D1.3: Algorithm Theoretical Basis Document (ATBD) for SAR Mode Processing Over Ocean

Project number
313238

Project title
LOTUS— Preparing Land and Ocean Take Up from Sentinel-3

Call (part) identifier
FP7-SPACE-2012-1

Funding scheme
Collaborative project

Deliverable Number D1.3
Title: “Algorithm Theoretical Baseline Document for SAR mode processing over ocean”
Nature: Report
Dissemination level: Public

Status: v1.0
Date: 11th June 2014

DOCUMENT CHANGE LOG				
Rev.	Date	Sections modified	Comments	Changed by
1	11th June 2014	All	First running version	Thomas Moreau, Pierre Thibaut, Lars Stenseng, Alejandro Egido
2	26th August	5	Updated Contribution	Ole Andersen

1. Introduction 5
1.1. Purpose and scope 5
1.2. Document structure 5
2. SAR Mode Altimetry Overview 6
3. Processing for Open Ocean 8
3.1. The Cryosat Processing Prototype (CPP) 8
3.2. The SAR-mode retracking 9
3.3. The Reduced SAR-mode processing 10
3.4. Algorithm description 12
3.4.1. RDSAR-mode algorithms definition 12
3.4.2. Retracking 15
3.4.1. Correcting estimates through adapted LUT 16
3.5. SAR-mode algorithms definition 16
3.5.1. SAR-mode retracking algorithm 17
3.5.1.1. To perform ocean retracking 21
3.5.1. Beam parameters algorithm 23
4. SAR-mode method over Coastal Zones 27
4.1. Scattering Geometry 29
4.2. Specular Point Position Determination 30
4.3. Stack delay/Doppler pairs geo-referencing 31
4.4. Delay/Doppler pairs selection 33
4.5. Waveform Retracking 36
5. SAR-mode method over Polar Ocean 39
5.1. Introduction to the Polar Ocean 39
5.2. Algorithm Description 40
*5.2.1. Overview of algorithm flow ..... **¡Error! Marcador no definido.***
5.2.2. Determine a power benchmark for the retracker 43
5.2.1. Identify leads (Classification) 41
5.3. Constraints and Limitations 46
6. References 47

1. Introduction

1.1. Purpose and scope

This document is the Algorithm Theoretical Basis Document (ATBD) for the methods used to process CryoSat-2 SAR mode data for the three ocean focus areas: Open Ocean, Polar Ocean and Coastal Zone. For each section a scientific background of each algorithm is provided, together with the description of the algorithms developed within the frame of the LOTUS project.

1.2. Document structure

This document is structured into four chapters describing:

- An overview of the SAR Mode Altimetry (section 2)
- the SAR-mode algorithms (section 3) that are applied for retrieving the different geophysical parameters over Open Ocean,
- and the SAR-mode method (section 4) that is applied for improving the observation of Coastal regions.
- the SAR-mode method (section 5) that is applied for retrieving the range from specular open water lead returns in the Polar Ocean,

2. SAR Mode Altimetry Overview

Synthetic Aperture Radar (SAR) altimetry was first described as Delay/Doppler Radar Altimetry in (K. Raney 1998). In this section, the terms Delay/Doppler Radar altimetry and SAR altimetry are used indistinctly. The key innovation of SAR altimetry is the addition of along track processing for increased resolution and multi-look processing. This technique requires echo delay compensation, analogous to range cell migration correction in conventional but unfocused SAR. Due to this innovation, spatial resolution is increased in the along-track dimension and Delay/Doppler mapping is provided. In turn, this allows for accumulation of equivalent looks of a scattering area, leading to speckle reduction and altimetric performance.

As detailed in previous deliverables [REF D1.1] SAR altimetry processes the data such that they could be seen as been acquired from a large synthetic aperture antenna, thus achieving a finer resolution than with conventional altimetry. This improved resolution, is achieved thanks to the relative movement of the platform with respect to the surface under observation. Differently to conventional altimeters, SAR altimeters use most of the power received, in fact in conventional altimetry the power contribution of the adjacent scattering points to the one of interest is lost.

Delay/Doppler radar altimeters benefit from the conventional pulse compression in the range dimension. Additionally, SAR altimetry introduces along-track processing for azimuth mapping. For that, an along-track FFT is applied to the data in order to obtain the Doppler spectral components of each echo. The main requirement for the use of the SAR processing in altimetry is the coherency within each burst of pulses (K. Raney 1998) which differs from conventional altimetry.

The Doppler frequency is related to the position of the scatter by the relative movement of the satellite with respect to the points on the surface. This has a geometric interpretation: the Doppler frequency is given by the dot product of the pointing or observation unit vector (depending on observation angle) and the spacecraft velocity vector. Therefore, there is a unique correspondence between the observed Doppler frequency f_D and the observation angle θ_i of the scatterer.

The coherent processing of the SAR altimeter echoes altimeter results in two independent dimensions: along-track (Doppler) and across-track (range). After SAR processing, these two variables describe an orthonormal data grid as shown in Figure 2.2. The actual nadir point is located in the Doppler bin correspondent to zero Hertz and ideally, it is equivalent to along-track position of the satellite. This coherent processing enables the improvement of the instrument final resolution. In contrast to conventional altimeters, where the footprint is a circle around the nadir point, in SAR altimetry the altimeter footprint corresponds to an elongated strip on the surface, with an across-track resolution of 4-5 km and an along-track resolution around 300 m.

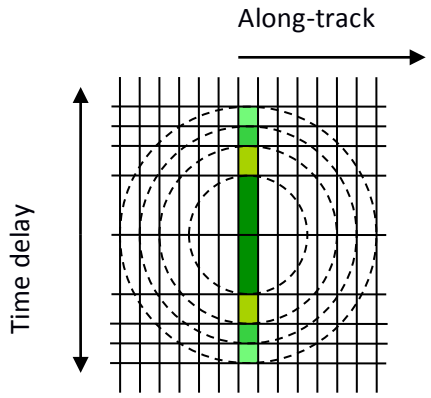


Figure 2.2: delay/Doppler altimeter footprint

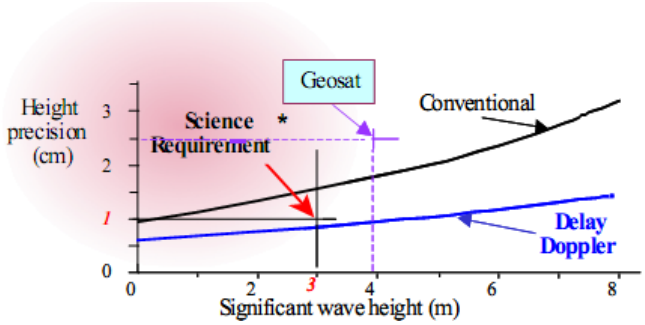


Figure 2.3: Sea Surface Height precision as a function of the Significant Wave Height, for a conventional and delay-Doppler altimeter [Raney, 2005.]

In conventional altimetry within each burst there will be several looks (one per pulse within the burst). SAR altimetry also maintains the previous looks, and due to its along-track geometry this technique will add additional looks with respect to a conventional altimeter, since the scatterer will be visible in different subsequent bursts. Using the Doppler shift the along-track looks will be detected, accumulated and averaged to reduce speckle noise effects. Compared to conventional altimetry the improvement in SNR as well as to the new waveform model (Phalippou and Enjolras 2007), the minimum attainable range accuracy is higher.

Figure 2.3 shows a plot of height precision versus SWH for a delay-Doppler altimeter and a conventional radar altimeter. The plot shows that a DDA is significantly better in precision than a conventional radar altimeter (RA) [Raney, 2005.]. The figure also shows that height measurements from a delay-Doppler altimeter are degraded less than those from a conventional radar altimeter in response to increasing SWH.

This technique is therefore expected to improve the estimation of geophysical parameter in both open ocean, coastal zones, and polar ocean. The following sections describe the algorithms developed within the LOTUS project regarding these main themes.

3. Processing for Open Ocean

3.1.The Cryosat Processing Prototype (CPP)

The processing chain that is implemented in the LOTUS project is inherited from the CNES CPP processing chain [Boy et al., 2012], for SAR and LRM mode data). The initial aim of the CPP was to contribute to expertise studies for the Cryosat-2 mission and for the future Sentinel-3 mission (and Jason-CS mission). In the frame of LOTUS, evolutions have been implemented in the CPP. They are described hereafter.

The CPP chain is shown in Figure 3.1. Two main functions are implemented. The first one generates L2 data from the Full Bit Rate (FBR) data sent by ESA. The second one generates LRM L2 data. This second function will not be used in the LOTUS project which only considers SAR mode processing.

For SAR mode data, SAR and reduced SAR (RDSAR) echoes, also called “Pseudo LRM”, are generated simultaneously. The SAR processor is based on the use of the aperture synthesis technique while a “conventional” processing on high PRF echoes (equivalent to on-board LRM processing) is applied to derive RDSAR-mode echoes. Then, the CPP retracks the echoes (L2 processing) using a dedicated numerical model for SAR data and using the classical Brown model for RDSAR data. This document will focus on particular processing : the SAR doppler retracking (Level 2) and the full RDSAR processing from raw data to Level 2.

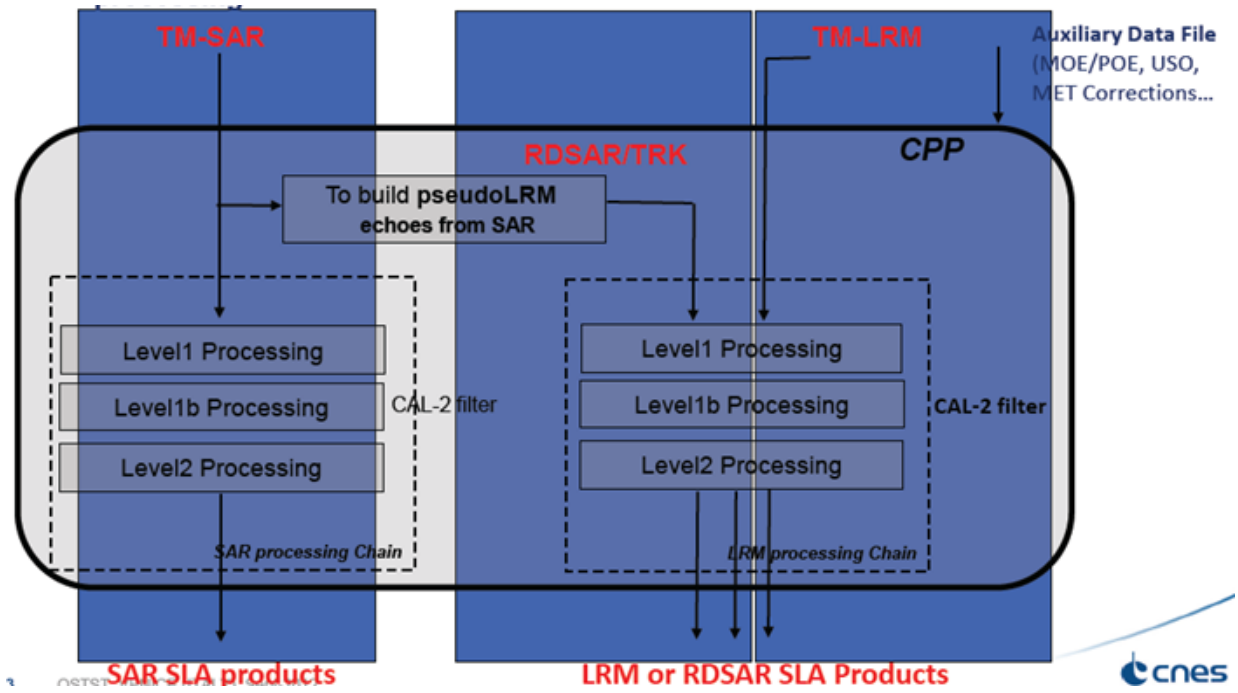


Figure 3.1: General flowchart of the CPP (from Boy et al., 2012).

3.2.The SAR-mode retracking

The CPP SAR-mode retracking algorithm is a standard least squares estimator (LSE) consisting in fitting a SAR-mode waveform, also called multi-looked echo, with a non-analytical echo model that is pre-computed off-line by an amplitude simulator. During the iterative process, this method computes numerically the derivatives of the mean return power by a finite difference involving the echo model database in which the sensitive parameters (sea-state, satellite parameters) vary, one parameter at a time, in a range of values and with a step size that have been chosen to ensure the accuracy and precision of the estimates. This approach is considered to be more robust than analytical ones, particularly when faced with atypical observations (e.g., elliptical antenna pattern, off-nadir mispointing angles, point target response) that are difficult to put into equations.

The SAR echo models are generated by an amplitude numerical simulator [Desjonquères et al., 2012] that mimics the Cryosat-2 altimeter response in SAR mode (taking into account the real elliptical antenna pattern and a real point target response). The simulator is fully numerical. It is based on a point-by-point radar response simulation on a gridded surface without limitation of resolution (fully adaptive). The satellite altitude and altimeter characteristics can be modified depending of simulated missions (Jason-2/3, AltiKa, Sentinel-3, CryoSat, Jason-CS…). A theoretical or measured antenna pattern can be used taking into account mispointing angle in both axes. Theoretical or measured Impulse Responses can be used as well. In addition, the surface height can be modified using a Digital Elevation Model and atmospheric attenuation or ground “reflection anomalies" can be introduced for specifics investigations. The Figure 3.2 illustrates the different steps of the SAR echo simulator. Major processes of the SAR simulator are in sequence:

1- The power return signals from each point of the gridded surface are computed then sorted by Doppler band and accumulated in the appropriate range gates of the waveforms.
2- The flat sea surface response is then convolved with the azimuth and range impulse response (AIR and RIR) of the radar (approximated by a sinc^2 function).
3- Prior averaging, the Doppler bands are corrected in range to compensate the slant range migration, i.e. to place all observations of a scatterer at the same radial distance from the satellite when the satellite moves along its orbit.
4- The Doppler beam waveforms (looks) from the same surface are summed (multilooking) to finally form the SAR echo model for a flat sea surface (the sea wave height is applied “on the fly” in the retracking process to provide an exact solution for the SAR waveform).
5- To finish, the simulator convolves the previous result with the PDF of the significant wave height.

Note that the simulator can also generate LRM numerical model. In that case, the step 3 is not performed (no range migration). This simulation option has been very useful since it was possible to cross compare and validate the simulator with the Brown model.

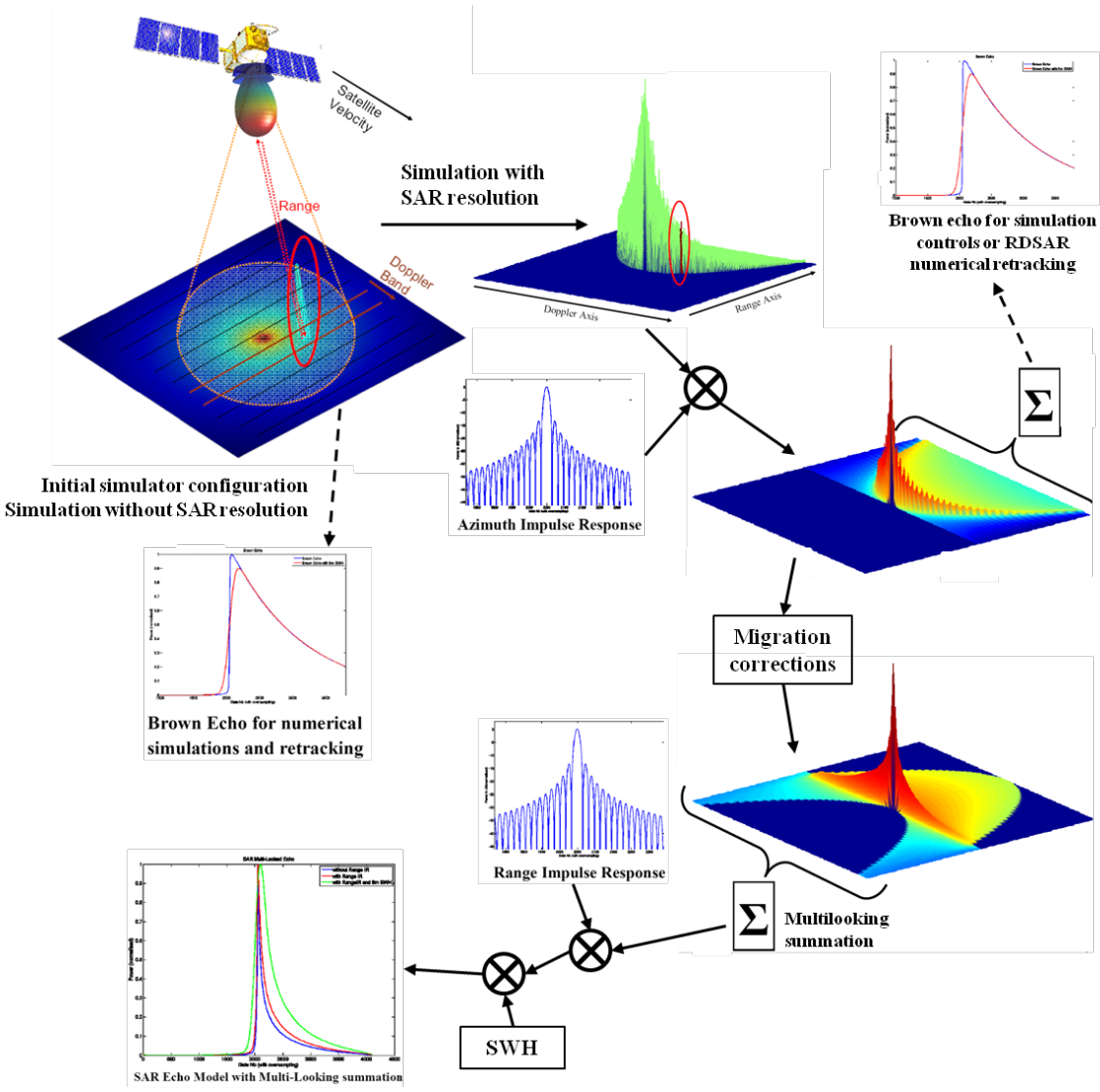


Figure 3.2: SAR echoes simulator (from Desjonquères et al., 2012).

3.3.The Reduced SAR-mode processing

Though the along-track improvement in sampling resolution is straightforward, it still remains some uncertainties in the SAR retrieved elevations as well as the other surface parameters accuracies. To allow the assessment of the in-orbit performances of the SAR mode data and in the same manner the quality of the processing method, a reduced SAR (RDSAR) methodology has been developed [Boy et al., 2012] that aims at emulating LRM echoes, similar to the conventional pulse limited waveforms, from SAR mode data to make quantitative comparisons of their measurements over identical sea state. SAR and RDSAR-modes CPP echoes are generated at the same along-track location allowing both retrievals to be directly subtracted without the need to apply any geophysical model (e.g., wet and dry troposphere correction, ionosphere correction, tidal correction, dynamic atmospheric correction) or orbit elements (like the orbital ephemerides to derive a precise altitude or also to align CryoSat-2 with other missions) that may contribute to differences and lead to unclear conclusions regarding the

comparison between the different processing approaches. This is especially true for the sea level anomaly (SLA) and other altimeter derived products like the wind measurements that account for corrections and/or models.

Figure 3.3 illustrates the differences in footprint geometry between the conventional pulse limited altimeter and the SAR altimeter: the footprint of a conventional altimeter is pulse-limited whereas the SAR altimeter uses the Doppler principle to achieve a small along-track footprint. The resulting RDSAR and SAR-modes power thus have an important distinction. Figure 3.3 shows the differences in the received echoes from a conventional altimeter (here a RDSAR-mode waveform) and the SAR-mode altimeter (waveform with a steep leading edge and fast decay trailing edge).

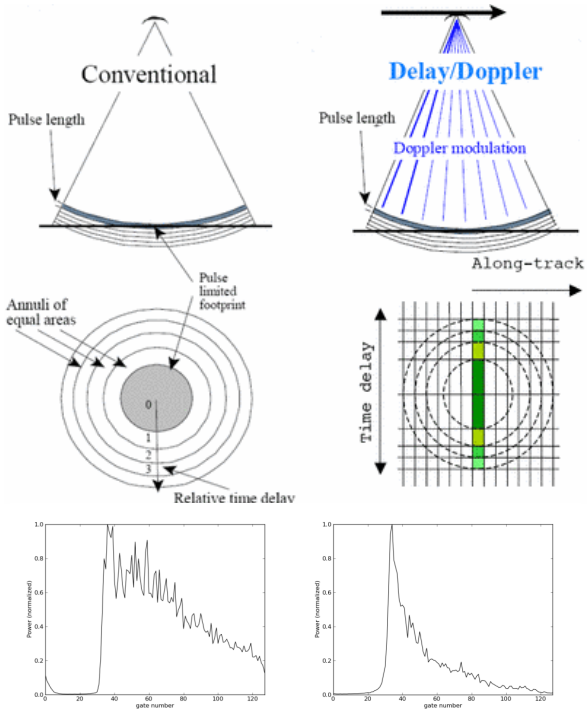


Figure 3.3: Comparison of the footprint geometry between (left panels) LRM conventional radar altimeter and (right panels) Delay/Doppler SAR altimetry (Credits R.K. Raney) and the associated waveforms (lower panel).

Another key objective of the RDSAR methodology was to build and maintain the data quality continuity between SAR and LRM modes. It is hence important to ensure that the RDSAR data are fully consistent and valuable to be considered as a well-known LRM reference during SAR mode.

3.4. Algorithm description

In this section, SAR and RDSAR–mode processing for open ocean are described. These algorithms are based on the existing CPP methods ([Boy et Moreau, 2013], [Boy et Moreau, 2013]) for which adaptations are specified and developed to improve the quality of the SAR and RDSAR-mode retrievals for open ocean. These algorithms meet adequately the key scientific requirements as specified in [LOTUS D1.2, 2013] and follow the overall strategy and technical approach that were recommended.

3.4.1. RDSAR-mode algorithms definition

The RDSAR mode echoes are obtained from the same individual returns that are used in SAR-mode (provided in FBR) but a different processing is used to generate the waveforms. The description of the RDSAR-mode algorithm in the subsequent sections is facilitated by the flow chart shown in Figure 3.8, which outlines the key parts in the way the algorithm works. In the following sections, a definition of the different block algorithms is given. They are listed in an order that follows the RDSAR mode processing scheme.

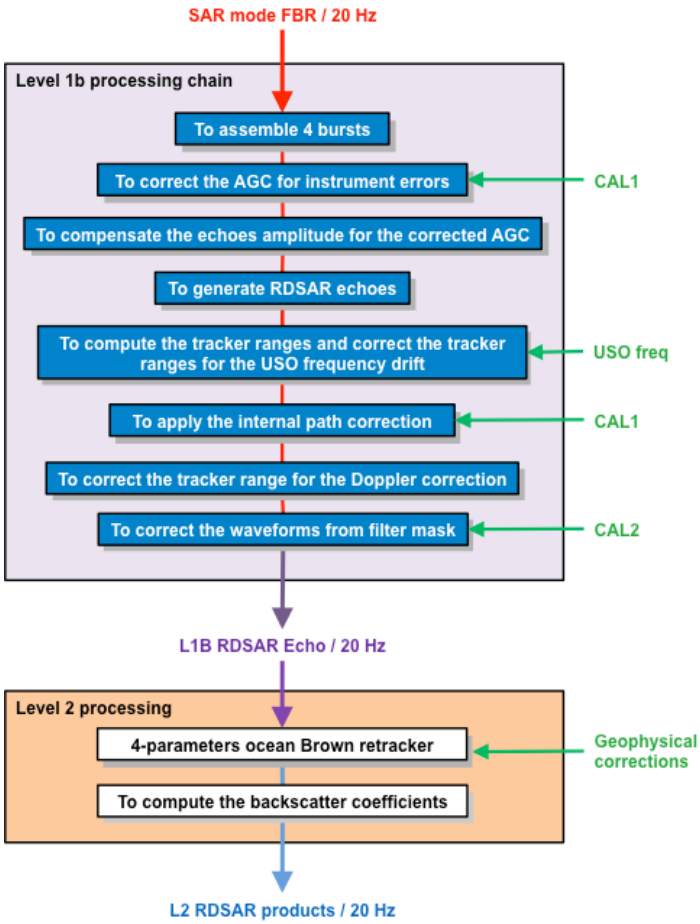


Figure 3.4: Flowchart showing the sequence of operations performed by the RDSAR processing scheme

To assemble 4 bursts

The FBR data form the input to the full RDSAR-mode processing. It is the highest-level data in which the full information content of the acquired data is retained. At this stage, the individual echoes are complex-valued, I and Q, data at high pulse repetition frequency, that are not calibrated, but deramped in time domain, and tagged with time and geolocation.

For each 20 Hz tracking cycle, 4 bursts of 64 echoes are processed through the L1b RDSAR-mode processing chain simultaneously as shown in the figure below.

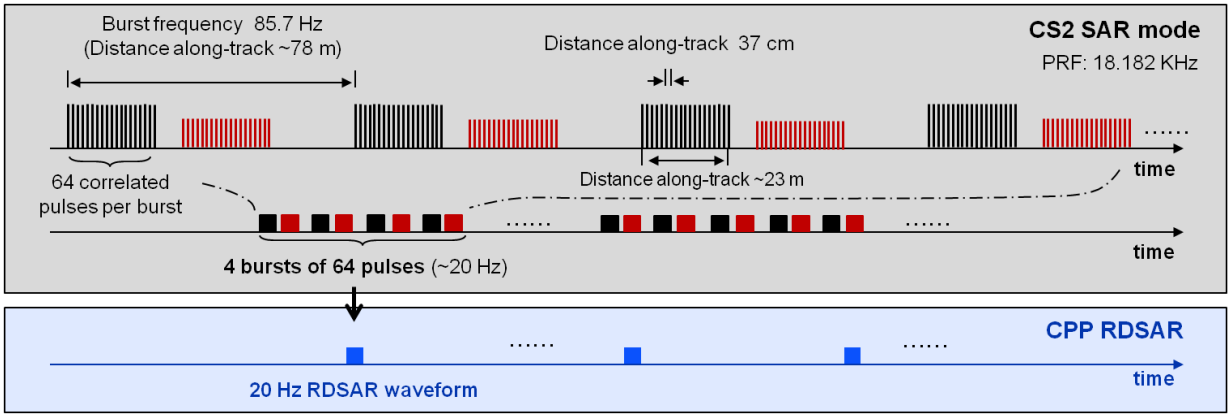


Figure 3.5: Timing of the SIRAL transmitted pulses and received echoes in SAR-mode and the processed RDSAR waveforms.

The obtained 20 Hz RDSAR data are assigned a time tag that corresponds to the time at the center of the 4 bursts. In turn, this time is corrected for a one-way satellite-to-surface travel of a radar wave in order to get the time corresponding to when the pulse hits the Earth surface. The 20 Hz point positions on the Earth surface are defined by the sub-satellite point location (the nadir point) at this time. Note that an equivalent on board timing will be implemented for SRAL for Sentinel-3.

To correct the AGC for instrument errors

An internal calibration monitors the instrument Range Impulse Response (RIR), like traditional altimeters, to identify any gain variation (during the life of the satellite) in the Transmitter/Receiver chain with respect to the nominal condition characterized on ground, which may be caused by some small instrument power changes during flight. This amplitude correction is related to the power level drift of this RIR. It is extracted from CAL1 measurements, and then added to the on-board AGC measurements (AGC1 and AGC2) to obtain the corrected AGC value *AGC_total*.

To compensate the echoes amplitude for the corrected AGC

This function applies the corrected AGC value to the complex raw echoes (*Wf_I_raw +jWF_Q_raw*), as follows:

$$Wf\_I(n) = Wf\_I\_raw(n)/\sqrt{10^{\frac{AGC_{total}}{10}}}$$

$$Wf\_Q(n) = Wf\_Q\_raw(n)/\sqrt{10^{\frac{AGC_{total}}{10}}}$$

n being the sample number of the echoes in the time domain (from 0 to 127)

The amplitude of each individual complex echo is scaled accordingly. This function is inherited from the SAR processing chain, and has been kept for convenience in the RDSAR chain too (even if the AGC setting is held constant during a tracking cycle of 4 bursts).

To generate RDSAR echoes

The purpose of this stage is to compute the LRM-like power waveforms from SAR mode pulses.

In a tracking cycle of 4 bursts, the 4*64 complex echo profiles may vary of position inside the range window due to the varying distance between the satellite and the Earth surface around its orbit.
Consequently, those individual echoes need to be first all aligned on the same reference gate before the averaging step is performed.
On-board, the altimeter performs a partial alignment of the echoes (in order to build the tracking echo). The (I,Q) echoes, as available in telemetry or FBR data, are corrected for the Burst to Burst shift but the shift inside the Burst is not managed. Moreover, the range shift is calculated from an on-board pre-computed radial velocity (COR2), which is very noisy.
So, the proposed method consists of:
1 – Undoing the partial alignment performed by the altimeter (using COR2 command)
2 - Applying a finer shift to the I and Q echo profiles, to make them aligned to a reference position corresponding to the first echo of the cycle, using the precise radial velocity provided by the orbit ephemeris (MOE or POE).
For each individual complex echo, the phase shift corresponding to the distance shift can then be written as:

$$\phi(n,t) = \exp(-2j\pi * \delta f * t)$$

where δf is the needed frequency translation, *n* the pulse number and *t* the time or abscissa of the time domain echoes. We apply the phase compensation term to each I and Q echoes converted in time domain by a Fast Fourier Transform (FFT).

For each of the 4*64 echoes (I and Q) corrected for distance, a 128-points FFT is performed and the I2+Q2 spectrum is derived. The 20 Hz Pseudo-LRM 128-points I2+Q2 spectrum is then computed by summing the 4*64 I2+Q2 waveforms.

To compute the tracker range and to correct the tracker ranges for the USO frequency drift

The tracker range is computed from the tracker range command of the corresponding cycle.
Then, as the RDSAR-mode echo has been built using the first pulse of the cycle as a reference and as the measurement is time-tagged at the center of the cycle, the tracker range is corrected for a distance corresponding to the radial satellite movement during half cycle duration.
To finish, the Ultra-Stable Oscillator (USO) clock period is needed to accurately convert the range delay counter values into time units (and then spatial units) in order to derive the altimeter range. Some

USO clock period variations (USO drift) may occur due to the ageing of the device and cause some error in range.
The correction is regularly updated in the ground processing via an auxiliary file. It has to be added to the altimetric retracked range.

To apply the internal path correction

The monitoring of the whole RIR shape is also performed to correct any internal range drift due to the offset of the RIR position with respect to the 0-frequency, accounting for instrumental evolutions.
This internal path correction extracted from CAL1 is to be added to the altimetric retracked range on ground.
This function is traditionally used in conventional altimetry. But the CPP RDSAR processing does not apply the internal path delay because this correction is a constant bias, identical for each mode (SAR, LRM and RDSAR). So, we can correctly cross compare each mode even if this correction is not applied.

To correct the tracker range for the Doppler correction

The tracker range is corrected for the Doppler correction, caused by the Doppler shift in the line of sight to the nadir direction.

To correct the waveforms from low-pass filter mask

A specific in-flight calibration mode (CAL2) allows the computation of the instrument filter mask needed to correct the echo waveform from the internal altimeter characteristics . It is obtained by accumulation of the noise spectra within the bandwidth of interest, which is then normalized and inverted to form the filter mask shape.
The *Filter* correction extracted from CAL2 (filter mask file) is to be applied to the averaged altimeter echo samples *Wf_Ampl_Raw* on ground. However, a unique filter is employed within the current CPP processing chain for this purpose considering that its evolution with respect to time is negligible.
The corrected waveform *Wf_Ampl* may be retrieved by:

$$Wf\_Ampl(n) = \frac{Wf\_Ampl\_Raw(n)}{Filter(n)}$$

n being the sample number of the echoes in the time domain (from 0 to 127)

3.4.2. Retracking

A conventional Brown ocean retracker based on unweighted least square estimations (also known as MLE) which are traditionally used with LRM echoes [Amarouche et al., 2004], is applied to the RDSAR-mode power waveforms for retrieving the different geophysical parameters (range, significant wave height, backscattering coefficient and off-nadir mispointing angle). The MLE-4 algorithm consists in estimating the 4-parameter set (range, significant waveheight, power and slope of the trailing edge) that optimizes the fit of the measured waveforms with the return power model. Its performances have been largely described in the literature [Thibaut et al., 2010].

3.4.1. Correcting estimates through adapted LUT

The 4 parameters are then corrected to account for the Gaussian approximation of the Point Target Response (PTR) in the Brown ocean retracker, the ellipticity of the CryoSat-2 antenna, and its particular speckle reduction property (different from conventional altimetry mode), through pre-computed Look-Up correction Tables (LUT) as follows:

$$Range = Range_{estimated} + LUT_{range}$$

$$SWH = SWH_{estimated} + LUT_{SWH}$$

The existing RDSAR CPP products are corrected applying a Lookup Table inherited from Jason-2, leading to possible bias in range and SWHs with the true values. This correction LUT has been updated for LOTUS to take into account not only the approximation of the PTR in the retracking algorithm but also the speckle reduction property of the RDSAR method. This RDSAR dedicated correction depends on the wave height and could be as high as 3 cm in range and 20 cm in wave at 4m-wave height as shown in Figure 3.6. Significant differences are observed with the Jason-2 correction LUT.

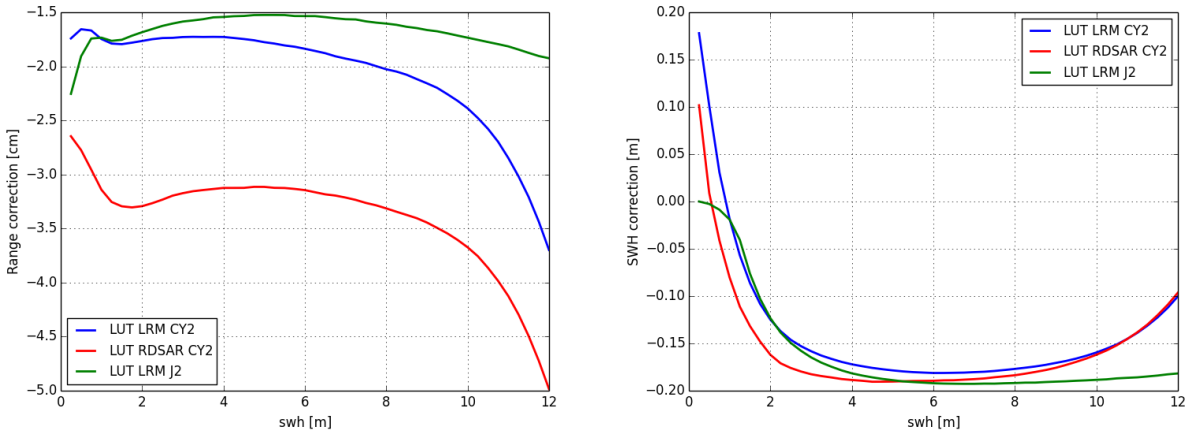


Figure 3.6: Range and SWH look-up table used for LRM-mode and RDSAR CPP products. Corrections are in cm for range and in m for wave height. The LRM Jason-2 look-up table is shown for comparison purposes only.

3.5. SAR-mode algorithms definition

In SAR mode, starting from the square-law detected Doppler beams (looks) of the stack (as shown in Figure 3.7) two types of signal are exploited:

- The multi-looked SAR power echo, which is simply obtained by an incoherent summation in azimuth direction of the beam waveforms of the stack. The numerical retracking algorithm that is currently operating in the CPP chain fits the SAR echo with pre-computed echo models to retrieve the range, SWH and sigma-naught.

- The distribution of power across the beams after integrating the power in range. It is expected that this distribution of power reflects the along-track antenna gain pattern, i.e. Gaussian. Any distortion of the Gaussian shape may not be seen in the multilooked echo, but would give amount of relevant information to be analyzed, in particular the along-track mispointing angle and the ocean surface (ocean surface roughness, slope of the surface, high reflectivity surface). By fitting the distribution with a Gaussian-like model, the estimated parameters will inform about the beam behaviour. The characterization of the ocean surface (e.g. stack width, kurtosis, skewness) from stack data is one of the priorities raised by the expert team of the mid-term CP4O project [Cotton, 2014]. It has been implemented in the LOTUS processing chain.

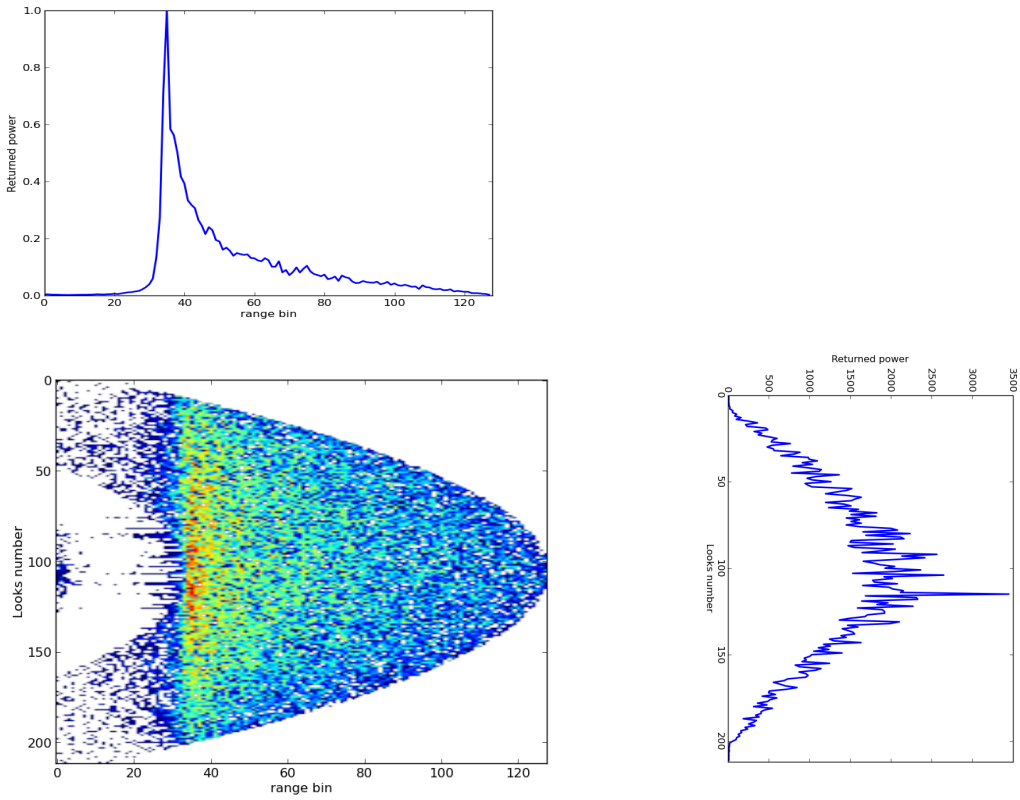


Figure 3.7: Stack of co-located Doppler beams over open ocean. Upper panel: Multilooked SAR power echo. Right panel: Distribution of power among looks in the stack.

3.5.1. SAR-mode retracking algorithm

The CPP retracking algorithm is a standard least squares estimator (LSE) consisting in fitting a SAR waveform with an echo model, also called multi-looked echo, that is pre-computed off-line by the simulator. As for conventional altimetry, the ocean parameters estimated from the numerical SAR retracking are obtained using a Newton-Raphson algorithm that is an iterative method as:

$$\theta_{n+1} = \theta_n - g(BB^T)^{-1}_{\theta_n}(BD)_{\theta_n}$$

where θ_{n+1} is the estimated parameter at iteration *n+1*; *B,D* are the partial derivatives and residuals matrix, and *g* is the loop gain (between 0 and 1).

As for unsolved analytical model, derivatives of the mean return power are computed numerically. The method consists in approximating the derivatives by a finite difference involving the echo model database in which the sensitive parameters (sea-state, satellite parameters) vary, one parameter at a time, in a range of values and with a step size that have been chosen to ensure the accuracy and precision of the estimates. At each iteration *n*, models using the current estimation vector θ_n are directly taken from the database. The performances of this method have been evaluated theoretically using simulated LRM waveforms and, have been statistically validated on real data by applying it on J2 raw measurements. The results are found to be consistent with those obtained from a classical MLE4 retracking [Thibaut et al, 2012].

The numerical SAR retracking is based currently on a 3-parameters model (range, significant wave height, amplitude) that accounts for varying off-nadir mispointing angles provided by the star tracker information (raw telemetry) T_x and T_y (roll and pitch angles). However this information is not directly usable and has to be first aligned on the altimeter electromagnetic axis using the mispointing angle ξ^2 estimated from the retracking of LRM data. To find the roll and pitch biases (*a,b*) we apply the least squares method that consists of minimizing the following squared residuals:

$$\xi^2 - [(Tx + a)^2 + (Ty + b)^2] = 0$$

This fitting method returns the duplet values a and b that correspond to the minimum mean squared error estimator. Then, the biases are added to the star tracker information to get the mispointing angle of the antenna in both axes (in along-track and cross-track directions) to be injected as input to the retracker.

The description of the SAR-mode algorithm in the subsequent sections is facilitated by the flow chart shown in Figure 3.8, which outlines the key parts in the way the algorithm works. In the following sections, a definition of the different block algorithms is given. They are listed in an order that follows the SAR-mode retracking scheme.

To load the SAR echo models

Prior processing the SAR echo waveforms, a set of 63 synthetized Doppler beams model is loaded into memory to be used as input to the retracker (see Figure 3.9). The selected beams correspond to a simulated model using configuration parameters that match the orbital and instrumental information for this series of data (notably the star tracker information). These sharpened beams (also called looks) are equally spaced in angle (through the plane defined by the satellite velocity and the Doppler axis). They are sampled in a range window of 256 bins that are oversampled by a factor of 64. Impulse responses in range and azimuth are used for generating the models.

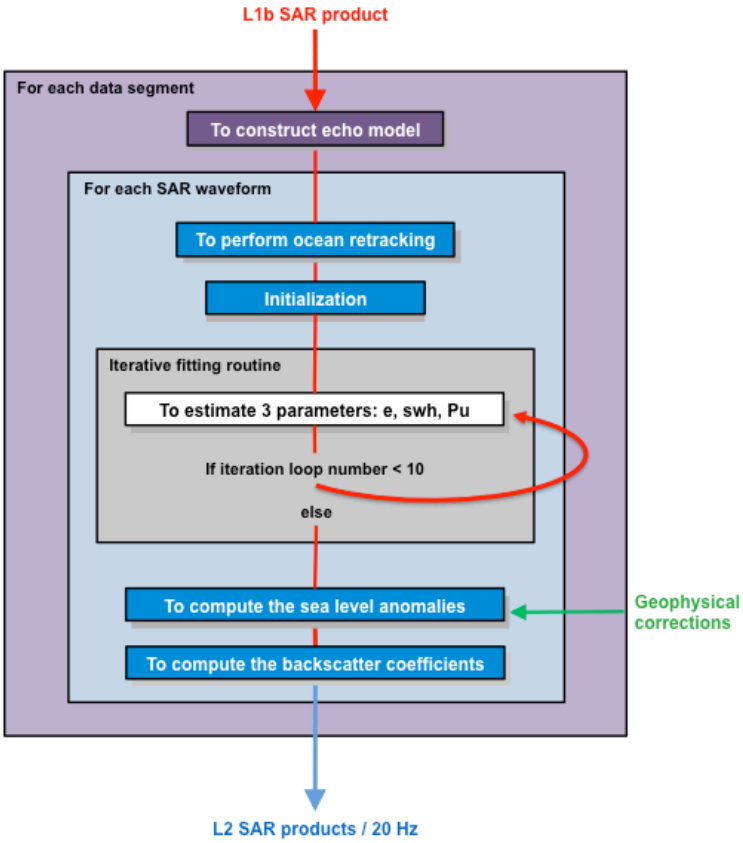


Figure 3.8: Flowchart showing the sequence of operations performed by the SAR numerical ocean retracker

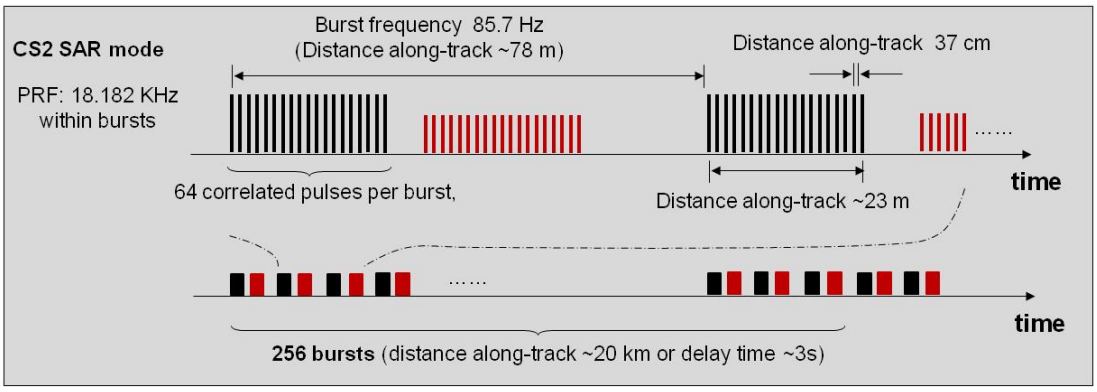


Figure 3.9: SIRAL transmission and reception timing in SAR-mode.

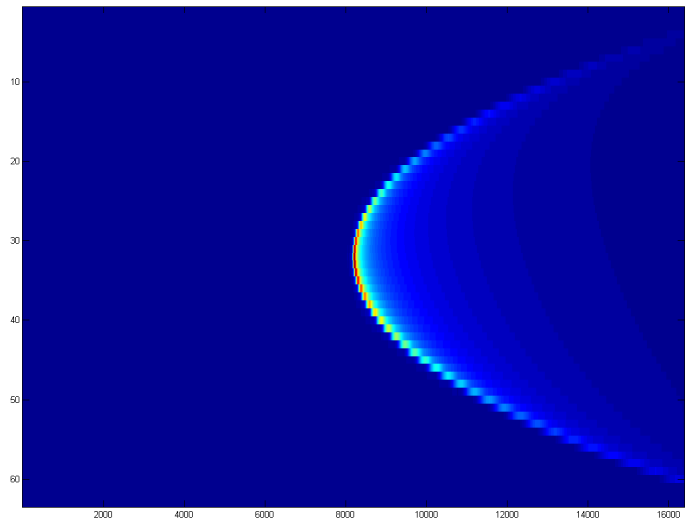


Figure 3.10: The simulated Doppler beams with 63x16384 samples. The apex of the hyperbola is placed at the centre of the range window.

To construct the multilooked waveform model

For each data segment (up to 6mn of maximum length where we may consider that the along track and across track mispointing are constant), there are four main steps in the process, following the sequence below:

- First, the centre of the echoes, E_{centre}, is evaluated based on the median of the estimated epoch from the reduced SAR echo.
- Secondly, the Doppler beams are truncated for those samples that are considered out of the altimeter range window, with respect to the E_{centre} value (see Figure 3.11).

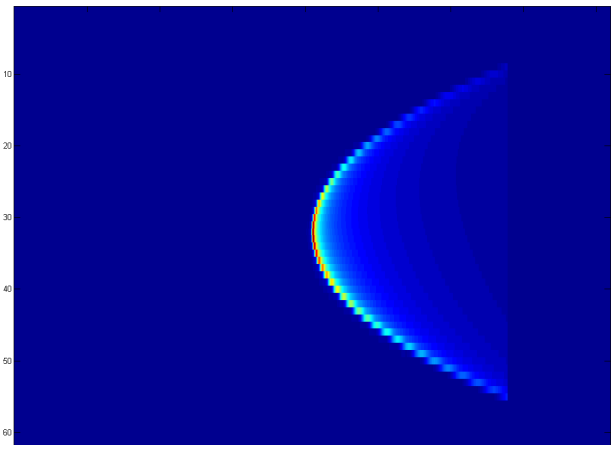


Figure 3.11: The Doppler beams have been shortened in range.

- The next stage consists of correcting in range the Doppler beams to align all looks viewing the same surface scatterer. Shifts in range are performed by a convolution operation with a Dirac delta function.

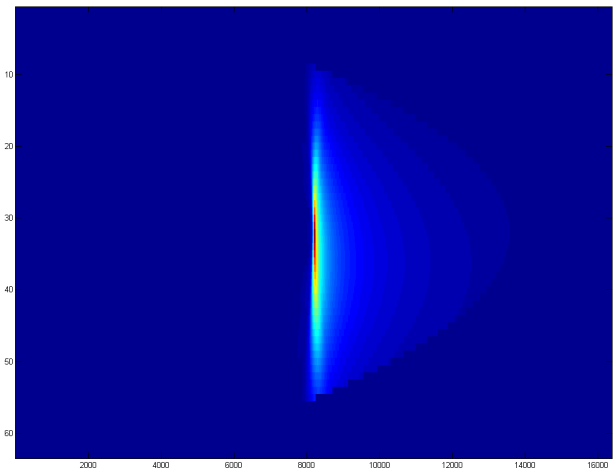


Figure 3.12: The Doppler beams have been aligned in range.

- Finally all squared Doppler beam waveforms of the stack are summed. The constructed multilooked echo is the numerical model that will be used to process the data segment.

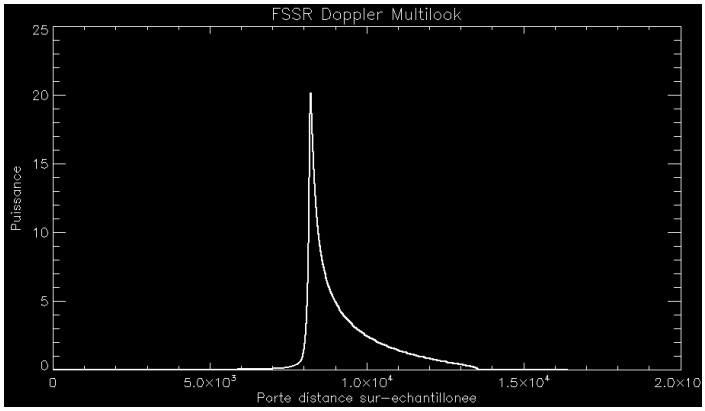


Figure 3.13: The multilooked echo model.

3.5.1.1. To perform ocean retracking

This function fits the numerical multilooked waveform model to one 20 Hz Level 1B SAR waveform using the LSE method (as for LRM) to retrieve geophysical variables and fitting quality information.

For each SAR echo waveform, a 3-parameter ocean retracker based on unweighted least square estimations (also known as MLE) that are traditionally used with LRM echoes (see RDSAR retracking section 3.4.2), is applied for retrieving the different geophysical parameters (range, significant wave height, backscattering coefficient). This method exploits the multilooked echo model as computed previously, for which different sea state parameters *E*, *SWH*, *Pu* are considered on each iteration step of the retracking process.

To compute the echo model for a given sea state

The echo model with the parameters *E*, *SWH*, *Pu* is calculated as follows:

1- the echo model is convolved with a Dirac delta function at the epoch *E*,
2- the shifted echo model is then convolved with a Gaussian sea surface height distribution (defined by a standard deviation equal to ¼ of the significant wave height),
3- *Pu* gain is applied,
4- The echo model is sub-sampled by factor 128 (keeping every 128-th sample) to provide a waveform with a resolution of 128 range gates.

Basic principle

The problem to solve is the estimation of a set of N_θ=3 parameters $\theta = \{\theta_1 = e, \theta_2 = swh, \theta_3 = Pu\}$. The system to solve results from the maximization of the logarithm of the likelihood function $\Lambda(\theta)$, i.e. from the system:

$$C(\theta) = \nabla[-Ln(\Lambda)] = 0$$

where *C* is the total cost function and ∇ is the gradient function.

This system is reduced to weighted Least Square Estimators, and is equivalent to set the Least Square function $\nabla\chi^2$ to 0, where the merit function χ^2 is defined by:

$$\chi^2 = \sum_i \left(\frac{V_i - Vm_i}{\sigma_i}\right)^2$$

where *V* represents the measured waveform, *Vm* the echo model, and where the weighting function is $\{\sigma_i\} = \{Vm_i\}$.

This system may also be represented by the following set of N_σ equations:

$$\sum_i \left(\frac{V_i - Vm_i}{{\sigma_i}^2}\right)^2 . \frac{\partial Vm_i}{\partial \sigma_k} = 0$$

An iterative solution is obtained by developing the total cost function in a Taylor series at the first order about an initial set $\theta_0 = \{\theta_{01} = e_0, \theta_{02} = swh_0, \theta_{03} = Pu_0\}$ of estimates:

$$\theta_{n+1} = \theta_n - g.\varepsilon_\theta$$

with: $\varepsilon_\theta = (BB^T)^{-1}BD$ (valued to the current values θ_n)

B,D are the partial derivatives and residuals matrix:

$$B_{ki} = \frac{1}{\sigma_i}.\frac{\partial Vm_i}{\partial \sigma_k},\ D_{i1} = \frac{1}{\sigma_i}.(Vm_i - V_i),$$

and where *g* is a loop gain (positive value, unique to the parameter being estimated).

Using $\{\sigma_i\} = \{Vm_i\}$, the Least Square Estimators method described above would put the most weight on the regions with the least power, i.e. on the regions with the least information regarding the parameters to be estimated. For this reason, the weighting function is superseded by a factor constant over a waveform ($\{\sigma_i\} = \sigma$). In order to normalize the residuals ($\{Vm_i - V_i\}$), this factor σ is set to the current estimate of the amplitude.

The derivatives of the mean return power *B* are computed numerically as following:

$$\frac{\partial Vm(swh_i, e_i, pu_i)}{\partial SWH} = \frac{Vm(swh_i + \delta swh, e_i, pu_i) - Vm(swh_i - \delta swh, e_i, pu_i)}{2\delta swh}$$

$$\frac{\partial Vm(swh_i, e_i, pu_i)}{\partial e} = \frac{Vm(swh_i, e_i + \delta e, pu_i) - Vm(swh_i, e_i - \delta e, pu_i)}{2\delta e}$$

$$\frac{\partial Vm(swh_i, e_i, pu_i)}{\partial pu} = \frac{Vm(swh_i, e_i, pu_i + \delta pu) - Vm(swh_i, e_i, pu_i - \delta pu)}{2\delta pu}$$

where models *Vm* using the current estimation vector θ_n are directly taken from the database.

Estimation (unweighted Least Square fit)

The fine estimates of the epoch (e), the significant wave height (swh), the amplitude (Pu) are derived from the iterative process defined previously, which is initialized from default values e_0, swh_0 and Pu_0 for each waveform.

This estimation process is stopped when 10 iterations is reached.

3.5.1.Beam parameters algorithm

Since the along-track antenna gain pattern can be accurately fit by a Gaussian function, the power distribution across the looks in a stack may be modelled as a Gaussian too. However if the satellite flies pitched, the antenna gain pattern will asymmetrically modulate the power distribution in the stack, leading to distort the pure Gaussian beam power distribution. The roughness of the ocean surface, such as significant wave height and mean square slope, may also impact the shape of the stack. The distribution is “peaky” over specular target (i.e. calm water) for which the angle of reflection is equal to the incidence angle, while a flatter curve determines a rougher water surface where the reflectance becomes more diffuse. Sea surfaces may be locally specular or very diffuse reflectors, but fall mostly somewhere between these two types. Statistical parameters of the shape stack may be useful in characterizing the surface roughness.

After averaging the power in each look over all ranges, we are finding the best-fitting curve to the resulting set of points by minimizing the sum of the squares of the residuals of the points from the curve. We propose to adopt a Gaussian-like model based on Gram-Charlier series [Kendall et Stuart,

1963] expanded up to fourth order (to include kurstosis of the density function), to fit the power distribution, as shown in Figure 3.14 for example.

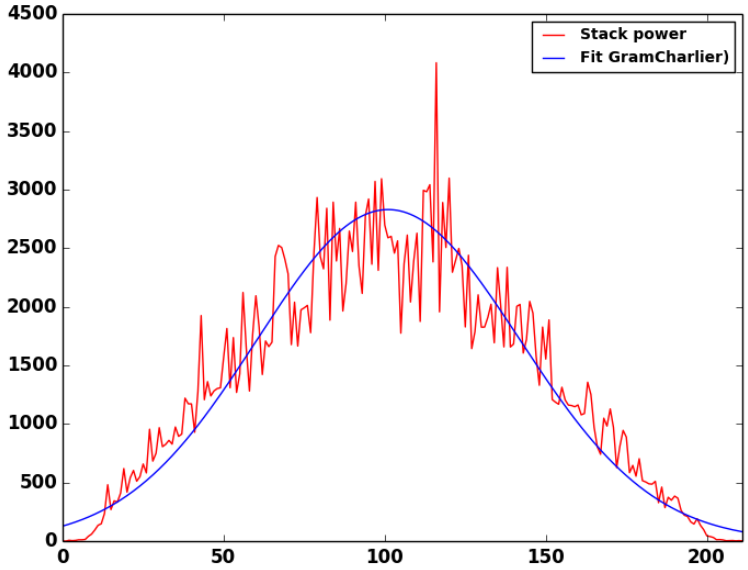


Figure 3.14: The distribution of power in a stack across the looks, and the best-fit Gaussian-like to the data.

The five retrieved moments are:

- the amplitude of that fit,
- the mean of that fit,
- the stack standard deviation of that fit which defines how widely spread the values are around the mean,
- the stack skewness which measures the asymmetry of the distribution,
- the stack kurtosis which measures the peakiness of the distribution. When greater than 3, the effect of the kurtosis is to peak the distribution.

The fitting parameters are determined resorting to a MLE iterative procedures as the one used in waveforms retracking. An example of plotting fit parameters is given in Figure 3.15

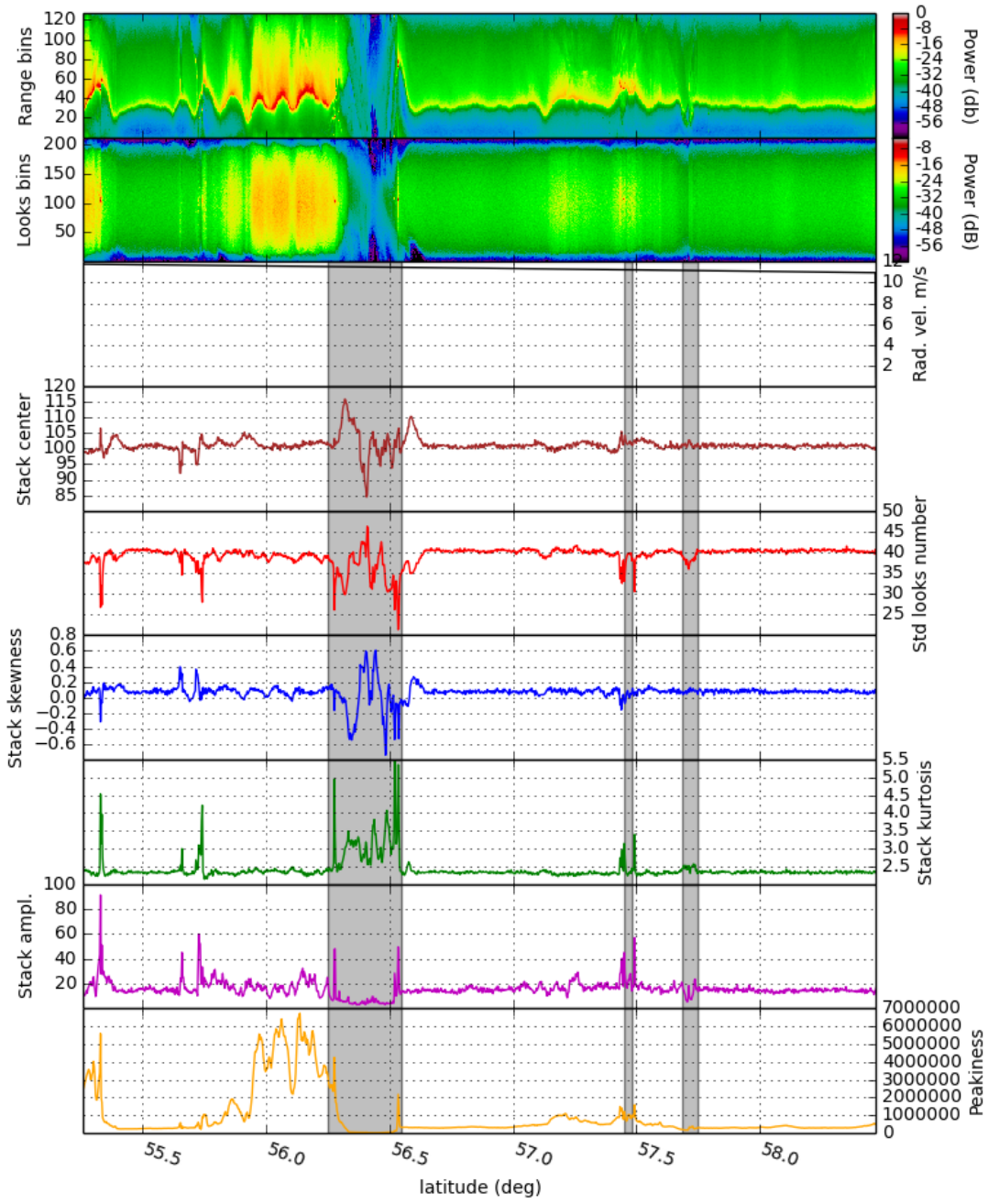


Figure 3.15: The beam parameters. The shaded data are from land surface, and from water bodies elsewhere.

To compute the sea level anomalies

For both SAR-mode and RDSAR data, some additional altimeter components are needed to be used in the Sea Level Anomalies (SLA) calculation as defined by this formula:

$$SLA = Orbit - Range - \sum_{i=0}^{N} C_i - MSS + bias$$

where the *bias* allows to refer the CryoSat-2 data to the Jason-2 reference level (and to take care of the systematic bias between LRM CryoSat-2 and Jason-2 measurements), *Orbit* corresponds to the distance between the satellite and the ellipsoid, *Range* is the distance measured by the altimeter between the satellite and the sea surface, *MSS* is the Mean Sea Surface of the ocean over a long period and $\sum_{i=0}^{N} C_i$ is the sum of all the corrections needed to take into account the atmospherically effects (wet and dry troposphere, ionosphere, inverse barometer) and the geophysical phenomena (ocean tides, high frequency atmospheric effects on ocean).

Note that the sea-surface state (electromagnetic sea-surface bias) is not considered in the equation since no SSB solution has been calculated yet.

To compute the backscatter coefficients

For each measurement, the 20 Hz Sigma0 is computed by combining the retracked amplitude of the waveforms with the scaling factors (for both SAR and RDSAR modes), as follows:

$$\sigma_0 = 30 * log_{10}(Orbit\_Alt) + 10 * log_{10}(Earth\_Rad + Orbit\_Alt) + 10 * log_{10}(Pu) + constant$$

To compute the wind speed

The Labroue and Tran's [2007] 2D model developed for Envisat mission (LRM) is used to compute wind speed for SAR-mode data.

4. SAR-mode method over Coastal Zones

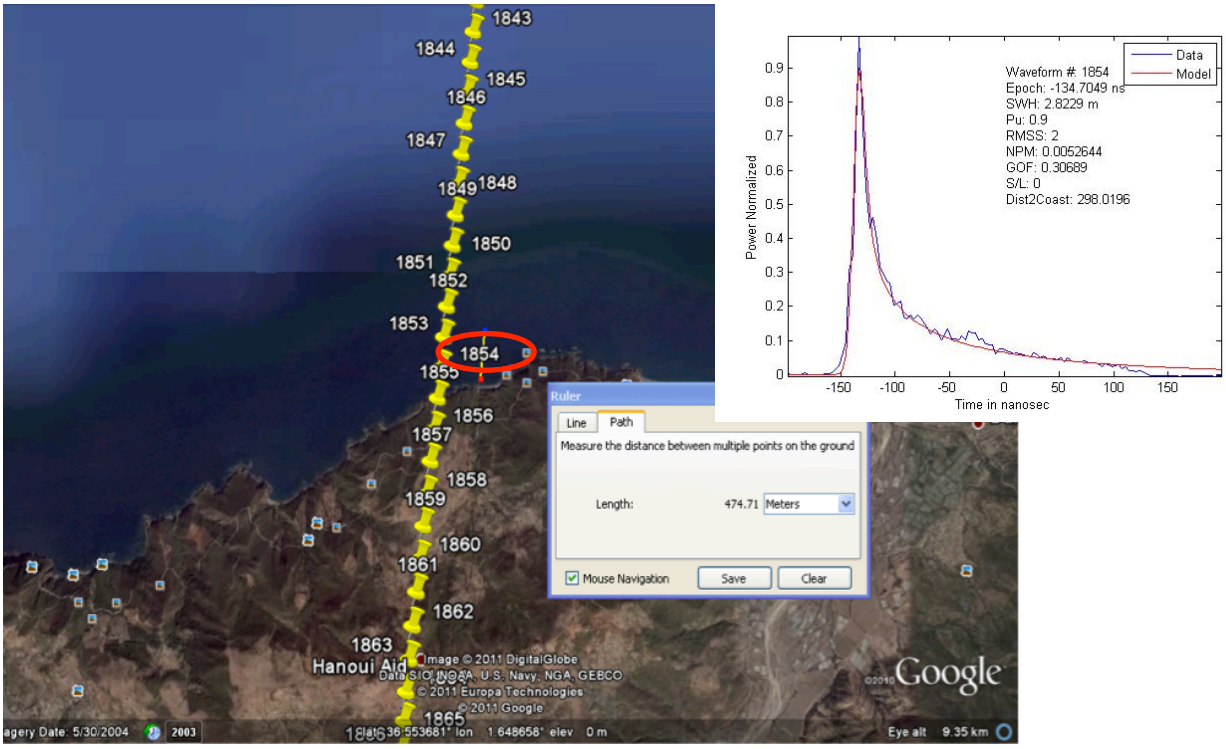
The increasing interest of SAR mode altimetry, especially in coastal areas, is linked to the higher resolving measurement capability in the direction of the platform velocity, i.e. along-track direction. The improved resolution is a consequence of the delay-Doppler processing of the radar echoes. The along-track resolution reaches values typically around 250 to 300 meters (R. K. Raney, Resolution and Precision of a Delay-Doppler Radar Altimeter 2005), which represents a remarkable improvement in comparison to the several kilometres currently achieved with conventional altimetry.

The shrinking of the radar footprint allows reducing the land contamination effect, which is the main limitation in the use of satellite altimetry for coastal areas. As a consequence of this, SSH, SWH, and WS measurements are expected to be provided at as close as several hundred meters from the shore. Various studies have already shown the capabilities of a delay-Doppler Altimeter to obtain useful radar echoes at a distance about 300 meters from the coast: see, for example, [(Dinardo, Lucas and Benveniste 2011)]. As can be observed in Figure 4.1 (a) waveform #1854, obtained from few hundred meters from the coast, was successfully retracked by the SAMOSA model, which allowed estimating the range delay and SWH for that area.

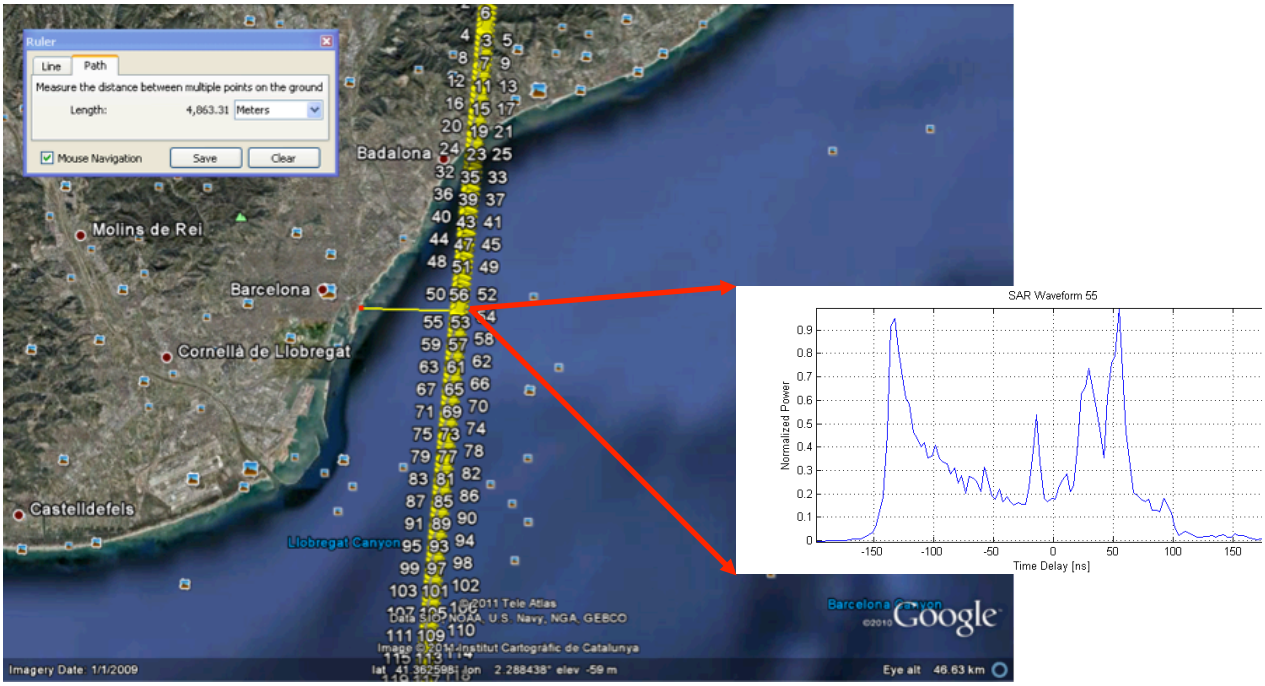
However, as pointed out above, in a delay-Doppler altimeter the resolution improvement occurs only on the along-track direction, while the across-track direction remains pulse-limited. This implies that the DDA's response in coastal regions depends on the relative orientation between the coastline and the spacecraft orbital plane. The 250 m dimension prevails when the coastlines are perpendicular the flight direction, whereas the longer pulse-limited direction domintaes when the altimeter is passing parallel to the shore. In [(Dinardo, Lucas and Benveniste 2011)] the authors provided examples of CryoSat-2 SAR mode waveforms in this situation. As can be observed in Figure 4.1 (b), despite the fact that the distance to the coast is almost 5 km, SAR waveform #56 is corrupted by a land effect in its tail, which prevents the use of this data for oceanographic applicatons.

Within the LOTUS project we have developed an algorithm to disregard the range bins of the SAR waveform which are likely to be affected by land contamination. This is achieved by geo-locating the delay and Doppler pairs of the SAR Altimeter stack, i.e. the 2-dimensional array for the range cell migrated Doppler looks aligned with respect to the nadir position. The SAR waveform retracker is then fed with this information and neglects the range bins subjected to land contamination from the geophysical parameter estimation process.

The following sections describe this algorithm and provide a first insight on the retracking improvement with respect to the situation in which the SAR waveforms close to coastal areas are not supervised for land contamination effects.



(a)



(b)

Figure 4.1: Example of SAR Altimetry waveforms for satellite tracks (a) perpendicular to the coast, and (b) parallel to the coast.

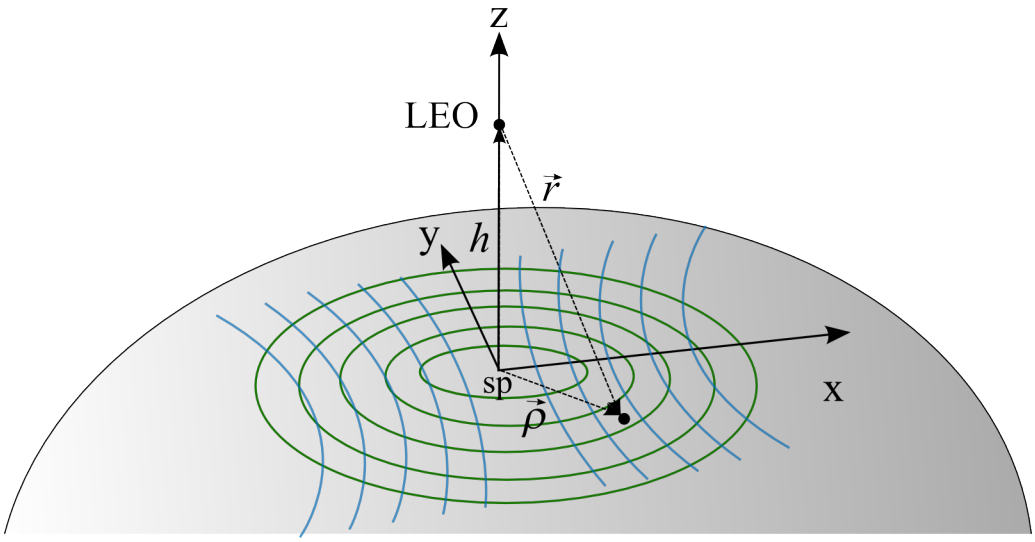


Figure 4.2: SAR Altimeter scattering geometry. In green, the iso-delay lines; in blue, the iso-Doppler lines. The x axis corresponds with the along-track direction of the platform.

4.1. Scattering Geometry

As mentioned above, the developed algorithm is based on the geo-location of the SAR Altimeter stack delay-Doppler pairs in order to be able to select the SAR waveform range bins that are not likely to be affected by land contamination effects. Such geo-location can be achieved as the different delay and Doppler bins define different regions on the surface.

The scattering geometry for a conventional Low Earth Orbit (LEO) altimeter is represented in Figure 4.2. The point on the surface with the shortest delay is defined as the specular point. In a nadir looking geometry, this corresponds with the sub-satellite position.

For simplicity and without loss of generality, it will be assumed that the origin of coordinates is located at the specular point position. For a point on the surface defined by $\vec{\rho}$ the delay of a signal transmitted from the altimeter with respect to the specular point, $\delta(\vec{\rho})$, can be calculated as [SAMOSA3-DPM]:

$$\delta(\vec{\rho}) = 2\,|\vec{r}(\vec{\rho})| - 2\,h \cong \frac{\alpha}{2\,h}(x^2 + y^2)$$

where h is the altitude of the satellite above the specular point, x and y are the along-track and across-track dimensions, and α is the orbital factor, defined as $\alpha = (R_E + h)/R_E$, in which R_E is the local radius of the Earth at the sub-satellite position.

The Doppler frequency is given by the dot product of the viewing unit vector and the satellite's velocity vector:

$$f_D(\vec{\rho}) = \frac{2\,\vec{r}(\vec{\rho})\,\vec{V}}{\lambda|\vec{r}(\vec{\rho})|}$$

where λ is the signal wavelength and V is the speed of the platform. Assuming that the speed of the platform is zero except in the along-track direction x, the previous equation can be approximated as:

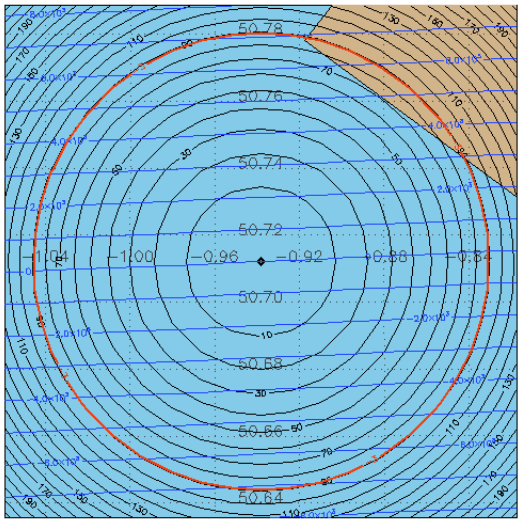


Figure 4.3: Scattering geometry; in black, iso-range circumferences around the specular point; in blue, iso-Doppler lines; in red, -3 dB antenna pattern.

$$f_D(\vec{\rho}) \cong \frac{2\,V_x\,x}{\lambda\,h}$$

The locus of points on the surface with equal delay and Doppler frequency with respect to the specular point are named iso-delay and iso-Doppler lines, respectively. As shown in Figure 4.3, in the case of the nadir looking altimeter, the iso-delay curves define circumferences around the specular point, whereas the iso-Doppler line define hyperbolas with the vertices located on the along-track direction. For a scattering geometry limited by the altimeter antenna radiation pattern, the iso-Doppler lines can be considered as parallel lines perpendicular to the along-track direction.

The previous equations define a quadratic system that should be solved for $\vec{\rho}$. By inverting this system is possible to geo-locate the points of the stack. Two solutions are obtained, corresponding to two points on the surface. However, in order to be able to geo-locate the delay-Doppler pairs of the stack, it is first necessary to determine to a certain extent, which range bin of the SAR waveform corresponds to the specular point.

4.2. Specular Point Position Determination

In order to perform the geo-location of the stack it is necessary to have a first estimation of the specular point position so that it is known which range bin of the SAR waveform could be adopted as the zero-delay reference. Since the full-waveform retracking cannot be performed due to the likely presence of land contamination an empirical algorithm is used to determine the range of the specular point, d_{sp}. As it is known, the specular point would be located at some point within the leading edge of the waveform. In this case, the middle point of the waveform leading edge was selected as the reference delay. To improve the precision in the estimation of the specular point, the SAR waveform was interpolated by a factor of 100 by means of a spline interpolation.

This estimation is not absolutely accurate and it would not be acceptable as an altimetric estimate, however, for the purpose of geo-locating the points of the stack on the surface it is a reasonable approximation. As an example, in Figure 4.4 an open ocean waveform is shown together with the approximation. The green diamonds represent the waveform maximum, the half power point, and the leading edge starting position. The waveform half power point was determined to be on the range gate 32.6, whereas the full waveform retracker provided a range estimation of 33.77 gates, i.e. 1 range gate difference, corresponding to 20 meters on the ground.

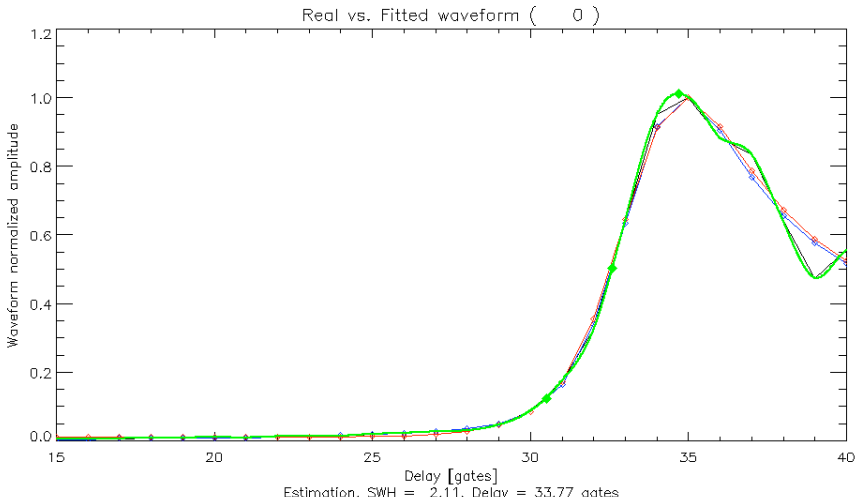


Figure 4.4: SAR-Altimeter waveform leading edge. In black, SAR waveform data; in green, interpolated waveform. Green diamonds represent the waveform maximum, half power, and leading edge starting positions.

4.3. Stack delay/Doppler pairs geo-referencing

Once the position of the specular point is known, it is straightforward to solve the delay and Doppler equations to obtain the points on the surface of the delay/Doppler pairs of the stack.

The Doppler frequency for beam n is obtained as:

$$f_{D,n} = \frac{2\,V}{\lambda}\cos(\theta_n)$$

where θ_n is the relative observation angle with respect to the nadir. This is obtained as:

$$\theta_n = \frac{\pi}{2} + \Delta\theta \cdot n$$

with $n \in \left[-\frac{N_{beams}}{2}, \frac{N_{beams}}{2}\right]$, and

$$\Delta\theta = \frac{V \cdot BRI}{\alpha\, h}$$

where BRI is the Burst Repetition Interval. The geo-location of the Doppler bands defines the along track position, x_n, of the points on the surface:

$$x_n = \frac{\lambda\, h\, f_{D,n}}{2\, V_x}$$

This defines the position of the along-track of the delay/Doppler map points. The actual position of the beams of the stack is always centred at nadir, however in order to calculate the range bins for each specific Doppler band, x_n should be used, and shift along track the Doppler band *a posteriori.*

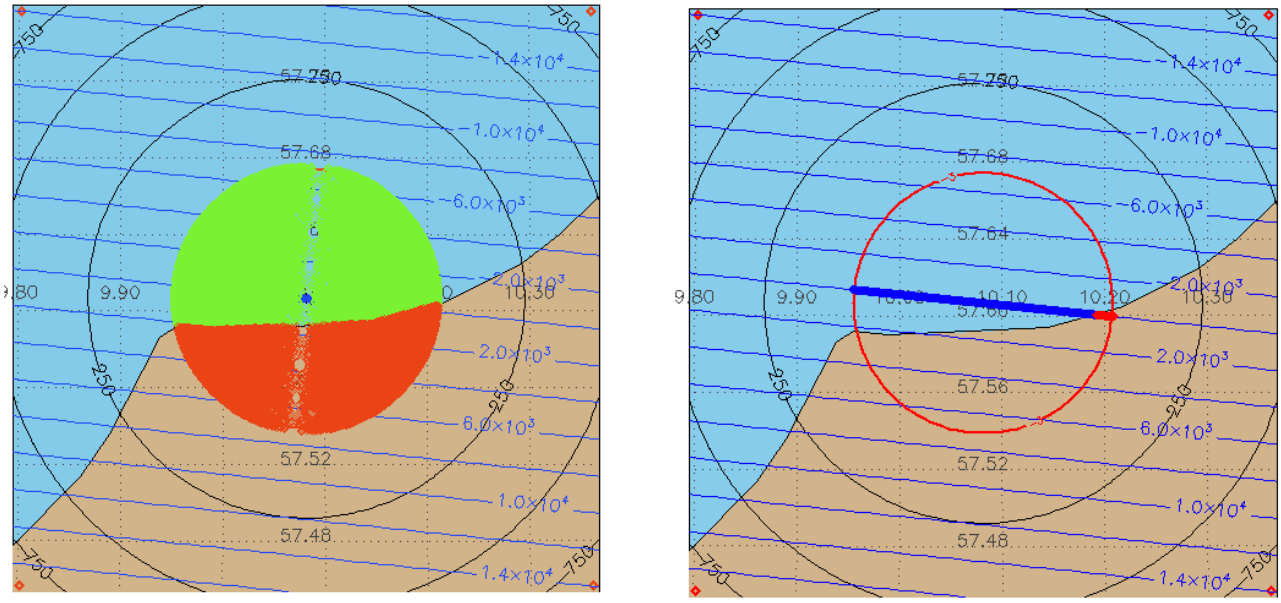
The across-track positions, y_m, are then obtained as:

$$\mathrm{y_m} = \pm\sqrt{\frac{2\, d_m\, h}{\alpha} - {x_n}^2}$$

where d_m is the range for the waveform gate m defined as:

$$d_m = \frac{c}{2\, f}\,(m - d_{sp})$$

where c is the speed of light, and f is the waveform sampling frequency. Note that $\mathrm{y_m}$ has two solutions for each range gate. This implies that a range ambiguity exists in the delay/Doppler stack, which means that it is not possible to distinguish whether the land contamination comes from one or the other side of the track. This forces to discard range gates that could effectively have energy coming from the ocean surface. This is better observed in Figure 4.5.



(a) (b)

Figure 4.5: Scattering geometry with delay/Doppler pairs projected on the surface. In green or blue, delay/Doppler points on the ocean surface; In red, delay/Doppler points on land, to be discarded from the analysis. (a) DDM delay/Doppler pairs. (b) Stack delay/Doppler pairs for zero Doppler.

In Figure 4.5(a) the delay and Doppler pairs of a burst have been represented on the surface, for a specific position of one CryoSat track off the coast of Denmark. The points over the ocean surface have been represented in green, whereas the points on land to be neglected from the analysis were depicted in red. To define the coastline, the NOAA Global Self-consistent, Hierarchical, High-resolution Geography (GSHHG) Database has been used.

To form the stack several Doppler beams are used from different bursts along the track. This is done in such a way that all beams point to the same position, i.e. the sub-satellite point of the zero-Doppler beam, in order to then be averaged together to form the SAR waveform. In our algorithms, the Doppler band positions need to be shifted along-track to the sub-satellite point. Once this is done it is possible to determine which of the range gates of each specific Doppler band lie on ground.

This is shown in Figure 4.5(b). For the zero-Doppler beam the points are represented on the surface. It is observed that only the outermost part on the right hand side of the track is on land. However, and despite the fact on the other side of the track the points corresponding to the same range position lie on the ocean, those will also be discarded as it cannot be warrantied that on that specific range gate there would not be land contamination.

4.4.Delay/Doppler pairs selection

The delay/Doppler pairs whose geo-located positions lie on land define a mask that is later applied to the stack in order to determine which of the range gates of the SAR waveform can be used reliably in the retracking without land contamination. An example of such a mask is shown in Figure 4.6. The parts of the stack represented in red are the ones subjected to possible land contamination and are then neglected for the retracking and geophysical parameter estimation.

Figure 4.7 shows the final SAR waveforms as the altimeter approaches the coast. In green we have represented the waveform lags that are free from land contamination and therefore will be used for the retracking. In red we show the points that have been neglected. Figure 4.7 (a) to (e) represent sequential 20 Hz observations. As it is observed, as the altimeter approaches the coast, more waveform range gates are susceptible to have land contamination, and are therefore discarded for later retracking processes. It is also observed that the waveform gates that do not follow a harmonic behaviour, such as the ones from the ocean, are well identified thanks to this algorithm.

In following stages the waveforms are retracked with the analytical SAMOSA model, and the SSH and SWH parameters are estimated.

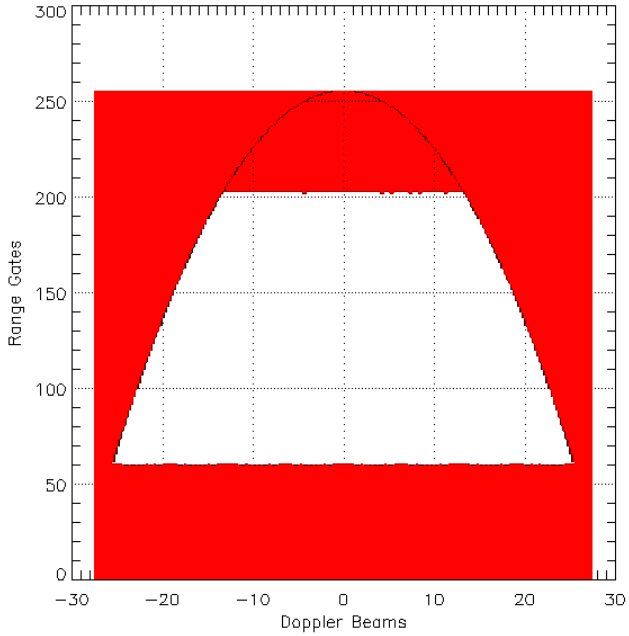
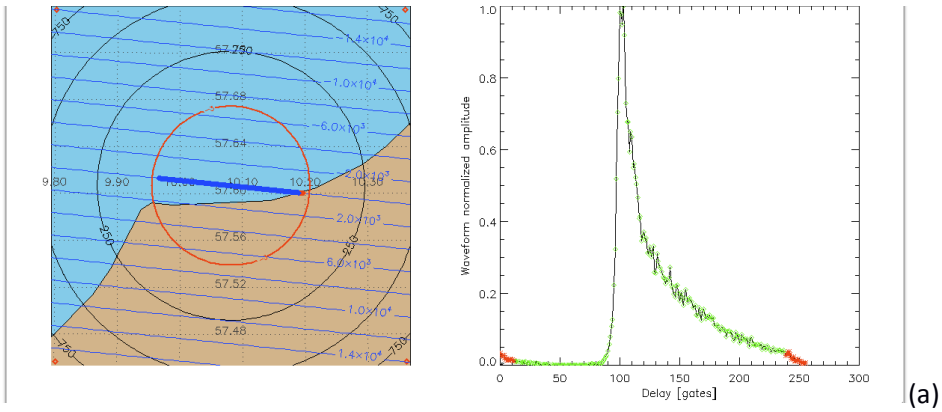
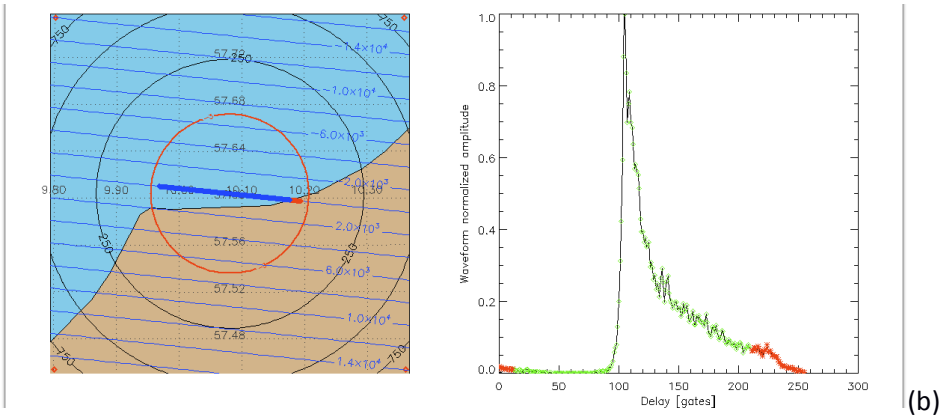


Figure 4.6: Stack mask to avoid land contamination.



(a)



(b)

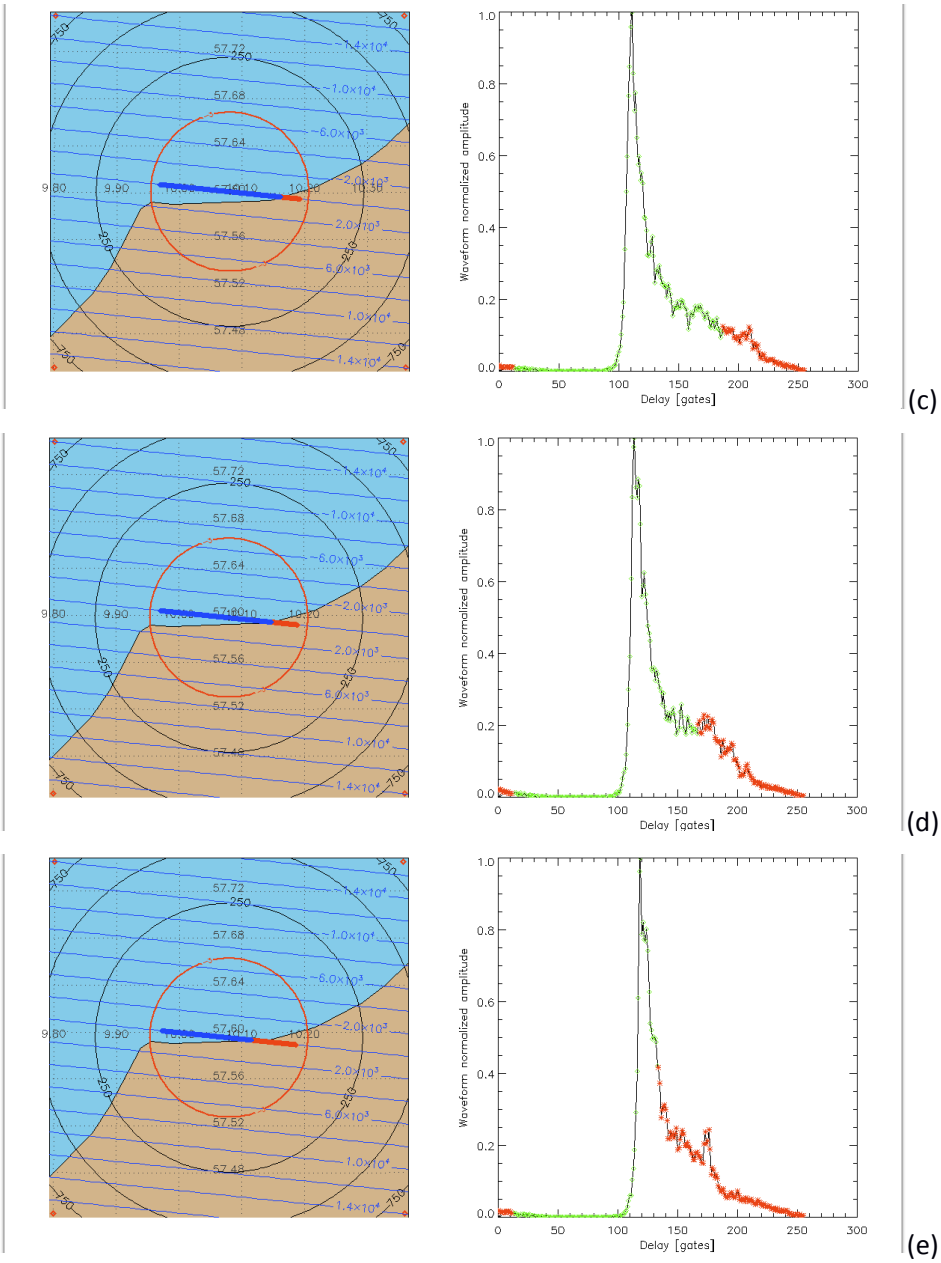


Figure 4.7: Scattering geometry and SAR waveforms. In green, waveform free from land contamination; In red, waveform lags discarded for the retracking.

4.5. Waveform Retracking

Within the frame of LOTUS WP1 and the CryoSat Plus for Ocean (CP4O) ESA project, the SAMOSA-3 fully analytical model [SAMOSA3-DPM], i.e. the current baseline implementation for the Sentinel-3 Level-2 processing, has been updated in order to be able to account for different Level-1 processing approaches. The model updates were focused on the appropriate handling of the energy distribution over the different echoes of the delay-Doppler stack, an application of a Look-Up Table (LUT) for the selection of a variable width Point Target Response (PTR) as a function of SWH, the complete implementation of the SAMOSA-2 model, and an appropriate estimation of the thermal noise from the SAR waveform.

The performance of the updated SAMOSA model can be seen by observing the goodness of fit of the model with respect to data. Figure 4.9 shows the best fit of the SAMOSA-3 model to actual Level-1b CryoSat-2 data. As can be observed, the SAMOSA waveform has a perfect correspondence with both the SAR data. The best fit is also compared to the best fit waveform provided by CNES within the Cryosat Processing Prototype (CPP) Level-1b product. As expected, both models, are mostly coincident, which demonstrate the validity of the proposed method.

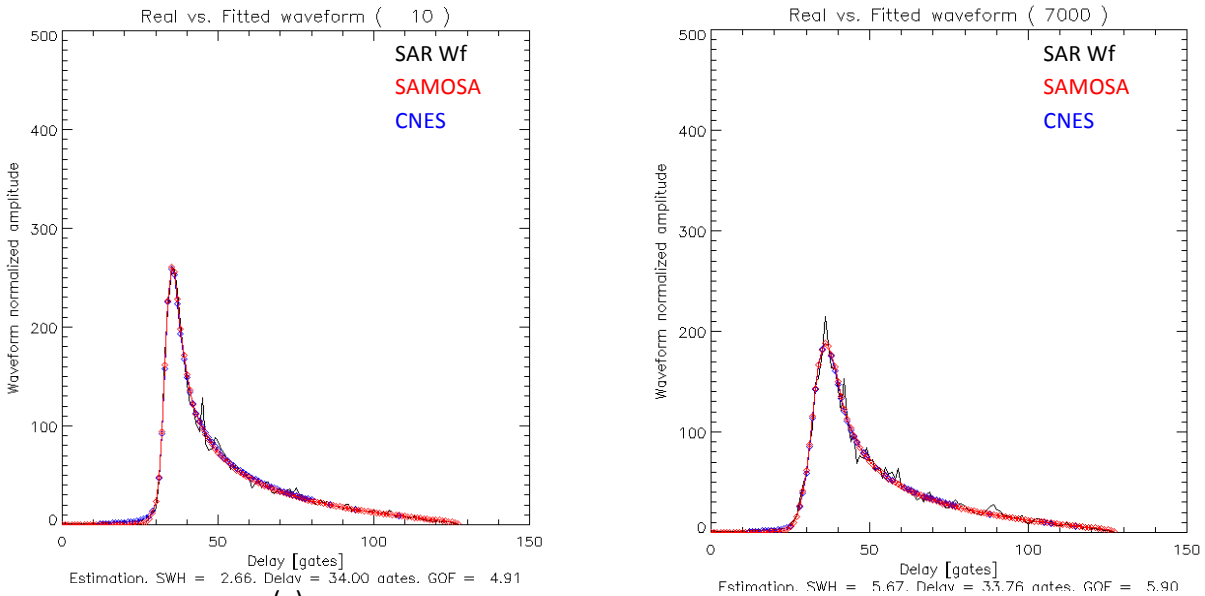


Figure 4.8: SAMOSA updated model best fit waveforms after stack trimming for (a) low SWH, and (b) high SWH.

The updated SAMOSA model was integrated within a full waveform retracker, which performs the joint estimation of Sea Surface Height (SSH), Significant Wave Height (SWH), and Sigma0, by means of an iterative Levenberg-Marquardt minimization algorithm. The search algorithm to be applied for the inversion problem is on a mean square minimization error approach. This means that the parameter vector $\vec{p}$ for a given waveform will be estimated by finding the that minimizes the mean square error

(MSE) between the waveform and the theoretical waveform from the SAMOSA model associated to the parameters in $\vec{p}$.

In the MSE method the minimization process is implemented by a Levenberg–Marquard optimization algorithm, which is a combination of the steepest-decent optimization method, i.e. minimization in the direction of the gradient, and the Newton Method, that uses a quadratic error model in order to speed up the optimum.

The search for $\vec{p}$ is started by an initial guess. Once the initial estimate has been made, the gradient of the MSE at that point $\nabla_p MSE(\overrightarrow{p_n})$ is found. The improved estimate for $\overrightarrow{p_{n+1}}$ is the previous estimate plus a small step dp in the parameter space in the opposite direction of the gradient. Mathematically, this expression can be summarized as

$$\overrightarrow{p_{n+1}} = \overrightarrow{p_n} - dp\ \nabla_p MSE(\overrightarrow{p_n})$$

The iterative process is continued until no further reduction in the MSE is obtained. The final value of $\vec{p}$ is the estimated value of the parameters for the given waveform. In our case, the parameter vector to be estimated would be the sea surface height (SSH), significant wave height (SWH), and waveform power (Pu), which is then translated to Sigma0.

Within the frame of CP4O, the SAMOSA model was adjusted to the Level-1 processing of the Cryosat Product Prototype (CPP) provided by CNES. The retracker was applied to these data in order to estimate the Level-2 geophysical parameters, which were cross-validated with the SSH, SWH, and Sigma0 provided within the CPP Level-1b product and calculated by means of a numerical retracker. In order to evaluate the retracker performance, a preliminary statistical comparison was performed with the retracking outputs of the SAR numerical model developed by CNES, and provided in the CPP Level-1b product.

A single track was selected as a benchmarking tool for the development of the algorithm. This track covers more than 20 latitude degrees and has a SWH range between 2 and 6+ meters, which was considered representative of a general situation.

The results for the comparison of the SWH and SSH estimations provided by both retrackers are shown in Figure 4.9. Figure 4.9 (a) shows the SWH difference between the updated SAMOSA retracker and the CPP solution. As can be observed, the SWH difference between both solutions is concentrated around cero with no significant slope along the whole track. This is verified if the SWH difference is represented against the CPP SWH estimation, Figure 4.9 (b). For the 1 Hz product, the SWH mean error and standard deviation yield 3 mm and 3.4 cm, respectively.

In a similar way, the SSH difference against the CPP product is shown in Figure 4.9 (c-d). A small trend is observed toward increasing SWH values, however this is still below 1 cm for the whole range. The SSH mean error and standard deviation equal 1 mm and 3 mm, respectively.

The processing scheme presented in this section will be used to produce the Coastal Ocean data products within *WP4 – Production of Demo Data and Assessment* of the LOTUS project, which will be used in later stages of the project.

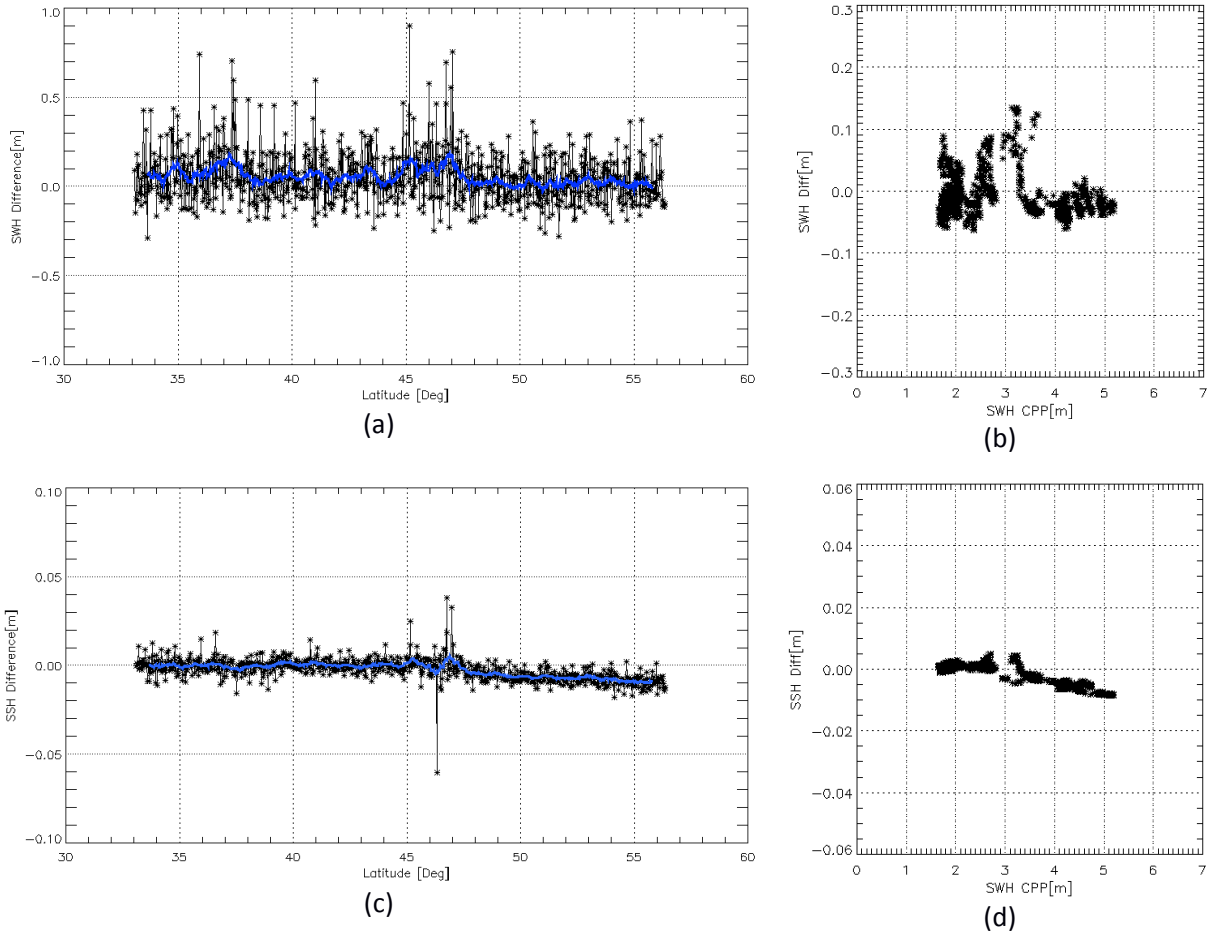


Figure 4.9: Difference between SAMOSA and CPP retracking solutions: (a) SWH difference along the track; (b) SWH difference vs CPP SWH estimation; (c) SSH difference along the track; (d) SSH difference vs CPP SWH estimation.

5. SAR-mode method over Polar Ocean

This section concerns the theoretical development and operational implementation of the system for classification and retracking of SAR mode altimeter waveforms obtained over Polar oceans where sea-ice is occurring. The content of this section is based on the deliverable D4.1 Algorithm Theoretical Baseline Document - Polar Ocean of the CryoSat plus for Ocean (CP4O) ESA project [Stenseng, 2014]. The current section represents a first iteration of the work to be developed within the frame of the LOTUS project.

The primary input to the system is the L1b and L2 products from the ESA Baseline-B processing, but additional L2 data have been considered as well.

5.1. Introduction to the Polar Ocean

For many surfaces, e.g. the open ocean, the surface can be assumed to be close to a diffuse reflector when observed from within a few degrees from nadir. These surfaces will reflect and spread the incoming radar energy almost equally in all directions, see Figure 1. A study of SAR altimetry data from the Airborne SAR/Interferometric Radar Altimeter System (ASIRAS) over open ocean show that within +/-6 deg. from nadir the return power is primarily attenuated by the antenna pattern [Stenseng, 2009].

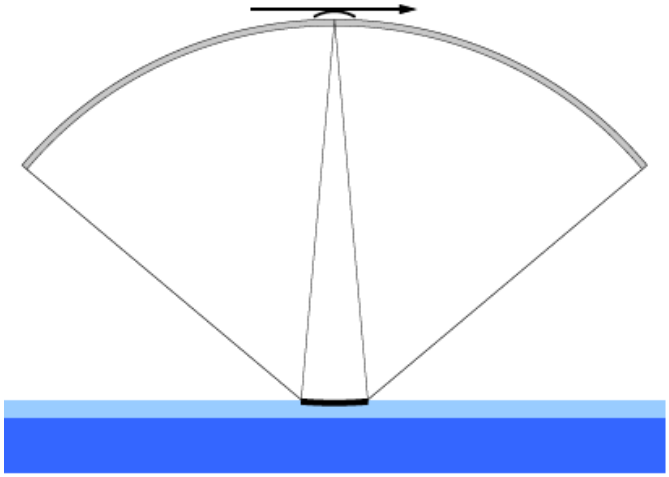


Figure 5.1: Reflected power from a diffuse surface e.g open ocean or a sea-ice floe.

In the sea-ice covered part of the Polar Ocean is a combination of several different surface types i.e. smooth young ice, rough old ice, pressure ridges, sea-ice melange, leads and open water. Here the primarily focus will be on the leads which can be used for deriving the sea level height. Secondary the open water and diluted sea-ice melange will be discussed.

Lead is typical defined as a fracture, with open water or thin newly formed ice, in the sea-ice cover with width up to several hundreds of meters and reaching several kilometers in lengths. Since the sea-ice cover attenuates the ocean waves, the thin ice or water surface in the lead will usually be smooth

with only small wind generated waves. This smooth surface will give a specular return where most power is reflected to the point closest to the burst position that is perpendicular to the water surface i.e. the nadir pointing beam, see Figure 5.2.

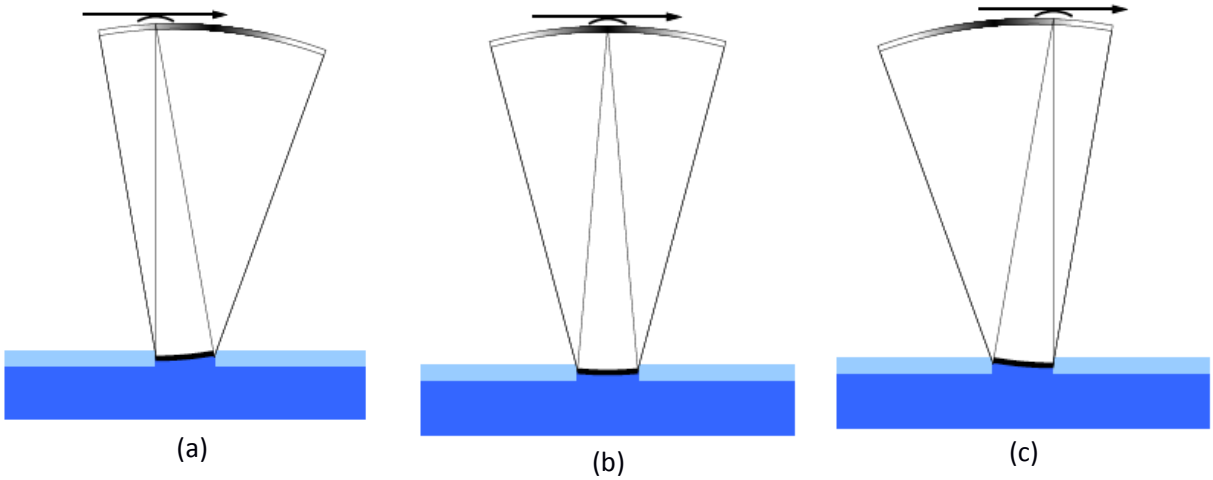


Figure 5.2: Reflected power from a lead observed from three different positions: (a) In front of the lead power is primarily reflected forward, away from the satellite. (b) Over the lead power is primarily reflected back to the satellite. (c) Past the lead power is primarily reflected backwards, away from the satellite.

5.2. Algorithm Description

The major challenges in obtaining sea surface height estimates in the Polar Ocean is many; the limited amount of sea-ice leads, areas with sea-ice melange, retracking of extreme peaky waveforms and off-nadir returns.

We have consequently taken the approach of working with 3 retrackers depending on the state of the Ocean as defined by the shape of the waveform.

1) SAMOSA-3 retracker for Open Ocean (normal ocean waveform)
2) SAMOSA-3 adapted retracker for Leads (specular waveforms)
3) Primary Peak empirical threshold retracker for sea-ice floes or melange (remaining waveforms)

The use of the Primary Peak empirical threshold retracker is performed for all waveforms that cannot be retracked by the SAMOSA-3 normal or adapted retracker. This included sea-ice floes/melange or even unclassifiable waveform (multipeakes). In principle these 3 retrackers enables us to retrack waveforms in most conditions in the Arctic Ocean.

It is important to notice that only the SAMOSA-3 physical retracker enables a determination of sea surface height, significant wave height and the wind speed. These cannot be determined by the

SAMOSA-3 adapted retracker for leads or the Primary peak empirical retracker which solely determines the sea surface height.

In order to process the data in the Polar Ocean, the first step is to perform a classification of the waveform (ocean, lead or sea-ice floe/melange) and to select the appropriate retracker. After the retracking, the output of the various retrackers are combined to give the final Arctic data product.

5.2.1. Classification

The classification is primarily based on a set of parameters describing the morphology of the waveform and the power distribution in the individual looks in the stack. Initial description of the technique is found in D1.1.

The parameters describing the power distribution is taken directly from the ESA Baseline-B L1b product and the parameters describing the morphology is described in the following. Furthermore, as described above the ESA Baseline-B L2 product is merged into the DTU dataset which allows a direct comparison of the ESA Baseline-B classification of the individual waveforms. A classic parameter for identification of specular returns is the Pulse peakiness (PP), which determines the ratio between the peak power and the integrated power in the waveform [Laxon, 1994; Peacock and Laxon, 2004]. In the classical formulation the first five samples are discarded to avoid wrap around effects and the ratio is scaled using the nominal tracking point bin. A pulse peakiness value is available in the ESA Baseline-B L2 product but the used formulation is unknown and results in values exceeding 500 as opposed to the classical values not exceeding 1. Equation 3 gives the formulation used in the DTU classification.

$$PP = \frac{65535}{\sum_{i=0}^{127} p_i} \tag{3}$$

To further characterize the peak a Gaussian (see Equation 4) is fitted to the most powerful bin including two bins on each side similar to approach by [Armitage and Davidson (2014)]. This will give additional information about the peak geometry and aid the identification of contaminated lead returns.

$$G_i = A_g \cdot \exp\left(-\frac{(i - E_g)^2}{2 \cdot W_g^2}\right) \tag{4}$$

By rearranging the classical radar equation [e.g. Skolnik (2001)], tacking the Earths curvature into account, and using the power estimated by the Gaussian fitted to the peak, a simplified version of $s0$ can be estimated as following:

$$\sigma_0 = 40 \cdot log_{10}(h) + 10 \cdot log_{10}\left(\frac{R_e}{R_e + h}\right) + 10 \cdot log_{10}\left(\frac{P_u}{P_{Tx}}\right) + C_{\sigma_0} \tag{5}$$

where Pu is the power in Watts derived from Ag, PTx is the transmitted power and Cs0 is a constant accounting for all other losses which will be assumed to be constant.

Using the stack standard deviation SSD as available in the ESA L1b product and the calculated pulse peakiness, returns with high stack standard deviation or low pulse peakiness are considered to be from the open ocean.

Single lead returns are selected if the pulse peakiness is very high, the stack standard deviation is low and the s0 is high. The values for the thresholds on these parameters can be relaxed if the number of lead returns is more important that the precision. This could be relevant in areas with permanent high sea-ice concentration and therefore fewer and smaller leads.

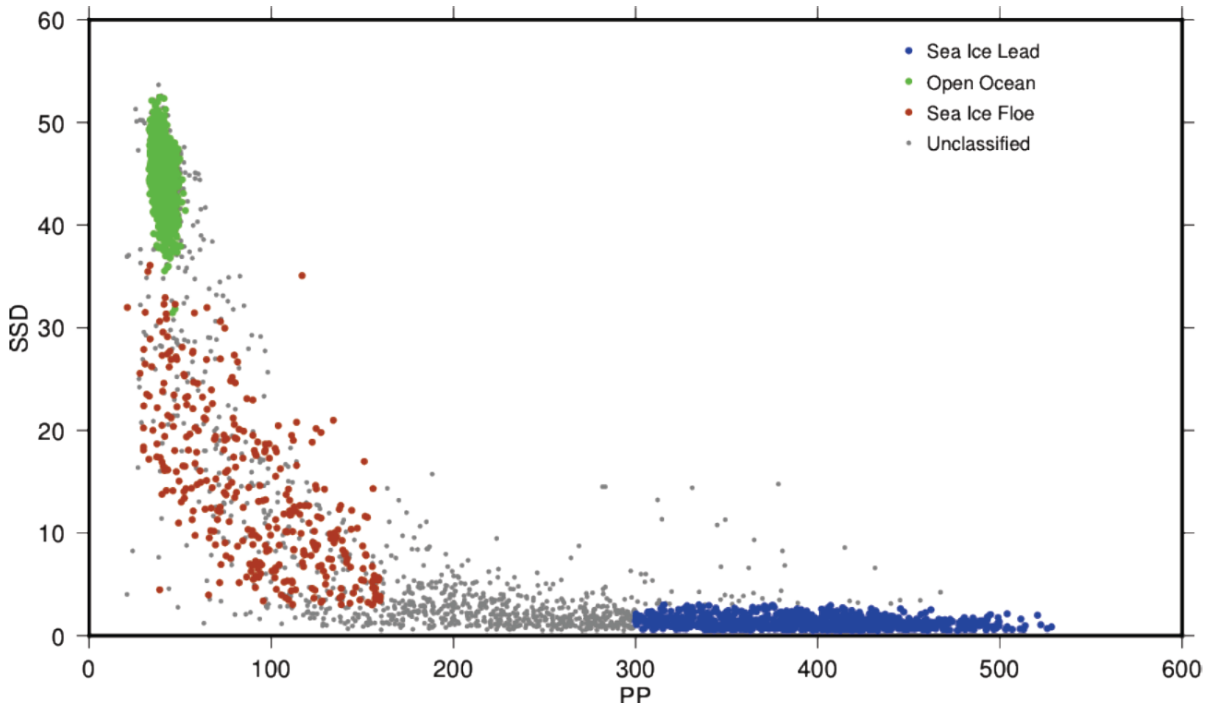


Figure 5.3 illustrating the various waveform classification as per DTU (from D1.1)

Following the recommendations by ESA we decided to perform the classification based on the PP and the SSD and using a standard scaling of the PP (Eq 3).

Sea ice leads if PP was greater than 300 and SSD smaller than 4

Open ocean if PP Is less than 9 and SSD is less than 36.

Sea ice-floes / melange/unclassified if PP Is less than 300 and SSD greater than 36.

Again the group of waveform falling outside the sea-ice lead and the open ocean will all be retracked by the Primary peak retracker.

5.2.2. SAMOSA-3 adapted retracker.

The general description of the SAMOSA retracker as applied for open ocean data in the Polar Regions is found in the section above and no further description is needed. Here the process of adapting the SAMOSA3 retracker for the Arctic is shown to very specular returns is shown in Figure 5.4.

For adaptation it is important to reduce the number of parameters which are being estimated in each iteration of the fitting procedure. There are multiple benefits of reducing the number of parameters being fitted. First, a reduced number of parameters being fitted reduces the computation time of the SAMOSA3 retracker as a model with less number of fitted parameters is simpler and faster. Also, for the purpose of this project, only the retracking position is of importance as it is used in sea surface height determination. Reducing the number of parameters being fit improves the precision on the parameter of interest. The adapted retracker is called the SAMOSA3A retracker.

In order to reduce the number of parameters being fit, the amplitude is normalized. This reduces the need to parameterize the amplitude. Further there is a need to deal with the varied types of waveforms present in the Arctic. Thus different modes are developed for dealing with the variation in the waveforms. The choice of parameters being fit would depend on the type of waveform being fitted. The different modes used are ocean mode and lead mode.

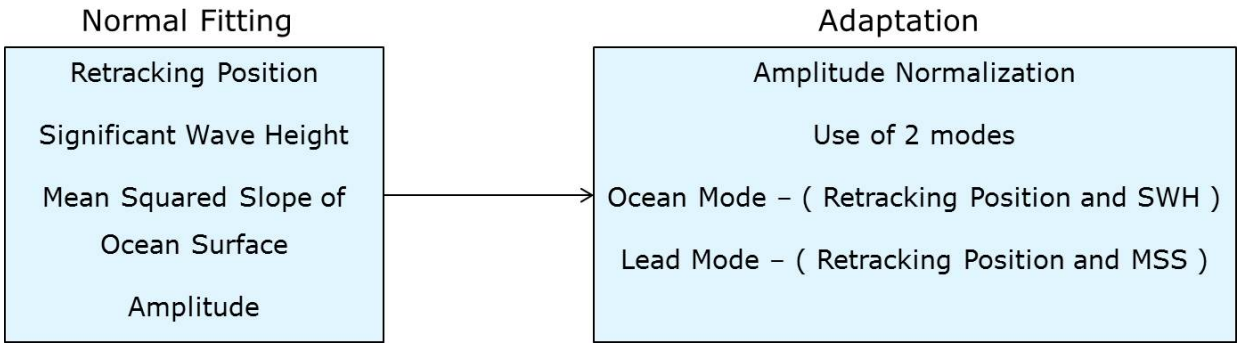


Figure 5.4: Adaptation of the SAMOSA3 retracker to SAMOSA3 adapted retracker by reducing the number of parameters being fit

5.2.3. Primary Peak empirical threshold retracker

The waveforms in that fails to be retracked using the SAMOSA3 or SAMOSA3-adapted retracker are generally complex or multi-peaked. The first step in the processing of these is to identify the primary Peak which we assume is the reflection from the sea surface. The algorithm to extract the primary peak uses the power in the bins of the reflected waveform to compute two thresholds (start and stop). These thresholds are to be compared with the power difference in consecutive bins of the reflected waveform. The start threshold is computed using the standard deviation of the power differences in alternate bins of the reflected waveform as described in Equation 1.

$$Th_{start} = \sqrt{\frac{(N-2)\sum_{i=1}^{N-2}(d_2^i) - (\sum_{i=1}^{N-2} d_2^i)^2}{(N-2)(N-3)}} \quad (6)$$

Pi is the value of power at the ith bin location. d_2^i as in Equation 7 is the power difference in alternate bins.

$$d_2^i = P_{i+2} - P_i \quad (7)$$

The stop threshold is computed using the standard deviation of the power differences in consecutive bins of the reflected waveform as described in Equation 8.

$$Th_{stop} = \sqrt{\frac{(N-1)\sum_{i=1}^{N-2}(d_1^i)^2 - (\sum_{i=1}^{N-2} d_i^i)^2}{(N-1)(N-2)}} \quad (8)$$

d_1^i as in Equation 8 is the power difference in consecutive bins.

$$d_1^i = P_{i+1} - P_i \quad (9)$$

In these equations N is the number of bins in the reflected waveform, Thstart is the start threshold used to compute the start point of the primary peak and Thstop is the stop threshold used to compute the stop point of the primary peak (Jain et al. 2014).

Next, a loop is run to check the power difference in consecutive bins (d_1^i) throughout the reflected waveform. The power difference in consecutive bins (d_1^i) at every location in the reflected waveform is compared with the start threshold. When the power difference exceeds the start threshold for the first time, the bin number is tagged as the start point of the primary peak. Within the primary peak, when the power difference of consecutive bins is less than the stop threshold, this is recorded as the

stop point of the primary peak It was found that the primary peak in SAR mode was only one to three bins wide. This narrow width of the primary peak is inadequate to apply the traditional OCOG and Threshold retrackers as these retrackers require thicker peaks to compute acceptable results. Thus the width of the primary peak must be increased by adding bins before the start point of the primary peak and bins after the stop point. Experiments were done by taking into account variable number of additional bins. When the retracker performance is evaluated as described later in the thesis, it is found that adding a total of 4 bins (2 bins before the start point and 2 bins after the stop point) to the primary peak gives the best results. Figure 1 shows the method of extraction of the primary peak.

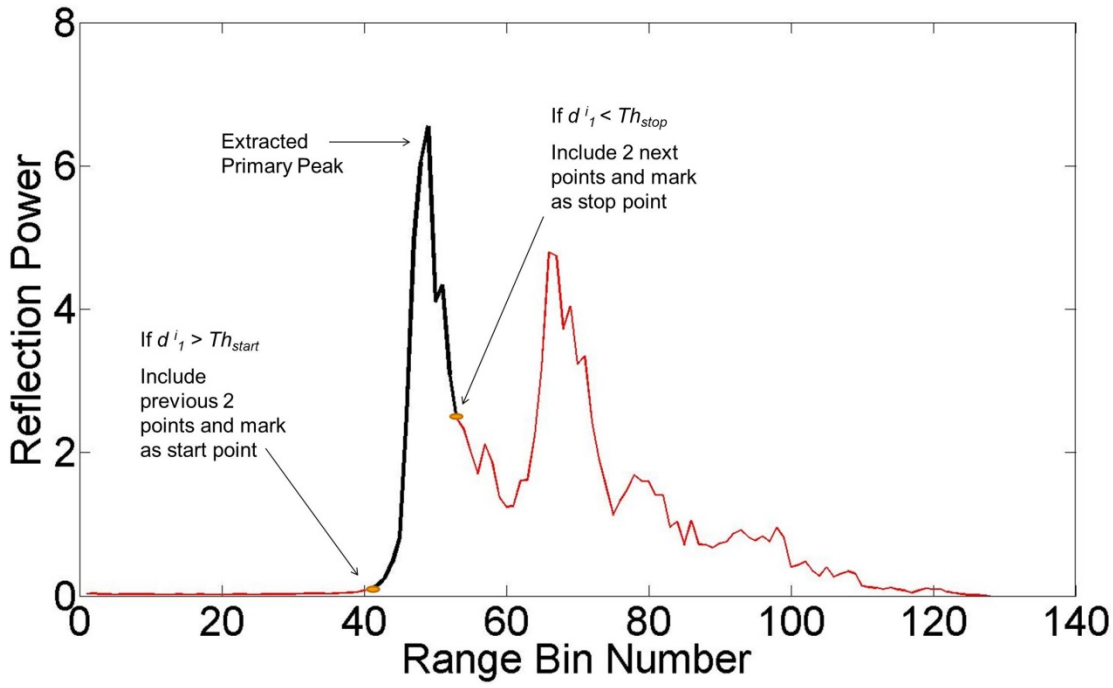


Figure 5.5: Definition of the primary Peak. The primary peak is shown in black and the rest of the waveform in red.

The subsequent retracking of the waveforms is performed using a simple leading edge threshold retracker similar to that of [Davis (1997)]. The description of this was done in D2.2 for the use over land where it is frequently used as it is a very robust retracker. The threshold value FT is chosen between 0 and 1 and indicates the fraction of the power benchmark to be used for the retracking point. To begin with the first bin position (j) with power higher than or equal to FT is found and a linear interpolation between bin j and bin j-1 is performed to obtain the fraction of a bin where the threshold power is found. Finally the epoch is derived, in bin counts, by adding j - 1 to the fraction as described in Equation 2.

$$E = \frac{F_T \cdot P_b - p_{j-1}}{p_j - p_{j-1}} + j - 1 \qquad (2)$$

5.3. Constraints and Limitations

The primary constrain of the presented algorithm is the natural variability in the number, size and cross-track location of the leads. Due to the nature of the sea-ice the number of leads changes both with region and season, allowing only few leads in the dense sea-ice north of Greenland and Canada and further limiting the general number and size of leads in the winter season.

The random cross-track location of the leads relative to the nadir point will introduce a negative bias in the height estimate of the water surface as the observed slant range to the specular lead will be grater than the range to the nadir point. The filtering of lead return groups will to some degree reduce the negative bias. Finally the limited bandwidth of the radar pulse will introduce noise in the height estimates as the highly specular peak is only captured in a very few gates (down to three gates) even after resampling at the power Nyquist frequency (i.e. zero padding the complex signal by size factor of two).

The use of different retrackers for different regions might lead to regional different retracker biases (depending on sea state). We have performed a calibration of the three selected retrackers to adjust for this bias but a full Arctic wide investigation of the bias is currently being performed.

6. References

[Amarouche et al., 2004]: L. Amarouche, P. Thibaut, O.Z. Zanife, J.-P. Dumont, P. Vincent and N. Steunou, "*Improving the Jason-1 ground retracking to better account for attitude effects*", marine Geodesy, Vol. 27, pp.171-197, 2004.

[Armitage et al, 2014]: Armitage, T. W. K. and Davidson, M. W. J. (2014). Using the interferometric capabilities of the esa cryosat-2 mission to improve the accuracy of sea ice freeboard retrievals. IEEE Transactions on Geoscience and Remote Sensing, 52(1):529–536.

[Boy et al., 2012]: F. Boy, T. Moreau, J-D. Desjonquères, S. Labroue, N. Picot, J-C. Poisson and P. Thibaut, *"Cryosat Processing Prototype, LRM and SAR Processing"*, presented at the 2012 Ocean Surface Topography Science Team Meeting. Available online:
http://www.aviso.oceanobs.com/fileadmin/documents/OSTST/2012/oral/02_friday_28/02_instr_processing_II/02_IP2_Boy.pdf

[Boy et Moreau, 2013]: F. Boy and T. Moreau, *"Algorithm Theoretical Basis Document (ATBD) of the CPP RDSAR processing for oceans)"*, S3A-NT-SRAL-00098-CNES, 1.0, 2013.

[Boy et Moreau, 2013]: F. Boy and T. Moreau, *"Algorithm Theoretical Basis Document (ATBD) of the CPP SAR numerical retracker for oceans)"*, S3A-NT-SRAL-00099-CNES, 1.0, 2013.

[Brown, 1977]: G. S. Brown, "*The average impulse response of a rough surface and its applications*", IEEE Transactions on antenna and propagation, 25(1), 67-74, 1977.

[Cotton, 2014]: D. Cotton, "*Mid Term Review – Executive Summary"*, CP4O-MTR-01, Ed.1, 2014.

[Davis, 1997]: Davis, C. H. (1997). A robust threshold retracking algorithm for measuring ice-sheet surface elevation change from satellite radar altimeters. IEEE Transactions on Geoscience and Remote Sensing, 5(4):974–979.

[Desjonquères et al., 2012]: J.D. Desjonquères, F. Boy and N. Picot, *"Altimeter SAR data over ocean – CNES processing strategy and continuity with LRM data"*, poster at the 2012 American Geophysical Union Meeting.

[Dinardo, S., B. Lucas, and J. Benveniste]: *SAR Altimetry in Coastal Zone: Performances, Limits, Perspectives.* San Diego: 5th Coastal Altimetry Workshop, 2011.

[Gommenginger et al., 2012]: C. Gommenginger, C. Martin-Puig , M. Srokosz, M. Caparrini, S. Dinardo, and B. Lucas, "*Detailed Processing Model of the Sentinel-3 SRAL SAR altimeter ocean waveform retracker*", SAMOSA3 WP2300 technical Note, ESRIN Contract No. 20698/07/I-LG "Development of SAR

Altimetry Mode Studies and Applications over Ocean, Coastal Zones and Inland Water", Version 2.1.0, 16 March 2012, 75 pages.

[Maulik Jain, Ole Baltazar Andersen, Lars Stenseng and Jørgen Dall, 2014] *"Sea Surface Height Determination In The Arctic Using Cryosat-2 SAR Data From Primary Peak Empirical Retrackers",* in press

[Kendall et Stuart, 1963]: Kendall M.G. and A. Stuart, *"The advanced Theory of Statistics, Volume 1: Distribution Theory"*, 2nd ed. London: Charles Griffin & Co., Ltd., 1963, sections 6.17-6.22.

[Labroue et Tran, 2007]: Labroue S. and N. Tran, *"Envisat RA2 and MWR Product and Algorithm and evolution studies – WP1200"*, Technical Report CLS-DOS-NT-07.133.

[Laxon, S. W. (1994)]: Sea ice extent mapping using ERS-1 radar altimeter. EARSeL Advances in Remote Sensing, 3(2):112–116.

[LOTUS D1.2, 2013]: "D1.2: *Scientific Requirements Consolidation*", CLS-DOS-NT-13-235, Oct. 2013.

[Raney, 1998]: R. K. Raney, "*The Delay/Doppler Radar Altimeter*", IEEE Transactions on geoscience and remote sensing, 36(5), 1578-1588, 1998.

[Raney, 2005]: R.K. Raney, "*Resolution and Precision of a Delay-Doppler Radar Altimeter".* Vol. 3. OCEANS, 2005. Proceedings of MTS/IEEE , 2005.

[Smith et al., 2012]: W. Smith and R. Scharroo, "*Pulse-to-pulse correlation in CryoSat SAR echoes from ocean surfaces: implications for optimal pseudo-LRM waveform averaging"*, presented at the 2012 Ocean Surface Topography Science Team Meeting. Available online: http://www.aviso.oceanobs.com/fileadmin/documents/OSTST/2012/oral/02_friday_28/05_instr_processing_IIb/03_IP2B_Smith1.pdf

[Stenseng, L. (2009)]: SAMOSA WP8 technical note. validation using airborne asiras data. Technical report, DTU-Space.

[Stenseng, L. (2014)]: L. Stenseng, *"D4.1 Algorithm Theoretical and Validation Document – Polar Ocean",* CP4O WP4000 Technical Note, ESA/ESRIN Contract No. 4000106169/12/I-NB "Cryosat Plus for Oceans", Version 1.0, 7 May 2014.

[Thibaut et al., 2010]: P. Thibaut, J. C. Poisson, E. Bronner, N. Picot *"*Relative Performance of the MLE3 and MLE4 Retracking Algorithms on Jason-2 Altimeter Waveforms*",* Marine Geodesy, 33: 1, 317 — 335

[Thibaut et al., 2012]: P. Thibaut, J.C.Poisson, F. Boy and N. Picot, *"*Numerical Solution for the Retracking Algorithm: Performances on Conventional Altimeter Waveforms*"*, presented at the 2012 OSTST Meeting, Venise.

[Thibaut et al., 2012]: P. Thibaut, T. Moreau, F. Boy and N. Picot, *"Coastal Altimetry : Evolution of measurement and retracking problems when switching from conventional (Ku, Ka) to SAR altimetry"*, presented at the 2012 Coastal Meeting.

[Wingham et al., 2006]: D. J. Wingham, C. R. Francis, S. Backer, C. Bouzinac, D. Brockley, R. Cullen, P. de Chateau-Thierry, S. W. Laxon, U. Mallow, C. Mavrocordatos, L. Phalippou, G. Ratier, L. Rey, F. Rostan, P. Viau, D. W. Wallis, "*CryoSat: A Mission to determine the Fluctuations in Earth's Land and Marine Ice Fields*", Advances in Space Research, Vol. 37, Jan. 2006, Issue 4, pp.841-871.

[END OF DOCUMENT]