackers are not publicly available. To show how a theoretical description might look like, another retracker is decribed here. The approximated theoretical waveform for this retracker is described accurately by [29].

Theoretical SAR waveform
It is convenient to introduce the arrival time t_s of the signal from the Point Of Closest Approach (POCA). Then the time τ_s with respect to t_s is defined as:

$$\tau_s = t - t_s \tag{2.65}$$

with t the time. All the equations in this section are described as a function of τ_s.

It is possible to define a function for the theoretical waveform $w_t(\tau_s)$ as a function of the source time function $P(\tau_s)$, the surface roughness function $O(\tau_s)$ and a flat surface response function $S(\tau_s)$ [29][30][31]. The source time function describes the response of incoming signals at a particular time due to deramping and applying a range FFT. The surface roughness function describes the probability density function of the surface, which is related to the SWH. The flat surface response function includes the effects of antenna gain and beamforming (i.e. the along-track FFT). The theoretical waveform is related to these three functions as:

$$w_t(\tau_s) = P(\tau_s) * O(\tau_s) * S(\tau_s) \tag{2.66}$$

with $*$ indicating a convolution.

In [29] a possible theoretical waveform is described that is optimized for the retrieval of elevation at cost of other parameters, like SWH. The derivation and theoretical background can be found in [29] and are based on the principles described in [30] and [31]. Ultimately the derivation yields a theoretical approximation of the waveform, which is a function of the amplitude A, the rise time σ, the time τ_s with respect to arrival of the signal from the POCA and trailing edge decay β. Figure 2.19 shows an example of the waveform with the parameters to be estimated. The function of the theoretical waveform is given as:

$$\boxed{w_t(\tau_s) \approx A\sigma^{1/2} e^{-\frac{\tau_s^2}{4\sigma^2}} D_{-1/2}(-\frac{\tau_s}{\sigma}) e^{-\beta\tau_s},} \tag{2.67}$$

with $D_{-1/2}(-\frac{\tau_s}{\sigma})$ the parabolic cylinder function [29]. The parabolic cylinder function is elaborated in Appendix A.

The physical quantities responsible for the trailing edge decay are mainly the antenna gain and illuminated area. The amplitude is a function of antenna gain, atmospheric effects and a backscattering coefficient. Then the rise time is a function of SIRAL's resolution and is related to the surface roughness, i.e. SWH over water [29]. Of particular interest for lake level measurements is the t_s, which is related to τ_s by Eq. (2.65) and to the range by Eq. (2.1).

Waveform retracking

Eq. (2.67) shows that the function for the theoretical waveform is non-linear with respect to parameters β, σ and τ_s. In order to fit the theoretical waveform using a least-squares solution, an iterative Newton-Raphson approach is used [32]. This method

is sometimes called the Levenberg-Marquardt method [33]. If all four parameters are estimated, the retracker is called the Most-Likelihood-Estimator 4 (MLE4)-retracker. Other solutions are also possible, in case of ocean surfaces for example β is sometimes held constant.

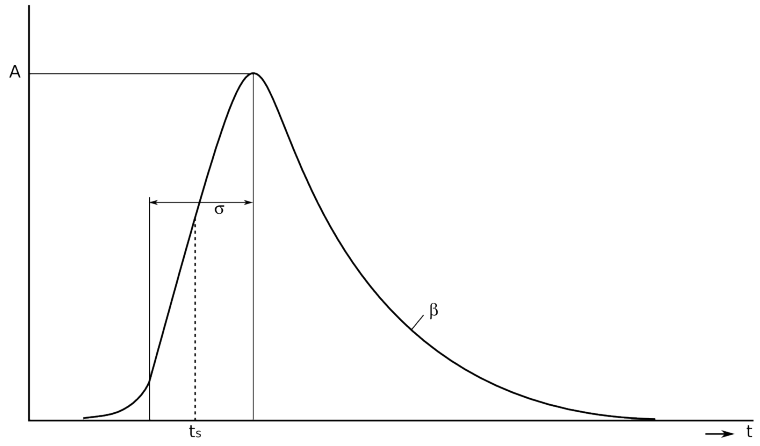


Figure 2.19: Theoretical waveform and visual respresentation of the parameters. The term τ_s is the time from t_s.

Let Y be the measurements, i.e. the power of 512 bins in the measured waveform, then a Taylor expansion can be made around Y_0 as:

$$Y = Y_0 + \frac{\delta Y}{\delta x}\Delta x \tag{2.68}$$

in which x is the vector with the parameters to be estimated: β, A, σ and t_s. Y_0 is computed, using initial values for the aforementioned parameters. The update for the parameters in x is computed using the least squares solution

$$\Delta\bar{x} = (L^T L)^{-1} L^T (Y - Y_0), \tag{2.69}$$

with

$$L = \frac{\delta Y}{\delta x}. \tag{2.70}$$

The estimates of the parameters are updated iteratively as:

$$x_{i+1} = x_i + \Delta x_i. \tag{2.71}$$

The method requires us to compute the derivates of the theoretical waveform with respect to the four estimated parameters. These are given by [29]:

$$\frac{\delta w_t}{\delta \tau_s} = A\sigma^{-3/2} e^{-\frac{1}{4}z^2} D_{-1/2}(z) e^{-\alpha\tau_s}, \tag{2.72}$$

$$\frac{\delta w_t}{\delta \sigma} = -A\sigma^{-3/2} e^{-\frac{1}{4}z^2} [\frac{1}{2} D_{-1/2}(z) - z D_{1/2}(z)], \tag{2.73}$$

$$\frac{\delta w_t}{\delta A} = \frac{w_t}{A}, \tag{2.74}$$

$$\frac{\delta w_t}{\delta \beta} = -\beta A\sigma^{-1/2} e^{-\frac{1}{4}z^2} D_{-1/2}(z) e^{-\beta\tau_s}, \tag{2.75}$$

with

$$z = -\tau_s/\sigma. \tag{2.76}$$

2.4 Conclusions

In this chapter the answer is given to the subquestion: **What are the measurement principles of Cryosat?**.

Cryosat is a satellite specifically designed to make measurements over the Cryosphere. It has two major mission objectives:

- Recovering regional and basin-scale trends in Arctic sea ice thickness and total mass.
- Estimate the contribution of shrinking of Antarctic and Greenland ice sheets to the global sea level.

In order to achieve these objectives it has radar altimeter onboard named SIRAL, which operates in three modes, depending on the terrain. LRM, with a footprint diameter of 15-20 km, is activated over open ocean and the middle of the ice sheets of Greenland and Antarctica, where the terrain is smooth. SARM, with an improved along-track resolution of 300 m, is activated over ice shelfs (which is floating ice), to measure the freeboard. SARInM, using two antennas, is activated over rough terrain like Tibet, to have better along-track resolution and cross-track slope measurements.

In all three modes a chirped pulse is transmitted. The received pulses are deramped. The frequency of the deramped signal depends on the range towards the scatterer. The deramped echoes are sampled into 128 complex values for LRM and SARM and into 512 complex values for SARInM. In LRM 91 consecutive burst are averaged, FFT'd into a

waveform and downlinked to the ground station. In SARM and SARInM the pulses are transmitted in burst of 64 pulses. The slight phase shifts between these pulses make it possible to distinguish 64 beams, corresponding to 300 m along-track strips. Per burst 64 pulses of 128 or 512 complex values are transmitted to the ground station. In SARInM this is performed for both antennas.

In level 1 processing instrumental correction are applied and measurement records are made in a specified form. The level 1b processor uses an along-track FFT to discriminate strips of 300 m along-track and 15-20 km cross-track. Before this is applied a beamsteering operation ensures that the beams are focussed to one ground location. Additionally, slant range corrections and Doppler correction are applied centered at one of the ground locations (measurement points). Then another FFT is applied to retrieve a waveform of 128 complex or 512 complex values. The final step in level 1b processing it multilooking, leaving only one averaged waveform per ground location.

The level 2 processor estimates a theoretical shape in the multilooked waveform. It optimizes several parameters, which are related to the elevation, surface roughness, etc. Finally, it computes geophysical parameters (like elevation) using geophysical corrections, the satellite location and the retracking parameters.

Chapter 3

Waveform Modelling Algorithm

This chapter describes a proposed algorithm to simulate a Cryosat waveform in SARInM over a water surface. It primarily answers the question: **How accurately can one model a Cryosat waveform?** The methods and results are presented in five sections. The first section addresses the relevant reference systems. Section 3.2 describes step-by-step the algorithm to simulate an echo and and proposed processing methods to get a multilooked waveform. Section 3.3 shows some special cases of simulated waveforms. Section 3.4 validates the algorithms and waveforms by comparing the latter to real level 1b waveforms. And finally, section 3.5 draws conclusions from the findings in the chapter.

3.1 Reference Systems

In order to use the equations described in chapter 2, two conversions of coordinates are necessary. The coordinates of the satellite and scatterers are given with respect to the WGS84 ellipsoid in geodetic coordinates. The satellite velocity vector is given in an Earth Centered Earth Fixed (ECEF) reference frame in Cartesian coordinates, which means that the origin is at the Earth's center and the x-axis is always pointing through the zero meridian at the equator, the z-axis points to the pole and the y-axis orthogonal to the xz-plane. The first step is converting the geodetic coordinates of the satellite positions and measurement positions into Cartesian coordinates. The second step involves converting the coordinates to a local Along-track Cross-track Up (ACU) coordinate system with the origin at the ground location, which is used in the equations in section 2.3.2.

3.1.1 Earth Centered Earth Fixed Cartesian system

Let us assume that the position of a scatterer in the WGS84 system has a latitude θ_s, a longitude λ_s and elevation h_s. The coordinates (X_s, Y_s, Z_s) in the ECEF system are computed as [3]:

$$X_s = (N(\theta_s) + h_s)\cos(\theta_s)\cos(\lambda_s), \tag{3.1}$$

$$Y_s = (N(\theta_s) + h_s)\cos(\theta_s)\sin(\lambda_s), \tag{3.2}$$

and

$$Z_s = (N(\theta_s)(1 - e^2) + h_s)\sin(\theta_s), \tag{3.3}$$

where

$$N(\theta_s) = \frac{a_e}{\sqrt{1 - e^2\sin^2(\theta_s)}}, \tag{3.4}$$

which depends on the semi-major axis a_e of the WGS84 ellipsoid and its eccentricity e.

3.1.2 Along-track Cross-track Up system

The conversion from the ECEF system to the ACU system involves two steps. The first step is a conversion from the ECEF system to a North East Up (NEU) frame with axis (x'_s, y'_s, z'_s) as shown in figure 3.1. In this figure, a_e and b_e represent the semi-major and semi-minor axis of the reference ellipsoid. In the second step, a rotation is applied around the z'_s-axis, in order to have the x-axis parallel to the local satellite flight direction.

A conversion from the ECEF system to the NEU system involves a translation and two rotations. The origin of the NEU system is located at the ground location. Therefore the position of the ground location in the ECEF system (X_g, Y_g, Z_g) is subtracted from the position of the current point in the ECEF system.

The coordinates are rotated around the z-axis over the longitude θ_g, so the y-axis will point East. The operation of the first step is rotation around the y-axis over the angle $\frac{\pi}{2} - \lambda_g$, where λ_g is latitude, to set the z-axis in the up direction.

Ultimately this leads to the relation between the ECEF coordinates and the NEU coordinates:

$$\begin{bmatrix} x'_s \\ y'_s \\ z'_s \end{bmatrix} = \begin{bmatrix} -\sin(\theta_g)\cos(\lambda_g) & -\sin(\theta_g)\sin(\lambda_g) & \cos(\theta_g) \\ -\sin(\lambda_g) & \cos(\lambda_g) & 0 \\ \cos(\theta_g)\cos(\lambda_g) & \cos(\theta_g)\sin(\lambda_g) & \sin(\theta_g) \end{bmatrix} \begin{bmatrix} X_s - X_g \\ Y_s - Y_g \\ Z_s - Z_g \end{bmatrix}. \tag{3.5}$$

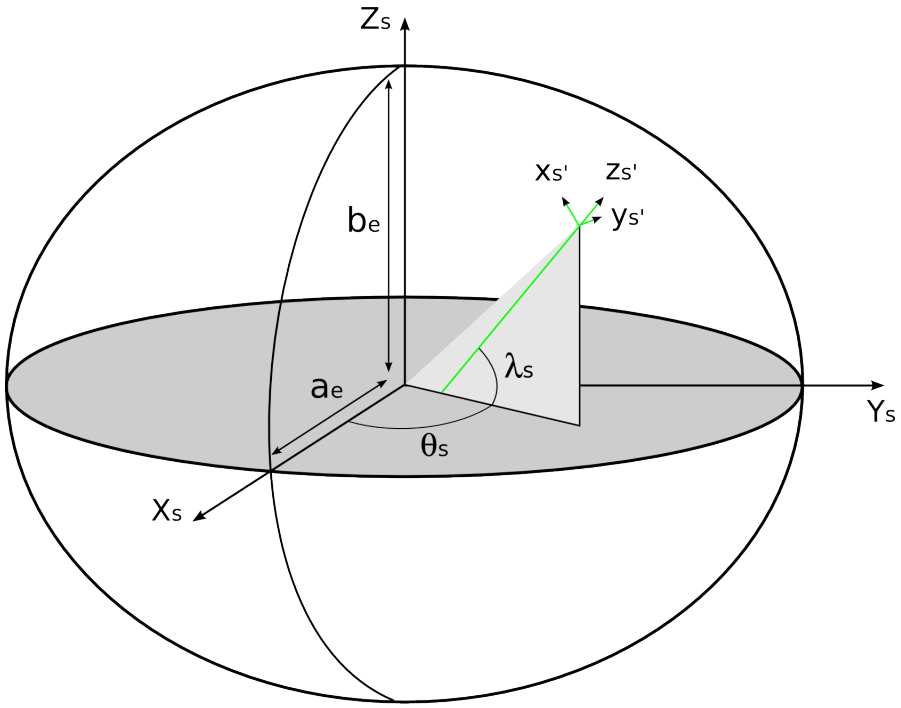


Figure 3.1: The NEU reference system with axes (x'_s, y'_s, z'_s) versus the ECEF reference system with axes (X_s, Y_s, Z_s). The parameters a_e and b_e represent respectively the semi-major (equatorial) and semi-minor (polar) axis of the reference ellipsoid.

The satellite is not flying exactly towards the North and therefore an additional coordinate conversion is required, which is performed in the second step. This involves a rotation around the z-axis through the flight direction angle ϕ_V measured clockwise from the North. The position in the ACU system is defined as (x_s, y_s, z_s). Computing the ACU coordinates from the NEU coordinates yields:

$$\begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} = \begin{bmatrix} \cos(\phi_V) & \sin(\phi_V) & 0 \\ -\sin(\phi_V) & \cos(\phi_V) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x'_s \\ y'_s \\ z'_s \end{bmatrix}. \tag{3.6}$$

3.2 Description of the algorithm

This section describes the designed Cryosat SARInM waveform modelling algorithm and software. The input for the program are physical constants, Cryosat parameters and the WGS84 ellipsoidal parameters, and a file containing the measurements locations and state vectors of the satellite from the level 1b data files. The output of the algorithm is a multilooked waveform for a specified ground location. The algorithm is based on the results and equations found in chapter 2. In appendix B a detailed description of the software is provided. An overview of the steps of the algorithm is shown in figure 3.2. The grey boxes indicate the input and output of the program.

Before starting the algorithm a ground location is determined over which the algorithm simulates a waveform. Using the measurement locations and state vectors of the satellite, a set of 64 'start-stop' locations per bursts closest to the ground location are computed in the ECEF system (described in section 3.1). Then at the vicinity of the latitude-longitude location a grid of scatterer positions is determined in the ECEF system. The elevation of these scatterers are either determined by the user or read from ASTER-GDEM data. In order to use the theory from chapter 2, the coordinates of the 'start-stop' locations, scatterers and bursts are converted to the ACU system (described in section 3.1).

3.2.1 Level 0 processing

Level 0 processing describes the process of simulating satellite measurement data. The satellite measurement data in SARInM consist of 512 samples long, complex deramped signals. The analytical description of these signals is given by Eq. (2.35) and Eq. (2.36). Since, no level 1 processing is taken into account in the program, this serves as input for the level 1b processing programs. Here it is described how to simulate a deramped signal

As shown in figure 3.2, from here on all measurements for every burst are treated seperately and eventually all the single look waveforms are stored in a stack. The loop runs through all the bursts, which is normally 57 bursts, since this is the average number of waveforms to be multilooked.

In the first step of the loop the along-track and cross-track angle of the scatterers with respect to the satellite antenna are computed. This is required to compute the antenna gain and the Doppler effect. Details on the procedure are provided in appendix B. Secondly, the travel time is computed, which is required to determine the deramped signal.

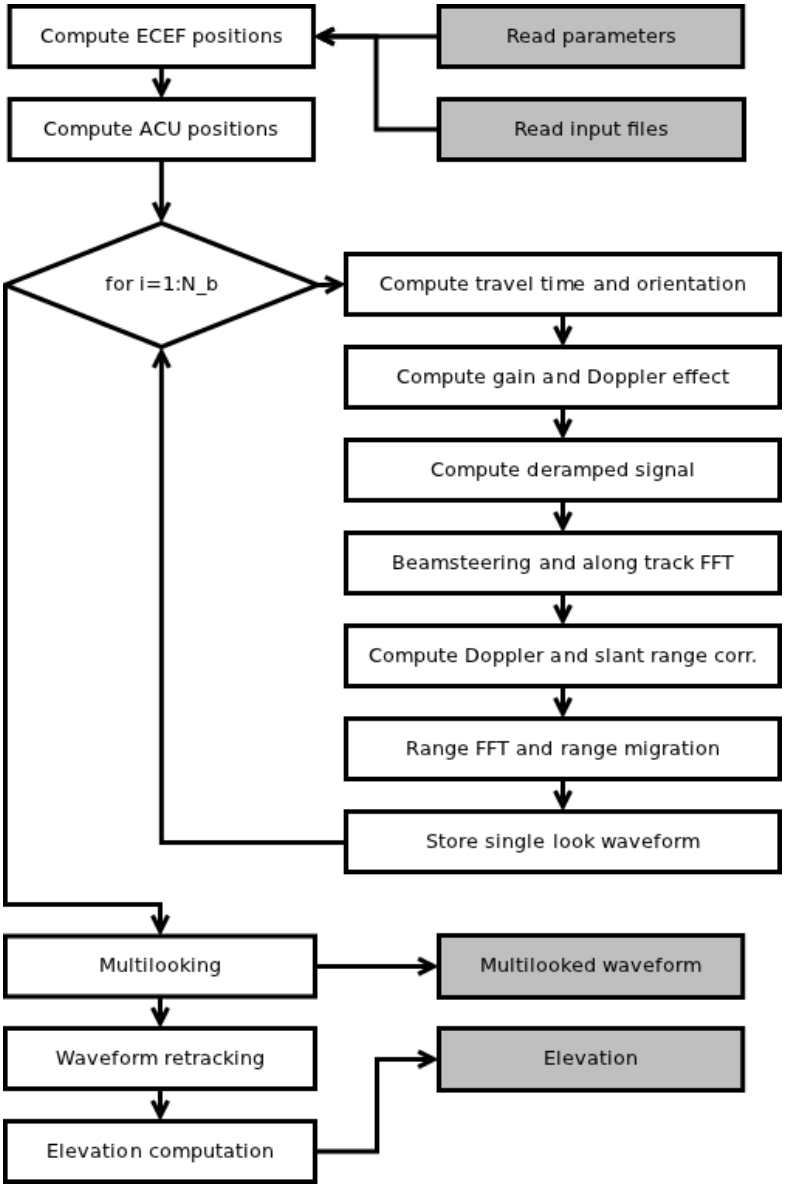


Figure 3.2: Flowchart of the simulation algorithm.

The computations of the antenna gain are based on Eq. (2.9). The amplitude of the signals depends on more parameters than only antenna gain as shown in Eq. (2.11). In this report modelling the shape of the waveform is more important than computing the actual amplitude, therefore path loss and transmitted power are not considered. The amplitude is adjusted in the retracking procedure described in chapter 4. However, ff-nadir scatterers especially on smooth surfaces, have generally a lower backscattering coefficient, which changes the shape of the waveform. This effect is considered and

discussed in sections 3.3 and 4.3.

Additionally, the Doppler effect is computed using the along-track angle and satellite velocity, which is computed from the satellite state vectors. Eq. (2.8) shows the computation of the Doppler effect.

The Doppler effect, antenna gain and travel time are input at the step of computing the deramped signal. The Deramped signal is computed for every 'start-stop' (pulse) location in the burst, so 64 signals per burst are stored in computer memory.

Figure 3.3 shows a deramped signal for a single scatterer, three scatterers and a set of scatterers. The deramped signals are supposed to be of constant frequency depending on their range. Therefore for one scatterer only a signal with one frequency is visible. For three and 25 scatterers the number of frequencies increase to respectively three and 25.

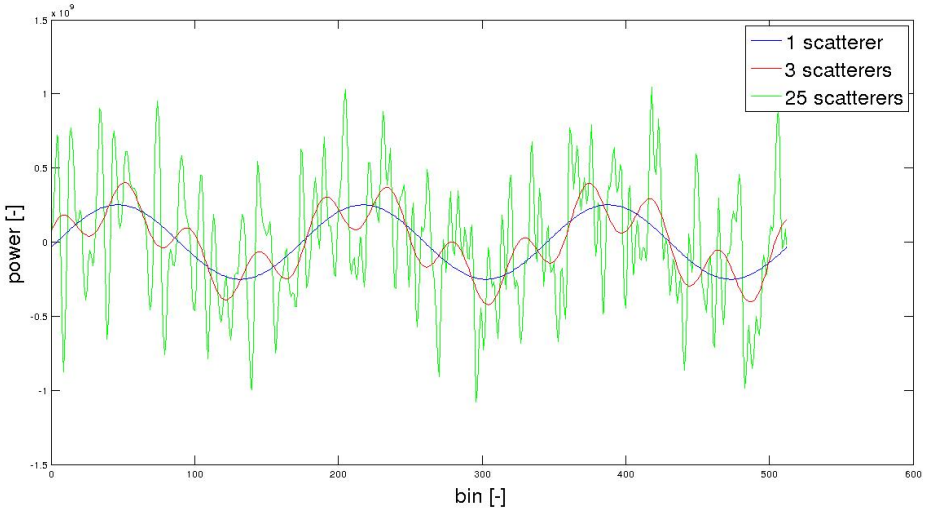


Figure 3.3: Deramped signals for one scatterer, three scatterers and 25 scatterers.

3.2.2 Level 1b processing

For every burst the level 1b step starts with a beamsteering procedure for the deramped signals and ends with storing a single look waveform. Eventually the stack of single look waveforms are combined into a multilook waveform after the for-loop ends. The part of the algorithm dedicated to level 1b processing follows the theory described in section 2.3.2.

In the beamsteering procedure a pre-FFT phase shift is applied to focus the beams created with the along-track FFT towards the ground location. The formula for computing

the phase shift per 'start-stop' location is given by Eq. (2.47). A more detailed description on the implementation of the procedure is described in appendix B. By applying the along-track FFT to the 64 'start-stop' location echoes, 64 signals corresponding to 64 strips of approximately 300 m along-track are created. After applying the FFT, only the signal corresponding to the beam that is directed to the ground location is kept.

In the next step, a phase shift is applied to compensate for the error of the Doppler effect on the range. Furthermore, a fine slant range correction is applied for the additional range due to the along-track position of the satellite with respect to the ground location. The coarse slant range correction, or range migration, is applied after the range FFT. Eq. (2.56) and Eq. (2.61) describe the computations for the phase shifts of respectively the Doppler effect and the slant range correction. Again a more detailed description on the implementation is found in appendix B.

The range FFT turns the signal into a single-look waveform. By applying range migration, discussed in section 2.3.2 the waveforms are ready to be multilooked.

Stacks of waveforms after range migration are shown in figure 3.4. In the figure the waveform at burst '0' shows the waveforms retrieved over a flat surface over a strip nadir the satellite. The bursts 1-15 show waveforms computed from bursts further along-track. The other bursts show waveforms computed from burst behind the nadir strip. Around burst '0' the waveforms have most power, due to the satellite gain pattern.

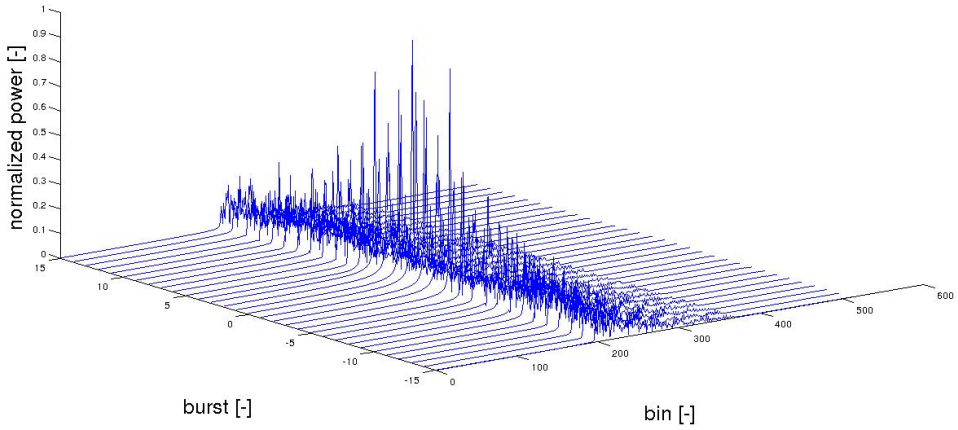


Figure 3.4: Stack of 31 single look waveforms.

After multilooking the single look waveforms, which is described by Eq. (2.63), only one waveform for the given ground location is produced. An example is shown in figure 3.5.

3.2.3 Level 2 processing

In level 2 processing an elevation is retrieved from the level 1b data. This part of the processing is described in the chapter 4, where we use the simulated waveform as a retracker for real level 1b waveforms.

How, retracking is applied to get Cryosat level 2 data is described in section 2.3.3. A theoretical waveform is fit into the simulated multilooked waveform, computed in the previous section. Four parameters, described in section 2.3.3, are optimized using a Newton-Rhapson iteration. The implementation of this procedure is sometimes called the Levenberg-Marquardt algorithm [33].

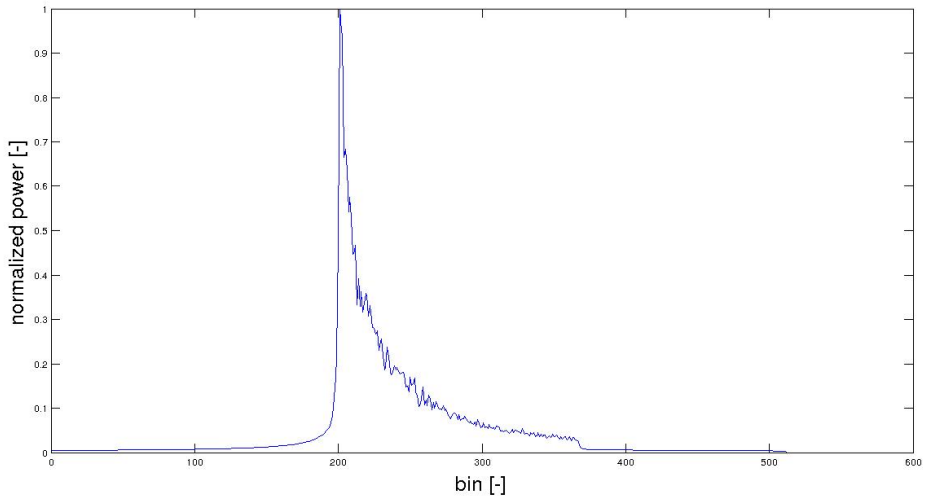


Figure 3.5: Multilooked waveform computed from 31 single look waveforms.

3.3 Waveform simulations

A waveform shape is affected by backscattering, processing procedures and the geometry of the surface. In order to get a better understanding of the response of the waveform to changes in these parameters, several waveforms in various situations are simulated. First, a waveform from a perfectly specular reflection is simulated, i.e. only one scatterer nadir the satellite. This particularly gives inside in the effects of the processing procedures on the waveform shape. Then, a waveform over a flat surface is computed. This particulary gives insight in the effects of choosing a scatterer resolution and the effects of speckle. Finally, the surface wave height and cross-track slope are varied to see the effects of the geometry of the surface on the resulting waveforms.

3.3.1 Single scatterer respronse

The single scatterer response is simulated from one scatterer exactly nadir the ground location. It particularly addresses the effects of the along-track and range FFT on the shape of the waveform.

Figure 3.6 shows a stack of 31 waveforms for one ground location. Every waveform is computed from 64 received signals from one bursts. The measurement window is opened such that the signal would be retrieved at bin 201 (if the first bin has number 1). In the figure it becomes clear that the signals are centered around bin 201 as expected.

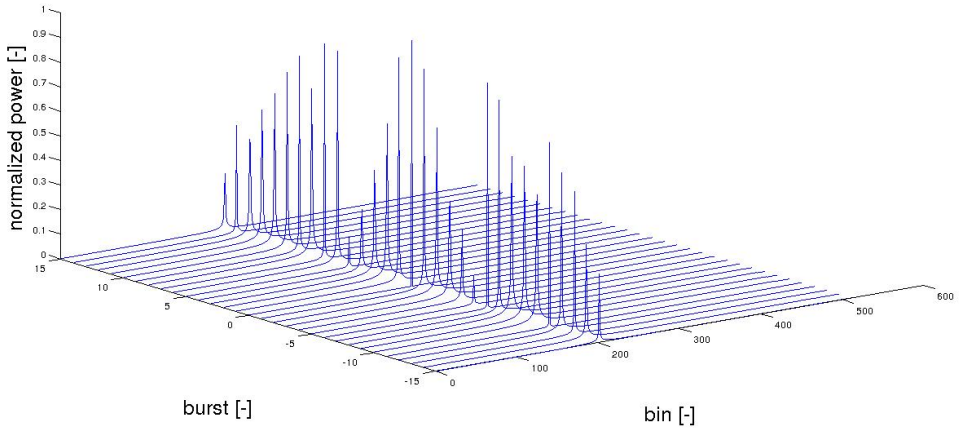


Figure 3.6: Stack of 31 waveforms for a specified ground location.

The position of the satellite at burst number 0 is exactly above the ground location. In figure 3.6 amplitude variations are visible over the received waveforms, with the highest amplitude at burst 0. There are two effects causing these amplitude variations: antenna gain and signal processing. Except for burst 0 the satellite is in along-track direction offset from the scatterer. This increases the polar angle with respect to the antenna boresight, and therefore the antenna gain decreases (Eq. (2.9)). As a consequence the amplitude decreases with increasing forward or backward along-track distance of the satellite with respect to the scatterer.

The other variations are explained by the signal processing effect. The response of an FFT in the frequency domain is a sinc-function. As explained in section 2.2.4, along-track positions of a scatterer cause a phase shift in the 64 signals. Therefore 300 meter along-track strips can be seperated using an FFT. If the scatterer is located exactly at a multiple of 300 meters, it falls in the middle of the strip and the maximum amplitude is retrieved. It could also be that the scatterer is located at the edge of a strip, therefore

the amplitude retrieved is less than the maximum. In order to clarify this, figure 3.7 shows the rock angles, i.e. the angles over which the beams are steered to the ground location. Since the scatterer is exactly at the ground location, the size of the rock angle can directly be coupled to the amplitude. At burst 6, the rock angle is large and therefore the amplitude in figure 3.6 is relatively low.

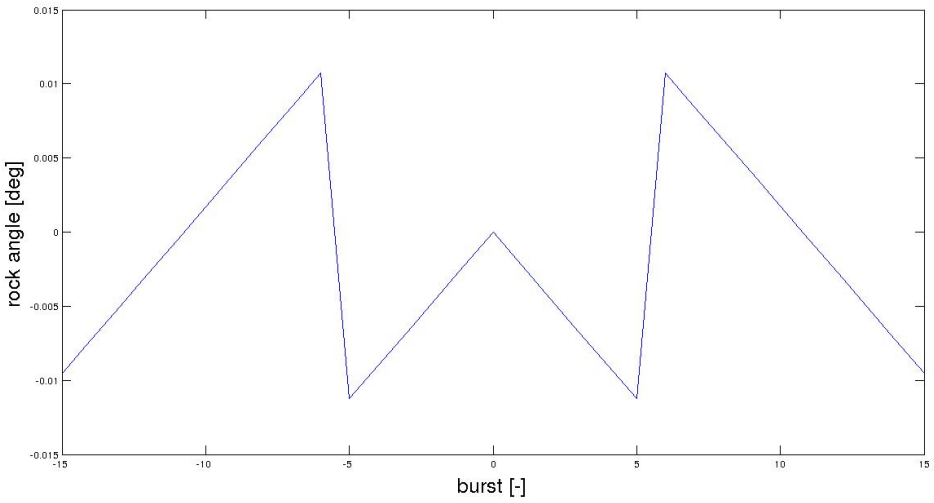


Figure 3.7: Rock angle for the 31 waveforms in the stack.

A similar effect occurs when applying the range FFT. The ranges correspond to a frequency in the deramped signal as explained in section 2.2.2. Scatterers that have a range, which corresponds exactly to one of the frequencies computed with the range FFT have a high amplitude. Scatterers with a slight frequency offset, will have a lower amplitude, due to the sinc-function response. Due to the sinc-function response, other bins in the waveform are affected as well, this is visible as a 'toe' or a trailing 'toe' in the waveform. In figure 3.6 no toe is visible at burst 0, but in the other bursts there is.

Ultimately, by multilooking this leads to a waveform shape as shown in figure 3.8. In the figure the effect of the range FFT becomes clear, by the presence of a starting toe and a trailing toe. The peak is like expected centered at bin 201.

3.3.2 Flat surface response

This paragraph investigates the shape of the waveform from a flat surface response. The effects of choosing a specific resolution of scatterers on the surface and the effects of multilooking for speckle reduction is shown. Furthermore, an explanation for the shape of the multilooked and single look waveforms is given.

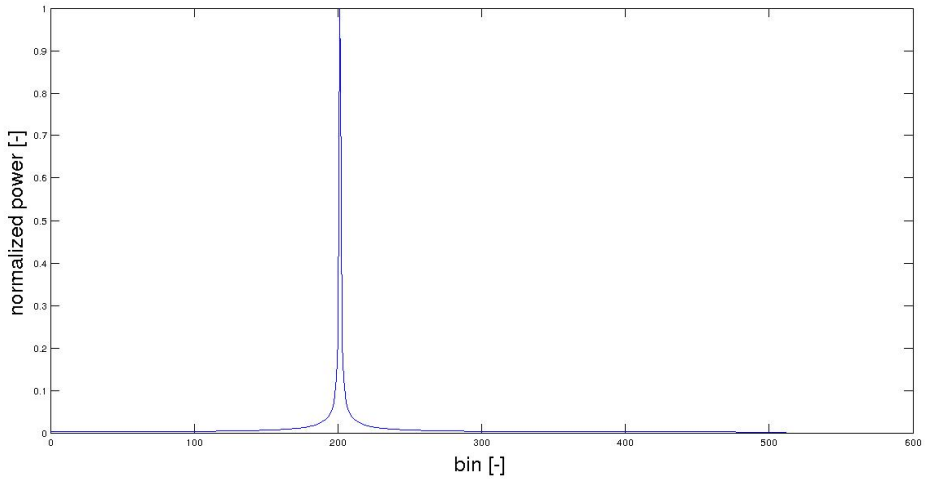


Figure 3.8: Image of a multilooked waveform of a single scatterer at the ground location.

Figure 3.9 shows a comparison of three single look waveforms with a multilooked waveform. The power of the waveform is normalized with respect to the peak power in the waveform. Waveform 0 respresent the acquisition at the point exactly above the ground location. Waveforms 6 and 12 are both at an along-track distance from the ground location.

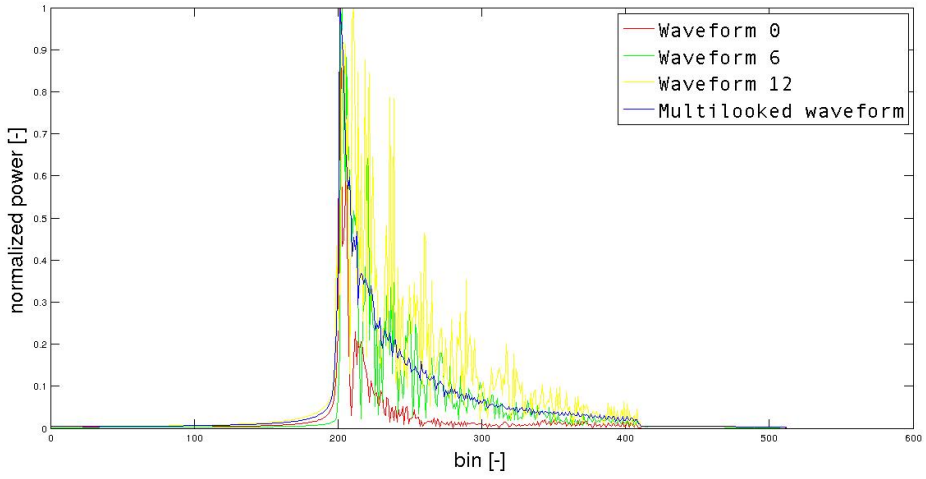


Figure 3.9: Image of a multilooked waveform of a single scatterer at the ground location.

Figure 3.10 illustrates which part of the ground is captured at bins indicated with numbers. In comparison with traditional altimeters, where the area captured per bin remains

approximately the same, for SAR altimeters the area significantly decreases. In the figure this becomes clear by looking at the size of the area (in the nadir strips) corresponding to bins 201, 202 and 203. Therefore a distinguishable peak is visible in the multilooked waveform. The area corresponding to bin 201, i.e. the peak, is called the Pulse Limited Footprint (PLF). Additionally, the antenna gain decreases the power for the scatterers cross-track and therefore the power decays even more after bin 201.

If the waveform 0 from figure 3.9 is compared to the along-track waveforms 6 and 12, it is visible that waveform 0 has a larger decay after bin 201. In figure 3.10 the strip for an along-track waveform is given as well. As is visible in the figure, the size of the area corresponding to the range bins does not always decrease fast. It depends on the width of the range rings (circles of equal range) and the position of the along-track strip with respect to the range rings. In this case it could be seen that the power at the range bins after bin 201 is higher than in range bin 201. As a result a slower decay is found for waveforms 6 and 12, but also more variability at the trailing edge, i.e. the along-track waveforms are less clean. By adding all the waveforms the multilooked waveform is computed. Due to the effects discussed above, the multilooked waveform will have a faster decay after the leading edge than most along-track waveforms, but a slower decay than the nadir waveforms.

The reason why multilooking is applied is speckle reduction. As visible in figure 3.9 the multilooked waveform is cleaner than the nadir waveform. If two scatterers within one range bin are π radian out of phase the power of both cancel, which is called destructive intereference. If two scatterers within one range bin are in-phase the powers are added, which is called constructive interference. Due to the uniform distribution of the probability of phases of scatterers, it is possible that constructive interference and destructive interference lead to peaks and troughs in single look waveforms. The probability density function for the power in a bin (assuming equal amplitudes for all scatterers captured within the bin) is a negative exponential for single look waveforms [34]. The probability density function p_t is related to the intensity I (or power) as [34]:

$$p_t(I) = \frac{1}{\bar{I}} e^{-\frac{I}{\bar{I}}}, \tag{3.7}$$

where $\bar{I}$ is the mean intensity. By applying multilooking over increasingly more waveforms the probability density function begins to resemble a Gaussian density function according to the Central Limit Theorem [34].

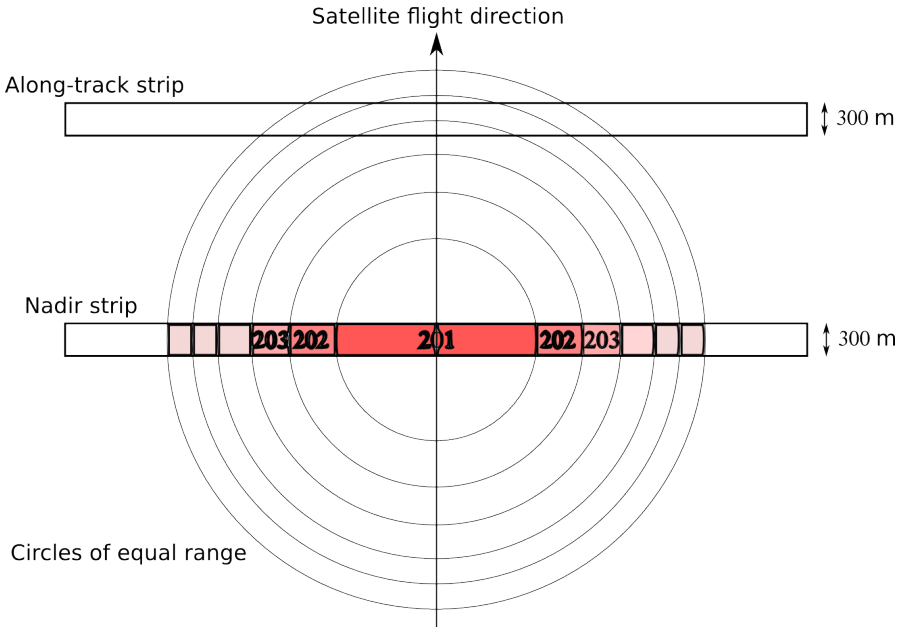


Figure 3.10: Illustration of what part of the surface is captured per bin at a flat surface. The first signal is received in bin number 201.

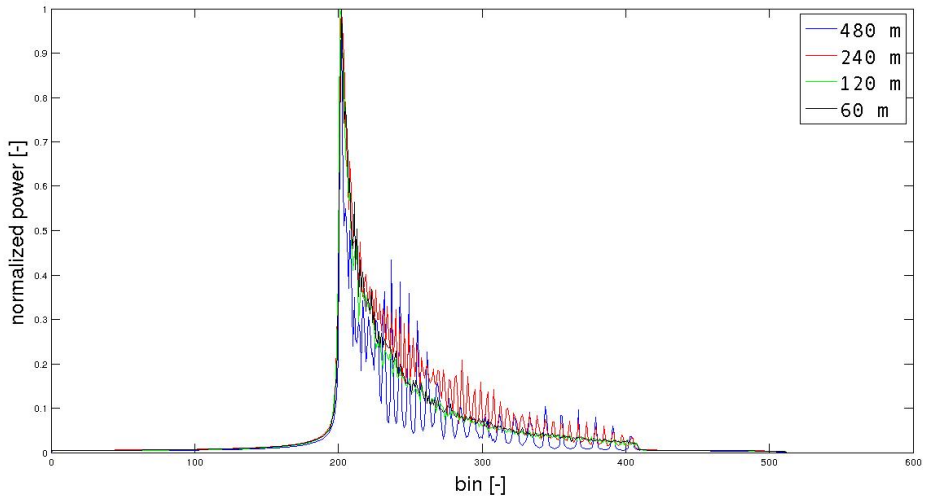


Figure 3.11: Comparison of multilooked waveforms simulated with various horizontal resolutions for scatterers.

Another variable to consider is the spacing of the scatterers. Figure 3.11 shows the effect on the multilooked waveform of taking various horizontal resolution of scatterers. If the

spacing of scatterers is decreased to 480 m, the quality of the multilooked waveform drops. It shows significant speckle after the leading edge. As already mentioned the area corresponding to the bins after bin 201, significantly decreases. Choosing a low resolution causes the number of scatterers in these bins to be small. Large constructive and destructive interference are therefore more likely to occur, which leads to increased speckle.

At horizontal resolutions of 60 and 120 m the waveform appears to be quite clean. A higher resolution increases the quality of the waveform, however this also increases the computational time. Therefore a resolution of 120 m is taken as the optimal one.

3.3.3 Effects of water waves

The surface of a lake is most of the time not completely flat. Water waves are causes by water dynamics, bathymetry and wind. The change in surface roughness due to water waves is modelled according to Eq. (B.2). Surface roughness influences the shape of the waveform and has therefore an effect on the computations of the retracker.

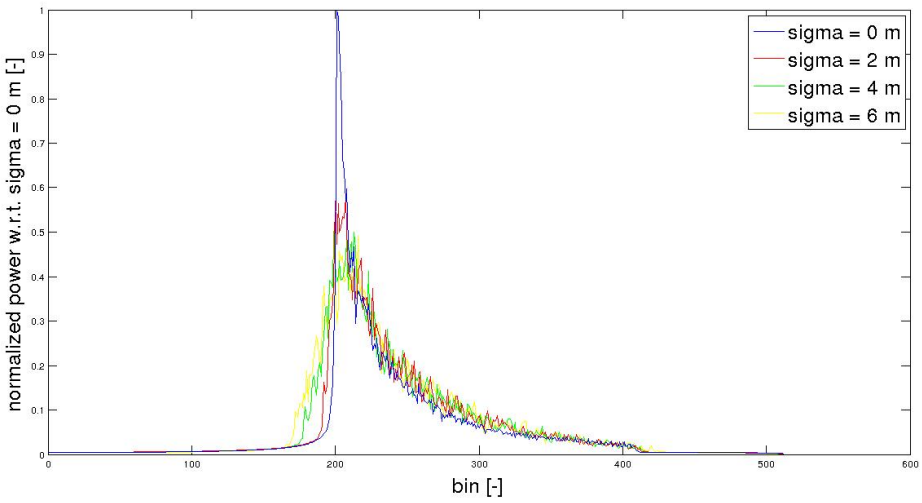


Figure 3.12: Comparison of multilooked waveforms computed over various surface roughnesses.

Figure 3.12 shows respectively waveforms corresponding to a surface roughness with standard deviations 0, 2, 4 and 6 m. For a sigma of 0 m the peak is clearly distinguishable, but with increasing surface roughness the waveform peak becomes more blunt. Over a flat surface an increasing range corresponds to an increasing cross-track angle and also corresponds to a decreasing area as shown in figure 3.10. On a rough surface the ranges within the circles of figure 3.10 vary. Since the range resolution of Cryosat is 46.9 cm, a

standard deviation of 2 m in elevation implies that the scatterers in the areas in figure 3.10 are captured in multiple range bins. Therefore the peak becomes more blunt and less high, and several scatterers are captured in the bins before bin 201.

3.3.4 Effects of cross-track slope

Cross-track slope has an effect on the estimated surface elevation. Figure 3.13 shows waveforms computed from smooth surfaces with different cross-track slope. A cross-track slope of 0.2 percent indicates that over 100 m the surface elevation increases by 0.2 m.

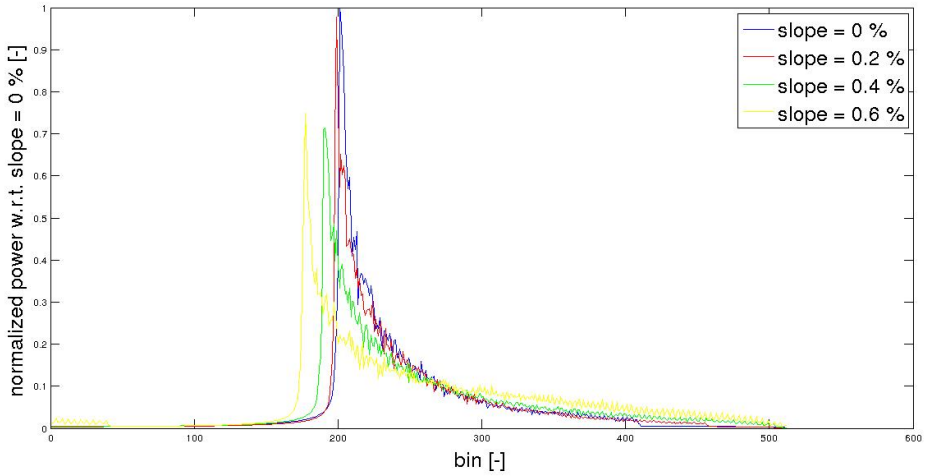


Figure 3.13: Comparison of multilooked waveforms computed with various levels of cross-track slope.

The figure shows that by increasing the cross-track slope, the shape of the waveform changes. Firstly, the peak of waveform moves forward, which means that the range to the surface must be decreasing. Figure 3.13 clarifies why this movement of the peak occurs. The nadir point of the satellite no longer represents the first scattering point. Due to the cross-track slope, the point at which the signal is orthogonal to the surface is the closest point. If a retracker is applied it will therefore give an incorrect elevation measurement. Cryosat uses two antennas to estimate cross-track slope and therefore is able to correct for it.

The second change is a decreased peak power. There are two causes for this. One is the satellite gain, which is higher at nadir. Since the peak in the waveform corresponds to an area off-nadir, the power is lower, however this effect is relatively small. The other one is that the number of scatterers in the area corresponding to the peak is lower. The resolution is taken horizontally and not with respect to the surface. The area size

corresponding to the first bin is equal for a flat surface and a sloped surface, but the number of scatterers is less.

A third change is the lengthening of the tail of the waveform. Since the difference between the highest and lowest elevation becomes larger, the difference in smallest and largest range increases. Therefore, more bins are affected.

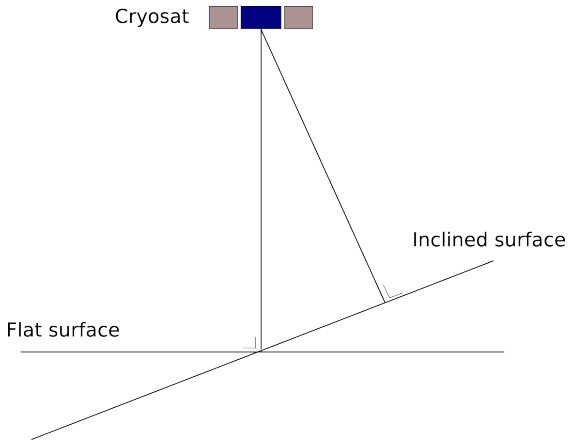


Figure 3.14: Geometry of cross-track slope compared to a flat surface.

Figure 3.13 also shows that retracking these waveforms gives an error in the elevation, since it measures an off-nadir point. Using a second antenna, like Cryosat does, helps to remove the error due to cross-track slope.

3.4 Waveform validation

This section is dedicated to validation of the simulated waveform shapes. Since the algorithm is designed to retrieve lake levels, the simulated waveforms are compared with actual level 1b waveforms of Cryosat over the Nam Co lake.

To compare and calibrate the simulated waveforms, level 1b waveforms retrieved over open water are used. The Nam Co lake is one of the few lakes on the Tibetan plateau, which is large enough to have unpolluted waveforms. In figure 3.15 Nam Co is shown with a box that represents the area in the lake over which waveforms are assumed to be clean. Eight Cryosat ground tracks are shown in green.

Waveforms from two of the tracks were still polluted, probably by small islands within the lake. Ultimately six tracks with unpolluted waveforms remained for validation of the waveform shape. All the waveforms in one track are averaged to have clean waveforms for validation. A comparison of single multilooked waveforms and an average over the track is shown in figure 3.16. In paragraph 3.3.3 it is shown that the tail of a waveform gets

thicker with respect to the power of the peak if the surface roughness increases. The red multilooked waveform is therefore probably measured over quiet water, while the green waveform is measured over a surface with more roughness. Combining approximately 30-40 measurements leads to the average, shown by the blue curve.

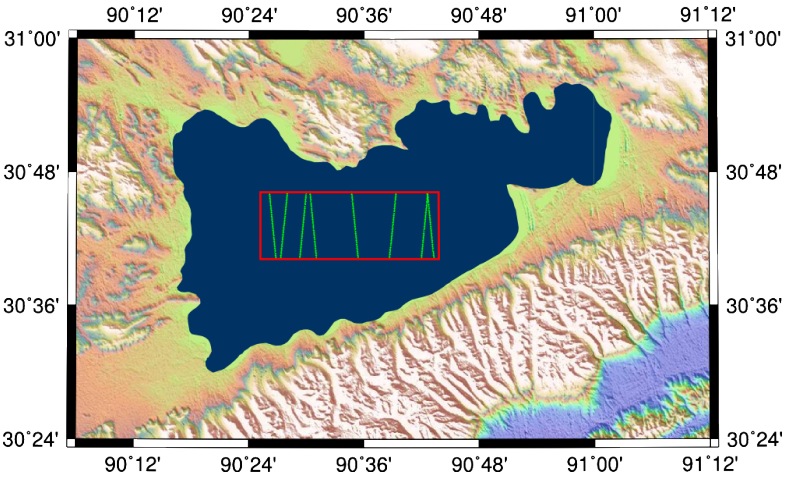


Figure 3.15: Within the red box the waveforms are not polluted by topography. In green the Cryosat ground tracks for 2012. Topographic map is made using ASTER GDEM2. ASTER GDEM is a NASA and METI product.

Since the waveforms since Februari 2012 in the level 1b data are oversampled by a factor of two [26] it is required to oversample the simulated waveform. The oversampled waveforms have 1024 samples, but only the middle 512 samples are stored in the level 1b data. They only represent a range window of 120 meters, because the binsize is reduced to 0.234 m from 0.469 m.

Oversampling of the simulated waveform is performed by first applying an FFT, then doubling the amount of frequencies in the frequency domain with all entries zero, and then applying an inverse FFT. If upsampling is done by interpolation, high frequency signals are introduced which are needed to be Low-Pass Filtered (LPF). However, by using the suggested approach this is not necessary since all the frequencies above the Nyquist frequency are directly set to zero.

First a simulated waveform over a flat surface is compared to the averaged waveforms. In order to compare them it is required that they are properly alligned. To do this on a subbin level, the averaged waveforms and the resampled simulated waveforms are oversampled five times. A built-in cross-correlation function in MATLAB is applied to find the offset between simulated and the averaged waveform. The results are shown in the top graph of figure 3.17.

In the middle graph of figure 3.17 the six averaged waveforms are plotted against the simulated waveform, which is the noisy blue curve. The simulated waveform, clearly

resembles the averaged waveforms, except for the presence of large speckle. This speckle occurs, due to a limited amount of scatterers taken to enhance the computational speed of the algorithm. However, single multilooked waveforms in real data, like the green and red curve in figure 3.16 also suffer from substantial speckle. The lower graph in figure 3.17 shows that relatively large errors around the peak (up to 30 percent) are present, probably because the surface roughness is not large enough.

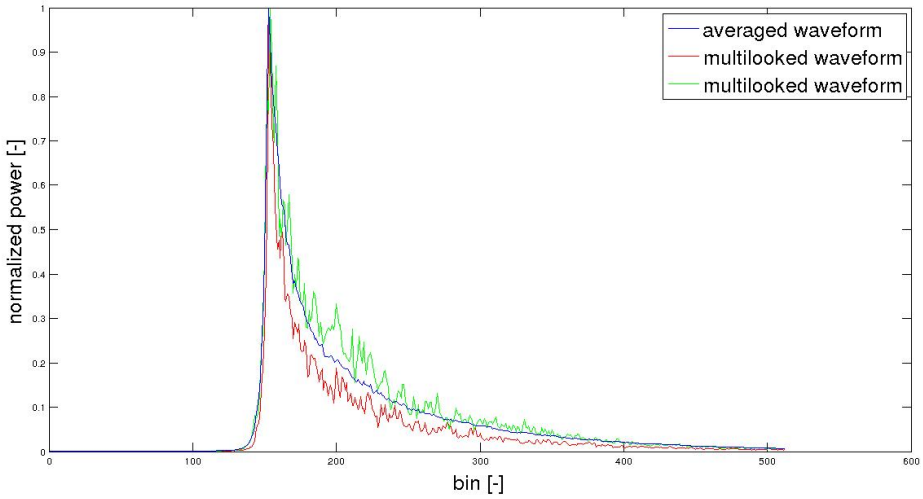


Figure 3.16: Two single multilooked waveforms, compared to an averaged waveform.

Track	Power RMS [-]	Power RMS [-]	Power RMS
	Flat surface	Rough surface	Smoothed waveform
1	0.0365	0.0987	0.0234
2	0.0399	0.1062	0.0286
3	0.0347	0.0950	0.0216
4	0.0373	0.0998	0.0266
5	0.0374	0.1022	0.0245
6	0.0361	0.0956	0.0241

Table 3.1: RMS of differences between simulated waveforms and averaged waveforms over six tracks. The surface roughness has a standard deviation of 0 for the flat surface response and 0.2 for the rough surface. The last column shows results for a flat surface response with a filtered tail.

Since the averaged waveforms are probably measured over an area with some surface roughness, another comparison is made in figure 3.18. From the middle and bottom graph it is visible that the errors increase substantially and that the waveform has substantial speckle. Furthermore, the dacay rate is not fast enough, and therefore the

simulated tail is thicker than in the level 1b averaged waveforms. Also there is no sign of a clear distinguishable peak anymore, due to broadening of the peak and due to large speckle.

In order to compare the results of the two simulations with each other and to find the accuracy of the simulation, the Root Mean Square (RMS) of the differences given in figure 3.17 and 3.18 is computed. With the averaged waveform w_a, the simulated waveform w_s and the number of samples N, the RMS is computed by:

$$RMS = \sqrt{\frac{1}{N}\sum_{n=0}^{N}(w_a(n) - w_s(n))^2}. \tag{3.8}$$

Table 3.1 shows RMS values of the differences between the simulated flat surface waveforms and the rough surface waveform on the one hand and the averaged level 1b waveforms. The simulation for the flat surface has an RMS between 3.5 and 4 percent. The simulation with a rough surface has an RMS between 9.5 and 11 percent. However, the RMS computation also takes into account the far end of the tail of the waveform and the part before any signal is received. Therefore the RMS for the flat surface waveform difference is around 7 or 8 percent for the relevant part of the waveform.

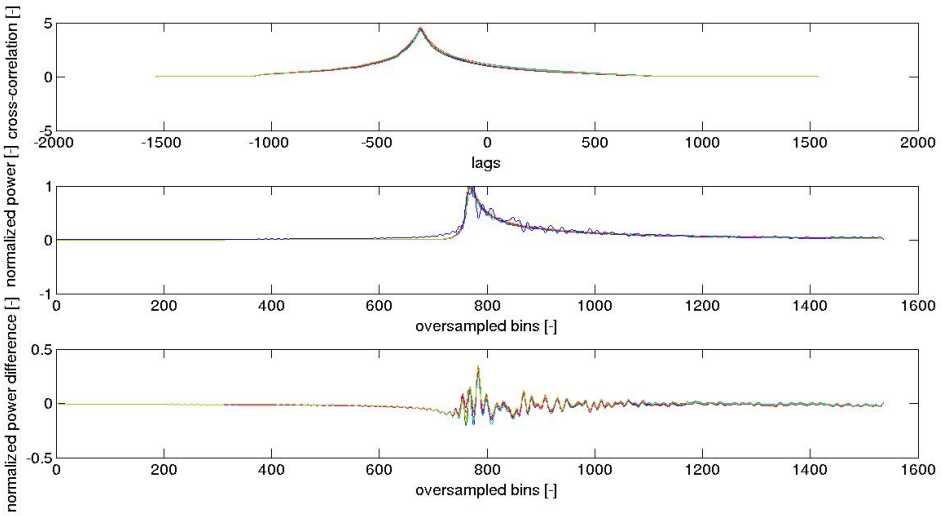


Figure 3.17: Top: cross-correlation between the simulated waveform and six averaged level 1b waveforms. Middle: simulated waveform (blue) vs. averaged level 1b waveforms. Bottom: difference between the averaged level 1b waveforms and the simulated waveform. The simulation assumed a flat surface.

Clearly, the waveform over the flat surface suffers from speckle, especially at the tail. In order to improve the waveform fit in the tail, but keep a sharp peak, a Gaussian filter is

applied after the peak. Let the distance $\Delta b_{i,n}$ be the number of bins difference between bin b_i an bin b_n and the standard deviation σ_b a property of the filter, then the filter is described as:

$$P_n = \sum_{i=-\infty}^{\infty} p_i \cdot e^{-\frac{\Delta b_{i,n}^2}{\sigma_b^2}}, \tag{3.9}$$

where the power p_i and the power P_n are respectively the power in the unfiltered signal at bin i and the power after filtering at bin n. Using $\sigma_b = 5$ bins and starting the filter procedure three bins after the peak showed the best results. These results are shown in figure 3.19. The maximum error around the peak descreased to less than 20 percent, compared to 30 percent for the unfiltered results. The RMS also dropped significantly to between 2 and 3 percent, as shown in table 3.1.

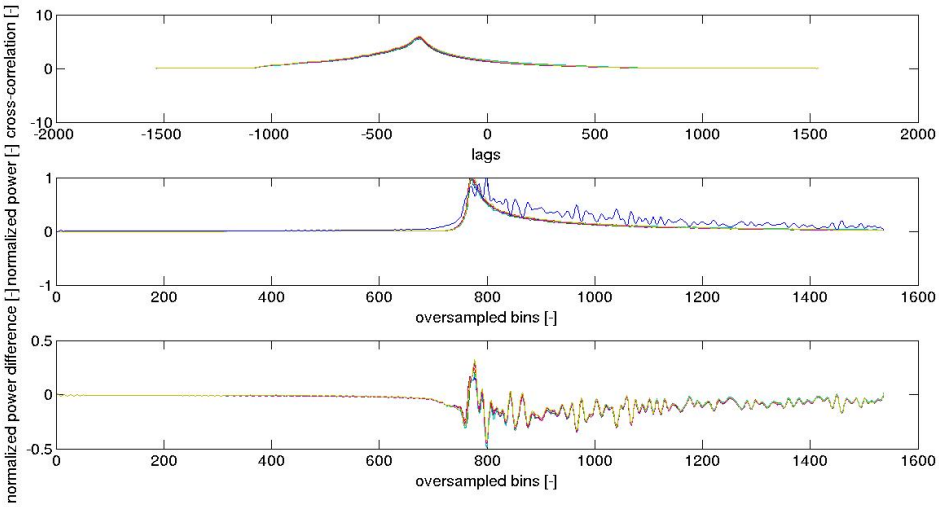


Figure 3.18: Top: cross-correlation between the simulated waveform and averaged waveforms. Middle: simulated waveform (blue) vs. averaged waveforms. Bottom: difference between the averaged waveforms and the simulated waveform. The simulation assumed a surface roughness with $\sigma = 0.2$ m.

3.5 Conclusions

This chapter provides the answer to the subquestion: **How accurately can one model a Cryosat waveform?**

In order to simulate a Cryosat waveform, using the theory of chapter 2 it is first required to compute the positions of the pulses, bursts and scatterers in the ACU reference

frame. Using the relative positions between satellite and scatterers a deramping signal is computed immediately. Since only the shape of the waveform is important for retracking, the absolute amplitude is discarded.

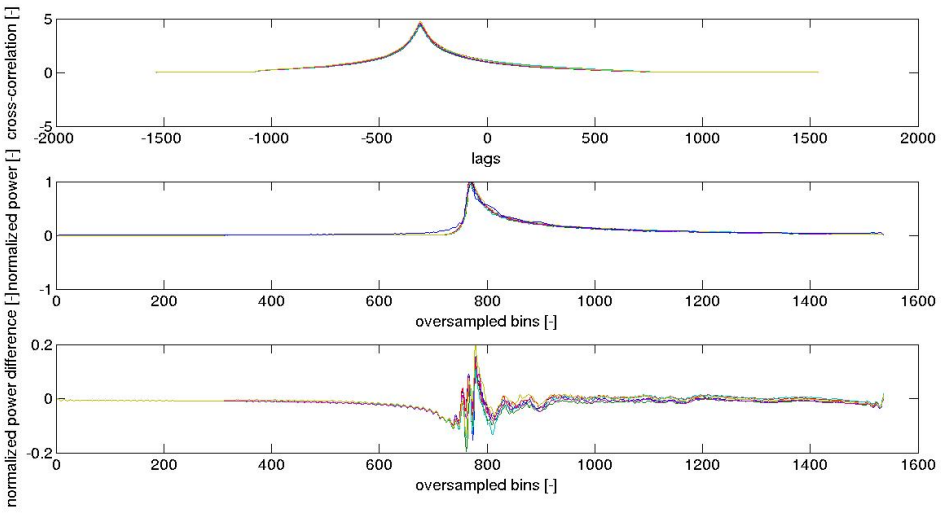


Figure 3.19: Top: cross-correlation between the filtered simulated waveform and averaged level 1b waveforms. Middle: filtered simulated waveform (blue) vs. averaged level 1b waveforms. Bottom: difference between the averaged level 1b waveforms and the filtered simulated waveform. The simulated waveform is filtered using a Gaussian filter, as explained in the text.

The shape of the multilooked waveform depends on the processing procedures, backscattering and geometry of the surface. In this chapter only a signal from a homogeneous surface, like water was modelled, not taking into account backscattering. The geometry was varied by changing surface roughness and cross-track slope. If the resolution of scatterers is improved (at the cost of processing time) the speckle in the waveform is reduced. Increasing the number of single look waveforms to be averaged decreases trailing edge decay and also reduces the speckle. Surface roughness increases the width of the peak and cross-track slope changes the position of the peak. Finally, over very smooth waters the backscattering coefficient of the off-nadir surface comes close to zero and therefore the waveform shows almost a specular response.

By comparing the simulated waveform shape over unpolluted level 1b waveforms at Nam Co the shape is validated. A RMS of differences of 3-4% is found for a flat surface response. Introducing surface roughness decreases the fit to approximately 10%. A further reduction of the RMS of differences is found by applying a Gaussian filter over the speckled tail of the simulated waveform. This results into an RMS of 2-3%.

Chapter 4

Lake level estimation

This chapter addresses a new method to estimate lake levels using Cryosat SARInM level 1b data. The new method uses a numerical retracker, based upon the waveforms which are modelled as described in chapter 3. In this chapter an answer is given to the subquestion: **How can a received waveform be retracked with a modelled waveform?**

In section 4.1 the computation of elevations is discussed in officially provided level 2 data. It shows how the results of the retracker, the geophysical corrections and the satellite position are used to compute a lake level. Furthermore, the accuracy of the elevation in level 2 data is discussed using an error budget.

Elevations in level 2 data are computed using a retracker based upon an analytical description of the waveform, which is called in this report an analytical retracker. It is also possible to use a cross-correlation between a modelled waveform and the level 1b waveforms as a retracker, which is called numerical retracker. Section 4.2 describes what the problems with analytical retracking are and why a numerical retracker is applied.

In section 4.3 the implementation of a numerical retracker is discussed. Besides the cross-correlation procedure, several other procedures or improvements are applied to get an optimal result.

In section 4.4 the procedure to determine the lake level from a number of measurements in one track is discussed. It includes the removal of the geoid height variations from the lake levels and the determination of the precision.

Ultimately in section 4.5 conclusions are drawn and an answer is given on the subquestion.

4.1 Cryosat level 2 data

In section 2.3.3 a description of an analytical retracker was given. This section describes how to compute an elevation after retracking. It also shows the error budget of the level 2 data over open water. Section 4.2 shows why this error differs over small lakes from open water.

4.1.1 Elevation computation

In order to compute a two-way travel time after retracking, an equation is set up using the window delay $t_{wd,L1b}$ (time between pulse transmission and a reference point in the generation of the deramping chirp, corrected for instrument errors), the time t_d between the reference point in the deramping chirp and receival of the signal (including retracker correction) and corrections t_c. Note that $t_{wd,L1b}$ is different from, but related to t_r used in Eq. (2.35) and Eq. (2.36). The timing t_r of the deramping is the time of the zero bin, and $t_{wd,L1b}$ is the time of the middle bin in the waveform, normally bin 255. The equation for the two-way travel time t_{2w} is:

$$t_{2w} = t_{wd,L1b} + t_d + t_c. \tag{4.1}$$

Using Eq. (2.1), this is rewritten as:

$$R = R_{wd,L1b} + \Delta R_d + \Delta R_c, \tag{4.2}$$

with,

$$R_{wd,L1b} = \frac{ct_{wd,L1b}}{2}, \tag{4.3}$$

$$\Delta R_d = \frac{ct_d}{2}, \tag{4.4}$$

and

$$\Delta R_c = \frac{ct_c}{2}, \tag{4.5}$$

in which $R_{wd,L1b}$ denotes the window delay in meters, ΔR_d the retracker correction in meters and ΔR_c the range correction in meters. So the range R is the range if a signal is received at the moment the generation of the deramping chirp starts and only instrumental corrections are applied. These corrections are applied in level 1 processing and mentioned in section 2.3.1. The value for $R_{wd,L1b}$ is read from the level 1b data.

The retracker correction ΔR_d is the additional range if the signal is received later. The value for ΔR_d is computed by waveform retracking. Then the geophysical corrections ΔR_c are applied in level 2 processing, these are available in the level 1b data.

Over three different surface types: ocean, sea ice and land, different sets of geophysical corrections are applied [35]. Inland waters are considered to be land type surface. The atmospherical corrections applied over land are for dry tropospheric delay, wet tropospheric delay and ionospheric delay. The tidal corrections applied over land are for the ocean tide, solid earth tide and geocentric polar tide [35].

For satellite altimeters tidal corrections and the dry troposheric delay correction come from models. Most satellite altimetry missions use dual frequency measurements to estimate the ionospheric delay and a radiometer to estimate wet troposheric delay respectively. Because of the absence of dual frequency measurements and a radiometer, Cryosat applies models for these as well. As a consequence, the range measurements are up to several centimeters less accurate.

Ultimately the elevation h_s is computed using the altitude z_0 of the satellite above the strip and the range R, which results in:

$$h_s = z_0 - R. \tag{4.6}$$

Since Cryosat in SARInM mode also computes phase differences between the antenna (interferometry), the elevation $h_{s,SARInM}$ in SARInM is computed as:

$$h_{s,SARInM} = z_0 - R \cdot \cos(\Upsilon_s), \tag{4.7}$$

with Υ_s the cross-track angle to the scattering point.

4.1.2 Error budget

Besides errors in geophysical models errors of other types play a role as well. This concerns orbital errors, instrumental correction errors, retracking errors as well as the cross-track angle error (in SARInM). Table 4.1 shows the error budget for Cryosat. The orbit is computed from DORIS measurements and has an accuracy of 0.02 m approximately. The instrumental corrections are applied in the level 1b processor. These corrections are for oscillator drift, pointing angle errors and internal calibrations, which add up to several centimeters. In an optimal situation the retracker has an error of approximately 0.02 m. However, over lakes smaller than the footprint of Cryosat the retracker performance decays due to land surface signals in the waveform. In SARInM Cryosat is able to measure the cross-track angle. The cross-track angle error is for the estimation of additional range due to scatterers that are off-nadir cross-track.

Source	RMS [m] [-]	Reference
Orbital errors	**0.02**	[36]
Instrumental corrections	**0.05**	[37]
Geophysical corrections	**0.04**	
Ionospheric correction	0.02	[36]
Wet tropospheric correction	0.03	[36]
Dry tropospheric correction	0.007	[36]
Tides	0.02	[36]
Retracker	**0.03**	
Cross-track angle	**0.015**	[1]
Total error	**0.10**	

Table 4.1: Error budget of Cryosat SARInM elevations.

These errors are estimated for a common situation of Cryosat measurements over open ocean if all processing problems of Cryosat are fixed. However, the current products of Cryosat still suffer from several processing problems, which lead to biases in the order of several decimeters. Various errors also depend upon the latitude, surface type and place in the orbit [38]. Therefore in most cases the errors are larger than shown in table 4.1. A more detailed discussion about the error sources is given in [36].

The accuracy of the retracker is of particular interest. This will descrease over lakes smaller than the footprint due to pollution of the waveform. More details on the accuracy are given in chapter 5.

4.2 Shortcomings of analytical retrackers

As mentioned in sections 1.2 and 4.1, waveforms acquired over small lakes or near the shores of large lakes are polluted by surrounding topography. This leads to less accurate measurements or no reliable measurements at all of the lake level when an analytical retracker is used. This section shows examples of polluted waveforms, and discusses why analytical retrackers are not able to find the correct lake level. The section also shows why it should be possible to find an accurate lake level, using an alternative approach. In order to explain the inaccuracies two cases are discussed: one where the waveform is slightly distorted, and one where the waveform has several peaks. ,

4.2.1 Distorted waveforms

Based on the theory of section 2.3.3, four parameters are optimized in the analytical retracker to fit the waveform: rise time, arrival time, trailing edge decay and amplitude. Figure 2.19 shows were these parameters apply and how they affect the shape of the waveform. For the Cryosat SAMOSA retrackers used to compute level 2 data, this set

is likely to be extended to six parameters [1]. One of the additional parameters is noise level. However, a theoretical description of the SAMOSA retrackers is not available, therefore we try to show the problems of analytical retracker in general, using the theory presented in section 2.3.3.

Changes in shape

In figure 4.1 a polluted waveform is plotted against an unpolluted waveform. Both of the waveforms are measured at the Nam Co lake in the same track, i.e. on the same day. Apart from possible different heights of water waves, one of the waveforms is polluted by surrounding topography of the lake. In addition, the change in shape is caused by the shape of the lake, i.e. even if the pollution was removed from the waveform, the waveforms still would have different shapes. This is because the percentage of water scatterers is less than 100 percent (this is explained in detail in paragraph 4.3.3). By comparing the waveforms four major differences are observed:

- The starting toe is much larger in the polluted waveform.
- The leading edge is less steep.
- The trailing edge is polluted and is not decaying rapidly.
- The peak is slightly shifted.

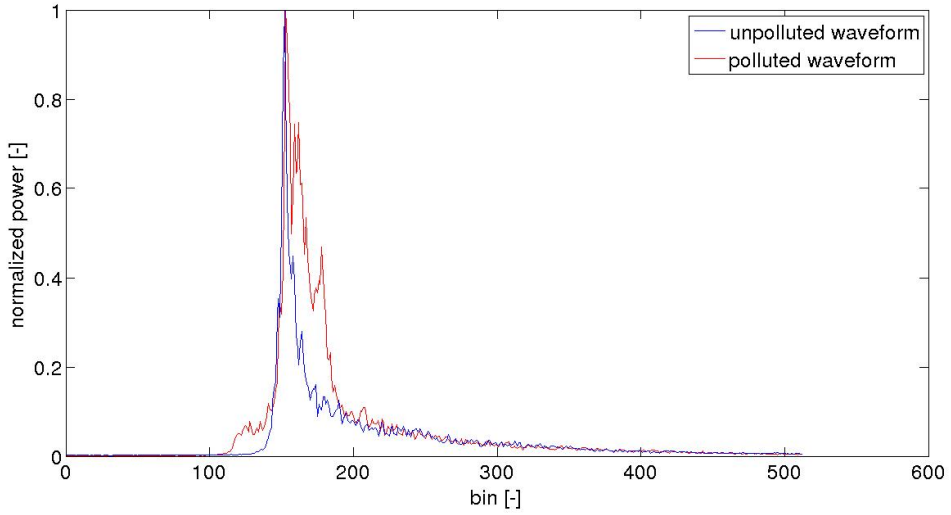


Figure 4.1: Unpolluted waveform versus a slightly polluted waveform.

Problems with analytical retrackers

There are several consequences for the retracker performance over polluted waveforms, like the one given in figure 4.1. Since the leading edge is less steep and the toe is larger, primarily the rise time increases, which is related to the roughness of the surface.

This is the not the roughness of the water surface, but an effect from the surrounding topography. Another problem is the estimation of the trailing edge decay, which will be completely different in the polluted waveform. And the final problem is the estimation of the time of arrival of the first signal from the lake. Over unpolluted waveforms the arrival time of the first signal is referenced to a point approximately halfway the leading edge. In a polluted waveform the signal of the water surface in the waveform changes shape, because only part of the footprint is covered by water. Therefore the computed reference point using analytical retracking might be wrong. And secondly, since all the estimated parameters are related to each other, the estimated rise time and leading edge decay influence the estimation of the arrival time.

Possible adaptations in numerical retrackers

In the numerical retracking the surface roughness is held constant. According to the findings in section 3.4 the best results are obtained when the roughness is almost zero. It is expected that better lake level estimates near the shores of large lakes are obtained, when the roughness in the retrackers is held constant. Furthermore, the waveform shape computed using a numerical retracker can be adjusted to the shape of the lake and therefore only includes water pixels. Therefore the fit is expected to be better than an analytical retracker, which works optimally over open water.

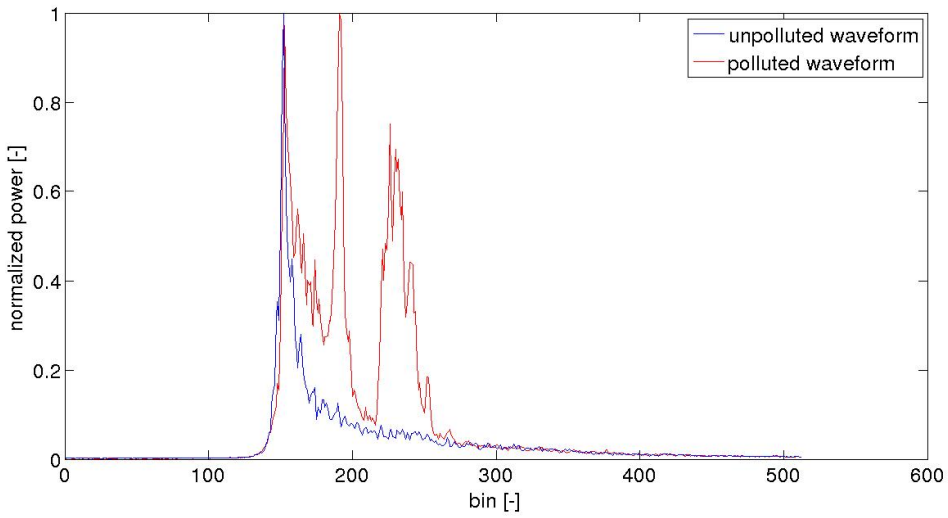


Figure 4.2: Unpolluted waveform versus severely polluted waveform with multiple peaks.

4.2.2 Multipeak waveforms

Changes in shape

In many cases at the shores of Nam Co, waveforms with multiple peaks are measured. An example of such a waveform is given in figure 4.2. These multipeak waveforms are

a result of multiple scattering surfaces, for example: water, a mountain side and a flat land surface close to the shore.

Problem with analytical retrackers
Analytical retrackers use the Levenberg-Marquardt approach to fit the waveform. In section 2.3.3 it is shown that this approach is essentially a first-order non-linear least-squares technique. The optimum fit it finds is dependent on the chosen initial values. In case of the waveform shown in figure 4.2, if the corresponding arrival time for bin 100 is chosen as the initial value, it will probably fit the waveform in the first peak since it is located close to bin 100. However, if the initial value for the arrival time is set at bin 300, the fit is made with the peak at bin 250, which leads to an incorrect range measurement.

Possible adaptations in numerical retrackers
There are two options to avoid this problem using a numerical retracker. Cross-correlation of the simulated waveform and the multipeak waveform will give three possible elevations, because three peaks are visible in the multiple peak level 1b waveform. Either the highest cross-correlation is taken, assuming that despite of the pollution the signal from the water surface gives the highest correlation. Or, especially when the signal from the water is relatively weak, all three elevations are computed and stored. These are compared to other elevations measured along-track. It is expected that a mountain signal changes along-track due to varying topography, but the lake level is constant. Therefore, the best elevation is chosen such that most measurements are close to this elevation.

4.3 Numerical retracker

In this section the proposed retracker is discussed, which is based upon the modelled waveform shape presented in chapter 3. As already shown in section 3.4, the fit of the waveform is improved by applying a Gaussian filter over the tail. The rest of the retracking procedure is described in this section.

Over open water the retracking procedure consists of only one step, which is cross-correlating the simulated waveform with the level 1b waveform to find the best fit. Close to the shore, i.e. within 7 km (the number given in section 1.2) the procedure is adjusted. In most cases within 7 km multiple peaks are found, and therefore multiple elevations are computed, hence the second step implements a procedure to find the optimal solution.

Next to the retracking procedure, there are two procedures to enhance the outcome of the elevation computation. The first one is weighting, which gives different weights to the waveform bins according to their importance. The second one is a procedure to adjust the retracker shape according to the lake shape, in order to only take into account lake pixels and not to compute a waveform over the complete surface.

The final implementation of these procedures is discussed in the chapter 5. It will be based on the results of the implementation of the retracker with different settings over open water and near shore.

4.3.1 Cross-correlation

As described in section 4.1 the range depends on the retracker correction. The retracker correction R_d is the arrival time of the signal with respect to the reference point to which the window delay is computed. In SARInM this reference point is bin 255, if the first bin is assumed to be 0. In the proposed algorithm this is bin 256, since the first bin is labeled 1, which is the outcome of the FFT procedure implemented in MATLAB. The numerical simulation is set such that the first return signal is computed exactly at bin 256. By finding a peak in the cross-correlation between the simulated waveform and a level 1b waveform, it is possible to see how far the simulated waveform must be shifted to fit the level 1b waveform best. The shift in bins is related to the retracker correction, which is explained in this paragraph.

To compute the cross-correlation ρ_x, the algorithm needs the level 1b waveform W_{L1b} and the simulated waveform W_s, which is based on the simulation described in chapter 3. Then the cross-correlation is computed as:

$$\rho_x(i) = \frac{1}{I-j} \sum_{j=0}^{I-j-1} W_s(i+j) W_{L1b}(j), \tag{4.8}$$

where i and j are the numbers of the bins involed and I denotes the total number of bins.

As discussed in section 3.4, the resampled waveforms have a bin size of 0.234 m. If cross-correlation is applied, the resolution to which the retracker delay can be resolved is the bin size. This resolution limits also the accuracy. Therefore the two input waveforms, the simulated one and the level 1b one, are both upsampled ten times. This is done by taking the FFT of the waveforms and adding at the tails. The upsampling factor is required to be an integer, since the inverse FFT is most efficient when the numbers of samples is a power of two.

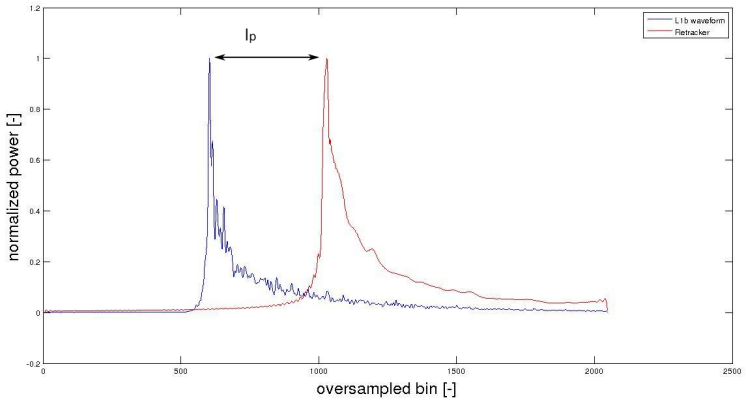


Figure 4.3: Finding the maximum cross-correlation between the two waveforms, gives a lag l_p which is related to the retracker correction.

Once the cross-correlation between the two upsampled waveforms is computed, the maxima in the cross-correlation are found and the corresponding lags are determined as shown in figure 4.3. The corresponding lag l_p along with an oversampling factor F_o are related to the retracker correction as:

$$R_d = \frac{l_p}{F_o} \frac{r_{res}}{2}, \tag{4.9}$$

with r_{res} the sampling frequencies and the division by 2 is included, because the waveforms in level 1b data are already oversampled by a factor of two (see section 3.4).

4.3.2 Multiple peak retracking

As mentioned in section 4.2.2 a problem that often occurs in topography polluted waveforms is the occurance of multiple peaks. An analytical retracker finds the signal parameters depending on the initial values entered in the Levenberg-Marquardt method. The numerical retracker presented so far finds the signal with maximum correlation, which not always corresponds to the water surface.

In this paragraph a solution is proposed to find the optimal lake level in the presence of multiple peaks and to reduce the number of erroneous measurements. Instead of computing one elevation per waveform, several elevations are computed from all the peaks found in the cross-correlation operation. This leads to results as shown in figure 4.4.

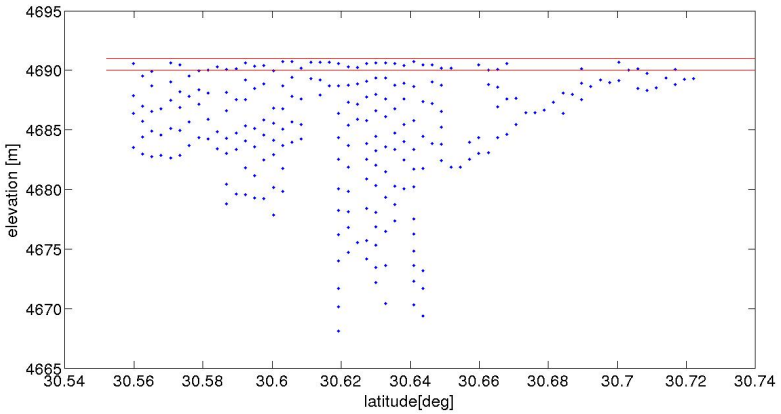


Figure 4.4: Approach to find the correct lake level in waveforms with multiple peaks. Sets of elevations are estimated per waveform. In this figure the first elevation is the best, but in many cases this does not occur.

In order to determine the most likely solution for the elevation, the measurements of at least 5 to 10 consecutive waveforms are considered. All the measurements are accumulated in bins of 40 cm elevation difference, which is approximately two times the error budget found in section 4.1. From this collection the mode (the bin with most measurements) is determined. Elevation measurements within four standard deviations from the mode are taken into account. It is possible that from one waveform two or more lake levels are computed which are within four standard deviation from the mode. Only the best one of these measurements is kept (closest to the mode).

4.3.3 Lake shape adjustment

Up to this stage the retracker uses a model that is based on a flat surface response. Most analytical retrackers are based on a theoretical flat surface model. For lake level retrieval in Tibet, this is in most cases not the optimal solution. Close to the shores of a lake, where land is within the altimeter footprint, the shape of the waveform from the water changes. As presented in figure 4.5, which shows the edge of the Nam Co lake and the pixels used in the algorithm, almost half of the pixels is not scatterering from a water surface. This means that the topography around the lake has effect on the waveform, but also implies that the response from the lake has another shape than in a flat surface model.

A better solution would be found by only incorporating the pixels over the water surface when matching the waveforms shape acquired over this area. To that end, one can use the MOderate-Resolution Imaging Spectroradiometer (MODIS) water mask [39]. The water mask used for this case is a binary image between latitudes 27 and 38 degrees and longitudes 72 and 105 degrees (this covers the entire Tibetan plateau). The resolution is

0.0021 degrees, which is over the Tibetan plateau is approximately 210 m in longitudinal direction and 230 m latitudinal direction. Since the resolution of the MODIS water mask is different from the pixels chosen in the algorithm, a nearest neighbour method is implemented to find weather a pixel corresponds to a water surface or not. This results in the reduction of the number of scatterers, visible in figure 4.6. It shows that only the pixels from the water surface are scattering. Near the edge, some pixels may still be seen scattering though they appear to be located over land. This has two reasons: one is that GMT (the software package used to generate the map in figure 4.6) uses lake bounderies that not necessarily coincide with those from the MODIS mask. The resolution of the GMT lake shape file is probably different from MODIS and the lake shape is computed from acquisitions over different years, so the lake shore could be relocated. The other reason is possible wetlands near the lake, which by the MODIS processing algorithms are interpreted as water.

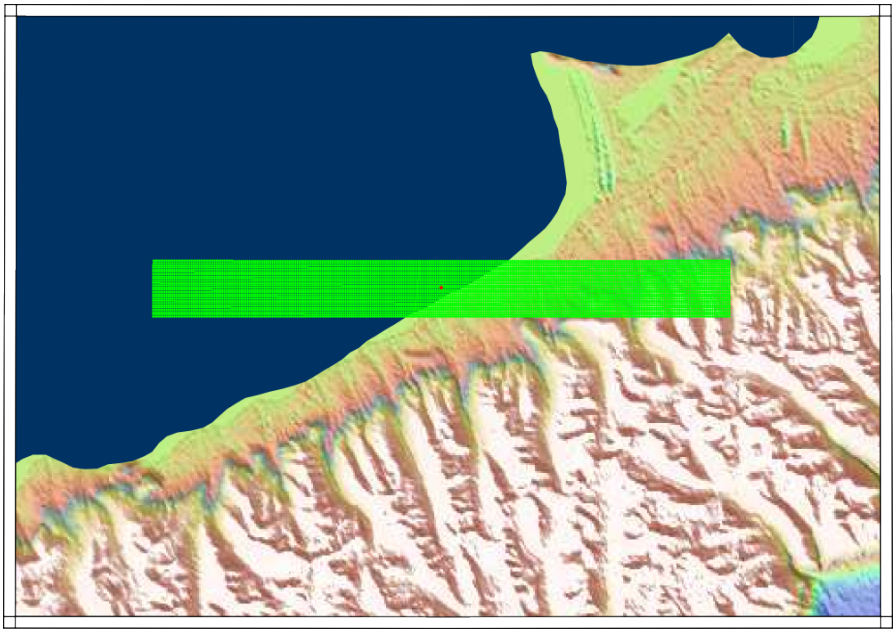


Figure 4.5: Nam Co lake, with in green the scattering pixels without applying shape adjustment.

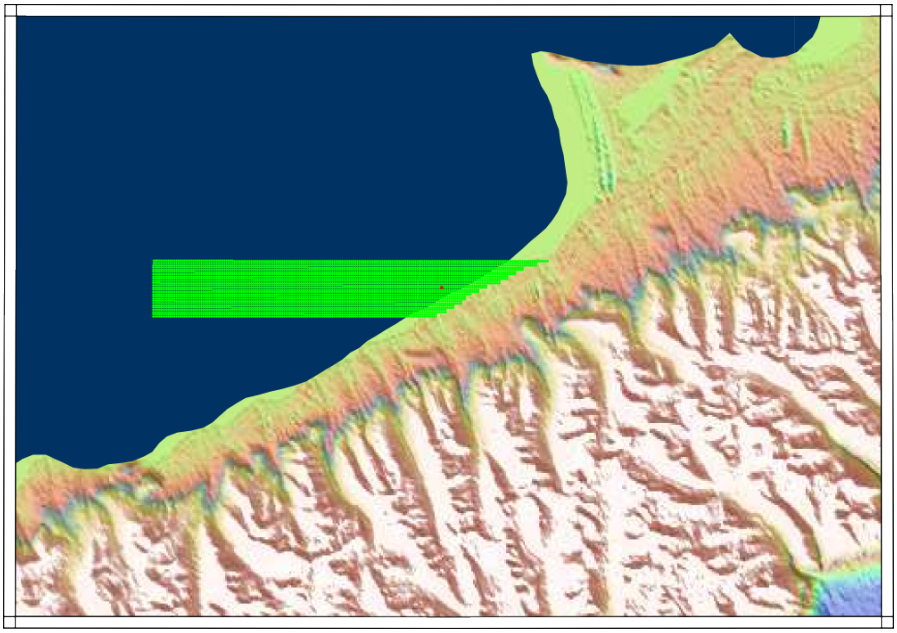


Figure 4.6: Nam Co lake, with in green the scattering pixels with applying shape adjustment. The topographic map is made using ASTER-GDEM2. ASTER-GDEM is a NASA and METI product.

In figure 4.7 a waveform computed from water pixels only is shown, compared to a flat surface response. It becomes clear that the peak is shifted slightly and the width of the peak is also a bit larger. For other places on the shorelines the change in shape of the waveform can either be more or less severe. More pollution in the level 1b waveform will make the flat surface response waveform (when the whole footprint is filled with water pixels) less suitable as a retracker, because the shape of the flat surface response does not resemble that the level 1b waveform anymore.

4.3.4 Weighting function

As mentioned in section 4.2 the arrival time of the signal from the lake corresponds to a point on the leading edge of the waveform. The trailing edge decay depends on the backscattering coefficient of off-nadir water scatterers. Also it was shown that the toe and the trailing edge of the waveform are possibly affected by pollution from land. Therefore the trailing edge decay can vary and have an effect on the accuracy of the measurements.

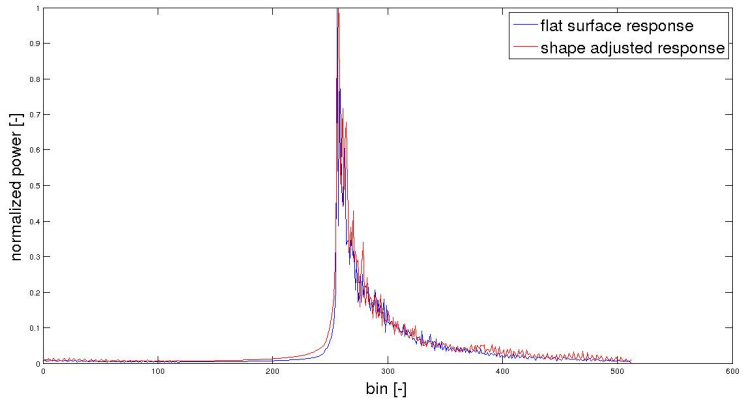


Figure 4.7: Flat surface response waveform, versus lake shape adjusted waveform.

As a consequence, a weighting function $K(i)$ with i the bin number, is introduced to give the bins around the peak of the simulated waveform more weight in the cross-correlation. It is expected that the implementation of the weighting function has a positive effect on the accuracy of measurements, especially at the edges of the lake. The weighting function is normally distributed as:

$$K(i) = e^{\frac{(i-i_{ref})^2}{\sigma_i^2}}, \tag{4.10}$$

with i_{ref} the reference bin close to the peak of the waveform and σ_i the standard deviation in number of bins. Setting σ_i to infinity yields the same results for the cross-correlation as without a weighting function. Setting σ_i very high causes a cross-correlation with almost a specular response simulation.

4.4 Lake level computation

After retracking and computation of the elevation at one measurement location using Eq. (4.6), one lake level is determined from all measurements in one track. This procedure involves two steps: subtraction of the geoid heights and an iterative removal of outliers.

4.4.1 Subtraction of geoid heights

Since the size of the lakes varies from 3 km to 70 km in diameter, the effect of the geoid is clearly visible in the measurements. An example of this is shown in the left graph of figure 4.8. The graph shows a linear trend of several decimeters over 0.2 degrees latitude,

which corresponds to approximately 20 km. As a consequence, it is not possible to take the average of this track. The mean of the measurements computed will vary several decimeters depending on the latitude-longitude location of the measurements.

In referencing the altimeter measurements to the geoid, Eq. (4.6) is extended to:

$$H_s = z_0 - R - N, \tag{4.11}$$

where H_s is the height above the geoid and N the geoid height. After removal of the geoid heights based on the EGM2008 model, which is done in the right graph of figure 4.8 this trend in the measurements from one track is strongly reduced. Since the geoid model EGM2008 is not very accurate in the Himalayas [40], a slight trend in still visible in the left graph of figure 4.8. In order to cope with this trend, the lake level in all computed tracks are computed with respect to the center latitude. This is discussed further in paragraph 4.4.2. The inaccuries in the geoid model may still affect the measurements from different tracks differently. This effect is not accounted for, but can later be mitigated by using a more accurate geoid model.

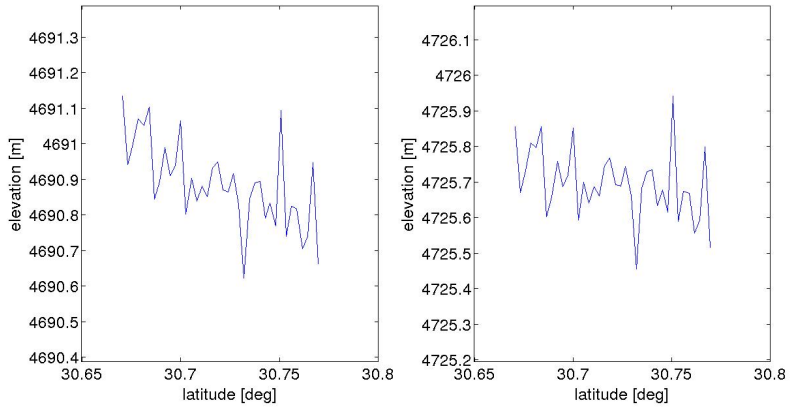


Figure 4.8: Elevations estimated over one track. Left: before the geoid height subtraction. Right: after the subtraction.

4.4.2 Outlier removal

As mentioned in the preceding paragraph, it is not possible to directly average the lake level measurement after geoid height subtraction, because of a small trend in the measurements presumably caused by inaccuracies in the geoid model. To remove the outliers more effectively, a trend is subtracted and a mean is estimated. Since Cryosat is orbiting almost in a polar orbit, the trend $\frac{dH}{d\theta}$ of the measurements is computed with respect to the latitude θ. Using the estimated elevations H and center latitude θ_m of the lake, we relate the mean elevation H_m to the estimates as:

$$\begin{aligned}\begin{bmatrix} H_1 \\ H_2 \\ : \\ H_s \end{bmatrix} &= L \begin{bmatrix} H_m \\ \frac{dH}{d\theta} \end{bmatrix} \\ &= \begin{bmatrix} 1 & \theta_1 - \theta_m \\ 1 & \theta_2 - \theta_m \\ : & : \\ 1 & \theta_s - \theta_m \end{bmatrix} \begin{bmatrix} H_m \\ \frac{dH}{d\theta} \end{bmatrix}.\end{aligned} \tag{4.12}$$

Values for the trend and mean elevation are solved using:

$$\begin{bmatrix} H_m \\ \frac{dH}{d\theta} \end{bmatrix} = (L^T L)^{-1} L^T H. \tag{4.13}$$

The RMS of the post-fit residuals is computed as:

$$RMS = \sqrt{\frac{(H - \hat{H})^2}{S - 2}}, \tag{4.14}$$

with S the number of measurements and $\hat{H}$ the estimated elevation from the model.

If the largest outlier is larger than three times the RMS, it is removed and the elevation and trend are computed again. This process continues iteratively. In order to keep as many measurements as possible at least 50 percent of the measurements is kept or in the case of small lakes five measurements.

4.5 Conclusions

The primary goal of this chapter is to give an answer to the subquestion: **How can a received waveform be retracked with a modelled waveform?**

A lake level is computed using three parameters, which are added. Firstly, the window delay is the range towards a reference point in the waveform (normally bin 255). Secondly, the retracker correction, which is the difference in range between the reference point and the retracker measurement. And thirdly, geophysical correction which are provided in the level 1b data.

Over open ocean the accuracy of the official Cryosat level 2 data using the SAMOSA retrackers is approximately a decimeter. However, due to signals from surrounding topography affecting the waveforms, the accuracy over most of the Tibetan lakes the accuracy decreases to several decimeters or meters. In various cases a lake level is not

found at all, but a signal from the surrounding mountains is retracked. This is especially the case with waveforms, which have multiple peaks representing signals other than the water surface. The elevation measurement depends on the initial values in the Levenberg-Marquardt approach.

Therefore a numerical retracker is proposed. The numerical retracker computes a waveform over a surface from several predefined scatterers, as shown in chapter 3. The waveform is oversampled and its shape is cross-correlated with Cryosat level 1b waveforms to find an elevation. If multiple peaks are found in the waveform, more along-track measurements are used to find the mode (the bin with most measurements) in all elevations measurements. The elevations within four standard deviations of the mode are kept, and labelled as valid measurements.

To enhance the performance of the numerical retracker, two additional steps are applied. One is a weighting function, which increases the weight of the measurements near the peak of the simulated waveform. And the other is lake shape adjustment, which only incoorporate pixels over the water surface.

In order to compute a lake level from measurements in one track, the geoid heights are removed first. Then the mean and the trend are subtracted from the observations. Finally, outliers are removed until all measurements are within three times the standard deviation. At least five measurements or 50 percent of the measurements are kept. Ultimately, for every track one elevation is computed.

Chapter 5

Retracking Results

In the previous chapters the processing routines, waveform simulations and water level measurement principles have been described. In this chapter the results of lake level measurements over Nam Co are shown and discussed. An answer is given to the subquestion: **How accurately can lake levels be retrieved using a numerical retracker?** In order to answer this question, the estimated lake levels using the proposed numerical retracker from section 4.3 are compared to Cryosat level 2 data and statistics are computed. Since the obtained estimates are not fully independent from those in Cryosat level 2 data, the resulting errors are probably under-estimated. A better approach would be to use in-situ water level measurements, however these are only available until 2011 and official Cryosat product only from February 2012.

The performance of the level 2 data and the numerical retracker are described in five sections. In section 5.1, I assess the performance of the numerical retracker using different settings over open water and compare the obtained results with the level 2 data. Section 5.2 desribes the performance of the numerical retracker with different settings near the shores of Nam Co. Section 5.3 provides recommendations on how to implement the retracker based upon results found in the previous sections. Section 5.4 discusses the potential lakes handled using Cryosat with the numerical retracker. Section 5.5 shows water level variations over the Nam Co lake. A trend is computed and seasonal variations are described. And in section 5.6 conlusions are drawn.

5.1 Open water lake levels

In this paragraph it is shown how well the proposed numerical retracker works compared to Cryosat level 2 data over open water. The numerical retracker is implemented based on the theory discussed in section 3.4 and 4.3.

In this section three different settings for the numerical retracker are used. The first one

uses an unfiltered simulated waveform, which is directly cross-correlated with the level 1b waveforms. The second setting uses a tail-filtered simulated waveform (based on the filtering in section 3.4). And the final one uses a tail-filtered simulated waveform which is weighted. The weighting function is set such that bin 256 has maximum weight and a standard deviation of 1 is selected. In order to avoid confusion, table 5.1 shows the name of each retracker are its settings.

Name	Retracker settings
N-retracker	Numerical retracker
FN-retracker	tail-Filtered (Numerical) retracker
WFN-retracker	Weighted (tail-Filtered Numerical) retracker

Table 5.1: Retrackers used to estimate elevations from unpolluted waveforms.

Figure 5.1 shows the resulting elevations, computed with Eq. (4.11) for one track. In the figure it is visible that the numerical retracker has a discrepancy of approximately half a meter with the level 2 data. For the weighted retracker this is slightly higher. The precision of FN-retracker seems to be the best, and even better than that of the level 2 data. The precision of the weighted retracker is lowest.

In this section several tables show the mean and precision of the obtained water levels as well as their discrepancy with level 2 elevation, for the three aforementioned retracker settings. Furthermore an interpretation of these results is given.

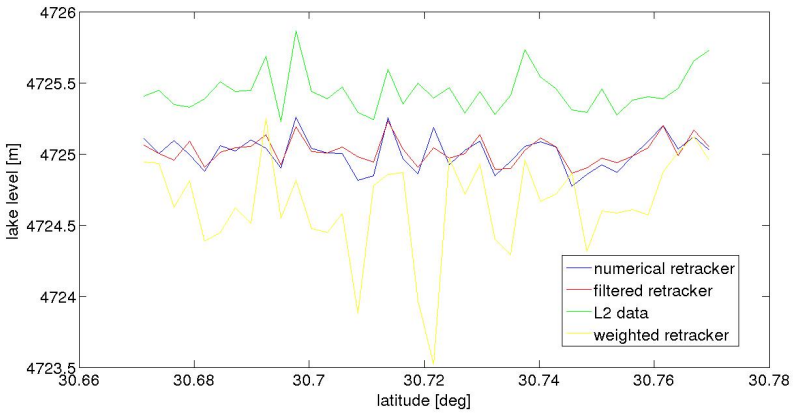


Figure 5.1: Elevations estimated over one track with various retrackers and taken from Cryosat level 2 data.

In table 5.2 the estimated average elevations are shown for six tracks, wherein waveforms are not polluted. The tracks are shown in figure 3.15. As was already visible in figure 5.1 there is a discrepancy of several decimeters between the numerical retrackers and the level 2 data. For the N- and FN-retracker this might have several causes.

A first possible cause is the behaviour of the cross-correlation method. It is designed to find a peak in the cross-correlation. Since the peak in the waveform has high power compared to the power in the surrounding bins, the peak is important in cross-correlation. On the other hand, the most important part of the waveform is actually a point on the leading edge of the waveform, which corresponds to the elevation. The cross-correlation is therefore likely to adjust its outcome towards the peak of the waveform instead of the leading edge. As shown in section 3.4, the largest discrepancies of the simulation with respect to the level 1b waveforms occur around the most important parts of the waveform, which is situated close to the peak. These errors have therefore a significant effect on the outcome of the computed elevation.

	Elevation [m]	Elevation [m]	Elevation [m]	Elevation [m]
Track	Level 2 data	N-retracker	FN-retracker	WFN-retracker
1	4725.47	4725.02	4725.00	4724.85
2	4725.73	4725.28	4725.27	4725.09
3	4725.73	4725.25	4725.25	4725.05
4	4725.79	4725.32	4725.29	4725.19
5	4725.57	4725.14	4725.17	4724.81
6	4725.39	4725.00	4725.00	4724.67

Table 5.2: Average elevation per track.

A second possible cause becomes more clear when showing the discrepancies in elevation, as done in table 5.3. The difference between both N- and FN-retracker with respect to the level 2 data are between 0.4 and 0.5 m. The binsize of Cryosat if 0.467 m, which is exactly in the order of the discrepancy. It is possible that either the processing of the level 2 data or the numerical retrackers induces a bias of in the order of the binsize. For lake level monitoring, a bias is not a problem as long as it is consistent. The discrepancy varies within 0.1 m, which is a result of variations of bias in the level 2 elevation and variations of bias in the elevation computed with numerical retrackers.

	Discrepancy [m]	Discrepancy [m]	Discrepancy [m]
Track	N-retracker	FN-retracker	WFN-retracker
1	-0.45	-0.47	-0.63
2	-0.45	-0.45	-0.62
3	-0.48	-0.48	-0.68
4	-0.47	-0.50	-0.60
5	-0.43	-0.40	-0.76
6	-0.39	-0.39	-0.72

Table 5.3: Discrepancy with level 2 data.

The WFN-retracker has a much larger error. This is mainly caused by the shape of the WFN-retracker, which is a peak with a width of approximately only one bin. This setting

is chosen, because it performs very robust in polluted waveform, which is discussed in section 5.2. It therefore retracks the peak of the waveform instead of a point on the leading edge of the waveform. In order to illustrate this problem the implementation of the WFN-retracker is shown in figure 5.2. The figure shows that there is a clear off-set present between the level 1b waveform leading edge and the WFN-retracker. Therefore the range is overestimated and the elevation too low.

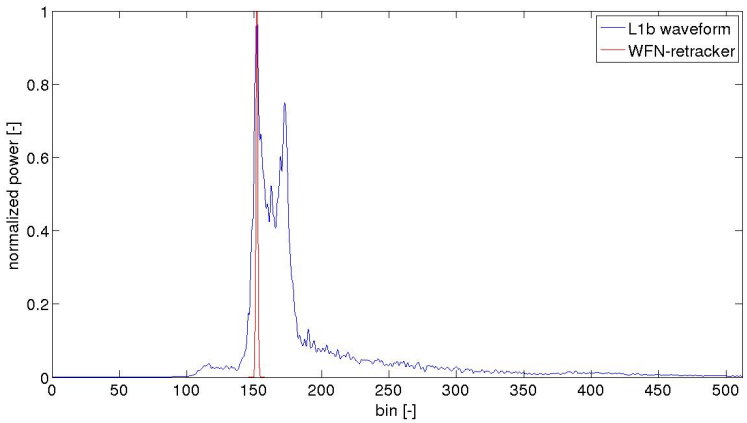


Figure 5.2: Elevations estimated over one track for various retrackers and Cryosat level 2 data.

Table 5.4 shows the precision over six unpolluted tracks over Nam Co. After filtering some outliers it becomes clear that the precision of the FN-retracker is higher than the unfiltered N-retracker. It even shows higher precision than the level 2 data. The WFN-retracker performs worst. This can be explained by the existence of speckle in the level 1b waveforms, which causes multiple peaks to exist close to each other in the waveform. The position of the highest power in the level 1b waveforms might therefore change due to the presence of speckle. Since the WFN-retracker retracks the position of the peak, its precision is worse than the other retrackers.

As a consequence of the filtering procedure, some lake level measurements are removed as outliers. The number of measurements kept is shown in table 5.1. The FN-retracker shows the highest precision, but also keeps most measurements. This indicates that the implementated FN-retracker makes consistent measurements over open water.

	Precision [m]	Precision [m]	Precision [m]	Precision [m]
Track	Level 2 data	N-retracker	FN-retracker	WFN-retracker
1	0.10	0.10	0.08	0.12
2	0.08	0.06	0.07	0.16
3	0.05	0.09	0.04	0.21
4	0.06	0.10	0.08	0.06
5	0.08	0.08	0.08	0.22
6	0.08	0.10	0.07	0.22

Table 5.4: Lake level precision over six tracks for the level 2 data, the numerical retracker and a filtered numerical retracker.

	# measurements [-]	# measurements [-]	# measurements [-]	# measurements [-]
Track	Level 2 data	N-retracker	FN-retracker	WFN-retracker
1	34/37	35/37	37/37	29/37
2	33/37	28/37	34/37	31/37
3	32/37	35/37	32/37	35/37
4	32/37	35/37	33/37	22/37
5	35/37	32/37	37/37	36/37
6	32/37	35/37	33/37	34/37

Table 5.5: Number of kept measurements and total number of measurements (37) in the six tracks for level 2 data, the N-retracker, the FN-retracker and the WFN-retracker.

5.2 Near shore lake levels

This section addresses lake level measurements from waveforms that are polluted by surrounding topography. Two additional retracking steps are applied: multiple peak retracking and lake shape adjustment, both described in section 4.3. In paragraph 5.2.1 polluted waveforms are retracked without the multiple peak procedure. In paragraph 5.2.2 polluted waveforms are retracked with the multiple peak procedure. The main purpose of both paragraphs is to show what the optimal retracking settings are, and what shortcomings the retrackers have. The ultimate aim is to find lake levels of lakes with a width of less than 7 km, as discussed in section 1.2.

The names of the retracker settings are listed in table 5.6. Since the FN-retracker showed the best performance over open water, therefore this is the only one extended and discussed in this section and not the N-retracker.

Name	Retracker settings
FN-retracker	tail-Filtered Numerical retracker
SFN-retracker	Shape-adjusted tail-Filtered Numerical retracker
MFN-retracker	Multiple peak tail-Filtered Numerical retracker
SMFN-retracker	Shape-adjusted Multiple peak tail-Filtered Numerical retracker
WMFN-retracker	Weighted Multiple peak tail-Filtered Numerical retracker

Table 5.6: Retrackers used to estimate elevations from polluted waveforms.

In order to compare the performance of the retrackers on polluted waveforms, a track close to the edge of Nam Co is selected (figure 5.3). The length of the track over water is approximately 40 km, and the track comprises 121 measurements. The distance between shore and the satellite nadir location varies between 0 and approximately 3 km. The measurements were acquired one day after track 6 from section 5.1. Statistics are therefore compared with the results for that track.

A plot is made of the fraction of land pixels along the latitude in figure 5.4. The fraction is computed over a 20 km wide strip of 300 m in length. The vertical black lines corresponde to the positions that are also visible in figure 5.3 to give an indication of the measurement locations. The pollution comes from the Western shore of Nam Co and is therefore single sided, except for a few islands. If the pollution is 50 percent, then everything to the West from the ground location is polluted by land pixels. If the pollution is 40 percent, then approximately the West 8 km of the strip is polluted, or all the pixels located further than 2 km west of the ground location.

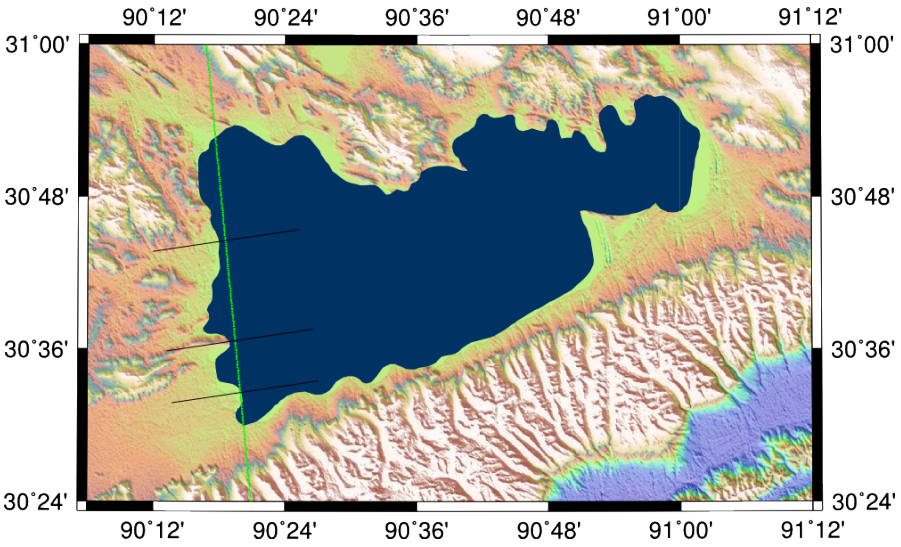


Figure 5.3: Track with polluted waveforms.

Indicated in red in figure 5.4 is the line of 42.5 percent pollution, or 1.5 km distance

from the shore. This corresponds to a lake with a cross-track width of 3 kilometers. This is approximately twice the Pulse-Limited Footprint (PLF), which is computed in section 5.4. It is expected that several valid lake levels can be found for lakes of around 3 kilometers in size.

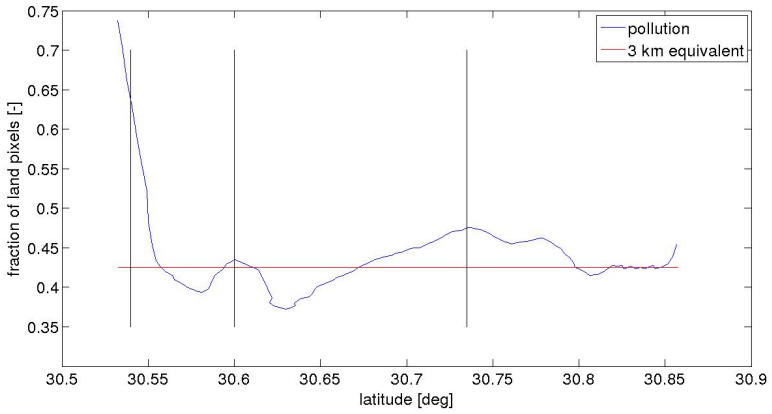


Figure 5.4: Fraction of land pixels in the footprints. Red indicates the line, that corresponds to lakes with a 3 km cross-track width.

5.2.1 Single peak retracking

In this paragraph the FN- and SFN-retracker are investigated. Both retrackers pick the largest peak in the cross-correlation to compute the correponding elevation. The results are compared to Cryosat level 2 data.

Figure 5.5 shows the level 2 elevations and the elevations computed using the FN- and SFN-retracker. The retrackers both have a discrepancy with respect to the level 2 data as is expected from the results in section 5.1. Furthermore, the numerical retracker elevations show some significant outliers, but more structured than the level 2 outliers.

First, the outliers indicated by the black boxes are discussed. A typical level 1b waveform from which the elevations are computed is shown in figure 4.2. Because we already know from track six in section 5.1 that the true water surface is around 4725 m, the peak in the waveform corresponding to the water surface is the first one. The second peak in the waveform is however larger and is therefore retracked. This causes the range to be larger and therefore the elevation to be too low. Since the initial values used in the SAMOSA retracker, which is used to compute the level 2 data, are such that the Levenberg-Marquardt algorithm finds the first peak, the elevations of the level 2 data become closer to the true value.

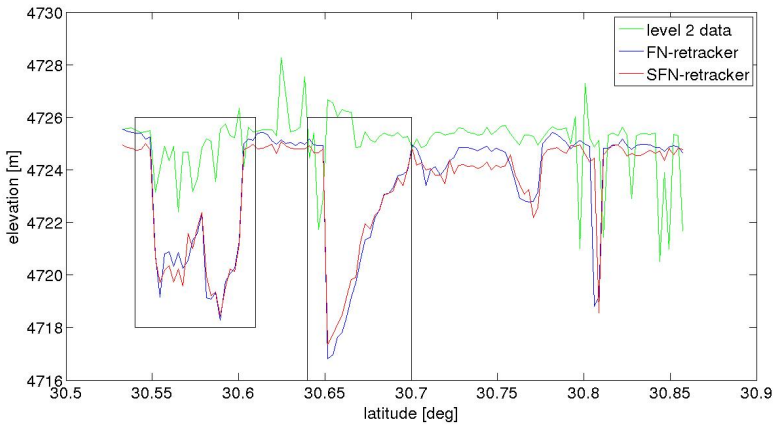


Figure 5.5: Elevations estimated over a track of polluted waveforms using single peak retracking. Indicated in the black boxes are the parts of the track where the single peak numerical retracker does not successfully retrieve lake levels.

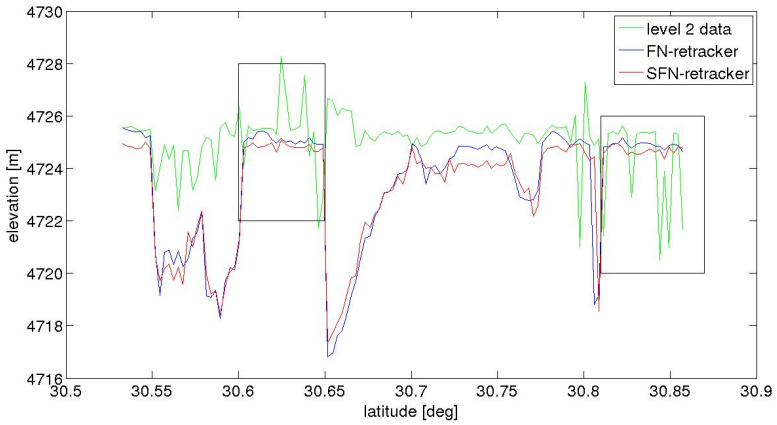


Figure 5.6: Elevations estimated over a track of polluted waveforms using single peak retracking. Indicated in the black boxes are the parts of the track where valid observations are obtained.

In other parts of the waveform, indicated with black boxes in figure 5.6, the FN- and SFN-retrackers perform better. An explanation for this is given with the help of figure 5.7. The figure shows a polluted level 1b waveform, in which the first peak correponds to the water surface. Since the peak of the water surface is relatively small compared to the second peak in the waveform, the SAMOSA retracker does not stay close to the first peak, but moves on to the second one. The shape of the simulated numerical retrackers however, is such that it better resembles the first peak.

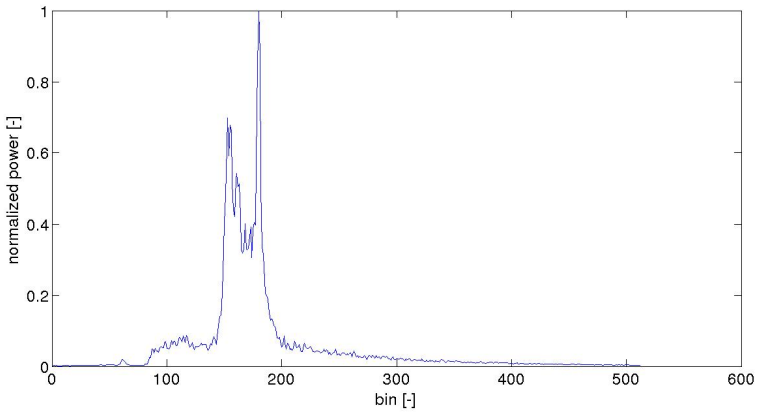


Figure 5.7: Polluted waveform of which the first peak is water power. Typical waveform between 30.80 and 30.85 degrees latitude.

Retracker	Parameter	value
level 2 data	Mean [m]	4725.35
FN	Mean [m]	4723.84
SFN	Mean [m]	4723.60
level 2 data	Discrepancy L2 open water [m]	-0.04
FN	Discrepancy L2 open water [m]	-0.55
SFN	Discrepancy L2 open water [m]	-0.79
level 2 data	Precision [m]	0.34
FN	Precision [m]	2.32
SFN	Precision [m]	2.07
level 2 data	# measurements	99/121
FN	# measurements	121/121
SFN	# measurements	121/121

Table 5.7: Comparison of mean, precision and discrepancy with unpolluted level 2 data of level 2 data, the FN- and SFN-retracker obtained elevations near shore.

Table 5.7 shows the a comparison between the FN- and SFN-retracker statistics and the statistics of the level 2 data. The outlier removal procedure applied here is the same as for open water and is described in section 4.3. The mean of the level 2 data is supposed to be close to the mean of the level data over open water. The discrepancy between those is four centimeters. The discrepancy between level 2 data over open water and the FN- and SFN-retracker is supposed to be 0.4 to 0.5 m, which is the discrepancy found over open water. However, the discrepancy is larger, especially for the SFN-retracker. Since many invalid measurements are present in the numerically retracker elevations, the precision becomes large and non of them are removed. Most of the invalid measurements show an elevation lower than the water surface, which is the cause for the increased

discrepancy. In general it is concluded that single peak retracking performs worse on polluted waveforms than the SAMOSA retrackers used to computed level 2 data.

5.2.2 Multiple peak retracking

The preceding paragraph discussed the implementation of single peak retrackers, where simply the peak with the highest cross-correlation with the simulated waveform is linked to the lake level. However, it was shown that only a relatively small fraction of the numerically retracked measurements were valid. Therefore multiple peak retracking is implemented, as described in section 4.3, which computes lake levels for all the peaks in the cross-correlation. Three retrackers are used: the MFN-retracker, the SMFN-retracker and the WMFN-retrackers, the names are explained in table 5.6.

In figure 5.8 a comparison between the single peak FN-retracker and the MFN-retracker is shown, along with level 2 data. From the figure it is concluded that most of the measurements of the numerical retrackers resemble each other, but that the MFN-retracker delivers better results in areas where the errors of the FN-retracker are large. The number of erroneous measurements is reduced, but significant errors are still present in areas with multipeak waveforms, as discussed in paragraph 5.2.1.

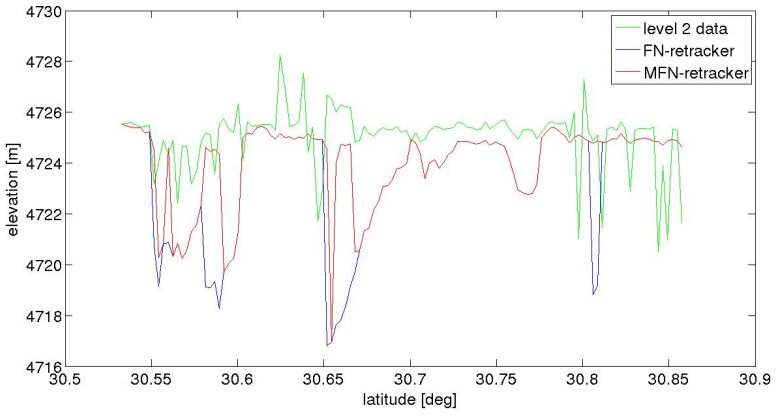


Figure 5.8: Comparison between the single peak FN-retracker, the MFN-retracker and level 2 data.

Figure 5.9 shows a comparison of the single peak SFN-retracker with the SMFN-retracker. Most of the invalid measurements of the SFN-retracker are improved by adding the multipeak procedure. The measurements are more noisy than in the case of the MFN-retracker, probably because of the tail-filtering procedure (discussed in paragraph 5.2.1) because it is optimized for the open water waveform simulation.

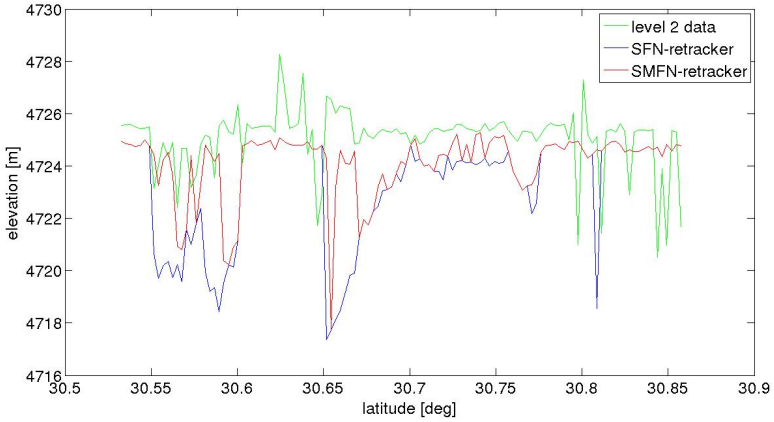


Figure 5.9: Comparison between the single peak SFN-retracker, the SMFN-retracker and level 2 data.

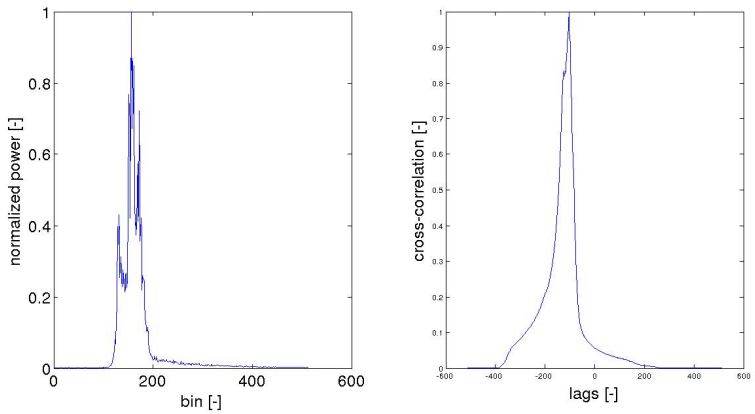


Figure 5.10: Left: a polluted level 1b waveform with three peaks. Right: the cross-correlation of this level 1b waveform with a simulated waveform with shape adjustment.

Even though multiple peak retracking is implemented, still a substantial amount of measurements with the numerical retrackers are invalid. To explain why this happens I show a typical level 1b waveform corresponding to an invalid measurement as well as its cross-correlation with the simulated waveform used in a SFN-retracker in figure 5.10. The polluted level 1b in the left plot shows three peak, while in the cross-correlation only one peak is visible. The peak corresponds to the second peak in the waveform, while the signal from the water surface is represented by the first peak in the waveform, and therefore an invalid measurement is made. The cause for this is the size of the second peak. After the first peak in level 1b waveform coincides with the simulated waveform, the cross-correlation is expected to go down, however because the second peak in the

level 1b waveform is relatively large, its multiplication with the tail of the simulated waveform increases the cross-correlation significantly. Therefore only one peak is visible in the cross-correlation.

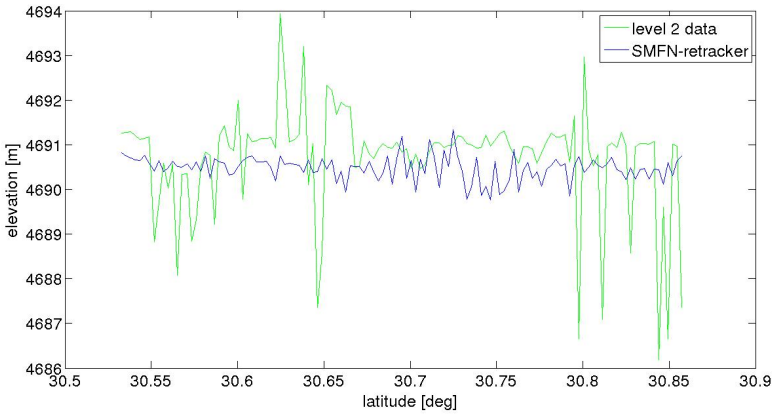


Figure 5.11: Comparison between a WMFN-retracker and level 2 data.

Retracker	Parameter	value
level 2 data	Mean [m]	4725.35
MFN	Mean [m]	4724.10
SMFN	Mean [m]	4724.49
WMFN	Mean [m]	4724.82
level 2 data	Discrepancy L2 ow [m]	-0.04
MFN	Discrepancy L2 open water [m]	-1.29
SMFN	Discrepancy L2 open water [m]	-0.90
WMFN	Discrepancy L2 open water [m]	-0.57
level 2 data	Precision [m]	0.34
MFN	Precision [m]	1.43
SMFN	Precision [m]	0.53
WMFN	Precision [m]	0.23
level 2 data	# measurements	96/121
MFN	# measurements	120/121
SMFN	# measurements	108/121
WMFN	# measurements	119/121

Table 5.8: Comparison of mean, precision and discrepancy with unpolluted level 2 elevations of level 2 data, the MFN- and SMFN-retracker computed elevations near shore.

Therefore, in strongly polluted cases we decide to use a more robust approach. The weighting function is implemented as described in section 5.1. As shown in figure 5.11 the implementation of the WMFN-retracker produces almost 100 percent valid mea-

surements over a strongly polluted track. This is at the costs of accuracy for the valid measurements that are obtained using the other retrackers. The reason for this decrease of accuracy is described in section 5.1. By comparing the figure 5.11 with figure 5.4 it is visible that the precision deteriorates around place with more pollution.

The statistics for the MFN-, SMFN- and WFN-retracker are shown in table 5.8. Outliers larger than three times the RMS are removed. Because of the high number of invalid measurements present in the MFN- and SMFN-retrackers the precision is worse than the level 2 data. Since most of the invalid measurements are below the actual water level, the discrepancy with level 2 data over open water (of track 6) is substantially larger than the discrepancy of the FN-retracker over open water. The performance of the SMFN-retracker is better than the MFN-retracker. The WMFN-retracker shows a comparable discrepancy with level 2 data over open as the WFN-retracker discussed in section 5.1. The precision over the near shore track of the WMFN-retracker is higher than the level 2 data. Furthermore, 119 of the 121 measurements are kept indicating robustness. In the level 2 data only 96 measurements are kept.

Alltogether, the best implementation over this polluted track is the WMFN-retracker. However, this is a track with severe pollution in the waveforms. Better results might be obtained with the MFN- or the SMFN-retracker over less polluted waveforms. For the SMFN-retracker much more computation time is required, but it likely to give slightly better results.

5.3 Retracker implementation

Since the performance of several retracker settings are discussed in the previous section, this section determines how and when to implement them. For example, the FN-retracker gave the best performance over open water, while the WMFN-retracker gave the best performance near shore.

In sections 5.1 and 5.2 I demonstrated the implementation of several retracking methods applied to measurements over respectively open water and near shore. It showed that different retracker settings were best for unpolluted and polluted waveforms. Over open water the single peak FN-retracker showed a consistent bias of 0.4 to 0.5 m with Cryosat level 2 data and a precision of approximately 0.07 m, which was better than level 2 data. Therefore it is suggested to implement the FN-retracker over areas where hardly any waveform pollution occurs. The cross-track distance to the shore should therefore be at least 3.5 km, which is in agreement with the minimal diameter of lakes (7 km) for measurements with conventional radar altimeters (see section 1.2).

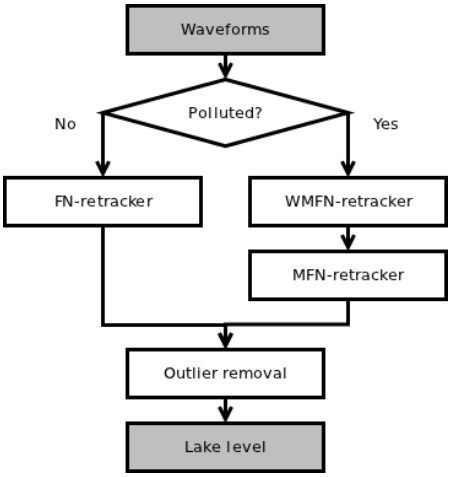


Figure 5.12: Flowchart of suggested retracking implementation.

In paragraph 5.2.2 I showed that only a limited number of valid observations were obtained using the MFN-retracker on polluted waveforms, and that the precision is deteriorated. The precision of the measurements improves by using the shape adjustment (SMFN), but the number of valid measurements remains limited. In paragraph 5.2.2 it is also found that the number of valid observations increased using the WMFN-retracker, using a weighting function and multiple peak retracking. The WMFN-retracker is robust, but is limited in precision. Over open water the retracker with a weighting function performed worse than the other retrackers.

Therefore it is suggested first to implement the WMFN-retracker over waveforms polluted by surrounding topography. And additionally the MFN-retracker, to estimate lake levels more accurately if the MFN-retracker finds enough valid measurements. Another option is to implement the SMFN-retracker over tracks with limited pollution, but the computation time is higher since a new waveform must be computed for every measurement.

Outlier removal as described in section 4.4 is set to three times the RMS and a minimum of at least 50 percent of the measurements is kept. In case of very small lakes the minimum number of measurements is five. A flowchart of the suggested implementation is shown in figure 5.12.

5.4 Target lakes in Tibet

Up to now approximately 50 lakes are monitored using conventional radar altimeters [41]. Based upon the results found in section 5.2, the number of potential target lakes is determined that could be measured with Cryosat if the ground track passes over the lake.

The numerical retracker is able to find elevations of lakes over which the waveforms are severely polluted by surrounding topography. As discussed in section 1.2 the lake should have a diameter of at least 7 km for proper retrieval of lake levels using a conventional radar altimeter. As discussed in section 2.2.4 the size of the footprint in along-track direction improves to approximately 300 m. Cryosat is orbiting in an almost polar orbit (section 2.1.2). Therefore we set the width of the lake as the difference between the minimum and the maximum longitude of the lake at a particular latitude. And the length of the lake as the difference between the minimum and maximum latitude at a particular longitude. The minimum width of a lake is therefore 7 km and the minimum length around 300 m in SARInM for which we at least expect to have appropriate results in Cryosat level 2 data. However, it is very difficult to determine an elevation from one single lake level measurement. If five consecutive measurements give consistent results the lake level estimate is more reliable. With a measurement every 300 m, the length of the lake is assumed to be at least 1500 m.

The goal of the numerical retracker is to find levels of lakes with a width less than 7 km. As shown in this chapter, the nadir location of the satellite can be as close as approximately 1.0-1.5 km for which the WMFN-retracker still gives appropriate results. This correponds to a lake of 2-3 km in cross-track width. By comparing figures 5.4 and 5.11 it is unlikely that decimeter level accuracy measurements are found smaller. This also comes close to the Pulse Limited Footprint. It is unlikely to find lakes smaller than the Pulse Limited Footprint (PLF), because it would mean that every bin in the waveform is at least to some extend polluted, even if the land next to the lake is practically flat. It is more likely that the proper results are found if the cross-track widht is twice the diameter d_{PLF} of the PLF is computed first. In order to clarify what the footprint size is, figure 5.13 shows the PLF's geometry in cross-track direction. The PLF is the area of which the signal corresponds to the first bin of the waveform. A top view of the PLF at bin 201 is given in figure 3.10. With range r and range resolution r_{res}, the diameter of the PLF is computed using Pythagoras as:

$$d_{PLF} = 2\sqrt{(r + r_{res})^2 - r^2}. \tag{5.1}$$

Using a range of $717 \cdot 10^3$ m and a range resolution of 0.469 m, the diameter of the the PLF is approximately 1.65 km. This is the mimimum footprint width for Cryosat under optimal circumstances. Water waves increase the peak width and therefore its footprint. Furthermore, the minimum diameter computed requires the centre of the footprint to be located in the middle of the lake, so it is more likely to find elevation of lakes larger than 3 km in diameter. The results in figure 5.11 for the WMFN-retracker showed that for lakes smaller than 3 km (or closer than 1.5 km to the shore) the estimated water levels started to deteriorate. Therefore the target lakes are assumed to have a width of at least 3 km and a length of 1.5 km.

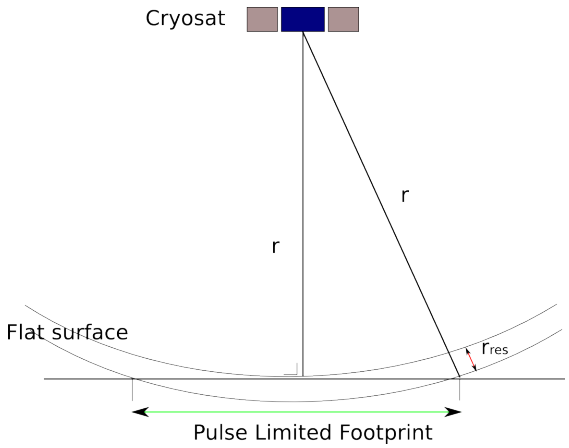


Figure 5.13: Geometry of a PLF. The term r_{res} indicates range resolution.

In order to identify the target lakes in Tibet the MODIS water mask is used, which is described in section 4.3. For all water regions in the water mask the maximum width and maximum length is computed in terms of the number of pixels. The number of pixels in the longitude direction is related to the width of the lake in kilometers depending on the latitude. The number of pixels in the latitude direction is related to the length of the lake in kilometers by a constant. A complete description of the computation of the number of potential target lakes is given in Appendix C. Ultimately, this leads to 488 potential lakes on the Tibetan plateau. A geographic overview of the potential lakes is given in figure 5.14.

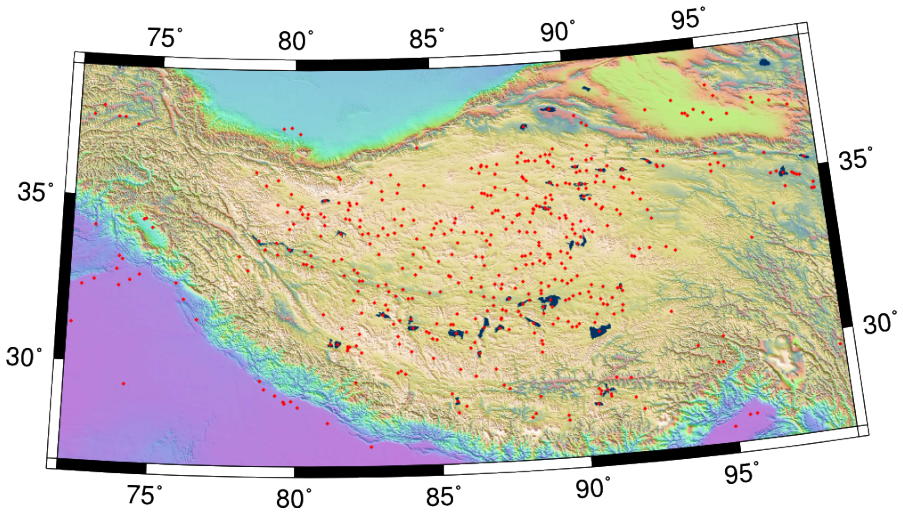


Figure 5.14: Topographic map of Tibet using a GTOP030 DEM, with target lakes indicated in red.

This number is an upperbound, because it includes several rivers and because it assume the minimal required width and length. If the lake only has a width of 3 km over a length of 250 m, then it is unlikely that five proper measurements are found. Therefore the expected number of targets is likely to be less. Furthermore, the ground-track spacing of Cryosat over Tibet is approximately 6.5 km, so for many lakes no measurements of water level are available.

5.5 Lake level timeseries

The retrackers are implemented as described in paragraph 5.3 to show timeseries of lake levels over the year 2012. This section is divided into a paragraph that shows a timeseries over a large lake Nam Co (paragraph 5.5.1) and a paragraph where an attemps is made to compute a timeseries over a small lake (paragraph 5.5.2).

5.5.1 Nam Co

A overview of the Cryosat tracks in the year 2012 over the Nam Co lake is given in figure 5.15. Over 2012, 17 tracks seemed suitable to monitor the lake level. One track is discarded, because for some unidentified reason there is a gap in the measurement over Nam Co. Possibly the gaps are the result of miscomputations of the tracker, which regulates antenna gain and the timing of the deramping chirp. One track did not have enough valid valid measurements (the minimum number taken into account is 10). And furthermore, the tracks in January and a part of February 2012 have been processed with a different level 1b processor, therefore these are discarded as well. The tracks over January and February are not displayed in the figure.

The retrackers are applied as described in section 5.3. Therefore the FN-retracker is applied if less than 32.5 percent of the footprint are not pixels from the water surface (this corresponds to 3.5 km from the shore). If the average pollution is more severe, the WMFN- and MFN-retrackers are applied. In this case only the WMFN-retracker gave best results for the near shore tracks. This leads to fourteen tracks over which the FN-retracker is applied and three tracks over which the other retrackers are applied.

The resulting timeseries computed with the settings described above are displayed in figure 5.16. In green an estimate of a trend and a seasonal variation is shown. The red lines indicate the 95 percent confidence interval, computed as two times the precision of measurements within the track. The actual accuracy is lower than the confidence interval, since the bias is not taken into, which is not a problem as long as the bias of the measurements remains constant. However, the biases between different retracker settings differ. Since only discrepancies with level 2 data are computed, nothing is done to remove these biases. Furthermore, measurements over one track are not completely

independent, since they are made under approximately the same conditions, which is another reason why the confidence interval is underestimated.

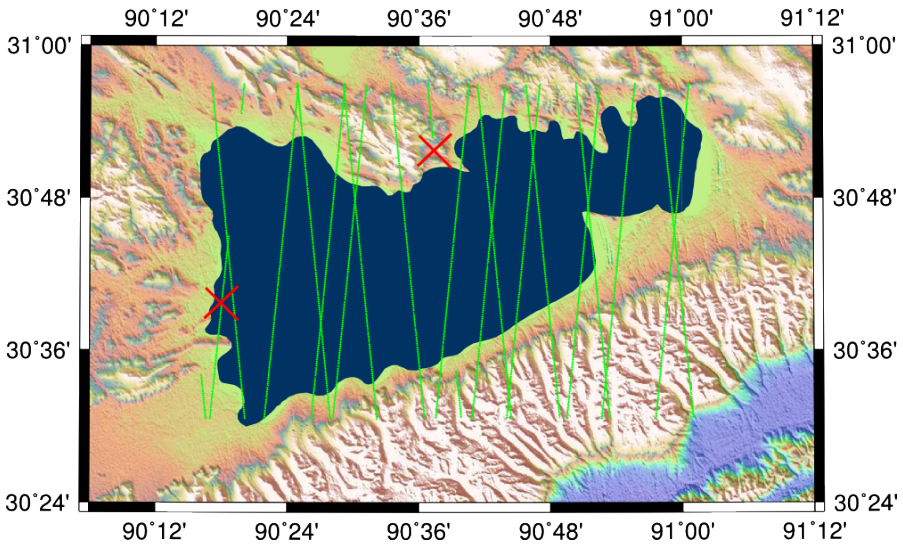


Figure 5.15: Cryosat ground tracks of 2012 over Nam Co.

Table 5.9 shows the amplitudes of the estimated trend and seasonal variations. The trend estimated from the timeseries is a few millimeters and its Signal-to-Noise Ratio (SNR) only 1.63. The estimated amplitude of the trend is therefore not very reliable. Estimation of a trend from longer timeseries after a couple of years of data acquisition by Cryosat will likely increase the SNR. The seasonal variations are clearly visible in figure 5.16. The estimated amplitude of the seasonal signal is 0.32 m and is more reliable because of the SNR of 3.20. Since the water level does not follow perfectly a sinusoid, better estimations might be obtained including higher frequency harmonics.

In conclusion it is possible to state that lake level estimates from Cryosat, using the approach suggested in this report, are accurate enough to find seasonal variations in water level for Nam Co. Since Nam Co is large, Cryosat passes many times over it, which could allow to even estimate subseasonal variations. The timeseries of measurements over Nam Co is not yet long enough to see a trend in lake level, but it might be visible after a couple of years of operation.

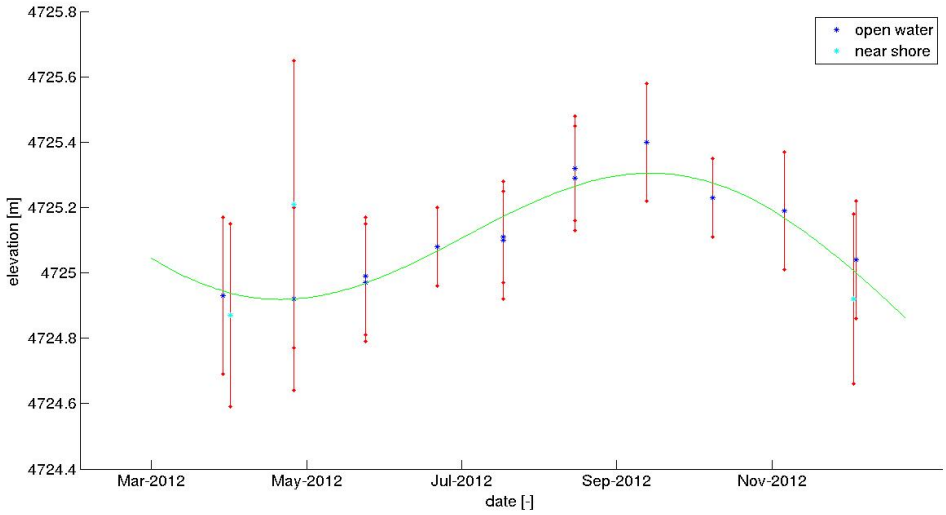


Figure 5.16: Cryosat ground tracks of 2012 over Nam Co.

Parameter	amplitude [m]	std. dev. [m]	SNR [-]
Trend	-0.00	0.00	1.63
Seasonal variation	0.32	0.10	3.20

Table 5.9: Amplitude, standard deviation and SNR of the trend and seasonal variations.

5.5.2 Small Lake

In order to study if the suggested numerical retracker is able to find water levels of small lakes, a lake of approximately 4 km in width is chosen. The lake is located at the latitude of 31.7 degrees North and the longitude of 90.7 degrees East. This is approximately 80 km North of the Nam Co lake. Since all the measurements are within 3.5 km from the shore the MNF- and WMFN-retrackers are applied only. A comparison is made with official level 2 data. If we show that a lake level of such a small lake can be measured with a SAR altimeter, not only timeseries can be made using Cryosat after a few years, but the same method could be implemented to the SAR altimeter measurement of the Sentinel-3 satellites, which are to be launched in 2014 and 2015 [42].

The tracks over the small lake are visible in figure 5.17. The first problem that occurs is that only one track is passing the lake during 2012. Therefore it is not possible to show a timeseries of elevations. Cryosat has an advantage of having a very dense track spacing, because of a repeat-track of 369 days. Which has a disadvantage that small lakes are only measured every 369 days (once per year), or if the lake is close to a crossover two times per year. Once the Sentinel-3 satellites are flying, which have a repeat-track of 27

days, some of the small lakes can be measured every month.

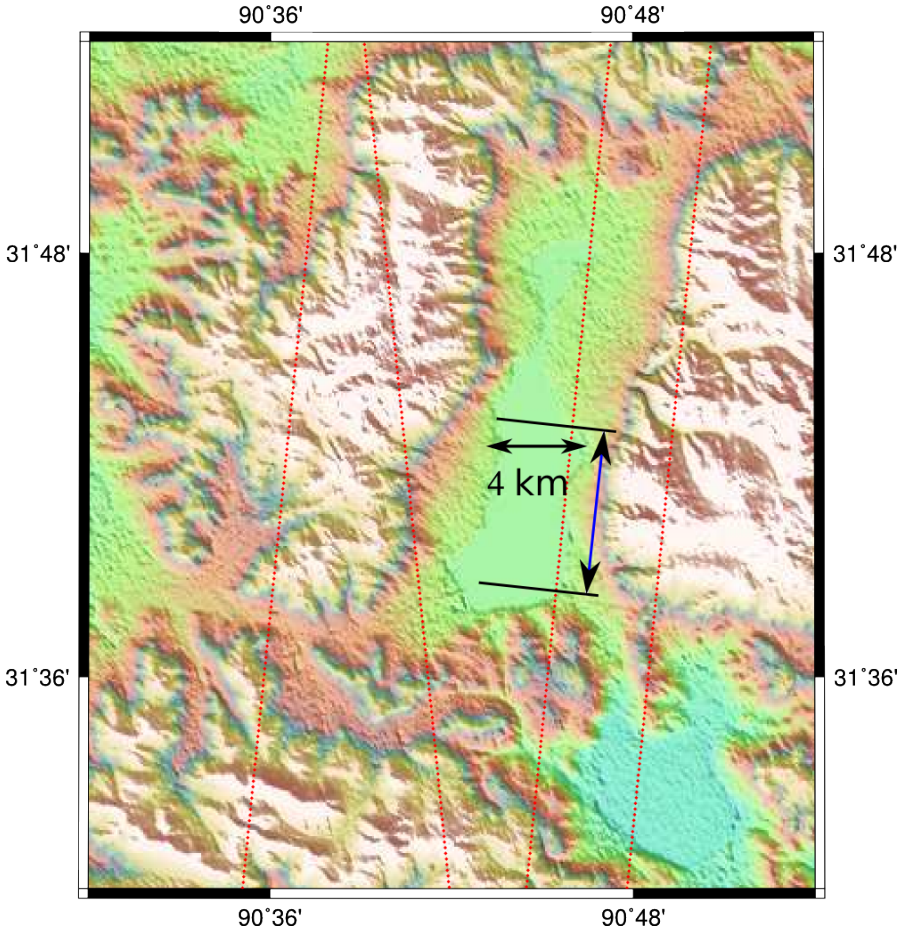


Figure 5.17: Cryosat ground tracks of 2012 over a small lake. Indicated with blue the part of the track were valid lake levels are obtained.

Parameter	Level 2	WMFN-retr.	MFN-retr.
Mean [m]	4686.44	4686.08	4686.15
Precision [m]	0.80	0.11	0.25
Included # measurements	36/41	32/41	31/41

Table 5.10: Comparison of mean, precision and number of measurements of level 2 data, the WMFN-retracker and the MFN-retracker.

In table 5.10 the results of the measurements of the only track over the small lake are shown. The table shows that there is a significant improvement when applying the suggested numerical retrackers, instead of using level 2 data. Since there are so many peaks in the waveforms it is necessary to use multiple peak retracking, which is not

implemented in level 2 data. The WMFN-retracker performs better than the MFN-retracker. This is caused by the extreme pollution of the level 1b waveforms as shown in figure 5.18. The MFN-retracker has problems finding peak in the cross-correlation in this case, therefore weighting is necessary, as discussed in section 5.2.

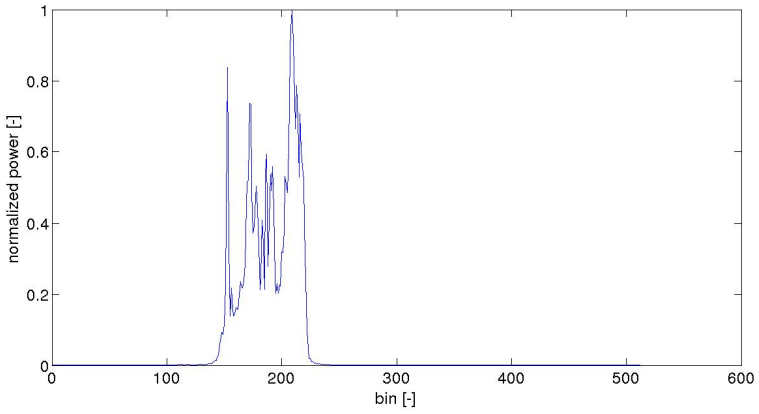


Figure 5.18: Waveform over a small lake, with high surrounding mountains.

Alltogether, the precision of the WMFN-retracker is as expected from the findings in section 5.2. The WMFN-retracker shows an off-set of 0.36 m from the level 2 data, which is smaller than found in the sections 5.1 and 5.2. However, the precision of level 2 over the small lake is much lower than over Nam Co, which suggests that some invalid measurements are in the computation of level-2 elevations.

5.6 Conclusions

The primary goal of this chapter was to give an answer to the subquestion: **How accurately can lake levels be retrieved with a numerical retracker?**

Several numerical retracker settings were applied in this chapter. In order to get the best results over open water, a tail-Filtered Numerical (FN)-retracker is applied. The precision of the FN-retracker is approximately several centimeters.

The results deteriorate closer to the shore, especially when the nadir location of the satellite is within 3.5 km of the shore cross-track. Therefore, a Weighted Multiple-peak Filtered Numerical (WMFN)-retracker was applied. It searches for the best peak in the waveform to get appropriate lake level measurements at cost of some accuracy. In some cases the unweighted Multiple-peak Filtered Numerical (MFN)-retracker gives better results. It is less robust than the WMFN-retracker, but has better accuracy when the pollution is not very substantial. The MFN-retracker is expected to yield better results as the nadir location of the satellite is cross-track gets further from the coast.

Over open water lake levels are retrieved with a precision of less than a decimeter, which is slightly better than in the case of level 2 data. A discrepancy with level 2 data is present over open water between 0.4 to 0.5 m. This is most likely a bias of the retracker, or a bin (0.47 m) offset caused by processing differences. Close to shores the precision deteriorates depending on the pollution of the waveform. If the amplitude of the signal from the water gets relatively low compared to that of the signal from the topography, the accuracy strongly deteriorates. If the nadir point of the satellite is more than 1.5 km from the coast (which corresponds to lakes larger than 3 km) the precision of the numerical retracking approach is around 0.1 to 0.3 m after filtering all outliers above three times the RMS. The accuracy is expected to be slightly worse than taken the precision per track into account. First of all, because the measurements in one track are not completely independent. Secondly, because the bias is not taken into account. Since no in-situ data is present no structural bias can be detected.

For Nam Co a set of 17 measurements is available of in 2012. This number is expected to be larger in 2013, because in 2012 during the first period of 2012 several measurements were discarded, as they were processed with a different level 1b processor. For small lakes of around 4 km in width the number of tracks reduces to 1 or 2 measurements per year. The cause for this is the 369 day repeat-track of Cryosat. The pattern of Cryosat ground-tracks has a 6.5 km cross-track spacing over Tibet and therefore most lakes in Tibet larger than 3 km are passed by a Cryosat track, but with low temporal resolution. With the upcoming Sentinel-3 satellites with their SAR altimeters, the temporal resolution may increase, because they have a cycle of ground-tracks of 27 days.

Chapter 6

Summary, Conclusions and Recommendations

In this thesis I have shown how water levels of Tibetan lakes can be estimated by applying a customly designed numerical retracker on Cryosat level 1b waveforms. With this chapter I furhtermore discuss to what extend the implementation the numerical retracker is a success. First a summary of the findings is provided in section 6.1. Then conclusions are given in section 6.1.4. Furthermore there is room for improvement and further research in the area of SAR altimeters for lake level monitoring, a detailed list of recommendations is provided in section 6.2.

6.1 Summary

The primary purpose of the Cryosat instruments is to provide geophysical parameters, like elevation, along the satellite track. In order to achieve this a signal is transmitted, then received and processed at a ground station to obtain geophysical parameters. Cryosat's altimeter SIRAL operates in three different modes, depending on the location on Earth. Over Tibet SIRAL operates in a interferometric SAR mode (SARInM), which has a better along-track resolution than a conventional radar altimeter of 300 m and it is able to compensate for cross-track slope.

The transmitted signal of SIRAL is a chirped pulse, which is transmitted in burst of 64 pulses in SARInM. The deramping procedure applied, causes the signals within the echo to get constant frequencies depending on their range or travel time. These signals are discretized into 512 complex samples and downlinked to a ground station.

At the data processing facilities the signals are processed to eventually retrieve geophysical parameters. Cryosat's processor consists of three stages: level 1, level 1b and level

2. In the level 1 processing stages, records are set up and instrumental corrections are added to the records. It also provides timing and geolocation of the measurements.

The level 1b processing is more complicated. The first step is to apply the instrumental corrections computed in level 1 and to compute geophysical corrections, like ionospheric and tropospheric corrections. Then an along-track FFT procedure is applied, which exploits the small phase shift of signal in the 64 pulses. It results in individual signal for 64 strips of 300 m along-track. A beamsteering procedure ensures that beams from several bursts are aligned at one ground location. After this, a Doppler correction and a slant range correction are applied to correct the ranges to a reference range. Then a range FFT procedure is applied, which uses the constant frequencies in the deramped echo to convert it to a single look waveform. Waveforms from different bursts are collected in a stack for one ground location. The waveforms are averaged to create a multilooked waveform.

In level 2 processing, the computed geophysical corrections computed in level 1b are applied. Then, a theoretical description of the waveform is fitted, by optimizing several parameters, into the multilooked level 1b waveforms. These parameters are related to geophysical parameters, like elevation.

6.1.1 Waveform modelling

Waveforms were simulated using an algorithm coded in a set of MATLAB programs. One input for the algorithm is a Cryosat level 1b file containing satellite burst positions, satellite velocity vectors and ground locations. The other input for the algorithm is a set of scatterer locations. The height of these scatteres are either computed using a DEM, or a flat surface is used.

Using the satellite velocity vector and burst locations the pulse locations are computed. Based on the relative positions and velocities of the scatterers with respect to the pulse locations, a deramped signal is computed and sampled at 512 points. This is the output of the satellite.

The level 1 processor operation are not applied, since no instrumental errors are simulated and the waveforms are already in right format. In the level 1b processor the geophysical corrections are ignored, since no geophysical errors or delays are introduced in the algorithm. The rest of the steps in the level 1b processor are applied, leaving one multilooked waveform per ground location.

In order to validate the waveform modelling algorithm, the simulated waveforms are compared to real level 1b waveforms over open water. Best results are obtained when setting the surface roughness, or water wave heights, to zero. The RMS difference between level 1b waveforms and the simulated waveform was on average, 2-3 percent. However, the largest differences were present around the peak, which is the most important part of the waveform.

The width of the level 1b waveform peak suggests a non-zero surface roughness, but increasing the surface roughness in the simulations resulted in higher RMS differences. Probably taking into account a dependancy of the backscattering coefficient on incidence angle and therefore the roughness of the surface, would increases the quality of the simulation. This dependence is also a possible explanation for the variety of shapes of the observed level 1b waveforms over open water.

6.1.2 Waveform retracking

In order to compute elevation in the level 2 product of Cryosat, the Cryosat processor uses analytical retrackers. These retrackers have several weak points when computing lake level, since they are designed for oceans. Once the shape of the waveform is polluted by surrounding topography, which occurs near the shore of a lake, the accuracy of the measurements decrease.

Up to now the minimal diamater of a lake over which a waveform could accurately be retracked in radar altimetry is approximately 7 km. In case of Cryosat in SARInM, the lake can be smaller in the along-track direction since the along-track resolution is improved. However, waveforms over lakes with a cross-track width smaller than 7 km are still polluted.

Waveforms with minor pollution are almost always retracked by the Cryosat retracker to an accuracy of several centimeters to decimeters. A more severe problem occurs when multiple peaks are present in the waveform, so that the lake level estimates depend on the initial conditions of the retracker optimization algorithm. Furthermore, the lake only covers a part of the footprint and therefore the signal over the lake is different from that over open water.

In order to mitigate the effects that occur due to pollution of the waveform, a numerical retracker based on simulated waveforms is applied. Over open water, the retracker is based on a flat surface response and tries to find the optimal fit by correlating the simulated waveform with the Cryosat level 1b waveforms. Over near shore regions, the retracker first uses a weighting function, putting large weight on the correlation around the peak of the retracker, to find the peaks in the waveform. In the near shore regions, lake levels corresponding to multiple peaks in the waveform are computed. The elevations over one track are accumulated and the mode of this stack of elevations is determined, after which the best solutions per waveform are kept. Additionally a retracker, adjusted to the part of the lake inside the footprint is applied to further improve the accuracy of the measurements. However, the shape adjusted retracker is less robust than the weighted retracker and is only able to provide accurate measurements over slightly polluted waveforms.

6.1.3 Accuracy of lake level measurements

Over lakes in Tibet two situations are distinguished: open water measurements and near shore measurements. Over the open water a filtered numerical retracker is applied based upon a flat surface response. Over near shore regions, a weighted multiple peak retracker is applied, which is more robust at the cost of accuracy. In several near shore cases, a unweighted retracker is applied to improve the accuracy.

Over open water the numerical retracker has a consistent bias with respect to Cryosat's level 2 data of approximately 0.4 to 0.5 m. This comes close to the bin-size of 0.47 cm, and is therefore likely a miscalculation by one bin. The precision over open water within a track is below one decimeter, a slight improvement over level 2 data, which has an precision of around one decimeter.

The absolute bias of the measurements is unknown, because of lack of (absolute) in-situ lake level measurements over 2012. A bias, if consistent, is however not an obstacle for monitoring lake level measurements. The accuracy (if we would remove the consistent bias) is likely to be larger than precision over one track, because the measurements in one track are not completely independent.

Near the shores of large lakes or over small lakes, the accuracy of level 2 measurements deteriorates. Many unvalid measurements are present due to the pollution of the waveforms. The weighted multiple peak retracker showed a much better performance there: practically all measurements were valid, but at the costs of accuracy. If the nadir point of the satellite is approximately 1.5 km away from the shore, the lake levels can be retrieved to approximately 0.24 m precision. This number is based upon the measurements taken into account after removng outliers. However, compared to the open water measurements, the bias with level 2 data over open water is less consistent and therefore the actual accuracy is probably several centimeters to a decimeter worse.

Over some parts of the polluted tracks, an unweighted multiple peak numerical retracker can be applied. It is less robust, but it gives better performance if the waveforms are less polluted. The bias with unpolluted level 2 data is less consistent than the open water numerical retracker and therefore the accuracy is several centimeters worse.

6.1.4 Further conclusions

The accuracy of the lake level measurements is dependent on several factors, like orbit accuracy, accuracy of instrumental corrections, accuracy of the geophysical corrections, etc. Especially, the accuracy of geophysical corrections is of interest when it comes to Cryosat, since this is worse than most other altimeters. For example, the altimeter onboard Envisat uses for example two frequency measurements to get rid of the ionosphere and a radiometer to estimate the delay of the signal due to wet troposhere, while Cryosat uses models for this which are less accurate. The absolute accuracy cannot be determined from the precision over one track.

The measurements of Cryosat are accurate enough to find seasonal signals in lak level. In a timeseries of the Nam Co lake level over 2012, annual fluctuations are visible with an amplitude of approximately 0.32 m. Since only data over only one year where used, it was not possible to find a reliable trend in the data. However in the upcoming years more data might be available, providing the possibility to extend the timeseries and find a trend.

Over small lakes the average accuracy of lake level measurements decreases compared to the large Nam Co lake, due to pollution. The accuracy increased strongly using the suggested approach in this thesis. Furhtermore, the fluctuations in small lakes are expected to be higher and therefore are likely above the noise level. The temporal resolution of water level measurements over a small lake is quite low, due to the 369 day repeat-track of Cryosat, which only causes the satellite to pass a few times per year.

Since there are still several problems to be solved with the Cryosat data, and every couple of months the data is reprocessed, the accuracy is likely to increase. The upcoming altimeters of the Sentinel-3 mission will incorporate a radiometer and collect dual frequency measurements, so the accuracy is expected to increase even further.

6.2 Recommendations

In this thesis I propose a numerical retracker designed to find water levels of Tibetan lakes. The proposed retracker shows a better performance than the analytic one used in the official level 2 data processing. However, there are still some unsolved problems. Also a validation of Cryosat data with in-situ measurements should still be performed. And since more SAR altimeter missions are to be launched in the coming years, the method should be adapted to those missions as well.

In this last section some improvements are suggested for waveform modelling and retracking techniques. Furthermore, some validations methods are proposed and an application to the upcoming mission is discussed.

6.2.1 Waveform modelling

Backscattering coefficient

Over open water the waveforms are not polluted, however they still differ in shape. This has one important reason: the surface roughness varies. Surface roughness has a direct effect on the shape of the waveform, because of varying travel time of the signal from surface scatterers. It however also changes the backscattering of the surface. When the water is smooth at the spatial scale comparable to the wavelength, water acts like a mirror, therefore the waveform consists of one thin peak (specular reflection). It

is therefore suggested to let the backscattering depend on the surface roughness and incidence angle.

Surface roughness

Since the backscattering coefficient, as well as the relative distance between the satellite and the scatterers depends on surface roughness, both effect the shape of the waveform. In this thesis a surface roughness with a standard deviation of zero was taken, since it gave the best fit with the level 1b waveforms. It is likely however, that after including the backscattering dependence on surface roughness and incidence angle, a better fit is found when setting the surface roughness higher than zero. As a suggestion, it is proposed to model waveforms over various surface roughnesses and use all of them as retrackers. The best fit will be then used to feine the value of surface roughness for furhter computations.

Averaging waveforms

Multilooking of a waveform is applied to reduce speckle. However, both in level 1b and simulated waveforms a substantial amount of speckle is still present. Inconsistencies between the observerd and simulated waveforms can lead to several centimeters or decimeters error in the resulting elevation. In this thesis the tail of the waveform is filtered to improve the fit. However, significant errors are still present around the peak, but if this part is filtered the distinctive peak in the waveform will dissappear.

It was shown that averaging unpolluted level 1b waveforms over a whole track, lead to a much smoother result. By varying the scatterer positions randomly within the footprint, different waveforms can be created with their own speckle pattern. By averaging these waveforms the modelled retracker waveform will be much smoother.

Improve lake shorelines

To determine if a satellite location is above water, a MODIS water mask with approximately 230 m resolution is used. However, this water mask is from 2010, and if lake levels varied, shorelines might have changed significantly. A solution to this problem is to use multispectral Landsat images to classify the surface. Not only will the temporal gap between the Landsat water mask and the Cryosat measurements be smaller, but additionally the accuracy will improve to approximately 30 m.

Reduce the amount of computational time

At this moment the algorithm is written in MATLAB. Modelling one waveform takes up to 6 minutes and therefore the algorithm is computationally expensive, especially in the near shore regions where lake shape adjustment is applied. After writing the algorithm in a compute language, like C or Fortran, the computational time might decrease substantially.

The computation of the deramped signal is the most computationally expensive part of the algorithm. It takes almost 50 percent of the computational time. Therefore

furhter attempts of code optimization should be focused, in the first instance on this particular.

Learning algorithm

In the algorithm, various settings for filters and retracking options are set. For example the tail of the modelled waveform is filtered using a Gaussian filter. And the weighting function, which is also based on a Gaussian around the peak.

The settings of one filter or option affects the results of the others. Now the settings are such that they give adequate results, but they can further be improved. A learning algorithm can be implemented to find the optimal settings automatically.

6.2.2 Retracking techniques

Cross-correlation improvements

Over open water the cross-correlation method showed a slight improvement over level 2 data. A weighting function had to be applied to a multiple peak retracker, the method used over strongly polluted tracks, in order to make the method robust and obtain valid results. Without the weighting function the accuracy increased over valid measurements, but the number of invalid measurements increased substantially. The primary cause for this problem is the implementation of the cross-correlation function. In several cases of multiple peak waveforms, only one peak is distinguished in the cross-correlation graph of unweighted retrackers. A large second peak in the waveform was the cause for the disappearance of the first peak.

A possible solution might be to only compute the cross-correlation locally and not over the whole waveform. This way other peaks in the waveform do not affect the cross-correlation.

Retrack leading edge

Another option is to use a completely different approach than cross-correlation. In multiple peak waveforms parts of the leading edges are clearly distinguishable. However, using the cross-correlation approach the optimum solution is based on the power. Since we are primarily interested in travel time it is better to use an approach to minize the distance between leading edges in the level 1b and simulated waveforms.

Therefore it is suggested first to detect leading edges in the level 1b waveform. Than remove all information from the waveform, except the leading edges. Than fit the leading edge of the simulated waveform on the leading edges in the level 1b waveforms.

6.2.3 Validation

Validation with Jason-2

In this thesis it was only shown what the precision of Cryosat lake level measurements is when the new retracker is used, and how strong the bias is compared to the Cryosat level 2 data. However, an actual accuracy and bias are not known. The bias is for lake level monitoring not important as it does not change over time.

During 2012 another altimetry mission, named Jason-2, was flying as well Jason-2 uses a conventional radar altimeter and is proven to provide accurate measurements over a few lakes in Tibet [41]. It is adviced is to use Jason-2 to compare Cryosat lake levels and validated if the trends and patterns found with Cryosat resemble those of Jason-2. This is only possible over large lakes, since Jason-2 carries a conventional radar altimeter.

Validation with in-situ lake levels

The lake level measurements of several altimeters, like RA-2 onboard Envisat and GLAS onboard IceSat, have already been validated over large lake like Nam Co and Qinghai on the Tibetan plateau. Lake level measurements over 2012 were not available at the time of this project, but may become available in the near future.

If the in-situ lake level measurements are absolute, the bias in the Cryosat-2 data can be determined and an accuracy can be found over a particular lake. If the in-situ lake level measurements are relative to some measurement point, the fluctuations can be compared, but the bias can not be determined.

6.2.4 Future missions

Sentinel-3 and Jason-CS

In the upcoming years two SAR altimeter missions are expected to be launched. Sentinel-3 is a two satellite mission, of which the first satellite is to be launched in 2014 and the second one in 2015. The satellites will have a repeat-track of 27 days and will therefore have more passes over several small lakes, at the costs of coverage. Jason-CS is a continuity mission of Jason series of satellites and will also have an altimeter almost similar to that of Cryosat. It is expected to be launched in 2017 as the follow-up of Jason-3. The Jason satellite will have a repeat-track of 10 days.

SWOT

A Ka-band Radar Interferometer will be launched in 2020 onboard the Surface Water Ocean Topography (SWOT) satellite. This is not a radar altimeter, but a completely new type of instrument. It is expected to resolve spatial samples of several hundreds of meters over continental surface waters. The satellite will have a repeat-track of 22 days. From 2020 on, this will be an additional source of information on water levels, next to satellite altimetry.

Appendix A

Parabolic Cylinder Function

In this appendix two specific forms of the parabolic cylinder function are described. These are required to compute the theoretical waveform described by Eq. (2.67). The parabolic cylinder function $D_{-a-\frac{1}{2}}(z)$ can be written in the form [32]:

$$U(a,z) = D_{-a-\frac{1}{2}}(z) \tag{A.1}$$

in which a is either 0 or -1 in our analysis and z is defined in Eq. (2.76).

For clarity the function $U(a,z)$ written in the form of three subfunctions $U(a,0)$, $u_1(a,z)$ and $u_2(a,z)$ as [32]:

$$U(a,z) = U(a,0)[u_1(a,z) + u_2(a,z)]. \tag{A.2}$$

The first subfunction is described as [32]:

$$U(a,0) = \frac{\sqrt{\pi}}{2^{\frac{1}{2}a-\frac{1}{4}}\Gamma(\frac{3}{4}+\frac{1}{4}a)} \tag{A.3}$$

with,

$$\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt \tag{A.4}$$

the Gamma function. It is not necessary to further evaluate the Gamma function, since the values for $a = 0$ or $D_{-\frac{1}{2}}$ is given by [32]:

$$\Gamma(\frac{3}{4}) = 1.2254167024... \tag{A.5}$$

and for $a = -1$ or $D_{\frac{1}{2}}$ by:

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}. \tag{A.6}$$

The other two subfunction are estimated by a power series, like [32]:

$$u_1(a,z) = e^{\frac{1}{4}z^2}[1 + (a - \frac{1}{2})\frac{z^2}{2!} + (a - \frac{1}{2})(a - \frac{5}{2})\frac{z^4}{4!} + ...] \tag{A.7}$$

and

$$u_2(a,z) = e^{\frac{1}{4}z^2}[z + (a - \frac{3}{2})\frac{z^3}{3!} + (a - \frac{3}{2})(a - \frac{7}{2})\frac{z^5}{5!} + ...]. \tag{A.8}$$

Ultimately, only two parabolic cylinder functions are required: $D_{-\frac{1}{2}}(z)$ in which $a = 0$ and $D_{\frac{1}{2}}(z)$ in which $a = -1$. The first function is given by:

$$\begin{aligned} D_{-\frac{1}{2}}(z) &= U(0,z) \\ &= U(0,0)[u_1(0,z) + u_2(0,z)] \\ &= \frac{\sqrt{\pi}}{2^{-\frac{1}{4}}\Gamma(\frac{3}{4})}[u_1(0,z) + u_2(0,z)] \\ &= \frac{\sqrt{\pi}}{2^{-\frac{1}{4}}\Gamma(\frac{3}{4})}e^{\frac{1}{4}z^2}[(1 - \frac{1}{2}\frac{z^2}{2!} + \frac{5}{4}\frac{z^4}{4!}...) + (z - \frac{3}{2}\frac{z^3}{3!} + \frac{21}{4}\frac{z^5}{5!}...)] , \end{aligned} \tag{A.9}$$

and the second one by:

$$\begin{aligned} D_{\frac{1}{2}}(z) &= U(-1,z) \\ &= U(-1,0)[u_1(-1,z) + u_2(-1,z)] \\ &= \frac{\sqrt{\pi}}{2^{-\frac{3}{4}}\Gamma(\frac{1}{2})}[u_1(0,z) + u_2(0,z)] \\ &= \frac{1}{2^{-\frac{3}{4}}}e^{\frac{1}{4}z^2}[(1 - \frac{3}{2}\frac{z^2}{2!} + \frac{21}{4}\frac{z^4}{4!}...) + (z - \frac{5}{2}\frac{z^3}{3!} + \frac{45}{4}\frac{z^5}{5!}...)] . \end{aligned} \tag{A.10}$$

It is possible to expand the term $e^{\frac{1}{4}z^2}$ also as a power series, which yields:

$$
\begin{aligned}
e^{\frac{1}{4}z^2} &= 1 + \frac{1}{4}z^2 + \frac{(\frac{1}{4}z^2)^2}{2!} + \frac{(\frac{1}{4}z^2)^3}{3!} + \frac{(\frac{1}{4}z^2)^4}{4!} + ... \\
&= 1 + \frac{1}{4}z^2 + \frac{1}{16}\frac{z^4}{2!} + \frac{1}{64}\frac{z^6}{3!} + \frac{1}{256}\frac{z^8}{4!}.... .
\end{aligned}
\tag{A.11}
$$

Appendix B

Description waveform simulation algorithm

In this appendix a set of MATLAB programs is described, that simulate SARInM level 1b waveforms and compute the elevations, as theoretically described in chapter 2. The programs are built such that they compute a waveform and elevation for one strip of approximately 300 m along-track, located around a specified ground location.

The first section provides the values of the parameters for Cryosat, the Earth and some physical constants. Section B.2 describes the programs to estimate relative positions of scatterers, the satellite and the ground location. Furthermore it discusses the Doppler effect, antenna gain and eventually shows how a deramped signal is computed. Section B.3 describes the programs, which processes the deramped signals into a SARInM level 1b waveform. The final section discusses the deramping and the computation of the elevation.

B.1 Initialisation

In order to simulate a waveform over a particular location on the the Earth, it is necessary to set a ground location and burst locations or to extract them from actual Cryosat data. These data are binary and need to be extracted with a program downloadable from ESA's Cryosat website: https://earth.esa.int/web/guest/software-tools/content/-/article/software-routines-7114. In this section it is shown how to extract these locations, and how to load them into MATLAB. Secondly, a list of parameters and constants are provided that are used throughout the program.

B.1.1 Input files

First a set of Cryosat level 1b files are selected for the area of interest. Cryosat data are stored in two files, a header file containing some general information which could be used for the selection process and a data file. The header file is in ASCII format, while the data file is binary.

In order to run the simulation, only six parameters for a set of consecutive bursts have to be extracted from the data file: latitude, longitude, the altitude of the satellite and the flight direction. First it is necessary to define latitude and longitude of the ground location. Then 31 to 57 bursts around the ground location are selected, depending on how many waveforms are averaged. Adequate results are already obtained with 31 bursts, while Cryosat in most cases averages 57 bursts.

However, first it is necessary to convert the binary data to extract the data from binary files. The program *cs_tools*, which is downloadable from ESA's Cryosat website is used. This program is written in C language and to extract the appropriate parameters, some input files need to be adjusted, so the appropriate data are extracted. Therefore, first go to the folder:

/PATHcs_tools/cs_tools_v1_1/c_code/examples/.

Then open the file:

./cs_l1b_io_example.c.

To read the latitude, longitude, the altitude and the multilooked waveform the file is adjusted as shown in figure B.1. In the figure it is shown that the for-loop starting at line 304 passes through all the records, the for-loop at line 305 through the number of waveforms in the record. Then in lines 308 to 310 the altitude (j_Alt), longitude (j_Lon) and the latitude (j_Lat) are read. Additionally the velocity vectors can be read from the files. Finally the file is compiled and parameters are piped to an ASCII file as:

./cs_l1b_io_example /PATHcs_tools/cs_tools_v1_1/c_code/examples/cs_l1b_io_example /PATH-file/file >/PATHfile/file.ascii.

The ASCII file is used as input for the simulation; from this file the burst locations are read into MATLAB.

```
299             if( pt_L1b_SARIN != NULL )
300             {
301                     int nor;
302                     int noh;
303                     int xv;
304                     for ( nor = 0; nor < j_records; nor++) {
305                             for ( noh = 0; noh < 20; noh++) {
306                                     for ( xv = 0; xv < 512; xv++ ) {
307                                             printf("%"PRId32", ""%"PRId32", ""%"PRId32", ""%"PRId32"\n",
308                                             pt_L1b_SARIN[ nor ].at_Time_Orbit[ noh ].j_Alt,
309                                             pt_L1b_SARIN[ nor ].at_Time_Orbit[ noh ].j_Lon,
310                                             pt_L1b_SARIN[ nor ].at_Time_Orbit[ noh ].j_Lat,
311                                             pt_L1b_SARIN[ nor ].at_Waveforms[ noh ].aul_wfm[ xv ]);
312                                     }
313                             }
314                     }
```

Figure B.1: Adjustments in the Cryosat l1b read file to read the latitude, longitude, altitude and the multilooked waveform.

B.1.2 List of parameters

Along with burst locations and velocity vectors additional input is required in the form of constants and parameters of the Earth and Cryosat. Table B.1.2 lists the required properties of the WGS84 ellipsoid and table B.1.2 lists the physical constants.

Parameter	Value
Semi-major axis (a_e)	6378127.0 m
Eccentricity (e)	0.081819218

Table B.1: Parameters of the WGS84 ellipsoid (according to [3]).

Parameter	Value
Speed of light (c)	299792458 m/s

Table B.2: Physical constants (according to [3]).

The properties of Cryosat and SIRAL are split up into general parameters and SARInM specific parameters. The general parameters, give in table B.1.2, apply to any mode of Cryosat is operating in. Table B.1.2 lists the SARInM specific parameters.

Parameter	Value
Central frequency (f_c)	$13.575 \cdot 10^9$ Hz
Wavelength (λ_c)	0.022084159 m
Sampled bandwidth (B)	$320 \cdot 10^6$ Hz
Sampled pulse length (τ)	$44.8 \cdot 10^{-6}$ s
Transmitted pulse length (τ_t)	$49 \cdot 10^{-6}$ s
Time resolution (t_{res})	$3.125 \cdot 10^{-9}$ s
Peak antenna gain (G_0)	42 dB
Antenna constant 1 (γ_1)	0.0133
Antenna constant 2 (γ_2)	0.0148

Table B.3: General parameters of Cryosat SIRAL (according to [1]).

Parameter	Value
Number of samples and bins (N_s)	512
Sampling interval (Δt_i)	$0.0875 \cdot 10^{-6}$ s
Number of pulses per burst (N_p)	64
Pulse repetition frequency (f_p)	$18.182 \cdot 10^3$ s
Burst repetition interval (Δt_b)	$46.7 \cdot 10^{-3}$ s

Table B.4: SARInM specific parameters of Cryosat SIRAL (according to [1]).

B.2 L0 processing files

The purpose of this section is to describe an algorithm that simulates a deramped signal from a set of scatterers on the Earth's surface. First the positions of the satellite and scatterers are loaded into MATLAB and their coordinates in the ACU system are computed. Then the arrival time and relative orientations are computed. Finally the Doppler effect and antenna gain are determined for each scatterer and a deramped signal is computed, as described in section 2.2.2.

B.2.1 Satellite and scatterer positions

The first program, *compute_groundlocation.m*, loads the ASCII file described in section B.1.1 and reads the latitude and the longitude of the ground location. An estimated elevation is set (normally this is provided by the retracker, but this is only available in level 1 files), this could be the average elevation of the scene. The computations for slant-range correction use the estimated elevation.

The program *compute_burstlocation.m* reads the longitude, latitude and altitude of the

satellite burst locations. It furthermore computes the position of the bursts in the ECEF reference system, as described in section 3.1, with coordinates (X_b, Y_b, Z_b).

Alongside the burst location, also the positions of the satellite at the 'start-stop' locations of the pulse are required. Therefore the program *compute_satelliteposition.m* loads the velocity vector (V_x, V_y, V_z) and uses it to compute the 'start-stop' locations. Per burst 64 'start-stop' locations are computed. If $(X_p(k), Y_p(k), Z_p(k))$ are the coordinates of 'start-stop' location k in the ECEF system, they are related to the burst location as:

$$\begin{bmatrix} X_p(k) \\ Y_p(k) \\ Z_s(k) \end{bmatrix} = \begin{bmatrix} X_b - \frac{V_x}{f_p} \cdot (33.5 - k) \\ Y_b - \frac{V_y}{f_p} \cdot (33.5 - k) \\ Z_b - \frac{V_z}{f_p} \cdot (33.5 - k) \end{bmatrix}, \tag{B.1}$$

with f_p the PRF.

The scatterer positions are loaded from the global DEM, ASTER_GDEM by the program *compute_scattererlocation.m*. Every pixels has a latitude, longitude and an elevation coordinate. The resolution of the pixels is 1/3600 degree. Since the coordinates are given in the WGS84 format, they are converted in a similar way as the burst locations to the ECEF system. The coordinates in the ECEF system are described as (X_s, Y_s, Z_s).

If a flat surface is modelled, the elevations in the WGS84 format are constant. Another option is to model surface roughness like waves, which is done by giving the surface a Gaussian distributed elevation Z_s, like:

$$Z_s \sim N(h, \sigma_h), \tag{B.2}$$

where σ_h is the surface roughness and h the elevation. Modelling cross-track slope would require to set Z_s as a function of cross-track position Y_s.

B.2.2 Coordinate conversions

The coordinates computed by the programs described above are all reference to the ECEF system. In order to use the equations in chapter 2, they have to be converted into the ACU system. Using Eq. (3.5) the program *compute_coordinatetransformation.m* converts them into the NEU system.

Eventually the coordinates are converted to the ACU system, however it first requires the computation of the flight direction ϕ_V with respect to the North, measured clockwise. The satellite is considered to fly in a straight line during the burst [1] and therefore the flight direction is computed as:

$$\phi_V = atan(\frac{y'_p(64) - y'_p(1)}{x'_p(64) - x'_p(1)}), \tag{B.3}$$

where the coordinates $(x'_p(k), y'_p(k))$ denote the 'start-stop' location k in the NEU system.

Then using Eq. (3.6) the coordinates of the burst, 'start-stop' and scatterer location are converted to the ACU system.

Additionally, the program *compute_coordinatetransformation.m* sets the coordinate z_p constant during the burst, so the satellite flies parallel to the x-axis. This geometry makes the computations for beamsteering more convenient, but induces a small change in Doppler effect and antenna gain.

B.2.3 Arrival time and orientation

A program called *compute_arrivaltime.m* computes the ranges and relative positions between the scatterers and the 'start-stop' locations. Let the coordinates $(x_s(j), y_s(j), z_s(j))$ describe the position of a scatterer j in the ACU system. Then the range $R(k, j)$ between the scatterer and 'start-stop' location k is computed as:

$$R(k,j) = \sqrt{(x_p(k) - x_s(j))^2 + (y_p(k) - y_s(j))^2 + (z_p(k) - z_s(j))^2}. \tag{B.4}$$

The arrival time $t_s(k, j)$ of the signals is related to the range $R(k, j)$ by Eq. (2.1).

The relative positions $(x_{rel}(k, j), y_{rel}(k, j), z_{rel}(k, j))$ between the scatterer and the 'start-stop' location are described by:

$$\begin{bmatrix} x_{rel}(k,j) \\ y_{rel}(k,j) \\ z_{rel}(k,j) \end{bmatrix} = \begin{bmatrix} x_s(j) - x_p(k) \\ y_s(j) - y_p(k) \\ z_s(j) - z_p(k) \end{bmatrix}. \tag{B.5}$$

For computing the Doppler effect it is necessary to find the along-track look angle $\alpha_{along}(k, j)$, which is computed as:

$$\alpha_{along}(k,j) = \frac{x_{rel}(k,j)}{R_{rel}(k,j)}. \tag{B.6}$$

Furthermore, to compute the antenna gain it is required to know the azimuthal $\upsilon(k, j)$ and polar angles $\theta(k, j)$ with respect to the antenna boresight, described in section 2.2.1. The azimuthal angle is given by:

$$\upsilon(k,j) = atan2(y_{rel}(k,j), x_{rel}(k,j)), \tag{B.7}$$

while the polar angle is computed as:

$$\theta(k,j) = asin(\frac{x_{rel}(k,j)^2 + y_{rel}(k,j)^2}{R(k,j)}). \tag{B.8}$$

B.2.4 Doppler effect and antenna gain

Before computing the deramped signal the influence of the Doppler effect $D(k,j)$ is computed, which is done by the program *compute_Doppler.m*. Signals from scatterers at an along-track look angle larger or smaller than zero experience a slight frequency shift, which is estimated with Eq. (2.8.

Then the program *compute_gain.m* computes the antenna gain $G(k,j)$ according to Eq. (2.9), with value for constants G_0, γ_1 and γ_2 given in table B.1.2.

B.2.5 Deramped signal

The final step in the L0 processor is to compute the deramped signal, as it is transmitted to the Earth. It is assumed that the amplitude is only affected by the antenna gain and not by path loss, scattering properties of the surface, etc. The deramped signal is computed as described by Eq. (2.35) and Eq. (2.36). It requires the input of antenna gain, Doppler effect, arrival time and the time t_r at which the deramping chirp is generated. Ultimately, the signal is sampled into 512 complex values at a sampling rate Δt_i described in table B.1.2. If the real part $m_{real,LPF}$ and the imaginary part $m_{imag,LPF}$ of the signal are combined with samples numbers $n = 1 : 512$, the deramped sampled signal m_s is described as:

$$m_s = (m_{real,LPF} + i \cdot m_{imag,LPF}), \tag{B.9}$$

sampled at $t_r + n \cdot \Delta t_i$.

B.3 L1b processing files

The level 1 procedures are not applied here, since the instrument calibration is discarded and the signals are already stored in the correct form. Therefore the l1b processor starts right after computing the deramped signal. The procedures are applied according to the descriptions in section 2.3.2. The goal of this section is to describe the programs to

process a deramped signal to a multilooked waveform.

B.3.1 Beamsteering

Beamsteering is applied by the program *compute_beamsteering.m* to adjust the range of beams that do not illuminate the exact ground location. A detailed explanation on the theory is given in section 2.3.2.

First it is determined which beam is steered, so that the angle $\Delta\alpha_a$ over which the beam must be steered (rock angle) can be computed. The beam number b_n is a function of the x-coordinates $x_p(1)$ and $x_p(64)$ of the first and last 'start-stop' location, the wavelength λ_c, the x-coordinate x_b and z-coordinate z_b of the burst. First the beam angle interval $\delta\alpha$ is computed as:

$$\delta\alpha = \frac{\lambda_c}{2(x_p(64) - x_p(1))}. \tag{B.10}$$

Then the beam number is computed as:

$$b_n = \frac{x_b/z_b}{\delta\alpha} + 1, \tag{B.11}$$

for which the value of λ_c is found in table B.1.2. The addition of 1 is because the first beam is nadir the satellite. With the beam number the rock angle is computed as:

$$\Delta\alpha_a = (b_n - 1) \cdot \delta\alpha - \alpha_g \tag{B.12}$$

in which α_g is the along-track look angle of the ground location with respect to the burst location.

Using the rock angle, the range adjustment $R_{bs}(k)$ for 'start-stop' location k is computed with Eq. 2.45 and the phase multiplier $\Phi_{bs}(k)$ with Eq. 2.47. In order to apply the phase multiplier, the phases of the deramped samples m_s are computed as well as the amplitude. Then the phase of the phase multiplier is added to the phases of the deramped samples. And finally, the updated phases and the amplitude are used to compute the steered sampled signal $m_{s,bs}$.

B.3.2 Along-track Fast Fourier Transform

An along-track FFT is applied by the program *compute_afft.m* to divide the signal into 64 signals corresponding to along-track strips of approximately 300 m. The principle is

explained in sections 2.2.4 and 2.3.2. The signal $m_b(n_b)$ for beam or strip number n_b produced from signals received at burst b is computed with:

$$m_b(n_b) = (1/N_p) \sum_{k=1}^{N_p} m_{s,bs}(k) \cdot e^{i2\pi \frac{(k-1)(n_b-1)}{N_p}}, \tag{B.13}$$

where N_p is the number of pulses per burst and $m_b(n_b)$ contains 512 samples.

The built-in program for MATLAB is iFFT. Care must be taken to select the correct beam. As shown under beamsteering the beam number was increased by 1, because the nadir beam was beam 1 instead of beam 0. The backward beams computed with the iFFT program are not beams -1, -2, -3, but instead the first backward beam is 64, and the second one 63.

B.3.3 Doppler correction

The Doppler correction is computed with the program *compute_Dopplercorrection.m* using Eq. (2.55). Like for beam steering a phase multiplier is computed, which is described by Eq. (2.56). The implementation is analog to that of beamsteering, by computing the phases and amplitudes of the samples in $m_b(n_b)$ and adding the phase of the phase multiplier. The Doppler corrected signal $m_{b,Doppler}(n_b)$ for strip n_b at burst b is than computed using the amplitude and the corrected phase.

B.3.4 Fine slant-range correction

The fine slant range correction is applied by the program *compute_slantrangecorrection.m*, which also computes the number of bins it needs to migrate in the program $range_migration.m$. The phase multiplier for the fine correction is computed with Eq. (2.61), in which z_0 is the z-coordinate of the burst above the ground location and the speed of light c has the value given in table B.1.2. After the fine correction, the slant range corrected signal $m_{b,src}(n_b)$ is FTT'd in range.

The number of bins $n_{b,mig}$ to migrate is computed as:

$$N_{bins,mig} = \frac{t_{src} - mod(t_{src}, t_{res})}{t_{res}}, \tag{B.14}$$

with the time resolution t_{res} as given in table B.1.2 and t_{src} as computed with Eq. (2.59).

B.3.5 Range Fast Fourier Transform

At this part of the processing the signal is ready to be turned into a waveform w_b by the program *compute_rfft.m*. This is performed by applying a range FFT, which is described as:

$$w_b(bin) = (1/N_s) \sum_{n=1}^{N_s} m_{b,src}(n) \cdot e^{i2\pi \frac{(n-1)(bin-1)}{N_s}}, \tag{B.15}$$

in which *bin* is the bin number and the number of samples N_s equal to 512. Typical single waveforms are shown in the section 3.3.

B.3.6 Range migration and stack creation

The program *compute_rangemigration.m* applies the coarse slant range correction. Waveforms are shifted by a number of bins $n_{b,mig}$ to compensate for the extra path length compared to the reference range as discussed in section 2.3.2. One waveform per burst is stored for the ground location by the program *create_stack*.

B.3.7 Multilooking

The final step in the L1b processor is multilooking, which is performed by *compute_multilooking*. Multilooking is described by Eq. (2.63), where the power of the samples from single waveforms is averaged. The multilooked waveform $W\bar{W}_{ml}$ is used by the level 2 processor to retrack and find an elevation.

Appendix C

Determination of potential target lakes

In this appendix the theory behind the computation of potential lake target from the MODIS water mask is discussed. The MODIS water mask has a spacing of 0.002083 degrees in both the longitude and latitude direction. The target lakes have a length in North-South direction of minimally 1.5 km and a width in West-East direction of 3 km. At a higher latitude, the spacing of pixels in longitude direction expressed in kilometers is smaller, since 360 degrees at the Equator corresponds ot 40.000 km circumference of the Earth, but at 30 degrees latitude approximately corresponds to 33000 km circumference. This effect must be accounted for. The spacing in the latitude direction expressed in kilometers is constant.

First the matlab command *regionprops* is used to compute a boundingbox around the lake and the centroid of the lake. Within this boundingbox the maximum width W_p and the maximum length L_p in pixels are found by counting the maximum number of lake pixels at one specific latitude of longitude.

By adapting Eq. 3.1 and 3.4, it is possible to computed the distance between the centroid of the lake and the rotation axis of the Earth, perpendicular to the rotation axis of the Earth. Let the latitude $\theta_{centroid}$ and longitude $\lambda_{centroid}$ be the position of the centroid of the lake, then the circumference C_θ at a particular latitude parallel to the Equator is:

$$C_\theta = 2\pi \cdot N \cos(\theta_{centroid}), \tag{C.1}$$

with

$$N = \frac{a_e}{1 - e^2}, \tag{C.2}$$

and the a_e the semi-major axis of the Earth and e the eccentricity. The circumference C_λ around the pole is computed by:

$$C_\lambda = 2\pi b_e, \tag{C.3}$$

with the semi-minor axis b_e of the Earth in WGS84 system.

For every lake it is possible to compute its maximum width $W_k m$ and maximum length $L_k m$ in kilometers, by relating them to the width and length in pixels. This results in:

$$W_{km} = \frac{W_p s_p}{360} C_\theta, \tag{C.4}$$

and

$$L_{km} = \frac{L_p s_p}{360} C_\lambda. \tag{C.5}$$

In here s_p is the pixel spacing of the MODIS water mask.

Bibliography

[1] D. Wingham, C. Francis, S. Baker, C. Bouzinac, D. Brockley, R. Cullen, P. de Chateau-Thierry, S. Laxon, U. Mallow, C. Mavrocordatos, *et al.*, "Cryosat: A mission to determine the fluctuations in earth's land and marine ice fields," *Advances in Space Research*, vol. 37, no. 4, pp. 841–871, 2006.

[2] R. Cullen, D. Wingham, P. Viau, C. Francis, and C. Mavrocordatos, "Esa's cryosat-2 multi-mode level 0 to level 1b science processors–algorithm design and pre-launch verification with asiras," in *Proceedings of Envisat Symposium*, 2007.

[3] J. R. Wertz, "Space mission analysis and design," 1999.

[4] G. Zhang, H. Xie, S. Kang, D. Yi, and S. F. Ackley, "Monitoring lake level changes on the tibetan plateau using icesat altimetry data (2003–2009)," *Remote Sensing of Environment*, vol. 115, no. 7, pp. 1733–1742, 2011.

[5] C. Lu, G. Yu, and G. Xie, "Tibetan plateau serves as a water tower," in *Geoscience and Remote Sensing Symposium, 2005. IGARSS'05. Proceedings. 2005 IEEE International*, vol. 5, pp. 3120–3123, IEEE, 2005.

[6] V. H. Phan, R. Lindenbergh, and M. Menenti, "Icesat derived elevation changes of tibetan lakes between 2003 and 2009," *International Journal of Applied Earth Observation and Geoinformation*, vol. 17, pp. 12–22, 2012.

[7] W. W. Immerzeel, L. P. van Beek, and M. F. Bierkens, "Climate change will affect the asian water towers," *Science*, vol. 328, no. 5984, pp. 1382–1385, 2010.

[8] S. Bergstrom, "Development and application of a conceptual runoff model for scandinavian catchments," 1976.

[9] X.-Y. Li, H.-Y. Xu, Y.-L. Sun, D.-S. Zhang, and Z.-P. Yang, "Lake-level change and water balance analysis at lake qinghai, west china during recent decades," *Water Resources Management*, vol. 21, no. 9, pp. 1505–1516, 2007.

[10] D. E. Alsdorf, E. Rodríguez, and D. P. Lettenmaier, "Measuring surface water from space," *Reviews of Geophysics*, vol. 45, no. 2, 2007.

[11] C. Birkett, "Radar altimetry: a new concept in monitoring lake level changes," *Eos, Transactions American Geophysical Union*, vol. 75, no. 24, pp. 273–275, 1994.

[12] C. S. Morris and S. K. Gill, "Variation of great lakes water levels derived from geosat altimetry," *Water resources research*, vol. 30, no. 4, pp. 1009–1017, 1994.

[13] G. Zhang, H. Xie, S. Duan, M. Tian, and D. Yi, "Water level variation of lake qinghai from satellite and in situ measurements under climate change," *Journal of Applied Remote Sensing*, vol. 5, no. 1, pp. 053532–053532, 2011.

[14] P. Berry, J. Garlick, J. Freeman, and E. Mathers, "Global inland water monitoring from multi-mission altimetry," *Geophysical Research Letters*, vol. 32, no. 16, 2005.

[15] F. Mercier, A. Cazenave, and C. Maheu, "Interannual lake level fluctuations (1993–1999) in africa from topex/poseidon: connections with ocean–atmosphere interactions over the indian ocean," *Global and Planetary Change*, vol. 32, no. 2, pp. 141–163, 2002.

[16] F. Frappart, S. Calmant, M. Cauhopé, F. Seyler, and A. Cazenave, "Preliminary results of envisat ra-2-derived water levels validation over the amazon basin," *Remote Sensing of Environment*, vol. 100, no. 2, pp. 252–264, 2006.

[17] R. K. Raney, "The delay/doppler radar altimeter," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 36, no. 5, pp. 1578–1588, 1998.

[18] J. R. Jensen, "Angle measurement with a phase monopulse radar altimeter," *Antennas and Propagation, IEEE Transactions on*, vol. 47, no. 4, pp. 715–724, 1999.

[19] M. R. Drinkwater, R. Francis, G. Ratier, and D. J. Wingham, "The european space agency's earth explorer mission cryosat: measuring variability in the cryosphere," *Annals of Glaciology*, vol. 39, no. 1, pp. 313–320, 2004.

[20] D. Wingham, "Cryosat science and mission requirements," *available on-line http://esamultimedia. esa. int/docs/Cryosat/MRD_21Nov*, vol. 99, 1999.

[21] C. Francis *et al.*, "Cryosat mission and data description," tech. rep., 2007.

[22] D. B. Chelton, E. J. Walsh, and J. L. MacArthur, "Pulse compression and sea level tracking in satellite altimetry," *Journal of atmospheric and oceanic technology*, vol. 6, pp. 407–438, 1989.

[23] H. AKLIOUAT, Y. SMARA, and L. BOUCHEMAKH, "Synthetic aperture radar image formaton process: Application to a region of north algeria," in *Proc.'Envisat Symposium*, 2007.

[24] D. Wingham, L. Phalippou, C. Mavrocordatos, and D. Wallis, "The mean echo and echo cross product from a beamforming interferometric altimeter and their application to elevation measurement," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 10, pp. 2305–2323, 2004.

[25] N. Galin, D. Wingham, R. Cullen, M. Fornari, and W. Smith, "Calibration of the cryosat-2 interferometer and measurement of across-track ocean slope," *AGU Fall Meeting Abstracts*, vol. 51, no. 1, pp. 57–72, 2013.

[26] ESA, "Cryosat product handbook," tech. rep., 2013.

[27] R. Cullen and D. Wingham, "Cryosat level 1b processing algorithms and simulation results," in *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*, vol. 3, pp. 1762–1764, IEEE, 2002.

[28] ESA, "Envisat ra2/mwr product handbook," *RA2/MWR products user guide*, 2002.

[29] E. S. Garcia, W. H. F. Smith, and S. D. T, "Retracking cryosat-2, envisat, and jason-1 radar altimetry waveforms for improved gravity field recovery," *Geophysical Journal International*, 2013.

[30] G. Brown, "The average impulse response of a rough surface and its applications," *Antennas and Propagation, IEEE Transactions on*, vol. 25, no. 1, pp. 67–74, 1977.

[31] G. Hayne, "Radar altimeter mean return waveforms from near-normal-incidence ocean surface scattering," *Antennas and Propagation, IEEE Transactions on*, vol. 28, no. 5, pp. 687–692, 1980.

[32] M. E. Abramowitz *et al.*, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, vol. 55. Courier Dover Publications, 1964.

[33] K. Levenberg, "A method for the solution of certain problems in least squares," *Quarterly of applied mathematics*, vol. 2, pp. 164–168, 1944.

[34] J. W. Goodman, *Speckle phenomena in optics: theory and applications*. Roberts and Company Publishers, 2007.

[35] ESA, "Geophysical correction application in level 2 cryosat data products," tech. rep., 2013.

[36] S. Vignudelli, *Coastal altimetry*. Springer, 2011.

[37] P.-Y. L. Traon, "Satellite altimetry," tech. rep., 2007.

[38] ESA, "Cryosat data product status," March 2013.

[39] M. Carroll, J. R. Townshend, C. DiMiceli, P. Noojipady, and R. Sohlberg, "A new global raster water mask at 250 m resolution," *International Journal of Digital Earth*, vol. 2, no. 4, pp. 291–308, 2009.

[40] H. H. Farahani, P. Ditmar, R. Klees, J. T. da Encarnação, X. Liu, Q. Zhao, and J. Guo, "Validation of static gravity field models using grace k-band ranging and goce gradiometry data," *Geophysical Journal International*, 2013.

[41] J.-F. Crétaux, W. Jelinski, S. Calmant, A. Kouraev, V. Vuglinski, M. Bergé-Nguyen, M.-C. Gennero, F. Nino, R. Abarca Del Rio, A. Cazenave, *et al.*, "Sols: A lake

database to monitor in the near real time water level and storage variations from remote sensing data," *Advances in Space Research*, vol. 47, no. 9, pp. 1497–1507, 2011.

[42] Y. Le Roy, M. Deschaux-Beaume, C. Mavrocordatos, M. Aguirre, and F. Heliere, "Sral sar radar altimeter for sentinel-3 mission," in *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, pp. 219–222, IEEE, 2007.