

---

# Estimation of significant wave height using Sentinel-3 data



Studienarbeit im Studiengang  
**Geodäsie und Geoinformatik**  
an der Universität Stuttgart

Junyang Gou

Stuttgart, August 2020

---

**Supervisor:** Dr.-Ing. Mohammad Tourian



# **Erklärung der Urheberschaft**

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

Ort, Datum

Unterschrift



# Abstract

Coastal area is one of the most important area for us. More than 600 million people (around 10% of the word's population) live in coastal areas that are less than 10 m above sea level. Nearly 2.4 billion people (about 40% of the world's population) live within 100 km of the coast. Therefore, monitoring of coastal waters is extremely important. Due to the limitation of the number and location, the tide gauge stations around the world cannot provide a sufficient amount of in-situ data. Therefore, satellite altimetry plays an increasingly important role, especially when the SAR altimeter is put into use. However, due to the complexity of the coastal water surfaces, the performance of the satellite altimeter over the coastal area is far worse than over ocean.

This thesis is dedicated to developing a method to determine one of the essential characters of the water surfaces - the significant wave height (SWH), using the Sentinel-3 data in the coastal area. The three primary steps of the method are extracting the thermal noise and the leading edge, fitting this part of waveform and determining the relationship between the new retracker and the physical model.

In the first step, an algorithm is developed to avoid the interferences of the noise on the trailing edge. Therefore, the peak of the leading edge could be determined more accurately. The condition for the start point of the leading edge of the PLRM waveforms is  $D_{wf} > 0.01$ , inherited from ALES, whereas a more appropriate threshold for the SAR waveforms has been found as  $D_{wf} > 0.03$ . In the second step, the limitation of the Gauss-Markov model for the waveform adjustment has been discussed. Thus, the Levenberg-Marquardt method has been chosen to adjust the waveform. In the third step, the relationship between the raising time ( $\Delta\text{Bins}$ ) and the  $\sigma_c$  ( $\beta_4$ ) has been found. Then, we could estimate the SWH directly from  $\Delta\text{Bins}$  which makes it possible to estimate the SWH from some complicated waveforms in coastal areas.

We have employed the developed methodology to determine the significant wave height in the coastal area near the Cuxhaven. The quality of the results has been proved by comparing with the in-situ data from the Elbe measuring station provided by the Federal Maritime and Hydrographic Agency of Germany. The validation showed that the proposed method can determine reliable SWH from approximately 1 km offshore, which is an improvement of earlier results.

**Keywords:** Sentinel-3, Significant wave height (SWH), Waveform retracking, Satellite altimetry, Coastal altimetry



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Satellite radar altimetry . . . . .	1
1.2	Significant wave height . . . . .	2
1.3	Satellite altimetry missions . . . . .	3
1.3.1	Sentinel-3 mission . . . . .	3
1.3.2	Other satellite altimetry missions . . . . .	5
<b>2</b>	<b>Satellite altimetry waveform</b>	<b>7</b>
2.1	Construction of a waveform . . . . .	7
2.2	The Brown-Hayne Theoretical Ocean Model . . . . .	9
2.3	SAR and PLRM . . . . .	11
2.3.1	Basic difference between SAR and LRM . . . . .	11
2.3.2	SAR and PLRM of Sentinel-3 . . . . .	12
2.4	Waveform types . . . . .	14
2.5	Empirical retrackers . . . . .	16
2.5.1	Offset Centre of Gravity retracker (OCOG) . . . . .	16
2.5.2	The $\beta$ -parameter retracker . . . . .	17
<b>3</b>	<b>Case study</b>	<b>21</b>
3.1	Choice of the virtual stations . . . . .	21
3.2	Relative distance . . . . .	22
3.3	SWH provided by Sentinel-3 . . . . .	22
<b>4</b>	<b>Methodology</b>	<b>25</b>
4.1	Basic idea . . . . .	25
4.2	Adjustment using Gauss-Markov model . . . . .	25
4.2.1	Principles . . . . .	25
4.2.2	Applying the 5- $\beta$ retracker using Gauss-Markov model . . . . .	28
4.2.3	Limitation of the 5- $\beta$ retracker . . . . .	29
4.2.4	Limitation of the Gauss-Markov model in the waveform adjustment . . . . .	29
4.3	Levenberg-Marquardt Method . . . . .	33
4.3.1	Principles . . . . .	33
4.3.2	Applying the 5- $\beta$ retracker using Levenberg-Marquardt method . . . . .	34
4.4	The 4- $\beta$ retracker . . . . .	37
4.4.1	Extracting the thermal noise and the leading edge . . . . .	37
4.4.2	Model of the 4- $\beta$ retracker . . . . .	37
4.4.3	Applying the 4- $\beta$ retracker using the Levenberg-Marquardt method . . . . .	41
4.5	Determining the begin and the end of the leading edge . . . . .	41
4.6	Quality of the 4- $\beta$ retracker . . . . .	46
4.7	Choosing the condition for the starting point of SAR waveforms . . . . .	47

4.8	Relationship between $\beta_4$ and $\Delta\text{Bins}$	48
<b>5</b>	<b>Results</b>	<b>51</b>
5.1	Oceanic area	51
5.1.1	Map and time series of the SWH	51
5.1.2	Discussion about the quality of the results	52
5.2	Coastal area	55
5.2.1	In-situ data	55
5.2.2	Map and time series of the SWH	56
5.2.3	Discussion about the quality of the results	57
<b>6</b>	<b>Conclusion and outlook</b>	<b>59</b>
6.1	Summary and Conclusion	59
6.2	Outlook	60

# List of Figures

1.1	The principle of altimetry [AVISO+] . . . . .	2
1.2	Statistical distribution of ocean wave heights [NOAA UCAR COMET Program] . . . . .	3
1.3	Ground track of Sentinel-3A (red) and Sentinel-3B (green) in Germany . . . . .	4
1.4	Past, active and future satellite altimetry missions [OpenADB, TUM] . . . . .	5
2.1	The process of constructing the returned waveform in an ideal situation [Tourian (2012)] . . . . .	7
2.2	(a) An ideal waveform from an ocean surface. (b) Actual waveform, showing wraparound and Rayleigh noise. [Quarly et al. (2001)] . . . . .	8
2.3	Theoretical Brown ocean waveform shape and corresponding retried ocean parameters [Vignudelli et al. (2011)] . . . . .	10
2.4	Pulse Transmission Scheme in case of Pulse-Limited Altimetry (LRM) and closed-burst SAR Altimetry [Dinardo (2020)] . . . . .	12
2.5	Evolution of the footprint's geometry with the time in case of pulse-limited altimetry (red) and in case of SAR altimetry (green) [Rosmorduc et al. (2011)] . . . . .	12
2.6	PLRM radar cycle transmitting pattern (left) and SAR radar cycle transmitting pattern (right) [ESA Sentinel Online] . . . . .	13
2.7	Prevalence of various waveform shapes for Jason-2 over one full cycle as a function of distance nearest coastline [Vignudelli et al. (2011)] . . . . .	14
2.8	Waveform of SAR (6993.18 m) and PLRM (6950.15 m) from the Virtual Station on Cuxhaven (53.8730, 8.7107), 2018-12-12 . . . . .	15
2.9	Schematic diagram of the OCOG method [Tourian (2012)] . . . . .	16
2.10	Schematic diagram of 5- $\beta$ Retracker with a linear trailing edge [Tourian (2012)] . . . . .	18
3.1	The coastal virtual station (blue star) at Cuxhaven (53.8730, 8,7107) and the oceanic virtual station (orange star) in the North Sea (57.3623, 6.7602) on the Sentinel-3B orbit (green line) . . . . .	21
3.2	The virtual station (blue star) at Cuxhaven (53.8730, 8,7107) on the Sentinel-3B orbit (green line) and the SAR data (orange points) in the searching circle with R = 10 km between 2018-12-12 and 2019-12-25 . . . . .	21
3.3	On-orbital Virtual station (red point) and relative distance . . . . .	22
3.4	Oceanic SWH provided by Sentinel-3 on 2018-12-12 . . . . .	23
3.5	Coastal SWH provided by Sentinel-3 on 2018-12-12, gray area is land . . . . .	23
3.6	Radargram of the normalized waveforms in the oceanic area on 2018-12-12 . . . . .	24
3.7	Radargram of the normalized waveforms in the coastal area on 2018-12-12 . . . . .	24
4.1	Iterative scheme of the Gauss-Markov model [Sneeuw et al. (2008)] . . . . .	27
4.2	Examples of 5- $\beta$ fitted oceanic PLRM waveform on 2018-12-12 (Gauss-Markov) . . . . .	30
4.3	Examples of 5- $\beta$ fitted oceanic SAR waveform on 2018-12-12 (Gauss-Markov) . . . . .	31
4.4	An unsuccessful case of Gauss-Markov model (Num. 36 on 2018-12-12 in the oceanic area, PLRM) . . . . .	32

---

4.5	A successful case of Gauss-Markov model (Num. 34 on 2018-12-12 in the oceanic area, PLRM) . . . . .	32
4.6	Examples of $5\beta$ fitted oceanic PLRM waveform on 2018-12-12 (Levenberg-Marquardt) . . . . .	35
4.7	Examples of $5\beta$ fitted oceanic SAR waveform on 2018-12-12 (Levenberg-Marquardt) . . . . .	36
4.8	An example of the algorithm. Num.1 oceanic PLRM waveform on 2018-12-12 . . . . .	38
4.9	Flowchart of the algorithm to extract the thermal noise and the leading edge . . . . .	39
4.10	Flowchart of the algorithm to eliminate the noise . . . . .	40
4.11	Examples of $5\beta$ and $4\beta$ fitted oceanic PLRM waveform on 2018-12-12 (Levenberg-Marquardt) . . . . .	42
4.12	Examples of $5\beta$ and $4\beta$ fitted oceanic SAR waveform on 2018-12-12 (Levenberg-Marquardt) . . . . .	43
4.13	Examples of $4\beta$ fitted coastal SAR waveform on 2018-12-12 (Levenberg-Marquardt), a) Sharp peak, b) Noise on the TE, c) Noise on LE and TE, d) Multi-peak . . . . .	44
4.14	Examples of the leading edge. a) Oceanic PLRM, b) Oceanic SAR, c) Coastal PLRM, d) Coastal SAR . . . . .	45
4.15	The scatter of the oceanic PLRM $\Delta$ Bins using the $4\beta$ and $5\beta$ retracker (left) and the histogram of $\Delta$ Bins (right) . . . . .	46
4.16	The expected values and the standard deviations of the $\Delta$ Bins . . . . .	47
4.17	The scatter of the oceanic PLRM and SAR $\Delta$ Bins using the $4\beta$ retracker (left) and the histogram of $\Delta$ Bins (right) . . . . .	48
4.18	An example of the non-physical model . . . . .	48
4.19	The relationship between $\Delta$ Bins and $\beta_4$ with two fitting lines . . . . .	49
5.1	Map of the significant wave heights in the oceanic area between 2018-12-12 and 2019-12-25. Left: SAR, Right: PLRM . . . . .	51
5.2	The median values and the standard deviations of the significant wave heights in the oceanic area between 2018-12-12 and 2019-12-25 . . . . .	52
5.3	The scatters of the SWH and the median values in the searching windows (oceanic area) . . . . .	53
5.4	The scatters of the SWH and the standard deviations of the $4\beta$ retracked waveforms (oceanic area) . . . . .	54
5.5	The measuring station Elbe (green star), the coastal virtual station (blue star) and the oceanic virtual station (orange star) . . . . .	55
5.6	Map of the significant wave heights in the coastal area between 2018-12-12 and 2019-12-25. Left: SAR, Right: PLRM . . . . .	56
5.7	The median values and the standard deviations of the significant wave heights in the coastal area between 2018-12-12 and 2019-12-25 . . . . .	57
5.8	The scatters of the SWH and the median values in the searching windows (coastal area) . . . . .	58
5.9	The scatters of the SWH and the standard deviations of the $4\beta$ retracked waveforms (coastal area) . . . . .	58

## List of Tables

1.1	Orbital information for Sentinel-3 [ESA Sentinel Online]	4
4.1	The results of different Dwf	47



# Chapter 1

## Introduction

Coastal area is one of the most important area for us. More than 600 million people (around 10% of the word's population) live in coastal areas that are less than 10 m above sea level. Nearly 2.4 billion people (about 40% of the world's population) live within 100 km of the coast. Therefore, monitoring of coastal waters is extremely important. Due to the limitation of the number and location, the tide gauge stations around the world cannot provide a sufficient amount of in-situ data. Therefore, satellite altimetry plays an increasingly important role, especially when the SAR altimeter is put into use.

Altimetry is a technique for measuring height. Satellite altimetry measures the time taken by a radar pulse to travel from the satellite antenna to the surface and back to the satellite receiver (two-way travel time). Combined with precise satellite location data, altimetry measurements yield sea-surface heights [Rosmorduc et al. (2011)].

Compared with conventional in-situ measurement, satellite altimetry has the advantages of high coverage. Nowadays, satellites cover almost every corner of the Earth. Therefore, it can provide data of the areas, which are difficult for in-situ measurements. Nevertheless, at the same time, the satellite altimetry is reaching the average accuracy of decimetre, which is far inferior to in-situ measurements with centimetre to millimetre. Moreover, satellite altimetry also has certain limitations. For example, reliable results could be generated over the spatially homogeneous surfaces like the ocean. For surfaces which are not homogeneous i.e. contain many disturbances, coastal areas, accurate analysis of the water surface becomes much more difficult. As mentioned in the abstract, the monitoring of coastal waters are very important to humans. Therefore, it is critical to improve the performance of satellite altimetry in coastal areas.

In this chapter, principles of satellite altimetry will be introduced.

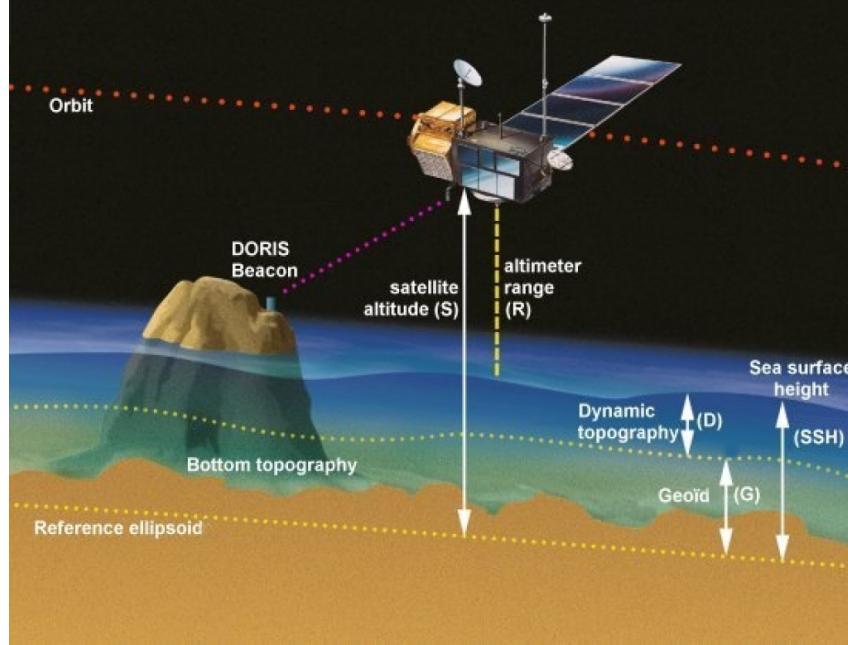
### 1.1 Satellite radar altimetry

The basic concept of satellite altimetry is deceptively straightforward. The principal objective is to measure the range  $R$  from the satellite to the sea surface (see Fig. 1.1). The altimeter transmits a short pulse of microwave radiation with known power toward the sea surface. The pulse interacts with the rough sea surface, and part of the incident radiation reflects back to the

altimeter. The range  $R$  from the satellite to mean sea level is estimated from the two-way travel time  $t$  by Eq. 1.1

$$R = \hat{R} - \sum_j \Delta R_j \quad (1.1)$$

where  $\hat{R} = \frac{ct}{2}$  is the range computed neglecting refraction based on the speed of light in vacuum and  $\Delta R_j$  are corrections for the various components of atmospheric refraction and for biases between the mean electromagnetic scattering surface and mean sea level at the air-sea interface [Fu and Cazenave (2001)].



*Figure 1.1: The principle of altimetry [AVISO+]*

The range estimate must be transformed to a fixed coordinate system and combined with the satellite altitude  $S$  relative to the same specified reference ellipsoid approximation of the geoid. Then, the sea surface height (SSH) can be calculated with Eq. 1.2

$$\begin{aligned} SSH &= S - R \\ &= S - \hat{R} + \sum_j \Delta R_j \end{aligned} \quad (1.2)$$

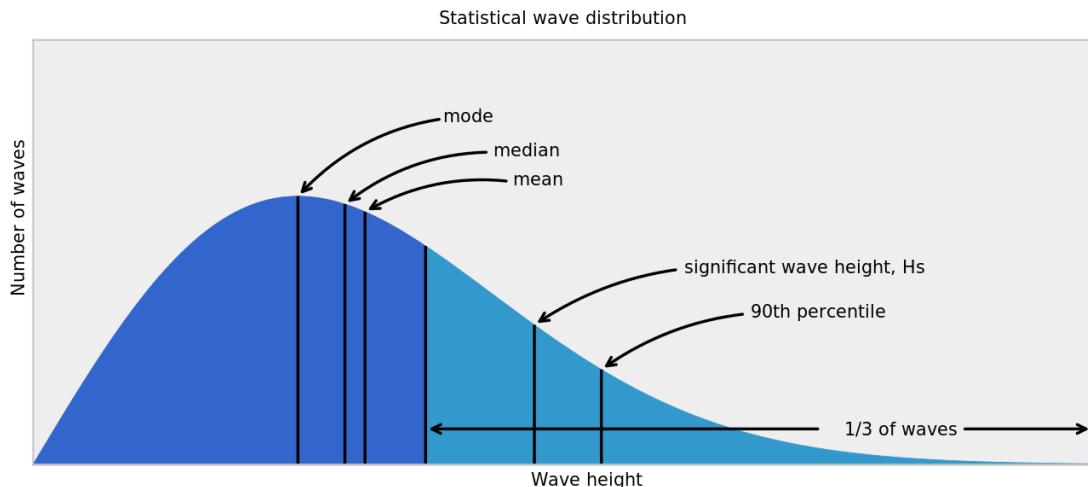
In addition to this, the magnitude and shape of the echoes (or waveforms) will also be generated. The waveforms contain massive important information about the characteristics of the surface, which caused the reflection. The details will be introduced in Chapter 2.

## 1.2 Significant wave height

Significant wave height (SWH or  $H_s$ ) is one of the most important parameter for the statistical distribution of ocean waves. The most common waves are lower than SWH. However,

statistically, it is possible to encounter a wave that is much higher than the significant wave (See Fig. 1.2).

Significant wave height is defined traditionally as the crest-to-trough height, which is third of the highest waves in the field of view and therefore denoted also as  $H_{1/3}$ . There is also a relationship between the SWH and the statistical characterization of the wave field in terms of the standard deviation  $\sigma_\zeta$  of the sea surface elevation  $\zeta$ . The sea surface elevation is equal to half of the wave height. For a wide range of bandwidths,  $SWH = 4\sigma_\zeta$  is a reasonable good approximation [Fu and Cazenave (2001)].



*Figure 1.2: Statistical distribution of ocean wave heights [NOAA UCAR COMET Program]*

## 1.3 Satellite altimetry missions

### 1.3.1 Sentinel-3 mission

Sentinel-3 is a dedicated Copernicus satellite delivering high-quality ocean measurements. In the marine environment, the primary objective of Sentinel-3 is to determine sea-surface topography, sea-surface temperature and ocean-surface colour parameters, offer EO data with global coverage every two days (with two satellites) in support of marine applications, and with near real-time products delivered in less than three hours.

Sentinel-3 includes currently two satellite: Sentinel-3A (launched in February 2016) and Sentinel-3B (launched in April 2018). In the longer term, the Sentinel-3 mission will have further satellites (Sentinel-3C and Sentinel-3D), extend this global monitoring. The orbit of Sentinel-3 is a near-polar, sun-synchronous repeat orbit with the ratio 385/27. The Table 1.1 contains a summary of useful orbital information for Sentinel-3. The Figure 1.3 shows the ground track of Sentinel-3 in Germany.

Altitude	Inclination	Period	$\beta/\alpha$	Ground-track deviation	LT at Des. Node
814.5 km	98.65 deg	100.99 min	385/27	$\pm 1$ km	10:00 hours

Table 1.1: Orbital information for Sentinel-3 [ESA Sentinel Online]

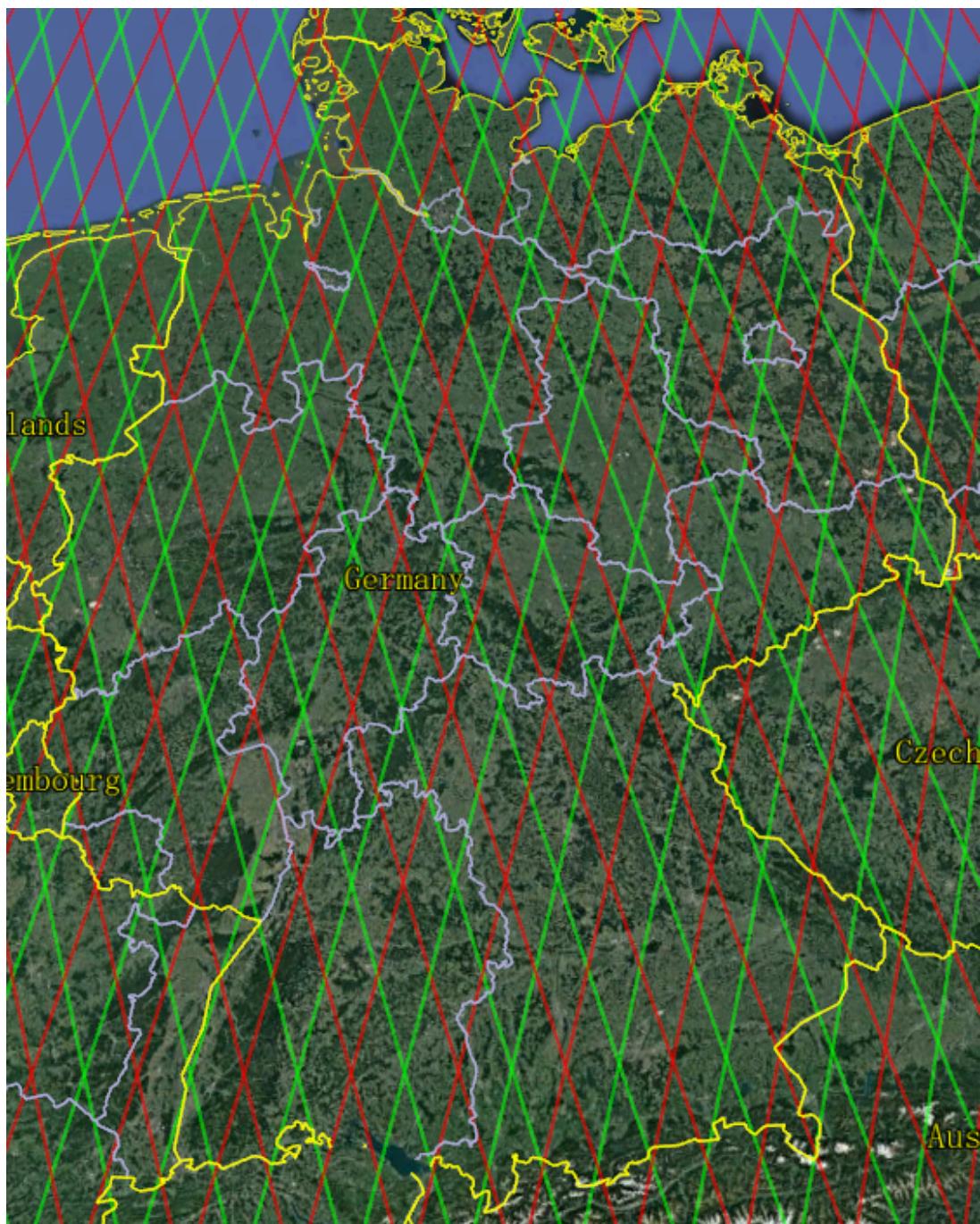
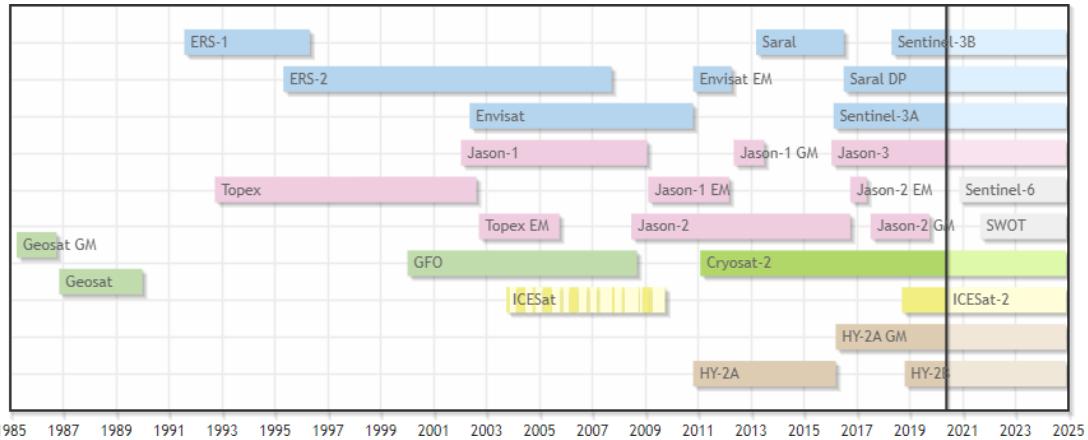


Figure 1.3: Ground track of Sentinel-3A (red) and Sentinel-3B (green) in Germany

The highlight of Sentinel-3 is that they carry a SARL altimeter, which has two possible radar measurement modes: LRM and SAR (operating 100% in SAR on orbit). The Sentinel-3 SRAL improved along-track resolution (approximately 300 m) in SAR mode facilitates sea surface height measurement close to the coast [EUMETSAT (2017)]. The details and the main differences between LRM and SAR will be introduced in Chapter 2.

### 1.3.2 Other satellite altimetry missions



*Figure 1.4: Past, active and future satellite altimetry missions [OpenADB, TUM]*

In addition to Sentinel-3, there are currently many other satellite altimetry missions in operation around the world, as shown in Fig. 1.4. Different satellite missions have different characteristics. Sentinel-3 and CryoSat-2 are currently the only satellite missions equipped with SAR altimeters. Thanks to the help of the delay-Doppler concept, the SAR altimeter dramatically improves the accuracy of conventional satellite altimetry, while significantly reducing noise. These missions will mainly benefit the coastal zone. Nonetheless, these two satellite altimetry missions have many differences.

The main limitation of conventional nadir-pointing radar altimeters is the space-time coverage dilemma. The orbital data of Sentinel-3 is described in the Table 1.1. It provides repeat orbit to achieve better temporal resolution, but the spatial resolution is lost. In contrast, CryoSat-2 provides higher spatial resolution but not very good temporal resolution (repeat cycle for 369 days, 30 days sub-cycle).

In addition to the above two satellite missions, there are a series of satellite altimetry missions such as Jason-3, Saral and HY-2. In the future, these missions will provide us with higher precision data to develop satellite altimetry, especially over the coastal zone and inland water surfaces.



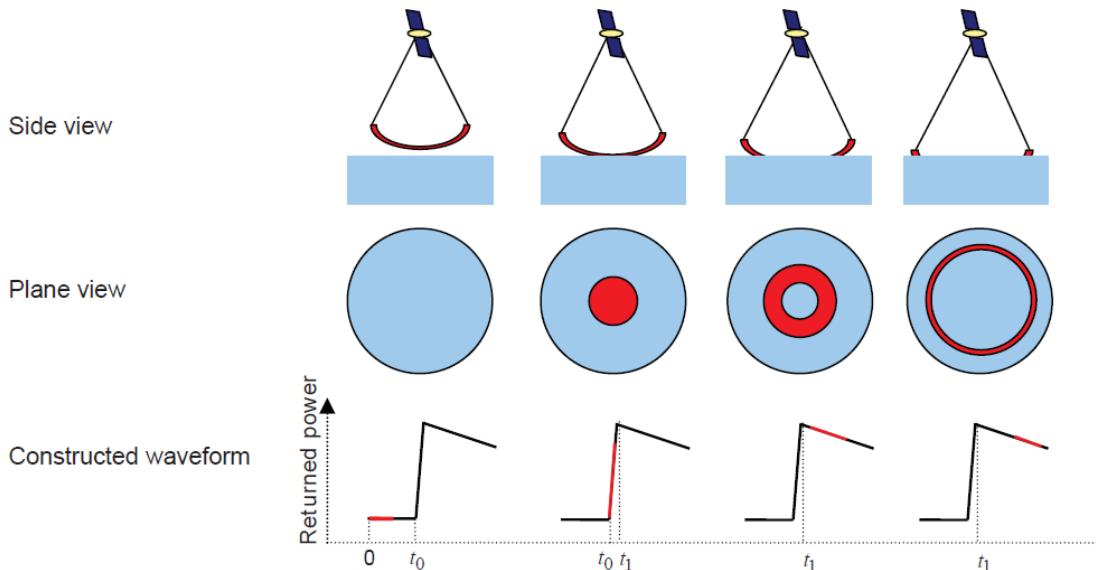
# Chapter 2

## Satellite altimetry waveform

### 2.1 Construction of a waveform

The conventional method of waveform construction consists of transmitting radar pulses through a pulse-limited radar altimeter which is reflected by the water surface and then received by the altimeter again. The waveform contains various information, such as range to the at-nadir surface, reflective backscatter the roughness of the water surface.

The Fig. 2.1 shows the interaction of the radar pulses emitted by the pulse-limited radar altimeter with an ideal assumed water surface without any roughness, which corresponds to constructing a waveform. The process could be separated into three main parts by two significant time  $t_0$  and  $t_1$ .



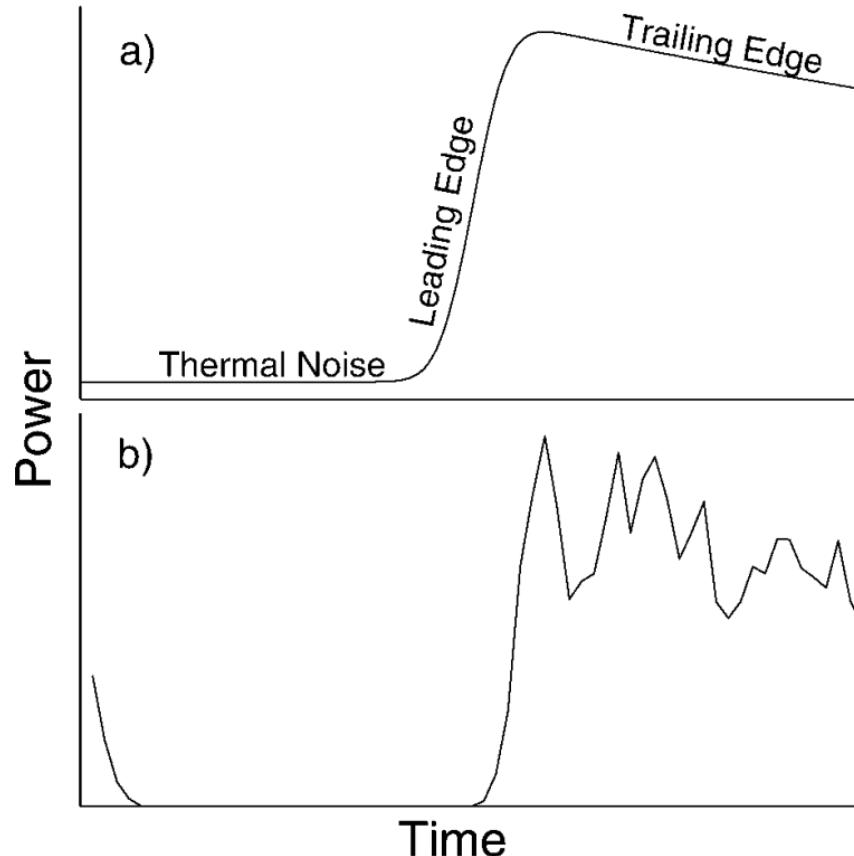
*Figure 2.1: The process of constructing the returned waveform in an ideal situation [Tourian (2012)]*

Before  $t_0$ , the on-satellite altimeter emits an electromagnetic pulse but still receives nothing. All the returned powers are thermal noise. Then, the spherical wavefront reaches the water surface directly under the satellite and is reflected by the water surface. The altimeter receives the first reflected signal at  $t_0$ , which leads to the rise of the returned power (Leading edge).

Between  $t_0$  and  $t_1$ , the interactive area between the wavefront and the water surface continues to increase, causing the returned power to continue rising, after which the spherical wavefront expands into a disc. The maximum returned power occurs at  $t_1$ , which is the time of transition to an annular ring.

After  $t_1$ , the returned power starts to decline because of the limitation of antenna beamwidth and fewer proper reflected facets (Trailing edge). In this part, the interactive area forms an annular ring with an increasing diameter and narrowing width. In general, the slope of the trailing edge is much smaller than the leading edge.

The constructed waveform contains noises. The altimeter reduces the proportion of noise by measuring multiple waveforms and averaging the returned power [Quarly et al. (2001)]. The constructed waveform is a time series of mean returned power, which contains three main parts [Brown (1977)]:



**Figure 2.2:** (a) An ideal waveform from an ocean surface. (b) Actual waveform, showing wraparound and Rayleigh noise. [Quarly et al. (2001)]

- **Terminal noise:** Before  $t_0$ , all the returned power is regarded as thermal noise. Thermal noise is often regarded as a constant in an ideal model. The power level of thermal noise is microscopic compared to the waveform (< 3% in PLRM mode, < 0.5% in SAR mode).
- **Leading edge:** The leading edge is the most crucial part of the waveform analysis in satellite altimetry. It contains most of the information we need. For example, the significant

wave height (SWH) and the range between the satellite altimeter and the mean sea surface at the nadir ( $R$ ) can be extracted from the leading edge [Tourian (2012)]. The specific details will be in the next sections elaborated.

- **Trailing edge:** As the returned power from the scattering surface is decaying, the trailing edge of waveform is constructed. It can be approximated by a straight line whose slope depends on the altimeter antenna pattern and the off-nadir angle [Fu and Cazenave (2001)].

## 2.2 The Brown-Hayne Theoretical Ocean Model

Starting from microwave scattering theory, the average return power as a function of time delay ( $t$ ) could be expressed as a convolution of three terms:

$$W(t) = FSSR(t) * PTR(t) * PDF(t) \quad (2.1)$$

where  $FSSR$  is the flat sea surface response,  $PTR$  is the radar point target response, and  $PDF$  is the ocean surface elevation probability density function of the specular point [Vignudelli et al. (2011)]. The  $PTR$  function is a  $(\frac{\sin x}{x})^2$  function which is usually approximated by a Gaussian function in order to perform the convolution of the three terms. Hence:

$$PTR(t) \approx \exp\left(\frac{-t^2}{2\sigma_p^2}\right) \quad (2.2)$$

where  $\sigma_p$  is the width of the radar point target response function. Barrick (1972) and Brown (1977) used:

$$\sigma_p = \frac{1}{2\sqrt{2\ln 2}}r_t \approx 0.425r_t \quad (2.3)$$

with  $r_t$  the time resolution. For example, in Sentinel-3:  $r_t = 3.125$  ns. The more common approximation of  $\sigma_p$  today is  $\sigma_p \approx 0.513r_t$  [Thibaut et al. (2004)].

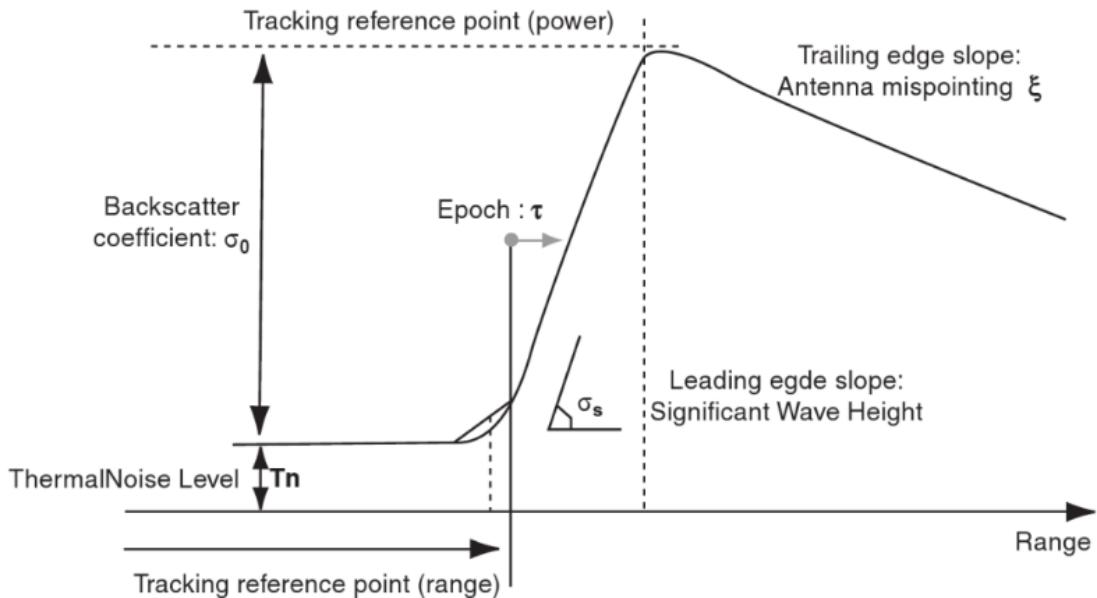
The formulation of the theoretical shape of returned waveform over the ocean surface is Brown (1977):

$$V_m = T_n + a_\xi P_u \frac{1 + \operatorname{erf}(u)}{2} \exp(-\nu) \quad (2.4)$$

where

$$\begin{aligned}
 a_\xi &= \exp\left(\frac{-4 \sin^2 \xi}{\gamma}\right) \\
 \gamma &= \sin^2(\theta_0) \frac{1}{2 \ln 2} \\
 \operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \\
 u &= \frac{t - \tau - c_\xi \sigma_c^2}{\sqrt{2} \sigma_c} \\
 v &= c_\xi \left(t - \tau - \frac{1}{2} c_\xi \sigma_c^2\right) \\
 c_\xi &= b_\xi a \\
 b_\xi &= \cos(2\xi) - \frac{\sin^2(2\xi)}{\gamma} \\
 a &= \frac{4c}{\gamma h \left(1 + \frac{h}{R_E}\right)} \\
 \sigma_c^2 &= \sigma_p^2 + \sigma_s^2 \\
 \sigma_s &= \frac{SWH}{2c}
 \end{aligned}$$

Fig. 2.3 shows the relationship between the waveform and the parameters implied in the waveform. The parameters mentioned in the Eq. 2.4 are described in detail.



**Figure 2.3:** Theoretical Brown ocean waveform shape and corresponding retrieved ocean parameters [Vignudelli et al. (2011)]

- $\xi$ : The off-nadir mispointing angle.
- $\theta_0$ : The antenna beam width. (Sentinel-3:  $\theta_0 = 1.28^\circ$  [ESA Sentinel Online])

- $\tau$ : The epoch or time delay i.e. the position of the waveform in the analysis window, with respect to the nominal tracking reference point. (Sentinel-3: The tracking point is gate 44 i.e. bin index 43 for Ku-band, gate 46 i.e. bin index 45 for C-band [ESA Sentinel Online]).
- $\sigma_s$ : Shows the slope of the leading edge, which relates to the significant wave height ( $\sigma_s = \frac{SWH}{2c}$  ).
- $P_u$ : The amplitude of the signal, which relates to the backscatter coefficient  $\sigma_0$ .
- $T_n$ : The thermal noise level.
- Non-waveform parameters:  $c$  is the speed of light.  $h$  is the satellite altitude and  $R_e$  is the Earth radius.

The Brown model is the basic model of satellite altimetry over the ocean and is also often known as physically-based retracker. The Brown model fits the ocean waveform very well, but in some complex cases such as coastal area or inland rivers, the Brown model is no longer applicable due to the significantly increased interferences. Nevertheless, there are some methods which are developed based on the Brown model that show acceptable performance in the coastal area, such as ALES [Passaro et al. (2014)].

Besides the physically-based retrackers, there are also many empirical retrackers, which are much simpler than the physically-based retrackers but also perform very well under some specific conditions. These empirical retrackers will be introduced in Section 2.5.

## 2.3 SAR and PLRM

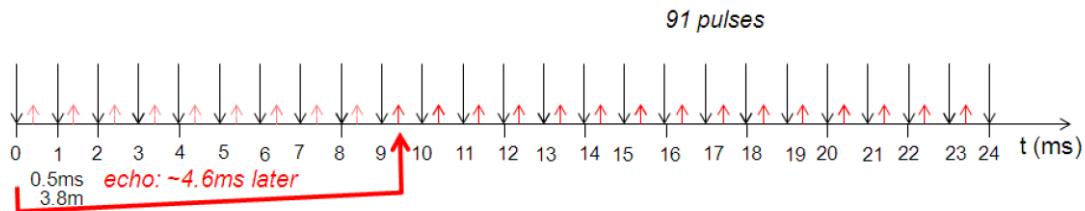
The waveform construction method described above is based on the conventional pulse-limited altimeter, which is commonly known as LRM (Low Rate Mode) and as PLRM (Pseudo-Low Rate Mode) in the Sentinel-3 mission. Besides, the Sentinel-3 mission also provides SAR (Synthetic Aperture Radar) mode, which significantly improves the along-track resolution. The main difference between these two measurement modes is related to the frequency used to transmit the pulses, which is called the Pulse Repetition Frequency (PRF).

### 2.3.1 Basic difference between SAR and LRM

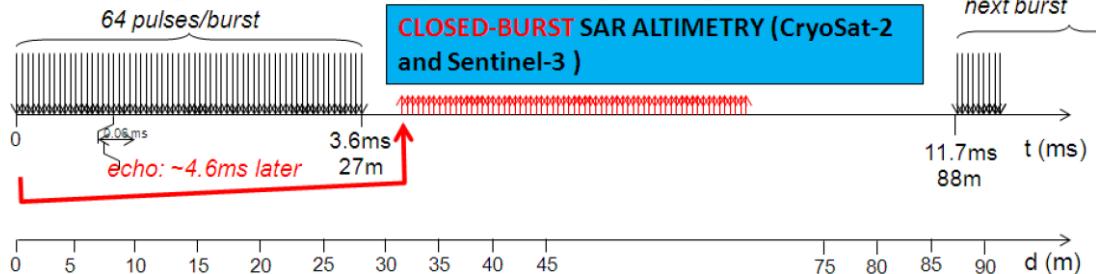
In pulse-limited mode (LRM), transmitted and received pulses are interleaved, i.e. pulses are received and transmitted continuously and reflections from the transmitted pulses are processed incoherently on a pulse-by-pulse basis [Fu and Cazenave (2001)].

In SAR mode, the pulses are transmitted and received in bursts with much higher PRF, so that successive received pulses in a burst are correlated. After the transmission of the burst, the altimeter exploits the empty inter-burst interval to receive the reflected pulses from the surface, see Fig. 2.4. The pulse-to-pulse coherence due to this high PRF allows the application of the delay-Doppler concept [Raney (1998)]. As result, a SAR altimeter will have a finer along-track resolution than a pulse-limited altimeter but, since the sharpening is just in the along-track direction, the SAR altimeter and the pulse-limited altimeter share the same across-track resolution [Dinardo (2020)].

**Classic Altimetry (LRM): PRF @ 1970Hz, Posting Rate @ 21Hz (350 m)**

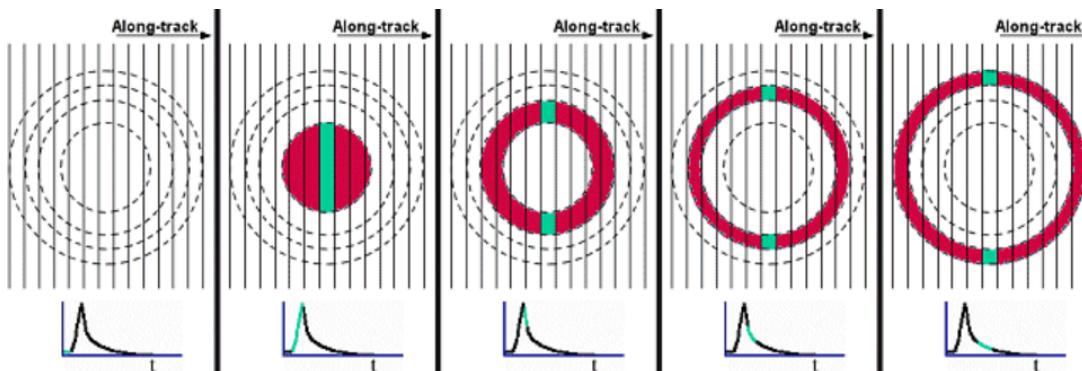


**SAR: PRF @ 17800Hz, bursts @ 84Hz, Posting Rate @ 21Hz (350 m)**



**Figure 2.4:** Pulse Transmission Scheme in case of Pulse-Limited Altimetry (LRM) and closed-burst SAR Altimetry [Dinardo (2020)]

Since the along-track resolution is developed, the shape of the constructed waveform is also changed, see Fig. 2.5. Compared with Fig. 2.1, the most impressing difference is that the descending of the tailing edge in SAR mode is much faster than in LRM mode. Because the SAR footprint is no longer constant with time delay, in a certain sense, a SAR altimeter footprint can be said to be beam-limited in along-track direction (and here the beam is the synthetic Doppler beam) and is pulse-limited in across-track direction.



**Figure 2.5:** Evolution of the footprint's geometry with the time in case of pulse-limited altimetry (red) and in case of SAR altimetry (green) [Rosmorduc et al. (2011)]

### 2.3.2 SAR and PLRM of Sentinel-3

In the Sentinel-3 mission, the pulse-limited mode is known as PLRM, not LRM. Because it is not real pulse-limited altimeter, but using the algorithm to compute both LRM-like Ku and C

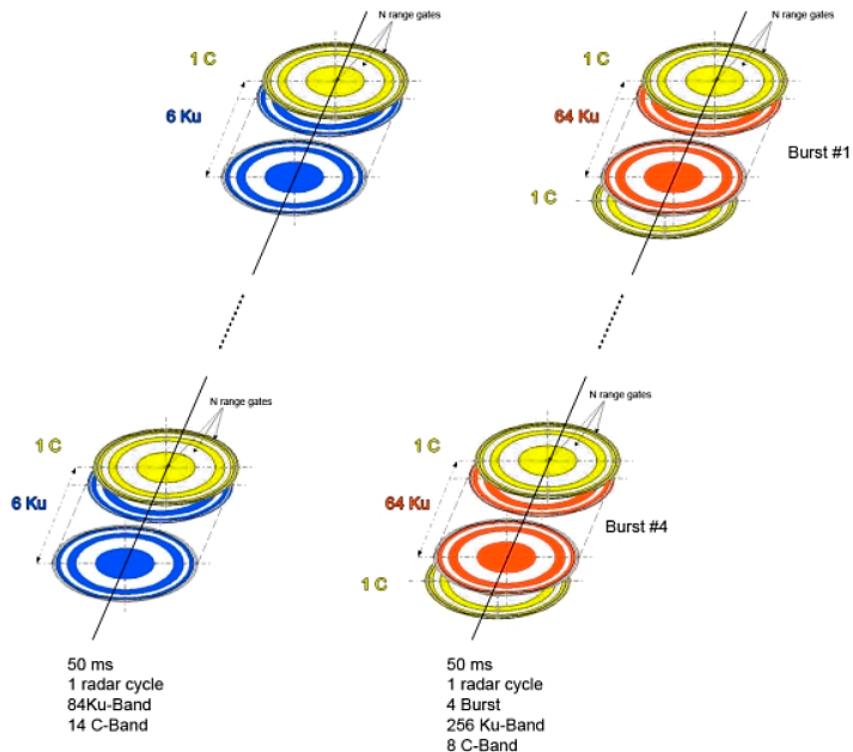
band power waveforms from SAR pulses, taking into account the onboard tracker to align the waveforms before power summation [EUMETSAT (2017)].

As mentioned above, the Sentinel-3 mission is using two frequencies:

- Ku Band (13.6 GHz): The main measurement band because it is the best compromise between the capabilities of the technology (relating to power emitted), the available bandwidth (determined by international regulations for specific applications), sensitivity to atmospheric perturbations, and perturbation by ionospheric electrons.
- C Band (5.3 GHz): The auxiliary band because it is more sensitive than Ku to ionospheric perturbation, and less sensitive to the effects of atmospheric liquid water. Its main function is to enable correction of the ionospheric delay in combination with the Ku-band measurements. To obtain the best results, an auxiliary band like this must also be as far as possible from the main one.

In PLRM, SRAL operates as a conventional pulse-limited altimeter with regular transmitting and receiving sequences, at a PRF of 1920 Hz. Patterns of six Ku-band pulses preceded by one C-band pulse. C and Ku-band echoes are accumulated separately over a 50 ms cycle of the radar cycle (i.e. 84 Ku-band pulses and 14 C-band pulses accumulated over that cycle).

In SAR Mode, 64 coherent Ku-band pulses are emitted in a burst (PRF of 18 kHz) surrounded by two C-band pulses. The burst cycle duration is approximately 12.5 ms so that a four-burst cycle is equal to the PLRM cycle of 50 ms.



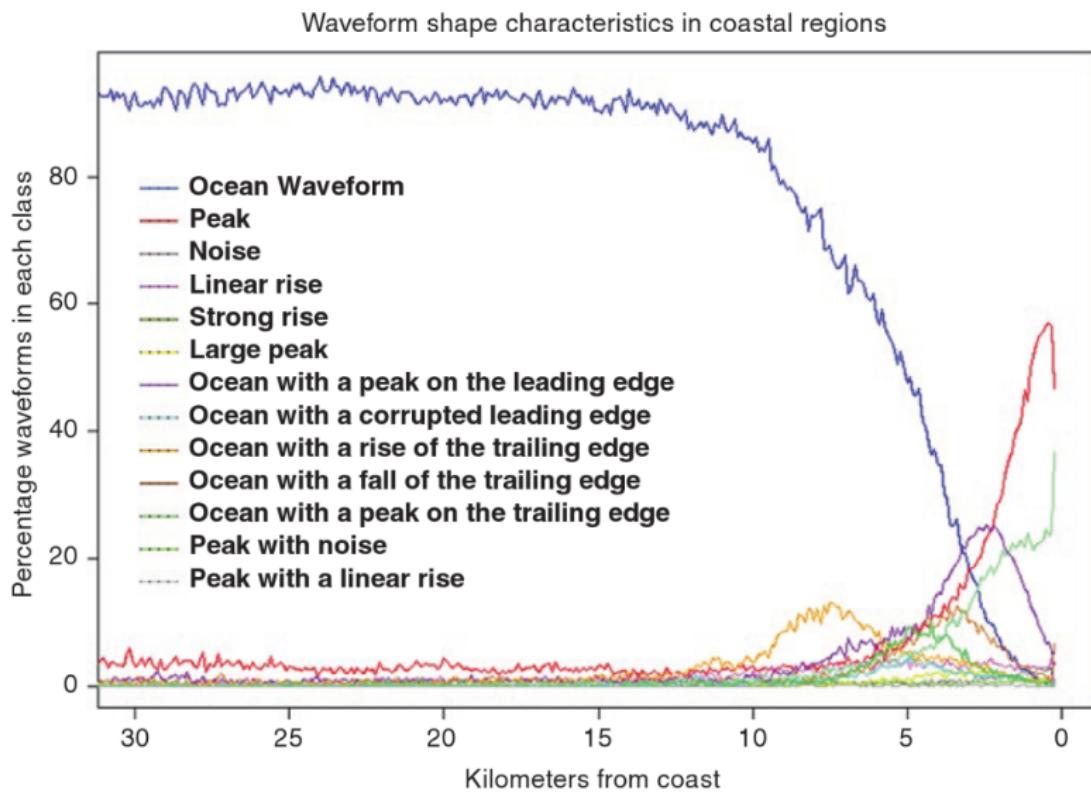
**Figure 2.6:** PLRM radar cycle transmitting pattern (left) and SAR radar cycle transmitting pattern (right) [ESA Sentinel Online]

## 2.4 Waveform types

The waveforms described above are all based on the assumption of water surface without any roughness. In the coastal area, there are various non-Brown waveforms. The most significant differences are the excess noise and the trailing edge that declines too fast.

How the waveform is affected depends on a weighted average of the surface area by the scattering coefficient of each surface [Vignudelli et al. (2011)]. If the land has a much less scattering coefficient than the ocean (typical case), the effect of land will be small, and the waveforms will remain unaffected until close to the coast. This situation often leads to a rapid decline of the trailing edge, because the direction of the nadir is still the water surface, but the reflection surface changes later from the water surface to the land, causing the returned power to drop or even turn to zero. In some environments (e.g. coral or atolls), however, land can be more highly reflective than the ocean, so that even a small area of highly reflective land in the footprint can have a significant impact on waveforms. This case leads to more noise.

Fig. 2.7 shows the prevalence of various waveform shapes for Jason-2 as a function of distance to the nearest coastline. The case of Sentinel-3 should be similar. Within 5 km from the coastline, Peak, Peak with noise, Ocean with a peak on the leading edge and Ocean are the most common waveform shapes.



**Figure 2.7:** Prevalence of various waveform shapes for Jason-2 over one full cycle as a function of distance nearest coastline [Vignudelli et al. (2011)]

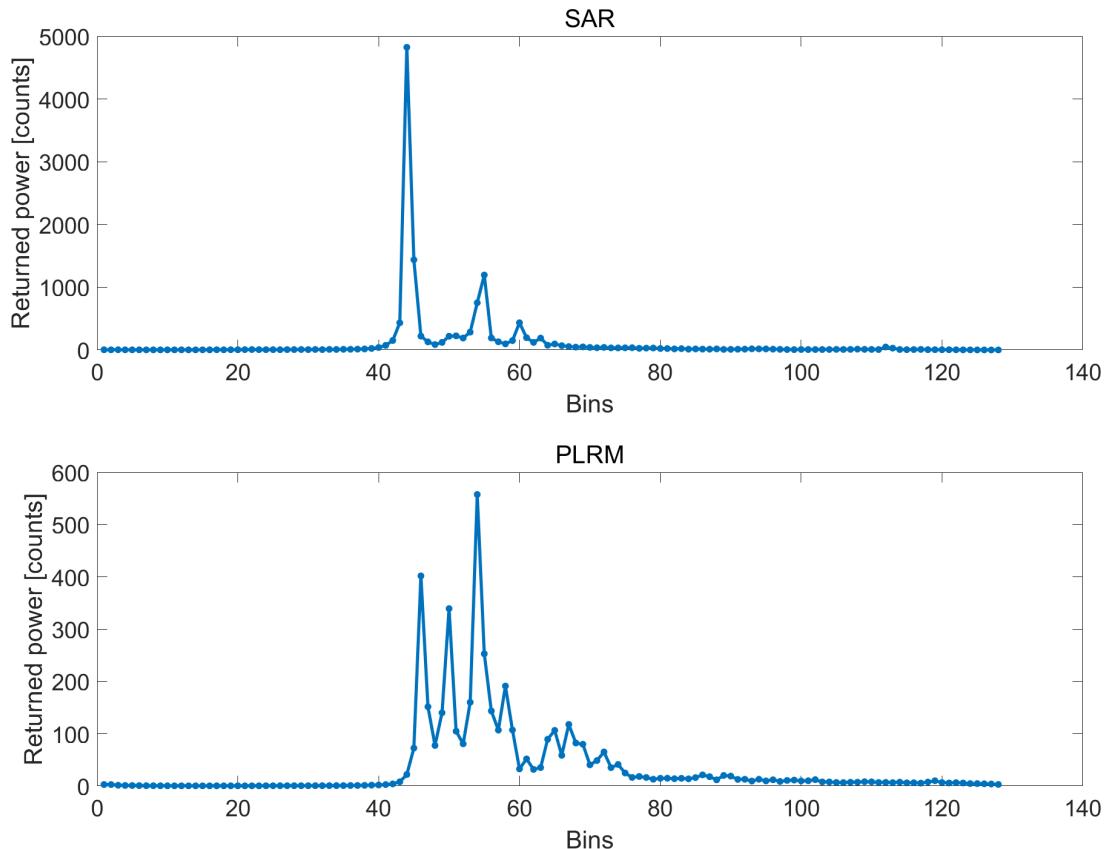
However, the actual situation is often not as clear as the theoretical classification. The

actual coastal waveform could be more complicated than a single class because various influencing factors could appear at the same time.

As shown in the Fig. 2.8, there are three peaks in the PLRM waveform. We cannot merely determine which part is the real leading edge that we need based only on the waveform. At this time, the SAR mode has demonstrated its advantages, as mentioned in Section 2.3.1. The difference between the positions of the two waveforms is about 40 m due to the difference in sampling time. Nevertheless, they can also be regarded as the waveform in the same area. Due to the smaller effective footprint, SAR mode avoids the loss of power at leading edges significantly and dramatically reduces the noises that may occur at the trailing edge.

Compared these two waveforms, we can determine that the first peak of the PLRM waveform is that we need. This conclusion is somewhat surprising because the returned power of the third peak is even higher than the first peak. This also shows how severe the PLRM waveform is affected in the coastal area.

For the above reasons, SAR is obviously the better choice. In the latter work of this thesis, SAR waveform will be the main research objective. The primary goal is to develop an algorithm to determine SWH from the SAR waveform. The algorithm will also be applied to determine the SWH from the PLRM waveform and compare it with the SWH from the SAR waveform, i.e. PLRM waveform will play a supporting role.



**Figure 2.8:** Waveform of SAR (6993.18 m) and PLRM (6950.15 m) from the Virtual Station on Cuxhaven (53.8730, 8.7107), 2018-12-12

## 2.5 Empirical retrackers

In this section, some waveform retracking methods developed over many years based on empirical observation and practical experience will be introduced. The empirical methods of waveform retracking may be classified into two categories: those based on the statistical properties of the waveform data and those based on fitting empirical functional forms [Vignudelli et al. (2011)].

Currently, there are some well known empirical retrackers: the offset centre of gravity retracker (OCOG), the threshold retracker, the improved threshold retracker and the  $\beta$ -parameter retracker. Only the OCOG retracker and the  $\beta$ -parameter retracker will be introduced in this section because the other two retrackers have not been used in this thesis.

### 2.5.1 Offset Centre of Gravity retracker (OCOG)

The Offset Centre Of Gravity (OCOG) retracking algorithm is a purely statistical approach which does not depend on a functional form. This algorithm estimates centre of gravity (COG) of a rectangular box. The twice of COG height is called waveform amplitude (A), and it determines the length of the box. The width (W) of the rectangular box determines the retracking gate, which is shown in Fig. 2.9:

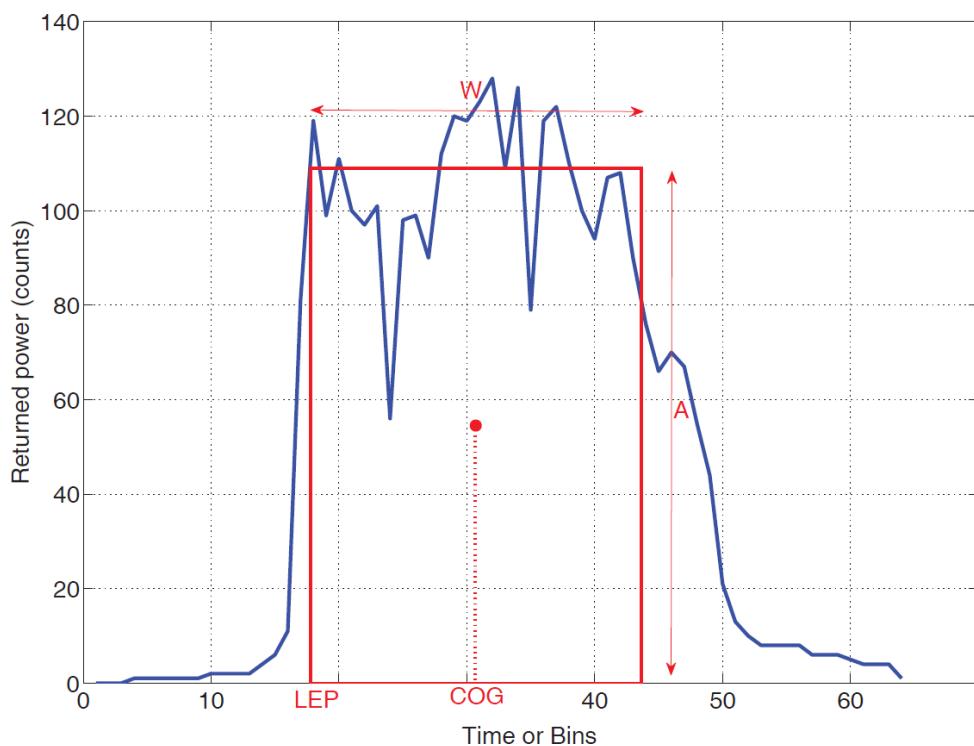


Figure 2.9: Schematic diagram of the OCOG method [Tourian (2012)]

These three parameters could be estimated:

$$A = \sqrt{\frac{\sum_{i=1+n_1}^{N-n_2} P_i^4(t)}{\sum_{i=1+n_1}^{N-n_2} P_i^2(t)}} \quad (2.5)$$

$$W = \frac{\left[ \sum_{i=1+n_1}^{N-n_2} P_i^2(t) \right]^2}{\sum_{i=1+n_1}^{N-n_2} P_i^4(t)} \quad (2.6)$$

$$COG = \frac{\sum_{i=1+n_1}^{N-n_2} iP_i^2(t)}{\sum_{i=1+n_1}^{N-n_2} P_i^2(t)} \quad (2.7)$$

where,  $P_i$  is the power of the  $i^{th}$  bin,  $N$  is the total number of samples in the waveform (128 for Sentinel-3),  $n_1$  and  $n_2$  are the numbers of bins affected by aliasing at the beginning and end of the waveform. Then, the position of the leading edge (LEP) is given by:

$$LEP = COG - \frac{W}{2} \quad (2.8)$$

OCOG is a very simple and robust waveform retracker, because it is not related to any physical properties of the reflecting water surface. Therefore, the result of OCOG itself is not very accurate, but it can be used as the initial value for other methods, such as the  $\beta$ -parameter retracker.

### 2.5.2 The $\beta$ -parameter retracker

The  $\beta$ -parameter retracking algorithm is known as 5- $\beta$  or 9- $\beta$  functional form to fit single or double-ramped waveforms. Unlike OCOG, although the  $\beta$ -retracker is still an empirical retracker, we can already extract some physical information from it, especially when comparing the 5- $\beta$  retracker with the Brown model. It will be described in detail later.

The 5- $\beta$  Retracker is used to fit the single-ramp waveform as shown in Fig. 2.10. The general expression for the 5- $\beta$  parameter functional form of the returned power  $y(t)$  is [Zwally and Brenner (2001)]:

$$y(t) = \beta_1 + \beta_2 (1 + \beta_5 Q) P\left(\frac{t - \beta_3}{\beta_4}\right) \quad (2.9)$$

where

$$Q = \begin{cases} 0 & \text{for } t < \beta_3 + 0.5\beta_4 \\ t - (\beta_3 + 0.5\beta_4) & \text{for } t \geq \beta_3 + 0.5\beta_4 \end{cases} \quad (2.10)$$

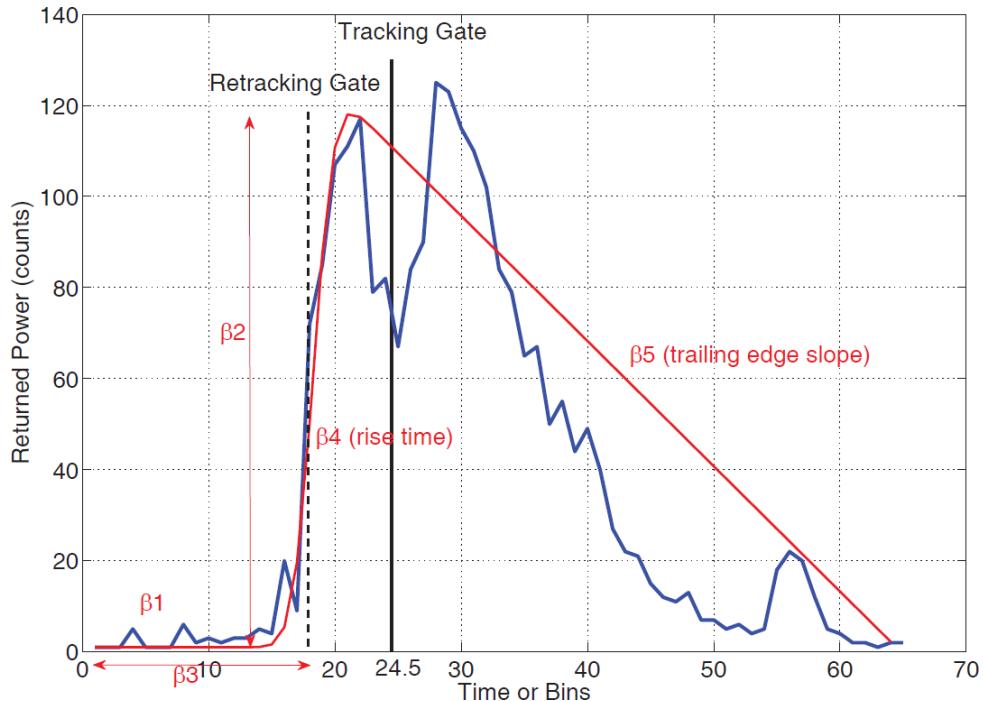
$$P(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-q^2}{2}\right) dq \quad (2.11)$$

$$= \frac{\left[ 1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) \right]}{2} \quad (2.12)$$

The unknown parameters  $\beta_1$  to  $\beta_5$  describe the different properties of the waveform:

- $\beta_1$ : The thermal noise level of the return waveform.
- $\beta_2$ : The amplitude of the returned power.
- $\beta_3$ : The mid-point on the leading edge of the waveform.
- $\beta_4$ : describes the slope of the leading edge, i.e. the waveform rising time.
- $\beta_5$ : describes the slope of the trailing edge.

An example of the 5- $\beta$  retracker waveform with a linear trailing edge is shown in Fig. 2.10:



*Figure 2.10: Schematic diagram of 5- $\beta$  Retracker with a linear trailing edge [Tourian (2012)]*

Besides, another way of expression is to use an exponential decay term instead of the linear trailing edge. A slightly different expression of the 5- $\beta$  retracker with an exponential trailing edge is given in [ Deng and Featherstone (2006)]:

$$y(t) = \beta_1 + \beta_2 \exp(-\beta_5 Q) P \left( \frac{t - \beta_3}{\beta_4} \right) \quad (2.13)$$

where

$$Q = \begin{cases} 0 & \text{for } t < \beta_3 - 2\beta_4 \\ t - (\beta_3 + 0.5\beta_4) & \text{for } t \geq \beta_3 - 2\beta_4 \end{cases} \quad (2.14)$$

The 9- $\beta$  retracker for the double-ramped waveform with an exponential trailing edge is given:

$$y(t) = \beta_1 + \sum_{i=1}^2 \beta_{2i} \exp(-\beta_{5i} Q_i) P \left( \frac{t - \beta_{3i}}{\beta_{4i}} \right) \quad (2.15)$$

where

$$Q_i = \begin{cases} 0 & \text{for } t < \beta_{3i} - 2\beta_{4i} \\ t - (\beta_{3i} + 0.5\beta_{4i}) & \text{for } t \geq \beta_{3i} - 2\beta_{4i} \end{cases} \quad (2.16)$$

The waveform created with the empirical  $\beta$ -parameter function show marked similarities with the waveforms produced with the theoretical Brown model. Applying Eq. 2.12 to Eq. 2.13 gives:

$$y(t) = \beta_1 + \beta_2 \frac{\left[1 + \operatorname{erf}\left(\frac{t-\beta_3}{\sqrt{2}\beta_4}\right)\right]}{2} \exp(-\beta_5 Q) \quad (2.17)$$

The Brown model is (applying the expression of  $u$  into Eq. 2.4):

$$V_m = T_n + a_{\xi} P_u \frac{\left[1 + \operatorname{erf}\left(\frac{t-\tau-c_{\xi}\sigma_c^2}{\sqrt{2}\sigma_c}\right)\right]}{2} \exp(-\nu) \quad (2.18)$$

Comparing the Eq. 2.17 and the Eq. 2.18, we can get the following relationships:

$$\beta_1 = T_n \quad (2.19)$$

$$\beta_2 = a_{\xi} P_u \quad (2.20)$$

$$\beta_3 = \tau + c_{\xi}\sigma_c^2 \quad (2.21)$$

$$\beta_4 = \sigma_c \quad (2.22)$$

$$\beta_5 = \frac{\nu}{Q} \quad (2.23)$$

It should be noted that the above equations do not mean complete equality because the parameters in the Brown model are physical quantities with specific physical meanings. However, the parameters in the  $\beta$ -parameter retracker are only unitless parameters, so the above equations just give us an intuitive feeling.



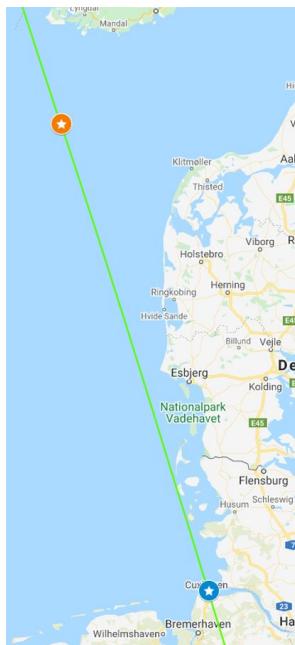
# Chapter 3

## Case study

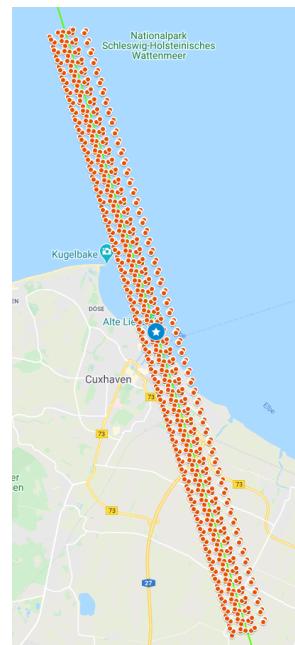
### 3.1 Choice of the virtual stations

This thesis is devoted to developing an algorithm to determine SWH, with a focus on the coastal area. The selected track is the relative orbit 328 of Sentinel-3B, pass number 655. The virtual station in the coastal area is selected close to Cuxhaven (53.8730, 8.7107). The virtual station in the oceanic area is on the same track over the North Sea, between Norway and Denmark (57.3623, 6.7602), see Fig. 3.1.

We will compare the estimated SWH with the SWH provided by Sentinel-3 in the oceanic area. Since Sentinel-3 cannot provide any SWH data in the coastal area, the data of the measuring station Elbe will be used for comparison and verification in the coastal area.



**Figure 3.1:** The coastal virtual station (blue star) at Cuxhaven (53.8730, 8.7107) and the oceanic virtual station (orange star) in the North Sea (57.3623, 6.7602) on the Sentinel-3B orbit (green line)

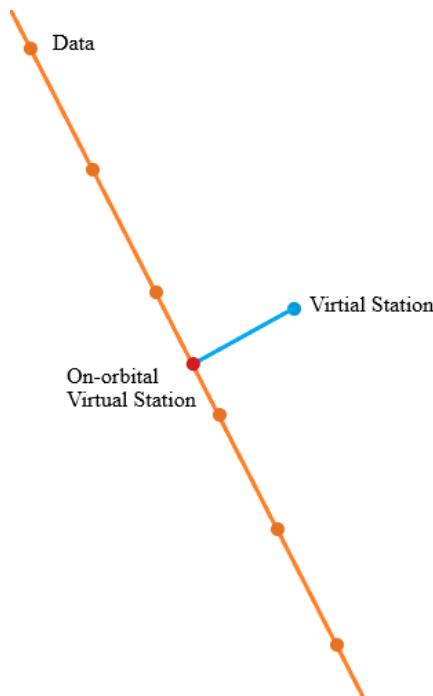


**Figure 3.2:** The virtual station (blue star) at Cuxhaven (53.8730, 8.7107) on the Sentinel-3B orbit (green line) and the SAR data (orange points) in the searching circle with  $R = 10 \text{ km}$  between 2018-12-12 and 2019-12-25

## 3.2 Relative distance

As shown in Fig. 3.2, due to the small vibration of the satellite orbit, the data position collected by the satellite is not always on the same line as the virtual station. If the distance between the data position and the virtual station is directly calculated, the distance will not be uniform. This problem is resolved in this thesis by making a vertical line perpendicular to the line of data through the virtual station, name the intersection point as the on-orbital virtual station, and calculate the distances between the data and the on-orbital virtual station as the relative distances (see Fig. 3.3). The difference between adjacent relative distances is basically the same, about 331 m.

Considering the flight direction of the satellite, the sign of relative distances before the virtual station is defined to be negative, and the sign after the satellite passes the virtual station will be positive.



*Figure 3.3: On-orbital Virtual station (red point) and relative distance*

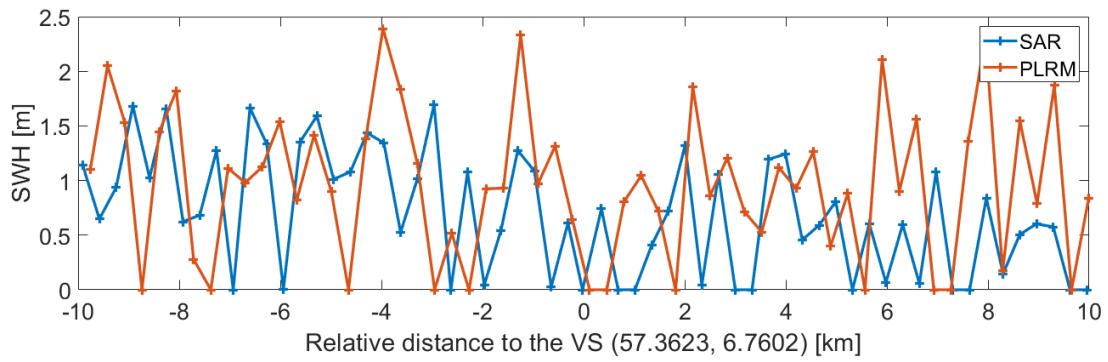
## 3.3 SWH provided by Sentinel-3

Sentinel-3 itself provides SWH generated using the ocean retracking algorithm. The algorithms used by SAR and PLRM modes are different [ESA Sentinel Online]:

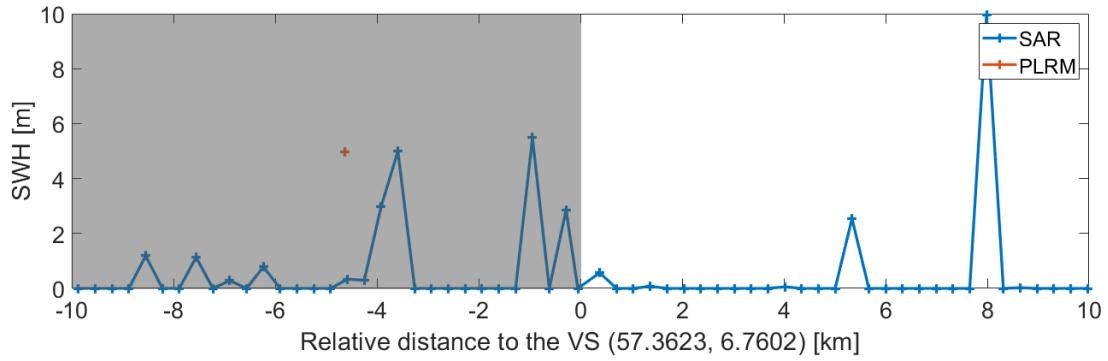
- PLRM: **Ocean-3**, inherited from JASON-2 mission, is used to perform a fitting of the waveform with a 4-th parametric model estimation (epoch, composite sigma, amplitude and mispointing angle) based on a weighted least square error (LSE) minimization derived from a Maximum Likelihood Estimator (MLE).

- SAR: **Fully analytical re-tracker**, operation over open ocean and coastal zones inherited from SAMOSA project. It tries to fit the theoretically modelled multi-look L1B waveform to the real L1B SAR waveform, providing estimates of the epoch, composite sigma, amplitude and mispointing angle.

The example of oceanic results are shown in Fig. 3.4 and of the coastal results are shown in Fig. 3.5. By comparing these data, we can find that the SWH in the coastal area cannot be determined using the on-board retracking algorithm. SWH of SAR mode in the coastal area showed a large number of 0 m and a very abrupt 9.953 m, while the results of PLRM are directly NaN. At the same time, for oceanic waveforms, Sentinel-3 can provide better results, but the fluctuation itself is still too large. Although the sampling positions of SAR and PLRM waveforms are not always the same, the position difference between two adjacent data is mostly tens of meters, and a small amount of them reach about 150 m. Therefore, the difference between the SWH of SAR and PLRM should not be so much as shown in Fig. 3.4.



*Figure 3.4: Oceanic SWH provided by Sentinel-3 on 2018-12-12*



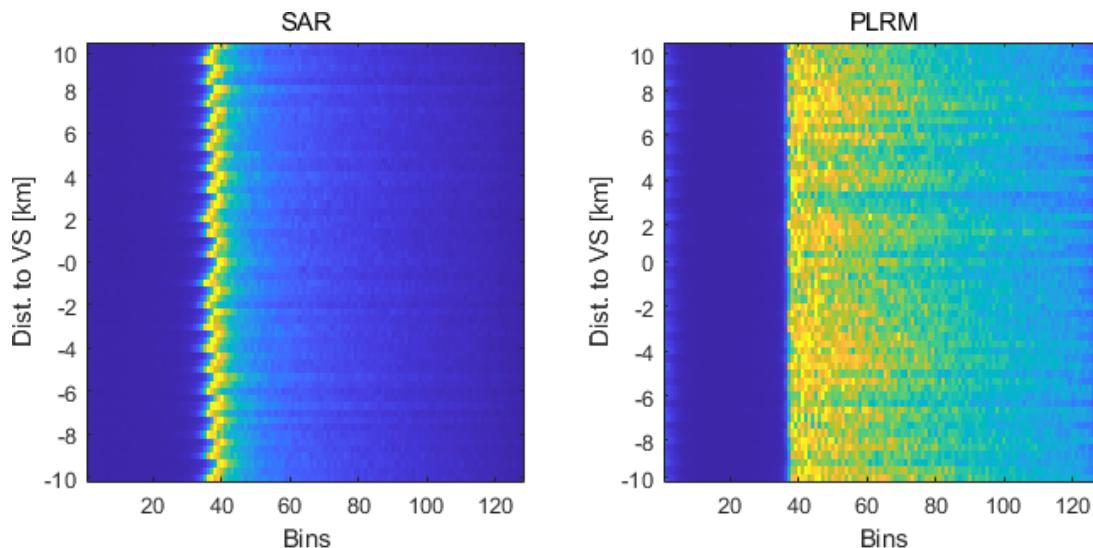
*Figure 3.5: Coastal SWH provided by Sentinel-3 on 2018-12-12, gray area is land*

The reason why Sentinel-3 cannot provide the SWH data in the coastal zone is that the onboard retracker of Sentinel-3 is still based on the Brown model. As shown in Fig. 3.6, the waveforms in the oceanic area are very uniform, which means that these waveforms conform to the theoretical Brown model (Fig. 2.3). However, the waveforms in the coastal area are much more complex, and there are many scattered high-power bins, which means that the waveform have multiple discontinuous peaks, see Fig. 3.7. Therefore, the retracker based on the Brown model is no longer applicable. Thus, the SWH data near the shore provided by Sentinel-3 are

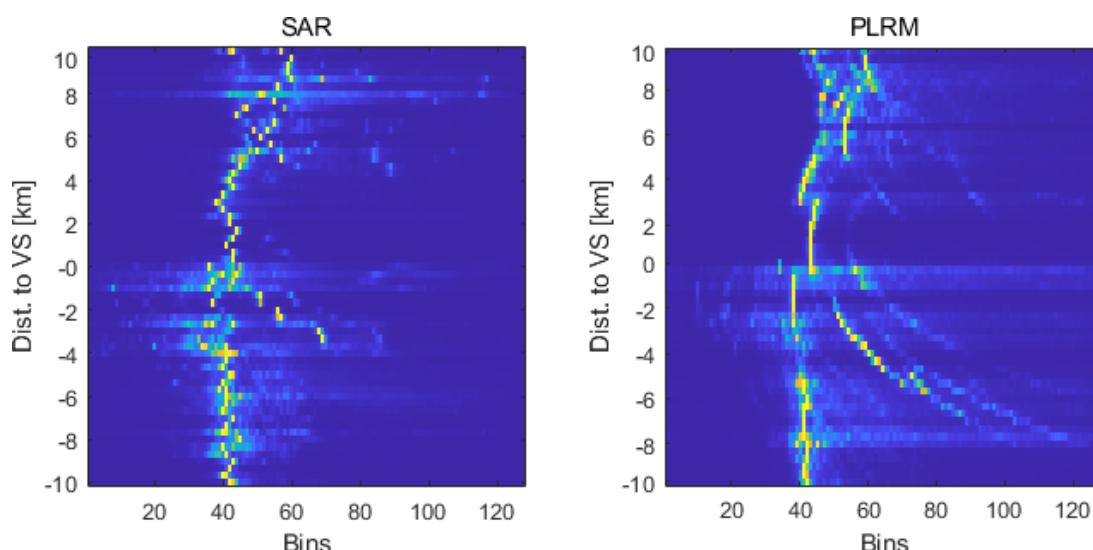
not reliable.

Also, by comparing oceanic waveforms in SAR and PLRM, it can be found that there are fewer high-power bins in SAR mode, often 2-3 bins and a large number of high-power bins appear in PLRM. The reason is as explained in Section 2.3.2, the noise of the trailing edge of SAR mode is far less than that of PLRM mode.

In Chapter 4 and Chapter 5, in order to conveniently compare the results of PLRM and SAR, such as calculation of the difference, all the PLRM data will be resampled based on the position of the SAR data, which means the PLRM data will be aligned to the nearest SAR data. In the end, for the map of SWH, the original position data will be used, i.e. not resampled.



*Figure 3.6: Radargram of the normalized waveforms in the oceanic area on 2018-12-12*



*Figure 3.7: Radargram of the normalized waveforms in the coastal area on 2018-12-12*

# Chapter 4

## Methodology

### 4.1 Basic idea

Significant wave height is defined traditionally as the third of the highest waves in the fields of view (explained in Section 1.2). Typically, the SWH can be determined by using the coefficient  $\sigma_s$ , which can be calculated from  $\sigma_c$  of the Brown model (Eq. 2.4) or the  $\beta_4$  of the 5- $\beta$  retracker (Eq. 2.13).

The problem is much more complicated in the coastal area. Due to the complicated situation in the coastal area, most of the waveforms do not conform to the Brown model. Therefore, the fitting algorithms based on the Brown model are no longer applicable, and we cannot extract the parameter  $\sigma_c$  directly from the waveform. But  $\sigma_c$  is related to the raising time of the leading edge, i.e. the difference between the starting point and the end point of the leading edge ( $\Delta\text{Bins}$ ). Therefore, this thesis attempts to extract the  $\sigma_c$  by determining the  $\Delta\text{Bins}$ . Considering the relationship between the  $\sigma_c$  and the  $\Delta\text{Bins}$  we can determine the  $\sigma_c$  as well. In the end, we could estimate the SWH from the non-Brown model.

The most important task now is to determine the  $\Delta\text{Bins}$ . To achieve this goal, we need to fit the waveform by using retracking algorithms. For oceanic zone, the existing retracking algorithm (5- $\beta$  retracker) works well on fitting the waveforms. However, for the coastal waveforms, they are no longer applicable due to more noise (especially in the trailing edge). This thesis attempts to determine the peak of the waveform to extract the thermal noise and the leading edge, and then only fits this part to avoid the influence of the noise included by the trailing edge on the fitted leading edge, see Section 4.4.

### 4.2 Adjustment using Gauss-Markov model

#### 4.2.1 Principles

Gauss-Markov model is known as the adjustment with observation equations. The model is as follows:

$$\mathbf{y} = \mathbf{Ax} + \mathbf{e} \quad (4.1)$$

where  $\mathbf{y}$  is a vector of observations,  $\mathbf{A}$  is the design matrix,  $\mathbf{x}$  is a vector of unknowns and  $\mathbf{e}$  is a vector of measurement errors. Define the *Lagrangian* or *cost function*:

$$\mathcal{L}_a(\mathbf{x}) = \frac{1}{2} \mathbf{e}^T \mathbf{e} \quad (4.2)$$

Then, the adjusted observations can be estimated by using least square criterion, which means find the  $\hat{\mathbf{x}}$  to minimize the cost function. We can get the following equations when we solve the minimization problem:

$$\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y} \quad (4.3)$$

$$\hat{\mathbf{y}} = \mathbf{A} \hat{\mathbf{x}} = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y} \quad (4.4)$$

$$\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}} = [\mathbf{I} - \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T] \mathbf{y} \quad (4.5)$$

In praxis, many observation equations are non-linear, like Eq.2.17. Therefore, we cannot directly derive the design matrix  $\mathbf{A}$  but need to use Taylor's theorem to linearize the model:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \quad (4.6)$$

$$= f(x_0) + \left. \frac{df}{dx} \right|_{x_0} (x - x_0) + \underbrace{\left. \frac{1}{2} \frac{d^2 f}{dx^2} \right|_{x_0} (x - x_0)^2}_{\text{negligible if } x - x_0 \text{ is small}} + \dots \quad (4.7)$$

Subtracting the  $f(x_0)$  terms:

$$f(x) - f(x_0) = y - y_0 = \left. \frac{df}{dx} \right|_{x_0} (x - x_0) + \dots \quad (4.8)$$

$$\Rightarrow \Delta y = \underbrace{\left. \frac{df}{dx} \right|_0 \Delta x}_{\text{Linear model}} + \underbrace{\mathcal{O}(\Delta x^2)}_{\text{Terms of higher order (model errors)}} \quad (4.9)$$

Now we can extend Eq. 4.9 to the high-dimensional case (with  $m$  observations and  $n$  unknowns) and neglect the terms of second and higher order:

$$\begin{bmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_m \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}}_0 \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \dots \\ \Delta x_n \end{bmatrix} \Rightarrow \Delta \mathbf{y} = \mathbf{A}(\mathbf{x}_0) \Delta \mathbf{x} \quad (4.10)$$

Jacobian matrix  $\mathbf{A}$

where  $\mathbf{x}_0$  is the vector of the initial values of unknowns. The influence of the initial values is also critical. If the differences between the initial values and the actual values are too huge, the cost function may converge to the local minimum instead of the global minimum, or even may not converge.

Combining Eq. 4.10 and Eq. 4.3 to Eq.4.5, the unknowns can be estimated by iteration, as shown in Fig 4.1.

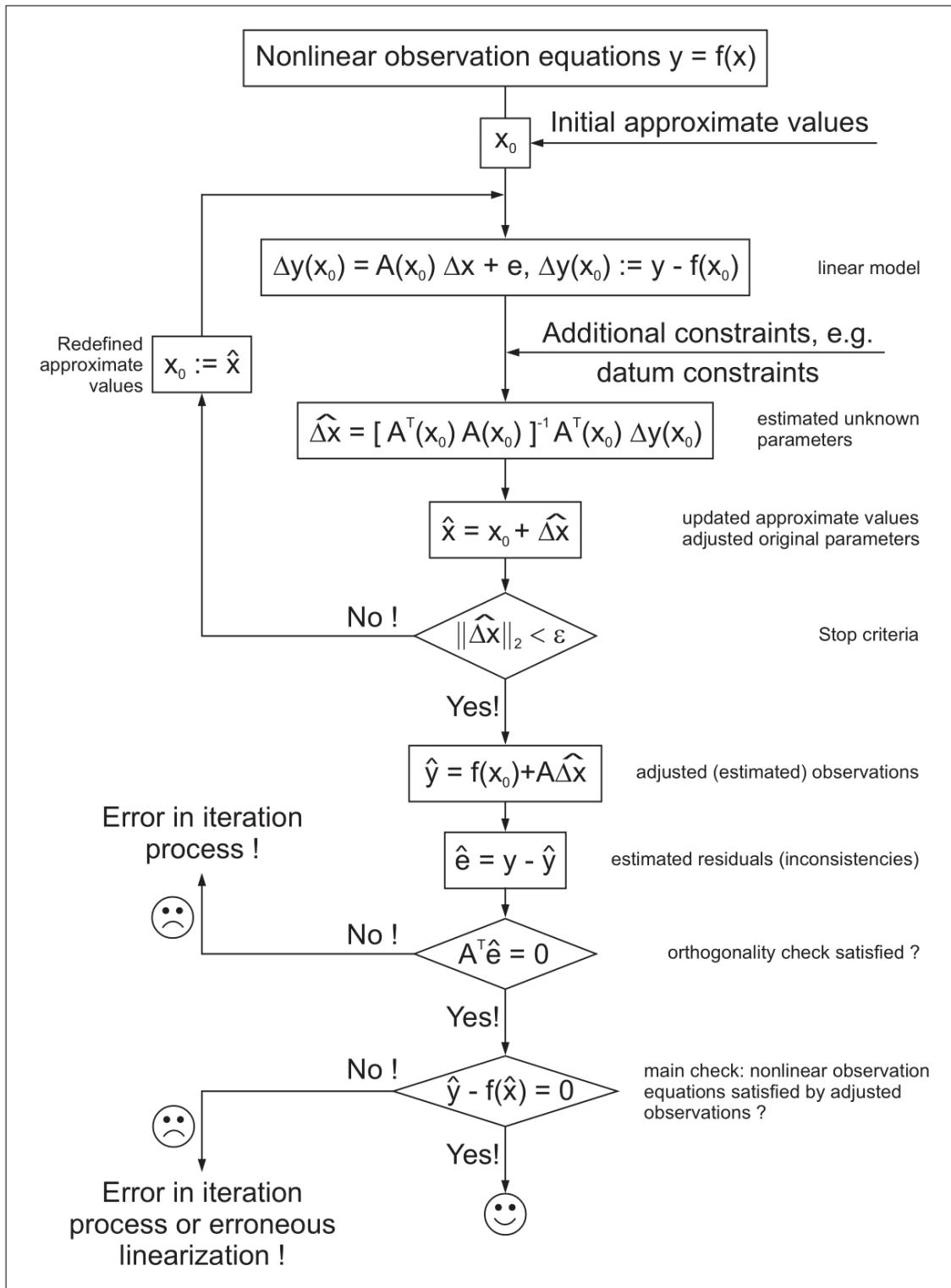


Figure 4.1: Iterative scheme of the Gauss-Markov model [Sneeuw et al. (2008)]

### 4.2.2 Applying the 5- $\beta$ retracker using Gauss-Markov model

The model of 5- $\beta$  retracker is shown in Section 2.5.2. We could compute the needed differentials as following (based on Eq. 2.13):

$$\frac{\partial y(t)}{\partial \beta_1} = 1 \quad (4.11)$$

$$\frac{\partial y(t)}{\partial \beta_2} = \exp(-\beta_5 Q) P(u) \quad (4.12)$$

$$\frac{\partial y(t)}{\partial \beta_3} = -\beta_2 \beta_5 \exp(-\beta_5 Q) P(u) \frac{\partial Q}{\partial \beta_3} - \frac{\beta_2}{\beta_4} \exp(-\beta_5 Q) \frac{\partial P(u)}{\partial u} \quad (4.13)$$

$$\frac{\partial y(t)}{\partial \beta_4} = -\beta_2 \beta_5 \exp(-\beta_5 Q) P(u) \frac{\partial Q}{\partial \beta_4} - \frac{u \beta_2}{\beta_4} \exp(-\beta_5 Q) \frac{\partial P(u)}{\partial u} \quad (4.14)$$

$$\frac{\partial y(t)}{\partial \beta_5} = -\beta_2 Q \exp(-\beta_5 Q) P(u) \quad (4.15)$$

$$(4.16)$$

where

$$u = \frac{t - \beta_3}{\beta_4} \quad (4.17)$$

$$\frac{\partial P(u)}{\partial u} = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right) \quad (4.18)$$

$$\frac{\partial Q}{\partial \beta_3} = \begin{cases} 0 & \text{for } t < \beta_3 - 2\beta_4 \\ -1 & \text{for } t \geq \beta_3 - 2\beta_4 \end{cases} \quad (4.19)$$

$$\frac{\partial Q}{\partial \beta_4} = \begin{cases} 0 & \text{for } t < \beta_3 - 2\beta_4 \\ -0.5 & \text{for } t \geq \beta_3 - 2\beta_4 \end{cases} \quad (4.20)$$

As mentioned above, the initial values have significant influences on the fitted results. Based on the relationship shown in Fig. 2.10 the initial values of  $\beta_1$  to  $\beta_4$  are selected as follows:

- $\beta_1^0$ : The median of the returned powers on the 5<sup>th</sup> to 15<sup>th</sup> bins.
- $\beta_2^0$ : The maximum of the returned power minus  $\beta_1^0$ .
- $\beta_3^0$ : The index (bins) of the peak of the waveform minus  $0.5 \cdot \beta_4^0$ .
- $\beta_4^0$ : The temporal resolution, i.e. 1 bin.

The choice of  $\beta_5$  is more difficult because the trailing edges of oceanic SAR, oceanic PLRM and coastal waveforms have different properties (slope). Therefore, the initial values of these three cases should be different. The solution is a little tricky. 2000 sets of  $\beta_5^0$  with a step size of 0.0005 within 0-1 have been tried. The standard deviation of the output results of different initial values haven been compared, and the following initial values have been selected:

$$\beta_5^0 = \begin{cases} 0.14 & \text{for Ocean, SAR} \\ 0.005 & \text{for Ocean, PLRM} \\ 0.5 & \text{for Coastal waveform} \end{cases} \quad (4.21)$$

In fact,  $\beta_5^0$  of the coastal waveform is not important, because the 5- $\beta$  retracker is not applicable in the coastal zone. The 4- $\beta$  retracker will be introduced in Section 4.4, and the peak of the waveform determination algorithm will also be introduced.

Now, Eq. 4.10 becomes:

$$\begin{bmatrix} \Delta y(1) \\ \Delta y(2) \\ \vdots \\ \Delta y(128) \end{bmatrix} = \begin{bmatrix} \frac{\partial y(1)}{\partial \beta_1} & \dots & \frac{\partial y(1)}{\partial \beta_5} \\ \vdots & \ddots & \vdots \\ \frac{\partial y(128)}{\partial \beta_1} & \dots & \frac{\partial y(128)}{\partial \beta_5} \end{bmatrix} \Big|_0 \begin{bmatrix} \Delta \beta_1 \\ \Delta \beta_2 \\ \Delta \beta_3 \\ \Delta \beta_4 \\ \Delta \beta_5 \end{bmatrix} \quad (4.22)$$

Now, we can use the method shown in Fig. 4.1 to estimate the 5- $\beta$  fitted waveform in the oceanic area. Some examples of the results are shown in Fig. 4.2 (PLRM) and Fig. 4.3 (SAR). The standard deviations haven also been computed using the following equation to evaluate the quality of the fitted waveform:

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{128} (V_m(i) - y(i))^2}{f}} \quad (4.23)$$

where  $V_m(i)$  and  $y(i)$  are the normalized returned power of the  $i^{th}$  bin and 5- $\beta$  fitted returned power of the  $i^{th}$  bin, respectively.  $f$  is the degree of freedom, which is equal to 123 here ( $f = m - n = 128 - 5$ ).

#### 4.2.3 Limitation of the 5- $\beta$ retracker

The 5- $\beta$  retracker has its own limitation. It can be seen by comparing Fig.4.2 and Fig. 4.3 that although the standard deviation of the SAR fitted waveform is not significantly larger than the standard deviation of the PLRM fitted waveform, the quality of the fitting results is significantly different. If we solve the least square problem by Levenberg-Marquardt described in Section 4.3, the quality of the PLRM fitted waveforms can be improved, but the problem of SAR fitted waveform still exists.

One possible reason is that due to the different property of SAR trailing edge and PLRM trailing edge, the 5- $\beta$  retracker cannot balance all the five parameters of SAR waveform, so a deviation occurs. However, if we look at these five parameters, we can find that although the position ( $\beta_3$ ) and the amplitude ( $\beta_2$ ) of the leading edge are shifted due to model limitations, the shape of the leading edge ( $\beta_4$ ) remains the same. So the results can still be used to determine the SWH. Besides, the SAR trailing edge contains much less noise than PLRM, which makes the peak of the waveform more accessible. Due to this property, the 4- $\beta$  retracker can better extract the leading edge of SAR waveform, which will be introduced in Section 4.4.

#### 4.2.4 Limitation of the Gauss-Markov model in the waveform adjustment

Comparing the results in Fig. 4.2, we can see that most of the fitting results are acceptable, with a standard deviation between 0.05 and 0.06. However, the result of the Num. 36 is not the best

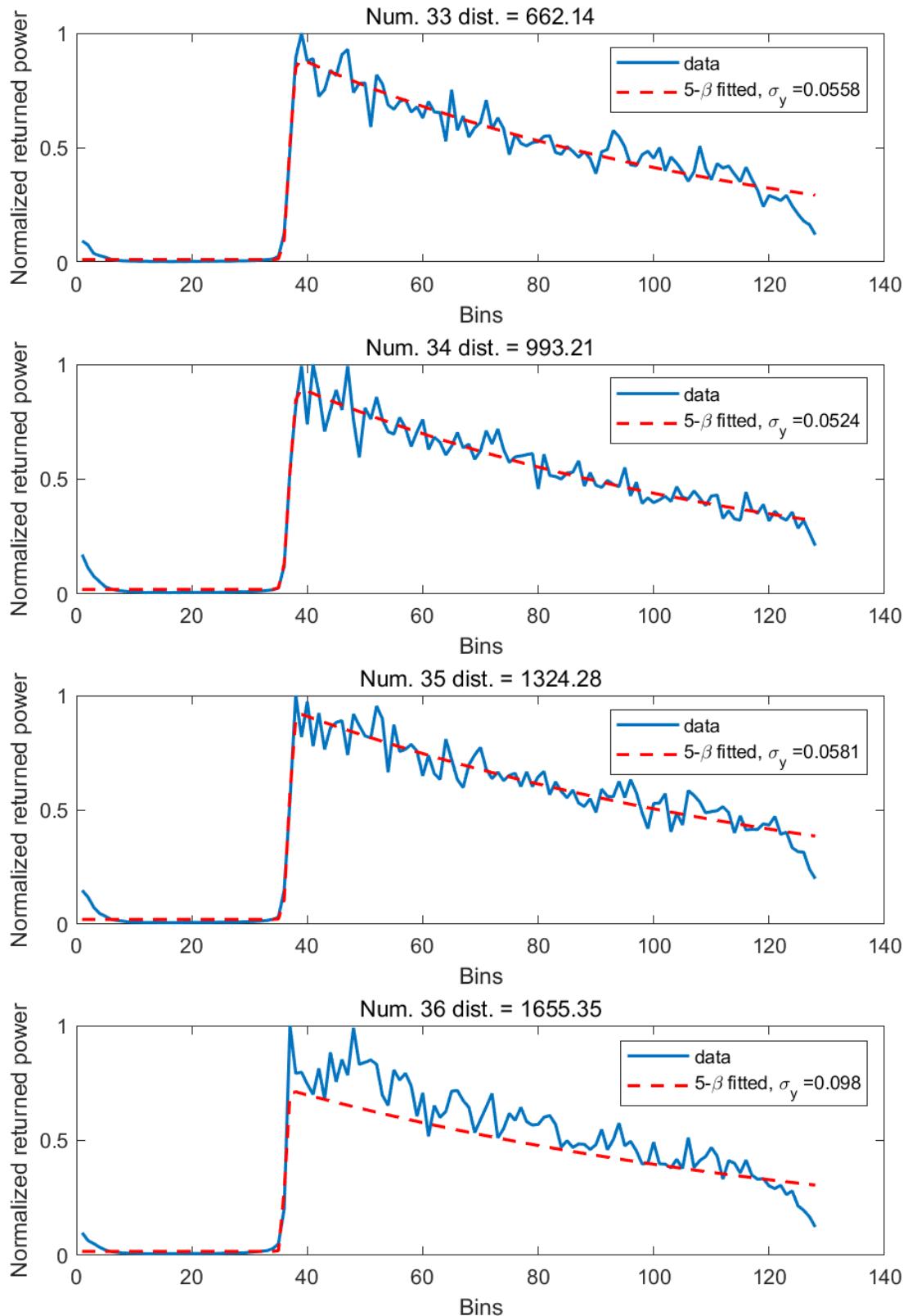


Figure 4.2: Examples of 5- $\beta$  fitted oceanic PLRM waveform on 2018-12-12 (Gauss-Markov)

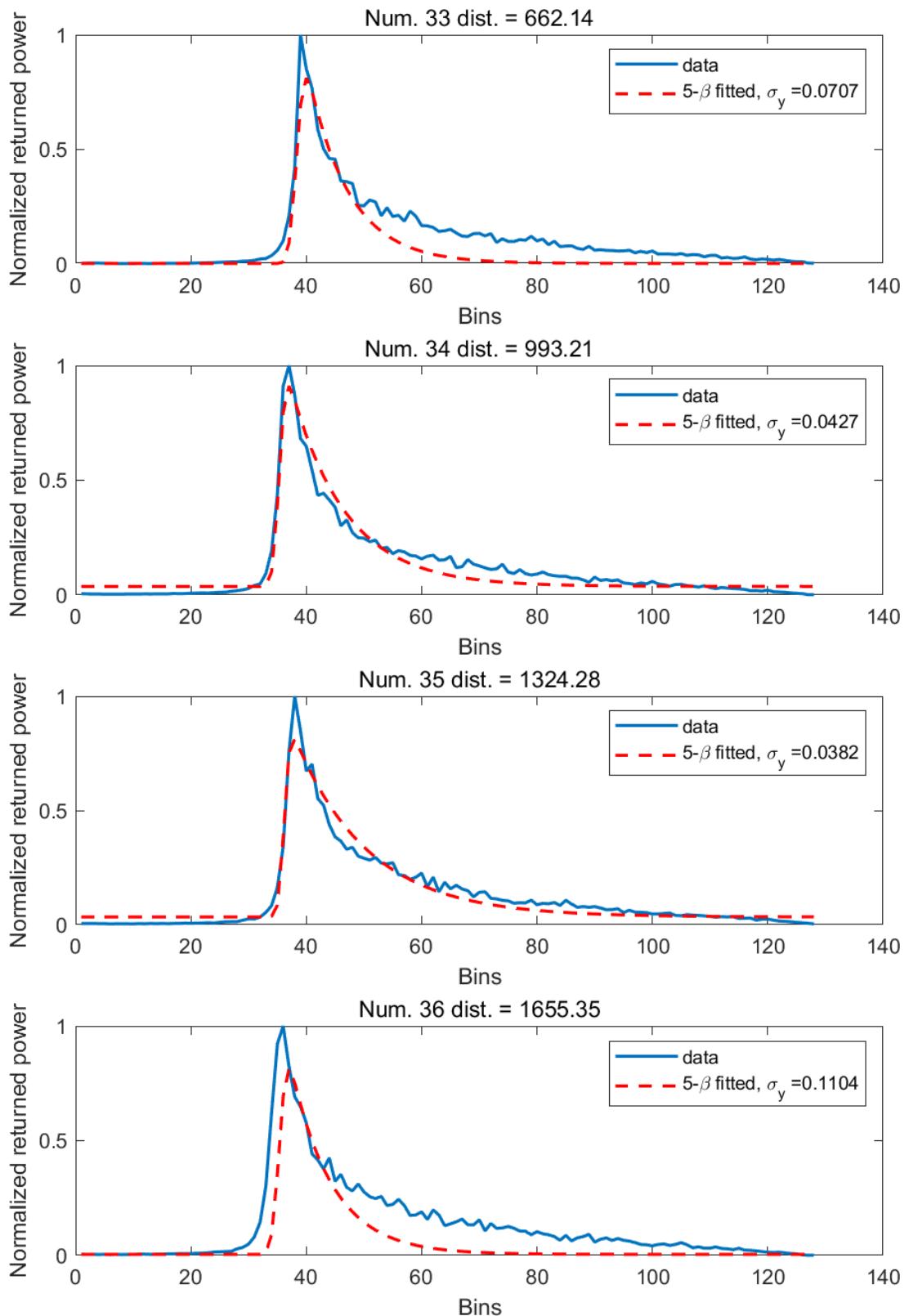
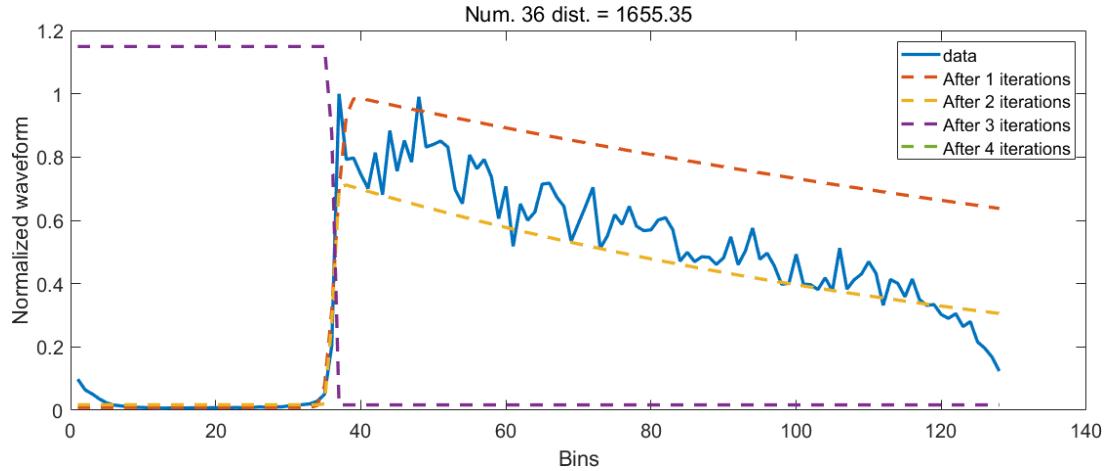
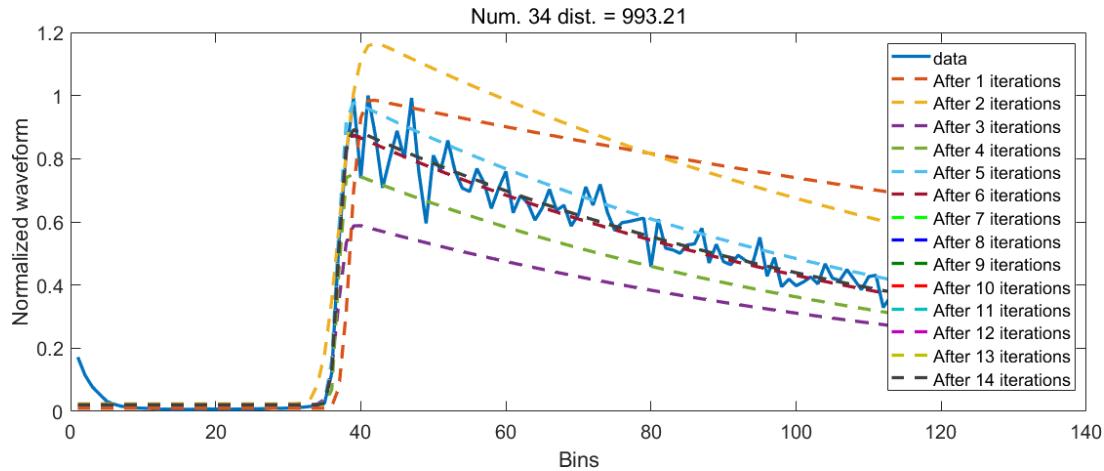


Figure 4.3: Examples of  $5\beta$  fitted oceanic SAR waveform on 2018-12-12 (Gauss-Markov)

fitting value, and the corresponding standard deviation is also close to 0.1. The reason is that although the waveform looks similar to the other waveforms, the selected initial values are not suitable for this waveform. Therefore, the iterative adjustment did not converge normally, and as the number of iteration increases, the magnitude difference of the values in the normal matrix is getting larger and larger. It will cause a rank deficient problem because the tiny values (relative) are typically considered as 0 within the PC processor. Thus, the inverse matrix of the normal matrix cannot be computed correctly. The final output is NaN, as shown in Fig. 4.4. As a comparison, Fig. 4.5 shows a convergence of the iterative adjustment (Num. 34).



**Figure 4.4:** An unsuccessful case of Gauss-Markov model (Num. 36 on 2018-12-12 in the oceanic area, PLRM)



**Figure 4.5:** A successful case of Gauss-Markov model (Num. 34 on 2018-12-12 in the oceanic area, PLRM)

Since the initial values cannot be determined individually based on each single waveform, the above problem cannot be avoided. Therefore, the standard deviation of each fitted waveform will be stored during the entire iterative process. If the final output is NaN, the fitted value with the smallest standard deviation in the iterative process will be output as the final result. Therefore, the output result may not be the best estimation.

## 4.3 Levenberg-Marquardt Method

### 4.3.1 Principles

In this section, Levenberg-Marquardt Method will be introduced to solve the least-squares problem [Levenberg (1944) and Marquardt (1963)].

The least-squares problem is defined as:

$$\min_{\mathbf{x}} f(\mathbf{x}) = \|F(\mathbf{x})\|_2^2 = \sum_i F_i^2(\mathbf{x}) \quad (4.24)$$

where

$$F(\mathbf{x}) = \begin{bmatrix} \bar{y}(\mathbf{x}, t_1) - \bar{\phi}(t_1) \\ \bar{y}(\mathbf{x}, t_2) - \bar{\phi}(t_2) \\ \vdots \\ \bar{y}(\mathbf{x}, t_m) - \bar{\phi}(t_m) \end{bmatrix} \quad (4.25)$$

where  $\bar{y}$  and  $\bar{\phi}$  include the weights of the quadrature scheme,  $\mathbf{x}$  is the vector of the unknowns in the model and  $t_i$  are scalars. Denoting the  $m$ -by- $n$  Jacobian matrix of  $F(\mathbf{x})$  as  $J(\mathbf{x})$ , the gradient vector of  $f(\mathbf{x})$  as  $G(\mathbf{x})$ , the Hessian matrix of  $f(\mathbf{x})$  as  $H(\mathbf{x})$  and the Hessian matrix of each  $F_i(\mathbf{x})$  as  $H_i(\mathbf{x})$ , we get:

$$G(\mathbf{x}) = 2J(\mathbf{x})^T F(\mathbf{x}) \quad (4.26)$$

$$H(\mathbf{x}) = 2J(\mathbf{x})^T J(\mathbf{x}) + 2Q(\mathbf{x}) \quad (4.27)$$

where

$$Q(\mathbf{x}) = \sum_{i=1}^m F_i(\mathbf{x}) \cdot H_i(\mathbf{x}) \quad (4.28)$$

The property of the matrix  $Q(\mathbf{x})$  is that when the residual  $\|F(\mathbf{x})\|$  tends to zero as  $\mathbf{x}_k$  approaches the solution, the  $Q(\mathbf{x})$  tends to zero. Thus when  $\|F(\mathbf{x})\|$  is small as the solution, a very effective method is to use the Gauss-Newton direction as a basis for an optimization procedure [ Mathworks (2020)].

In the Gauss-Newton method, a search direction  $d_k$  is obtained at each major iteration,  $k$ , that is a solution of the linear least-squares problem:

$$\min_{\mathbf{x}} \|J(\mathbf{x}_k) - F(\mathbf{x}_k)\|_2^2 \quad (4.29)$$

The direction derived from this method is equivalent to the Newton direction when the terms of  $Q(\mathbf{x})$  can be ignored. The search direction  $\mathbf{d}_k$  can be used as part of the line search strategy to ensure that at each iteration the function  $f(\mathbf{x})$  decreases. However, the Gauss-Newton method often encounters problems when the  $Q(\mathbf{x})$  is significant. A method that overcomes this problem is the Levenberg-Marquardt method [Levenberg (1944) and Marquardt (1963)]. The method uses a search direction that is a solution of the linear set of equations:

$$(J(\mathbf{x}_k)^T J(\mathbf{x}_k) + \lambda_k \mathbf{I}) \mathbf{d}_k = -J(\mathbf{x}_k)^T F(\mathbf{x}_k) \quad (4.30)$$

or, optionally, of the equations:

$$\left( J(\mathbf{x}_k)^T H(\mathbf{x}_k) + \lambda_k \text{diag} \left( J(\mathbf{x}_k)^T J(\mathbf{x}_k) \right) \right) \mathbf{d}_k = -J(\mathbf{x}_k)^T F(\mathbf{x}_k) \quad (4.31)$$

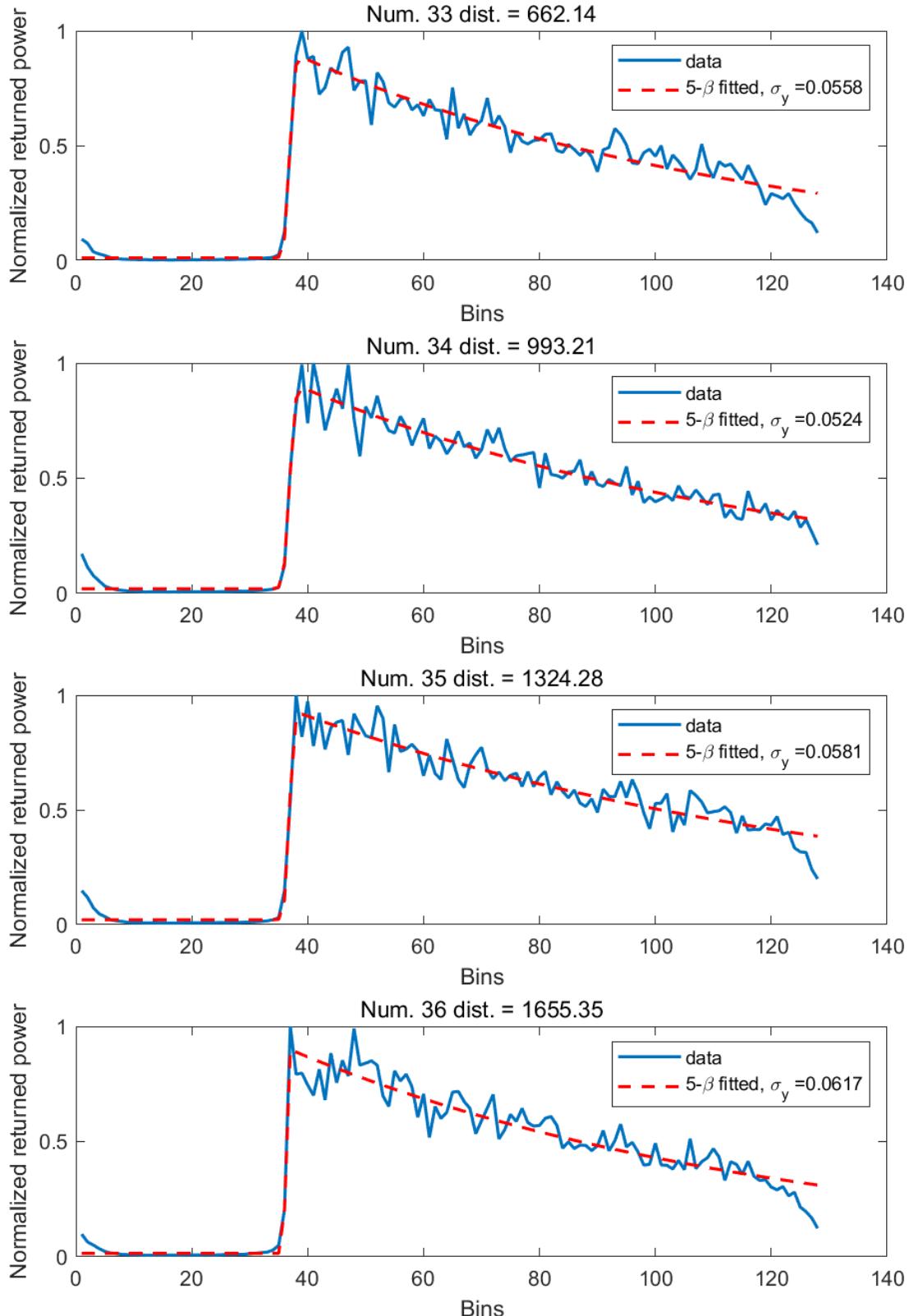
where the scalar  $\lambda_k$  controls both the magnitude and direction of  $\mathbf{d}_k$ .

### 4.3.2 Applying the 5- $\beta$ retracker using Levenberg-Marquardt method

There is one function named *lsqcurvefit* in the Optimization Toolbox of Matlab which uses the Levenberg-Marquardt method to solve the least-squares problem. This function will be used to solve the nonlinear model-fitting problems in this thesis. For more details, please read [lsqcurvefit] and [Least-Squares (Model Fitting) Algorithms]. Similarly, since the 5- $\beta$  retracker is not suitable for coastal waveform, here we only apply it on the oceanic waveform, the some examples of the results are shown in Fig. 4.6 (PLRM) and Fig. 4.7 (SAR), the standard deviations are calculated using Eq. 4.23.

Comparing Fig. 4.6 and Fig. 4.2, it can be found that for most PLRM waveforms, the fitted results using the two methods are the same, but for Num. 36, the improvement brought by the Levenberg-Marquardt method is noticeable. Due to the reasons mentioned in Section 4.2.4, the Gauss-Markov model cannot fit the Num. 36 waveform correctly, but the Levenberg-Marquardt method solves this problem. This is also the main reason why the Levenberg-Marquardt method is chosen as the fitting algorithm in this thesis.

Comparing Fig. 4.7 and Fig. 4.3, it can be found that the fitted SAR results using the Levenberg-Marquardt method are better (with smaller standard deviations), and the significant shifted problems have been solved (Num. 36). However, the problem of the thermal noise ( $\beta_1$ ) and the amplitude ( $\beta_2$ ) still exist. The possible reason is mentioned in Section 4.2.3.



*Figure 4.6: Examples of 5- $\beta$  fitted oceanic PLRM waveform on 2018-12-12 (Levenberg-Marquardt)*

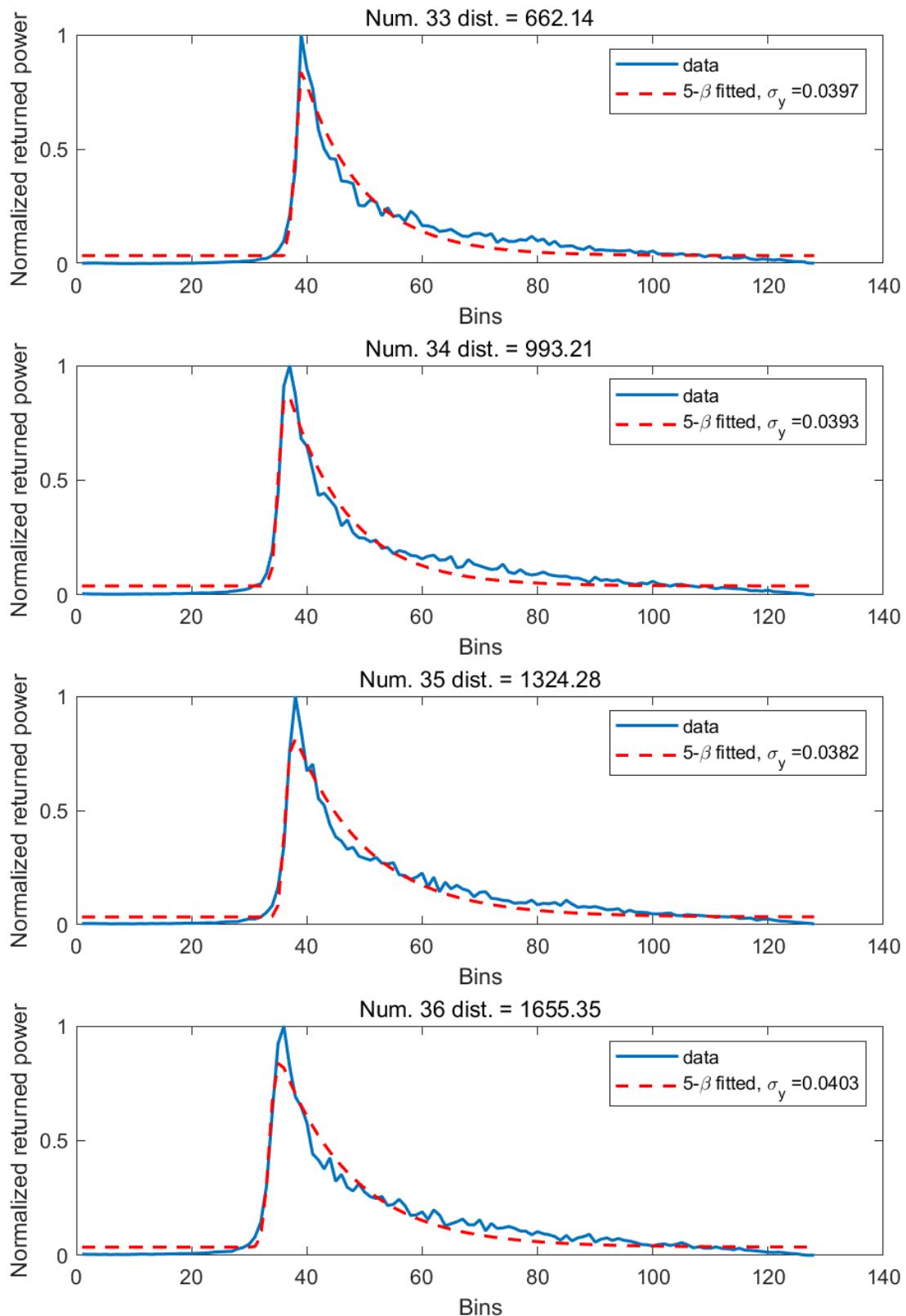


Figure 4.7: Examples of 5- $\beta$  fitted oceanic SAR waveform on 2018-12-12 (Levenberg-Marquardt)

## 4.4 The 4- $\beta$ retracker

Since the fitted waveform using the 5- $\beta$  retracker of SAR waveform is not perfect, and the 5- $\beta$  retracker is not suitable for coastal waveform at all, at the same time, as described in Section 4.1, only the information of the leading edge is needed to determine the SWH, this section considers a method to fit a part of the waveform. The goal is to fit the leading edge independent of the trailing edge. At first, the end of the leading edge, i.e. the peak of the waveform must be extracted as accurately as possible.

### 4.4.1 Extracting the thermal noise and the leading edge

The algorithm to extract the the thermal noise and the leading edge is shown in Fig. 4.9 and Fig. 4.10. This algorithm is based on the normalized waveform. The normalization is using the following equation:

$$\bar{y}_i = \frac{y_i}{\max(\mathbf{y})} \quad (4.32)$$

which means the normalized returned power is between 0 and 1. Then, the algorithm consists of four steps. The results are shown in Fig. 4.8.

1. Some local maximums will be defined as pre-peaks, where the normalized returned power is greater than 0.7.
2. Delete the normal points on the trailing edge.
3. Eliminate the noise on the leading edge.
4. Eliminate the noise on the trailing edge.

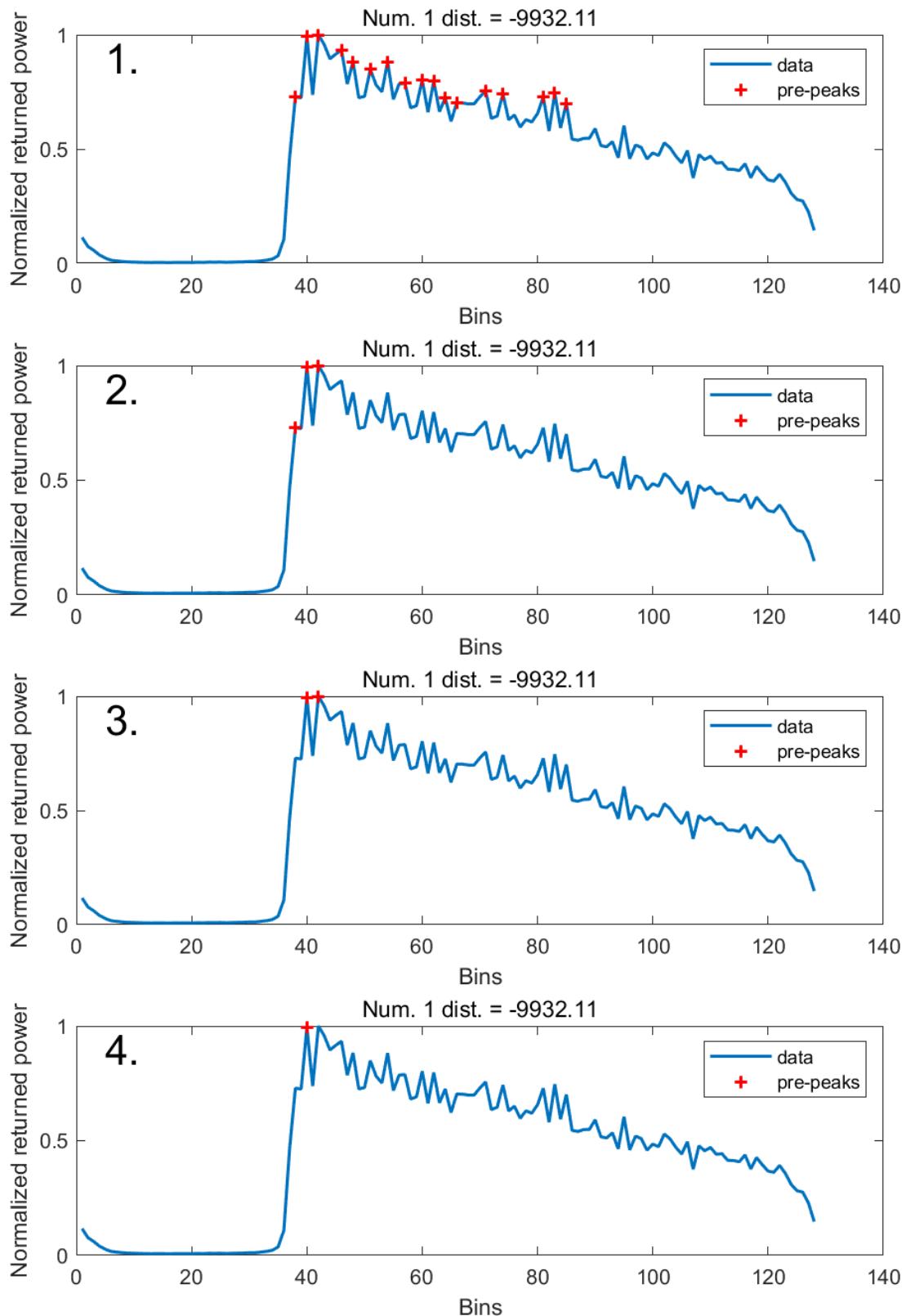
### 4.4.2 Model of the 4- $\beta$ retracker

Based on the 5- $\beta$  retracker, when only considering leading edge, most of  $Q$  is equal to 0 ( $t < \beta_3 - 2\beta_4$ ), so  $\beta_5$  has no effect on the fitted value at all. Before the bin of the peak, there are normally only 1 to 3 bins where  $Q$  is not equal to 0. The role of the entire term  $\exp(-\beta_5 Q)$  before the trailing edge is mainly to reduce the fitted value at the peak so that the transition to the trailing edge can be smooth. In the 4- $\beta$  retracker, the whole trailing edge is neglected. Therefore, we do not need to consider the problem of a smooth transition, so this term is no longer critical.

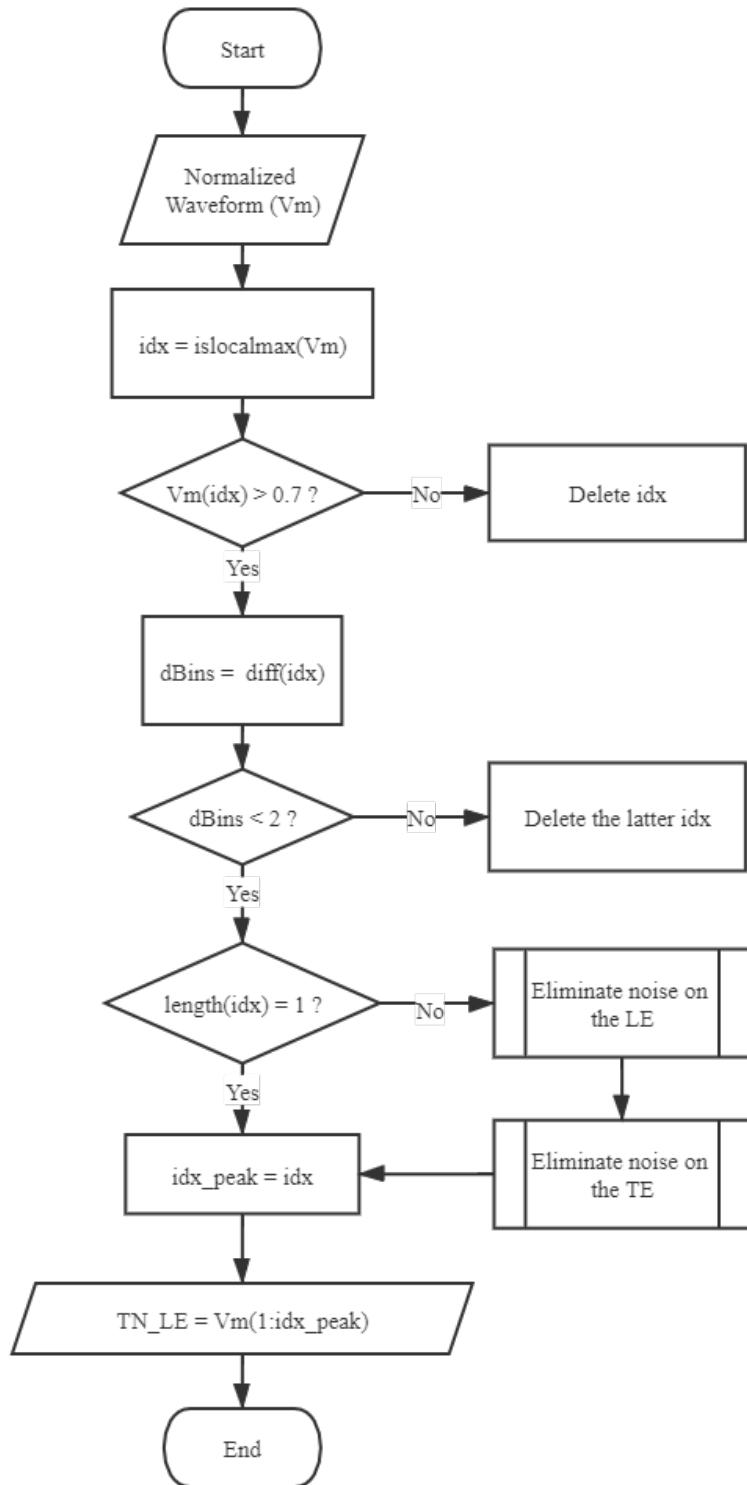
When we neglect the term of  $\beta_5$ , the model of the 5- $\beta$  retracker becomes:

$$y(t) = \beta_1 + \beta_2 P \left( \frac{t - \beta_3}{\beta_4} \right) \quad (4.33)$$

named the 4- $\beta$  retracker. The unknowns are  $\beta_1$  to  $\beta_4$ , and the dimension of  $y(t)$  is no longer equal to 128, but depends on the position of the peak (number of the bins in thermal noise and leading edge).



*Figure 4.8: An example of the algorithm. Num.1 oceanic PLRM waveform on 2018-12-12*



**Figure 4.9:** Flowchart of the algorithm to extract the thermal noise and the leading edge

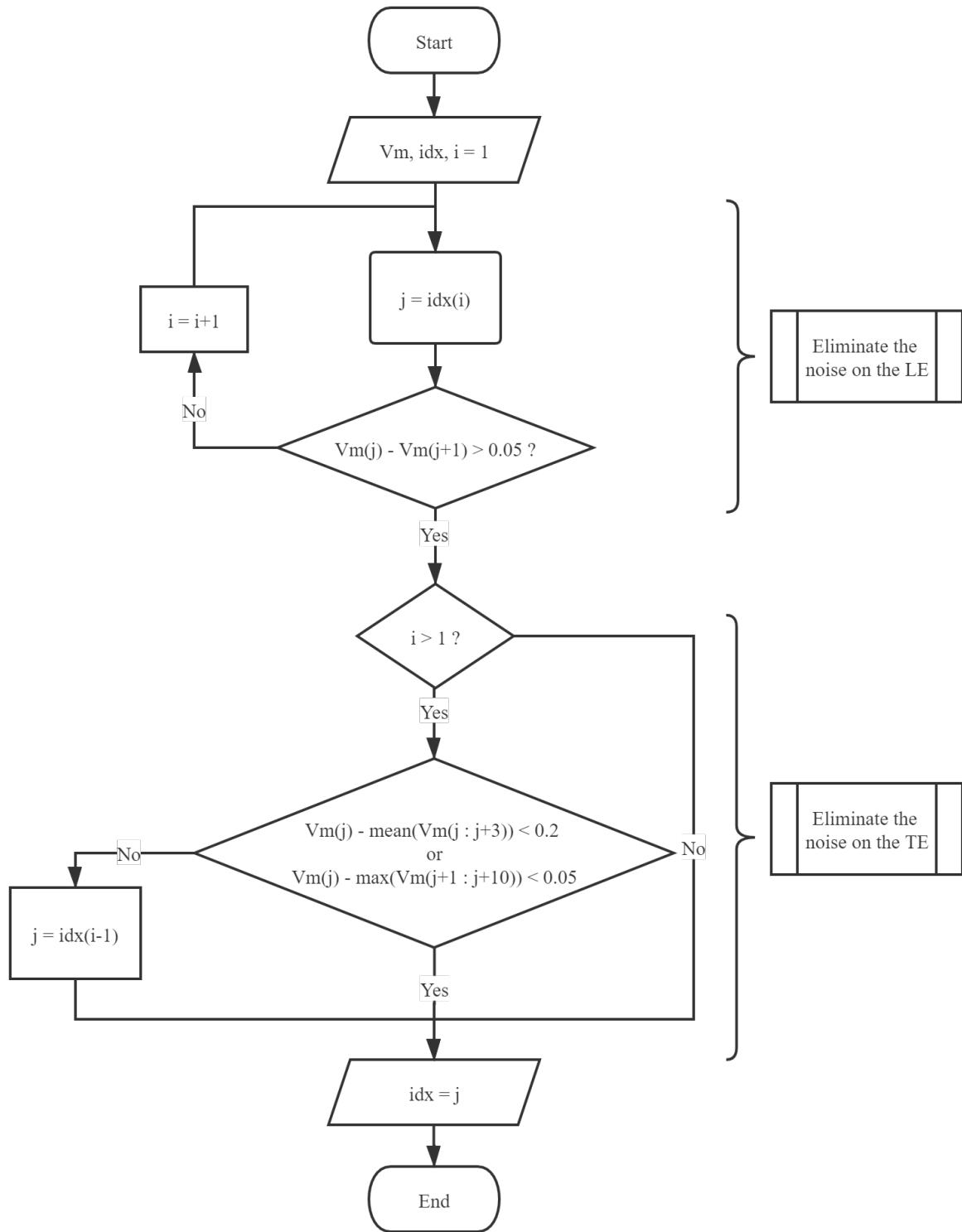


Figure 4.10: Flowchart of the algorithm to eliminate the noise

#### 4.4.3 Applying the $4\beta$ retracker using the Levenberg-Marquardt method

Now, we can use the  $4\beta$  retracker to fit the thermal noise and the leading edge by using the Levenberg-Marquardt method. The results compared with the  $5\beta$  retracker on the oceanic PLRM and SAR waveform are shown in Fig. 4.11 and Fig. 4.12, respectively.

As Fig. 4.11 shows, the fitted results using the  $4\beta$  retracker and the  $5\beta$  retracker on the leading edge of oceanic PLRM waveforms are basically the same. Only the maximal returned power of the  $4\beta$  fitted waveform is higher. The reason is, without the  $\beta_5$  term, the returned power of the leading edge is no longer reduced, as described in Section 4.4.2. In addition, the standard deviations of the  $4\beta$  fitted waveforms are also smaller, because the trailing edge has more noise than the leading edge, and this part of the noise has no effect on the  $4\beta$  fitted waveform.

The improvement of the  $4\beta$  retracker on the oceanic SAR waveform is pronounced, as shown in Fig. 4.12. Due to the reasons described in Section 4.2.3, the  $5\beta$  retracker cannot accurately fit the oceanic SAR waveform. The  $4\beta$  retracker has significantly improved the fitting result of the leading edge. Due to the working principle of SAR, returned power growth of the actual waveform is more gradual than the fitted waveform, which leads to the systematic error at the beginning of the leading edge. However, this error can be eliminated by choosing different thresholds, as described in Section 4.5.

The most significant contribution of the  $4\beta$  retracker is in the coastal zone. Because the coastal waveforms are very complicated, the retrackers based on the Brown model cannot provide reliable results. The  $4\beta$  retracker can better fit the leading edge. Fig. 4.13 shows four typical coastal waveforms. For the peaky waveform (a)) and the waveform with multi-peaks (d)), the  $4\beta$  retracker can fit both the thermal noise and the leading edge well. For the waveform with mild noise on the leading edge, the fitted thermal noise is too large. The reason is that the  $4\beta$  retracker treats the noise from other sources as thermal noise, but the fitted leading edge is still reliable. Since our primary goal is the estimation of the SWH, the thermal noise is also not essential for us. Therefore, the results are also acceptable. For the waveform with some abrupt noise on the leading edge (b)), the determined leading edge will be much more smoother than the real leading edge, which causes the determined  $\Delta$ Bins to be too large, see Section 5.2.

### 4.5 Determining the begin and the end of the leading edge

After we have successfully fitted the leading edge, the next step is to determine the starting and ending points of the leading edge as accurately as possible. The endpoint is well defined, which is the peak of the waveform. The starting point algorithm is to compare the difference with the forwards ( $D_{wf}$ ) of each bin. ALES [Passaro et al. (2014)] uses  $D_{wf} > 0.01$  for waveforms of Jason and Envisat, which will also be used in this thesis for PLRM waveform. For SAR waveform, as described in Section 4.4.3, the growth of the leading edge is more gradual. If the same  $D_{wf}$  is selected, the extracted  $\Delta$ Bins will be larger than the PLRM waveform, so we choose a larger threshold, i.e.  $D_{wf} > 0.03$ . This choice will be explained in detail in Section 4.7. Some examples of the results are shown in Fig. 4.14.

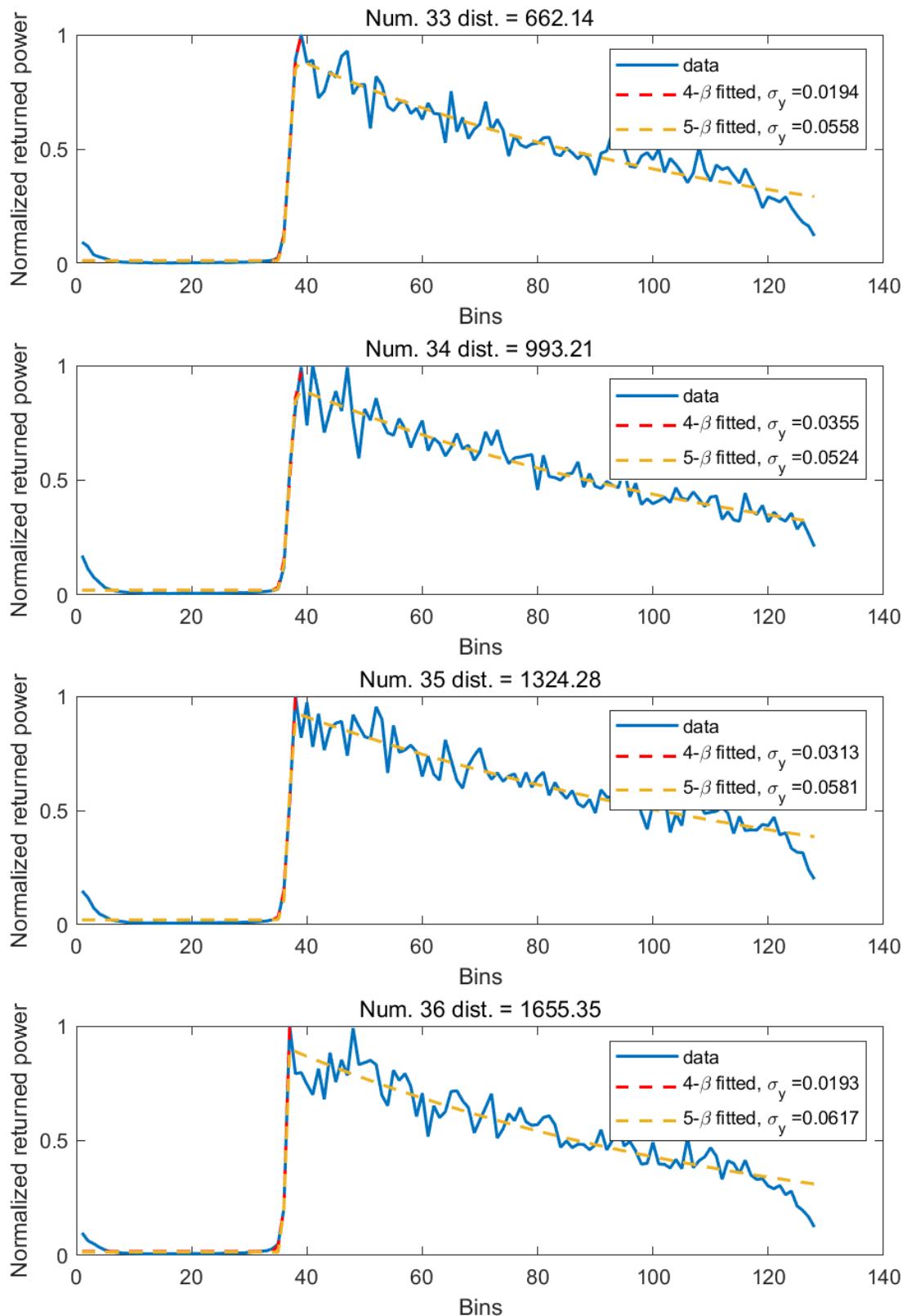


Figure 4.11: Examples of 5- $\beta$  and 4- $\beta$  fitted oceanic PLRM waveform on 2018-12-12 (Levenberg-Marquardt)

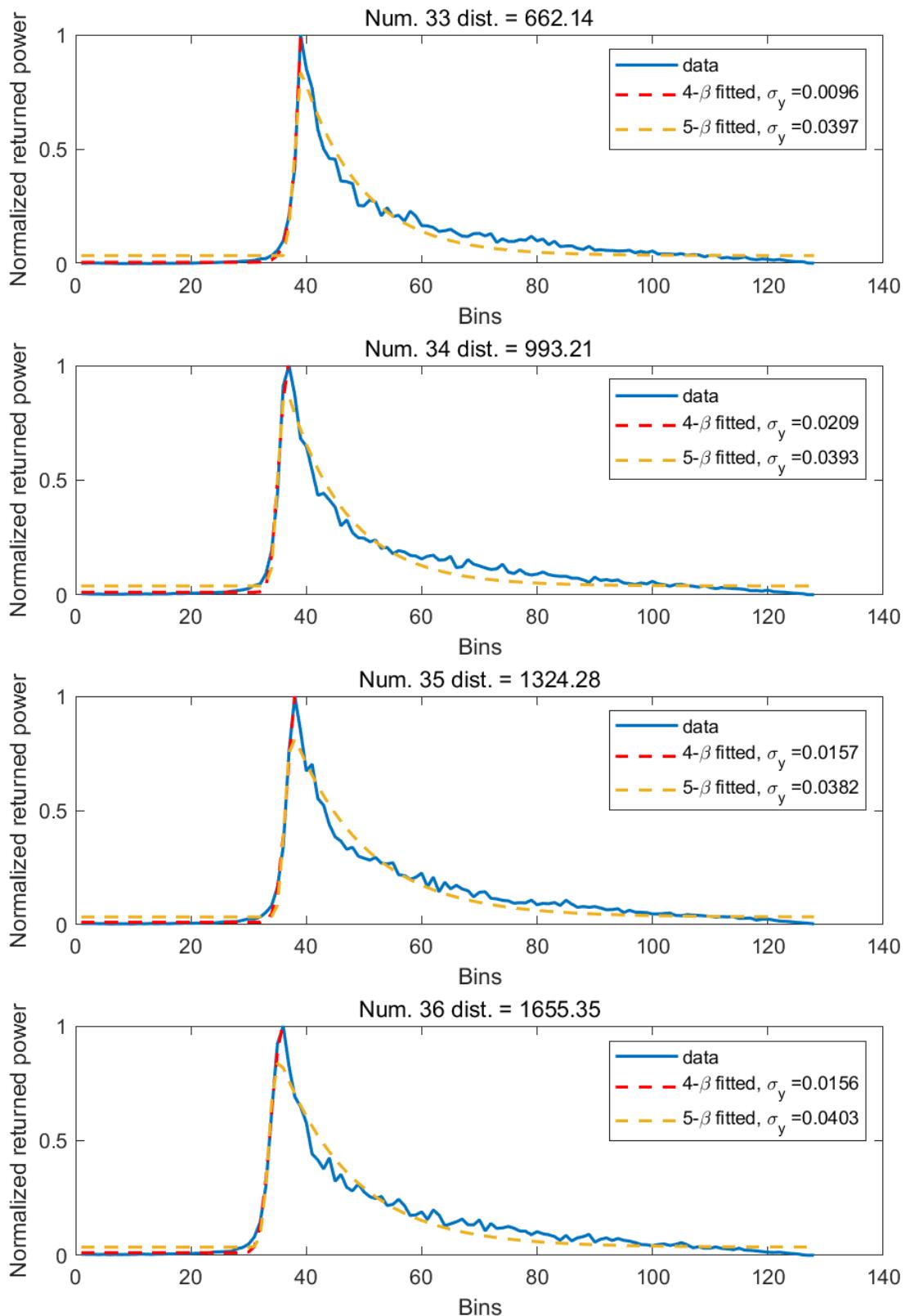
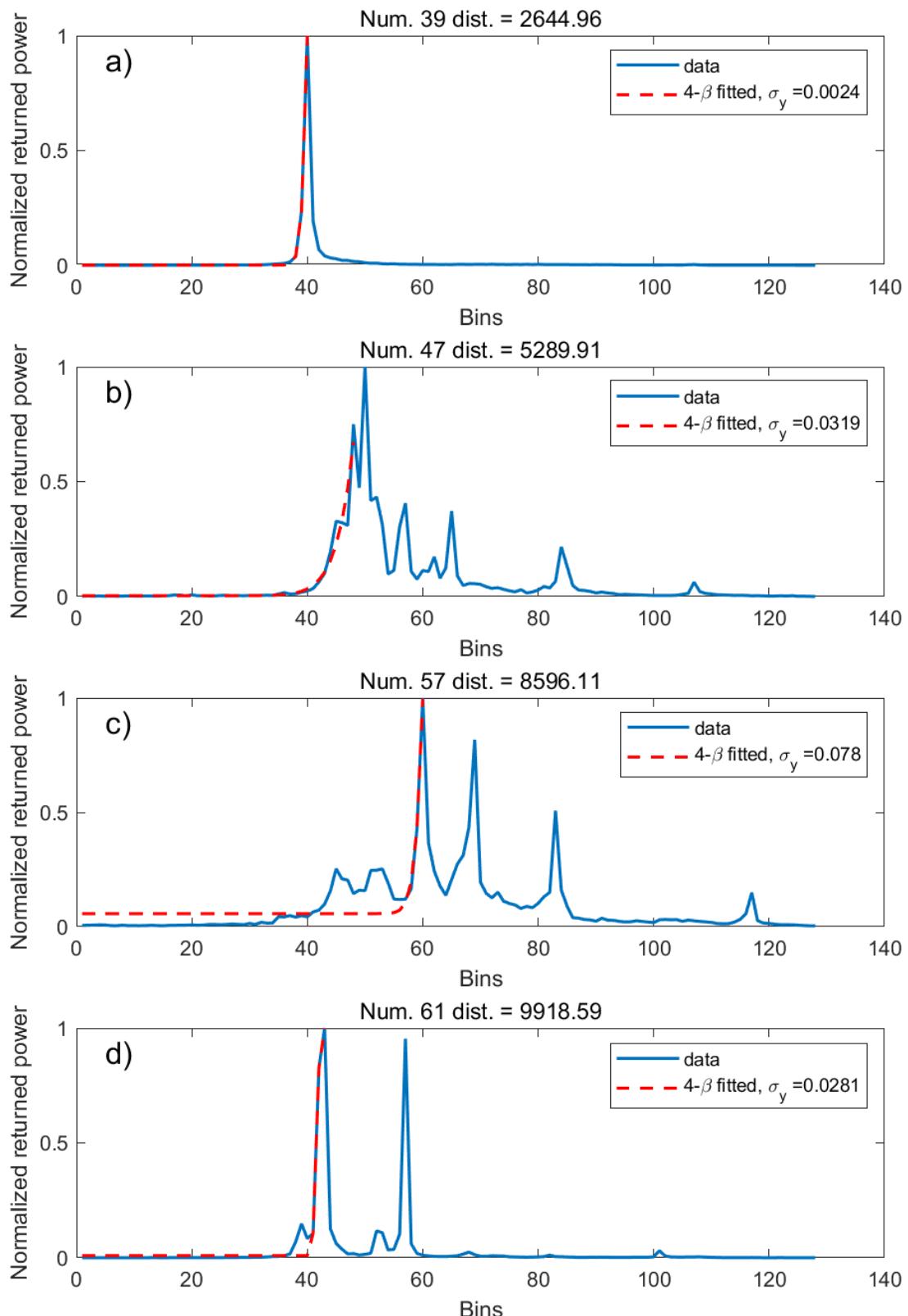


Figure 4.12: Examples of 5- $\beta$  and 4- $\beta$  fitted oceanic SAR waveform on 2018-12-12 (Levenberg-Marquardt)



**Figure 4.13:** Examples of 4- $\beta$  fitted coastal SAR waveform on 2018-12-12 (Levenberg-Marquardt), a) Sharp peak, b) Noise on the TE, c) Noise on LE and TE, d) Multi-peak

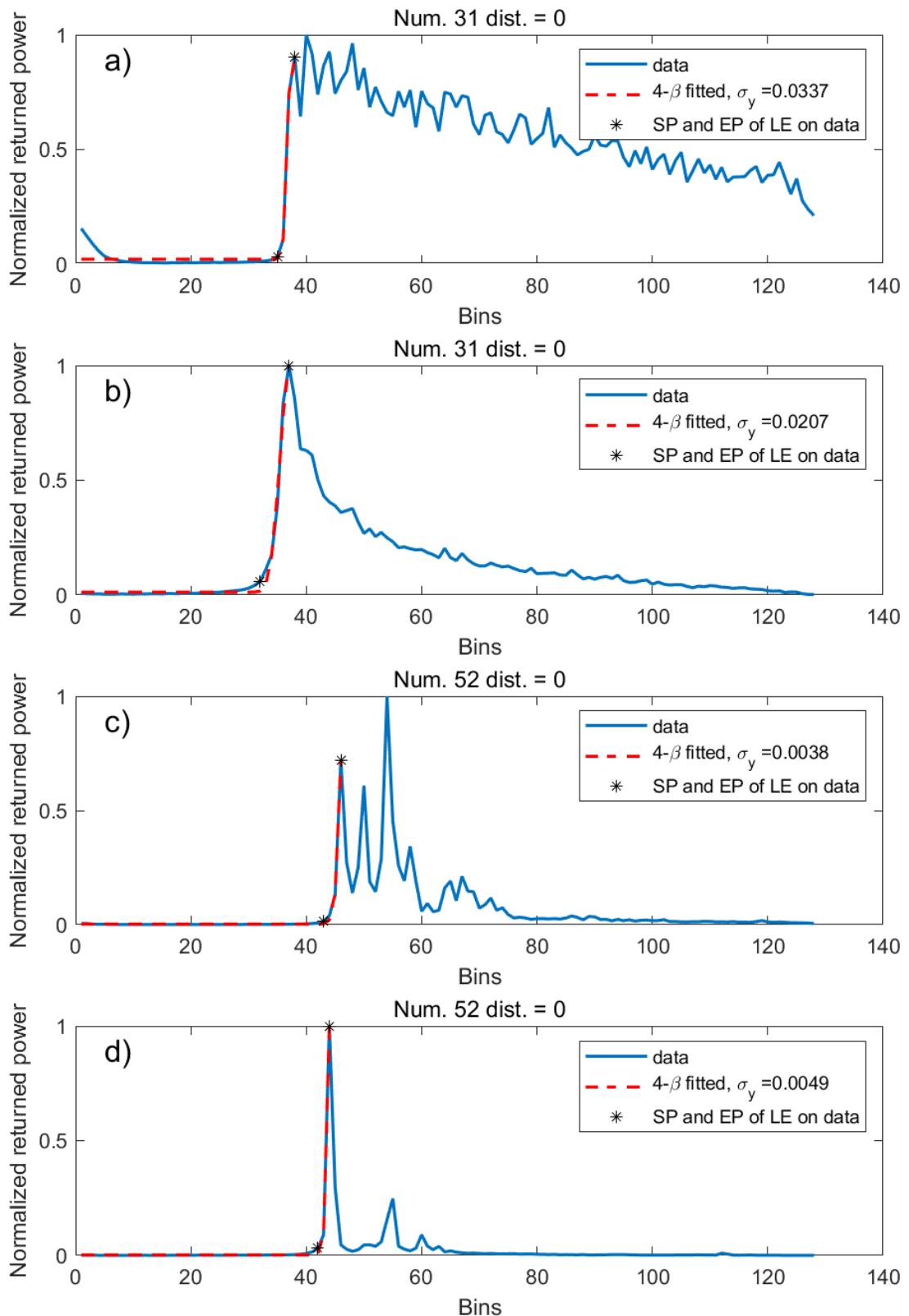
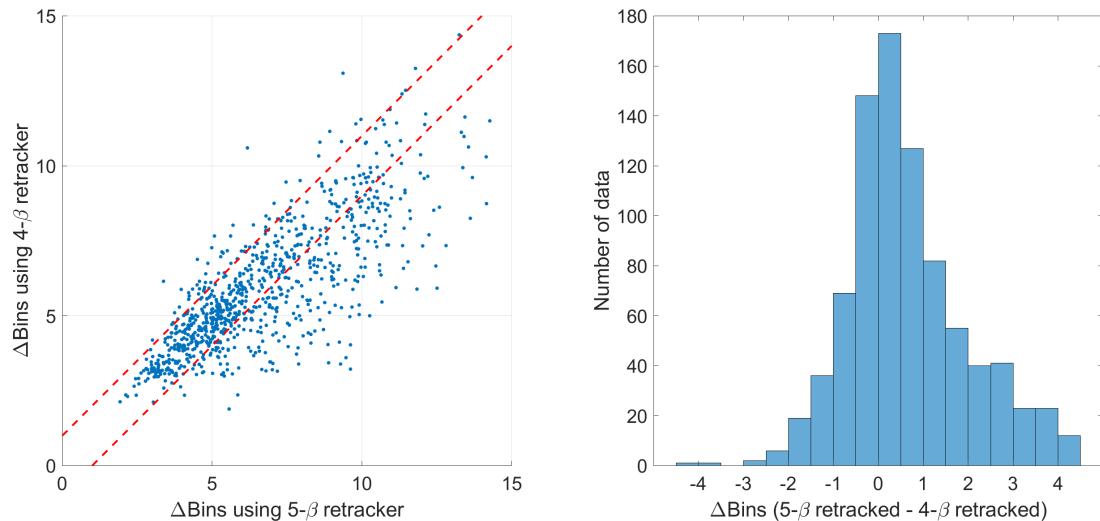


Figure 4.14: Examples of the leading edge. a) Oceanic PLRM, b) Oceanic SAR, c) Coastal PLRM, d) Coastal SAR

## 4.6 Quality of the $4\beta$ retracker

After we developed the  $4\beta$  retracker, we need to assess the reliability of the  $4\beta$  retracker. Since the  $5\beta$  retracker is suitable for oceanic PLRM waveforms, the reliability of the  $4\beta$  retracker will be analyzed by comparing  $\Delta$ Bins of the oceanic PLRM waveforms using the  $4\beta$  retracker and the  $5\beta$  retracker. All the oceanic waveforms between 2018-12-12 and 2019-12-25 have been used. The results are shown in Fig. 4.15.

Due to the limitation of resolution, the results whose  $\Delta$ Bins less than or equal to 1 are regarded as valid. There are a total of 882 data, of which 517 are valid. The valid ratio is 58.62%. The expected value of  $\Delta$ Bins is 0.6416, which means that there is a systematic error which causes the  $\Delta$ Bins using the  $5\beta$  retracker to be higher than the  $\Delta$ Bins using the  $4\beta$  retracker. The reason is that when the  $4\beta$  retracker is applied, the information of the trailing edge has been ignored, and sometimes due to the influence of the trailing edge, the end point of the leading edge using the  $5\beta$  retracker will move backwards. Nevertheless, the correlation of the  $\Delta$ Bins using the  $4\beta$  retracker and the  $5\beta$  retracker is 0.79, which means these results of these two retrackers are still highly correlated. Since the trailing edge of the SAR waveforms contains much less noise than the PLRM waveforms, the performance of the  $4\beta$  retracker on the SAR waveforms should be better than on the PLRM waveforms.



**Figure 4.15:** The scatter of the oceanic PLRM  $\Delta$ Bins using the  $4\beta$  and  $5\beta$  retracker (left) and the histogram of  $\Delta$ Bins (right)

In addition, Fig. 4.16 shows the expected values and standard deviations of  $\Delta$ bins on different dates. It can be seen from this figure that the expected values on different dates vary greatly. This may be due to different environmental conditions, such as wind speed. Therefore, we cannot directly add the above systematic error (0.6416) to the model. At the same time, the change in the expected values and the change in the standard deviations also show a correlation. A larger expected value is accompanied by a more significant standard deviation (e.g. 2019-01-08 and 2019-02-04). This means that the environmental conditions of these dates have caused larger errors between the  $4\beta$  fitted value and the  $5\beta$  fitted value. Some other environmental data are necessary for further analysis of specific reasons.

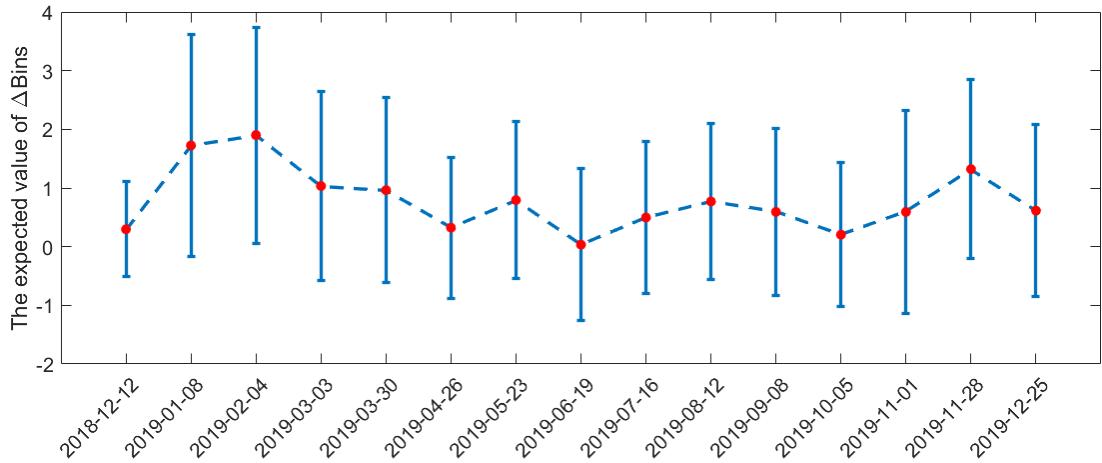


Figure 4.16: The expected values and the standard deviations of the  $\Delta$ Bins

## 4.7 Choosing the condition for the starting point of SAR waveforms

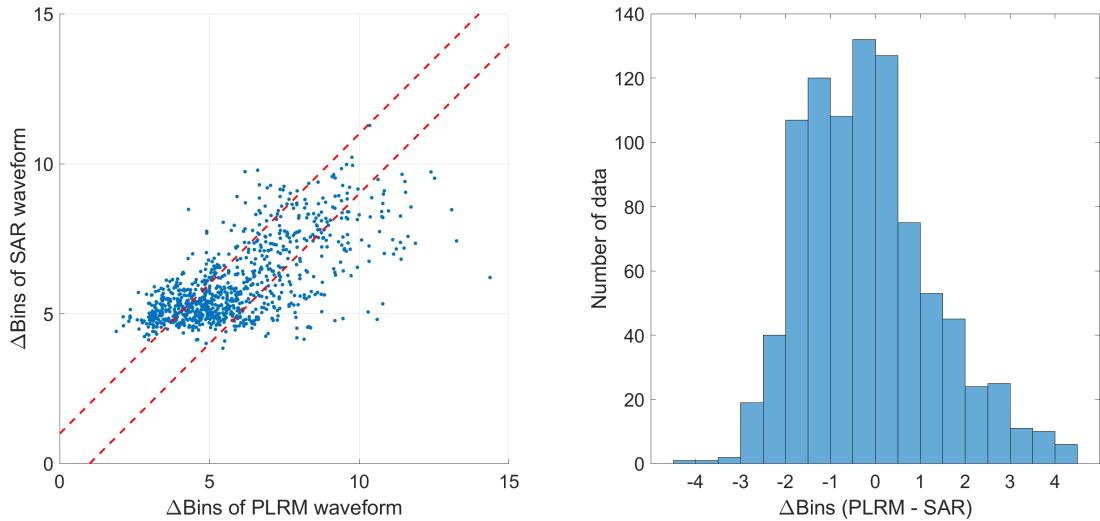
If we want to determine the  $\Delta$ Bins of the SAR waveforms by using the  $4\beta$  retracker, we need to discuss the condition for determining the starting point of the leading edge. The determined  $\Delta$ Bins of SAR waveforms using different thresholds will be compared with the  $\Delta$ Bins derived from PLRM waveforms. All the PLRM data used below are resampled based on the SAR data, as described in Section 3.3. Since the comparison here is not different retrackers of the same data, and the position of the PLRM data do not entirely match the position of the corresponding SAR data, we increase the confidence interval of the valid ratio from 1 bin to 2 bins, which gives a higher tolerance.

As described in Section 4.5, for PLRM waveforms,  $D_{wf} > 0.01$  is the condition for the starting point of the leading edge, while for SAR waveforms, a higher threshold is necessary. Table 4.1 shows the differences between  $\Delta$ Bins when choosing different  $D_{wf}$ . It can be seen from the table that different  $D_{wf}$  has little effect on the standard deviations, but has a significant impact on the expected values.  $D_{wf} > 0.03$  has the best expected value and the 2-bins valid ratio. Therefore, we choose  $D_{wf} > 0.03$  as the condition for the starting point of the SAR leading edge. The results are shown in Fig. 4.17.

$D_{wf}$	0.01	0.02	0.03	0.04	0.05	0.06	0.07
$E(\Delta\text{Bins})$	-1.22	-0.55	<b>-0.11</b>	0.22	0.49	0.73	0.94
$D(\Delta\text{Bins})$	<b>1.48</b>	1.50	1.54	1.58	1.61	1.64	1.68
valid ratio (2-bins)	64.70%	77.60%	<b>83.83%</b>	83.72%	81.31%	79.02%	76.72%

Table 4.1: The results of different  $D_{wf}$

We can also find in Fig. 4.17 that the shift of the PLRM  $\Delta$ Bins is more prominent, which means, when the SAR  $\Delta$ Bins are same, the PLRM  $\Delta$ Bins are not stable. The possible reason is mentioned above: The trailing edge will affect the determination of the leading edge, thereby affecting the performance of the  $4\beta$  retracker on the PLRM waveforms.

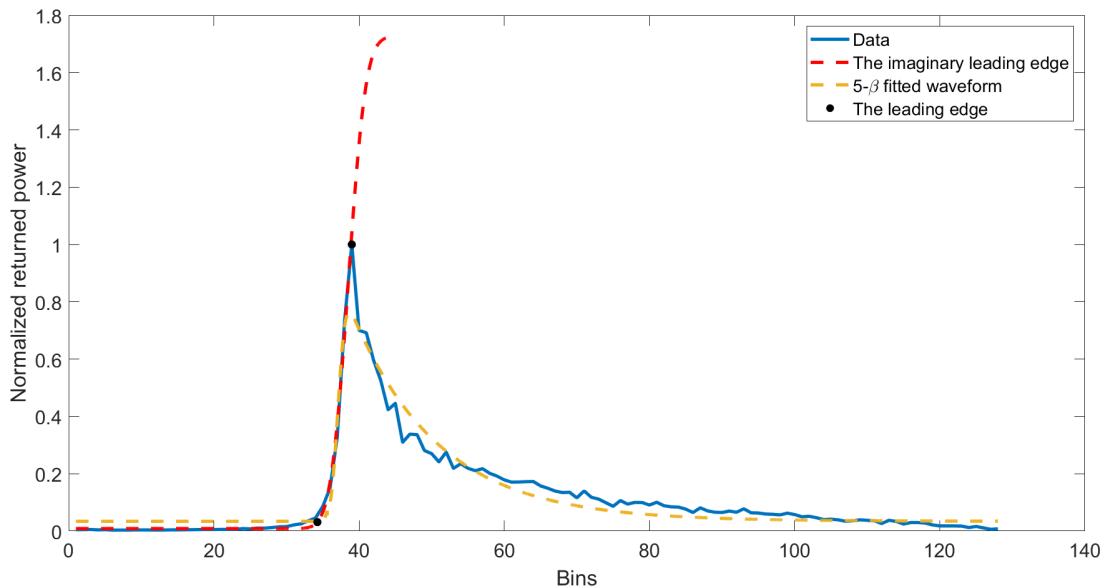


**Figure 4.17:** The scatter of the oceanic PLRM and SAR  $\Delta\text{Bins}$  using the 4- $\beta$  retracker (left) and the histogram of  $\Delta\text{Bins}$  (right)

## 4.8 Relationship between $\beta_4$ and $\Delta\text{Bins}$

Now we can use the 4- $\beta$  retracker to fit the leading edge independent of the tailing edge more accurately. However, we cannot use the estimated  $\beta_4$  to determine the SWH directly, because the 4- $\beta$  retracker fits the leading edge using a part of an imaginary ‘leading edge’ with much larger amplitude and longer raising time. Therefore, the parameters here are not related to the physical model as in the 5- $\beta$  retracker (Eq. 2.19 to Eq. 2.22).

Fig. 4.18 shows an example. The estimated parameters using the 4- $\beta$  and 5- $\beta$  retracker are also shown in the following:



**Figure 4.18:** An example of the non-physical model

The  $\beta$  parameters estimated by the 4- $\beta$  retracker are:

$$\beta_{41} = 0.0088 \quad (4.34)$$

$$\beta_{42} = 1.7173 \quad (4.35)$$

$$\beta_{43} = 38.5134 \quad (4.36)$$

$$\beta_{44} = 1.9108 \quad (4.37)$$

The  $\beta$  parameters estimated by the 5- $\beta$  retracker of the same waveform are:

$$\beta_{51} = 0.0342 \quad (4.38)$$

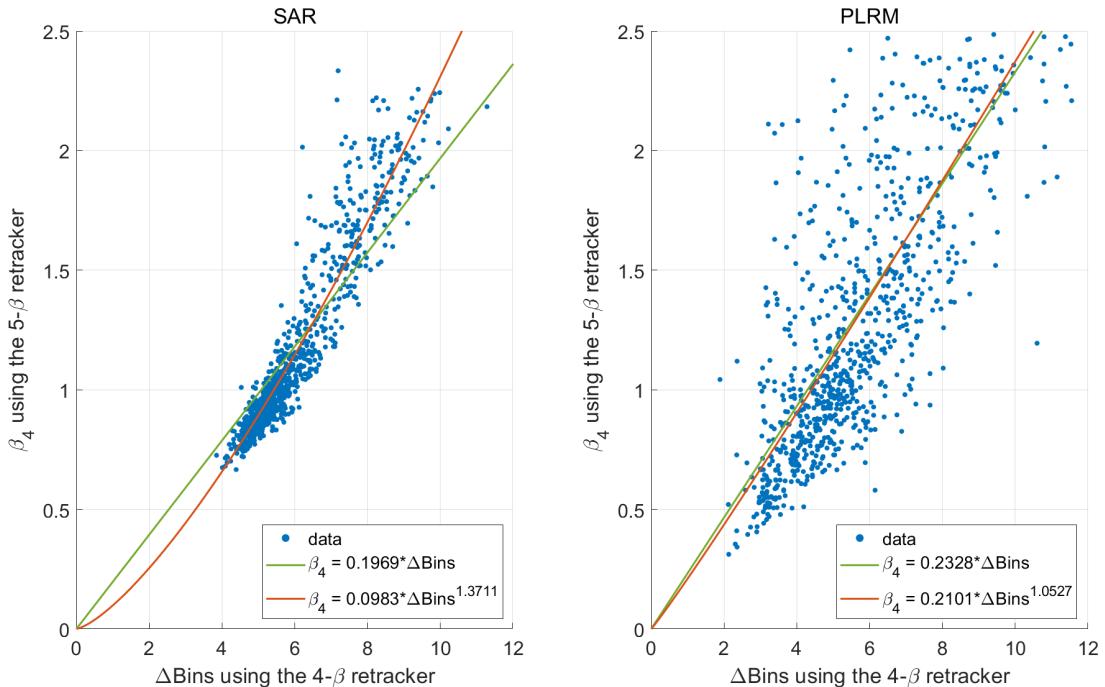
$$\beta_{52} = 1.6297 \quad (4.39)$$

$$\beta_{53} = 37.1740 \quad (4.40)$$

$$\beta_{54} = 0.7812 \quad (4.41)$$

$$\beta_{55} = 0.0840 \quad (4.42)$$

However, a part of the imaginary ‘leading edge’ fits the real leading edge better, and we can get the  $\Delta$ Bins more accurately from it. We could derive the real  $\beta_4$  according to the relationship between the  $\Delta$ Bins using the 4- $\beta$  retracker and the  $\beta_4$  using the 5- $\beta$  retracker. Then, we can determine the SWH. The relationship is shown in Fig. 4.19. We can assume a linear or power function model across the origin to fit the scatters in the figure to obtain the expression of the relationship between  $\Delta$ Bins and  $\beta_4$ .



**Figure 4.19:** The relationship between  $\Delta$ Bins and  $\beta_4$  with two fitting lines

From Fig. 4.19 we can find that the PLRM data are more dispersed than the SAR data, which means the  $\Delta$ Bins and the  $\beta_4$  of the SAR data show a stronger relationship. Therefore, the SAR fitting result is better than the PLRM fitting result. The reason is that the PLRM waveform contains much noise. Although we try to avoid the noise, this noise still has some influence on

the determination of the peak of the waveform. Therefore, we use SAR data to determine the relationship between  $\Delta\text{Bins}$  and  $\beta_4$ . The functional expression of the relationship is adjusted using the Gauss-Markov model.

The fitted line using the linear function with a RMSE of 0.1654:

$$\beta_4 = 0.1969 \cdot \Delta\text{Bins} \quad (4.43)$$

The fitted line using the power function with a RMSE of 0.1292:

$$\beta_4 = 0.0983 \cdot \Delta\text{Bins}^{1.3711} \quad (4.44)$$

The results show that the power function is a better model for the relationship between  $\Delta\text{Bins}$  and  $\beta_4$  because of the lower RMSE. This conclusion will be proved again later by comparing the estimated oceanic SWH.

When we get the relationship between  $\Delta\text{Bins}$  and  $\beta_4$ , we can determine the corresponding  $\beta_4$  by determining  $\Delta\text{Bins}$ . Considering Eq. 2.4 and Eq. 2.22, we can determine the SWH, even in the coastal area by using the following equation:

$$\text{SWH} = 2 \cdot c \cdot \sigma_s = 2 \cdot c \cdot \sqrt{(\beta_4 \cdot r_t)^2 - \sigma_p^2} \quad (4.45)$$

where  $r_t = 3.125 \text{ ns}$  by Sentinel-3,  $\sigma_p = 0.513 \cdot r_t$ ,  $c$  is the speed of light.

Eq. 4.45 shows a problem that when the  $\beta_4$  is smaller than 0.513, the  $\sigma_s$  is a complex number. Then, we cannot calculate the SWH correctly. The solution is that when the  $\beta_4$  is smaller than 0.513, we do not consider the  $\sigma_p$  term. The SWH will be calculated using the following equation:

$$\text{SWH} = 2 \cdot c \cdot \beta_4 \cdot r_t \quad (4.46)$$

This is a temporary solution because the  $\sigma_p$  should still have some influences. The SWH values derived using Eq. 4.46 are bigger than the real values. Some research similar to the Jason 1 altimeter ground processing look up correction tables [Thibaut et al. (2004)] for Sentinel-3 maybe valuable to solve this problem.

# Chapter 5

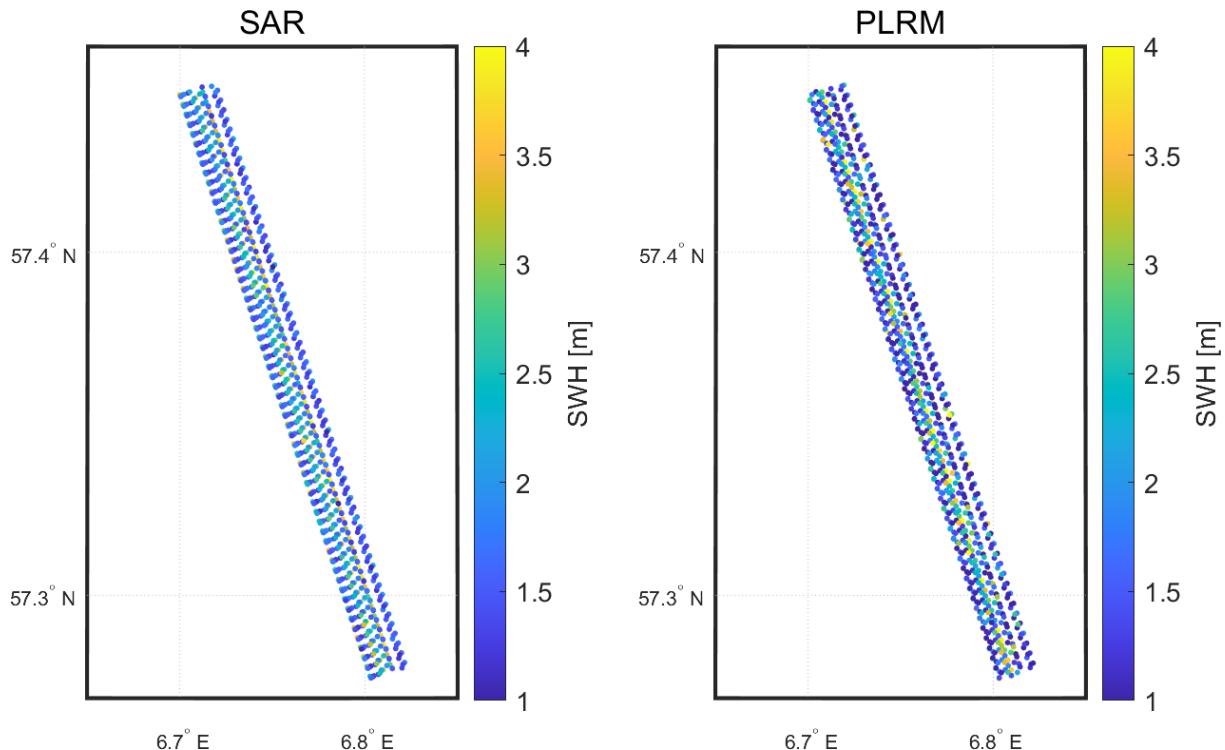
## Results

Now we can use our method to determine SWH. Firstly, we apply the  $4\beta$  retracker and the both fitting algorithms using the linear function and the power function in the oceanic area. The advantage of the power function will be proved again. Then, the results in the coastal area will be reported.

### 5.1 Oceanic area

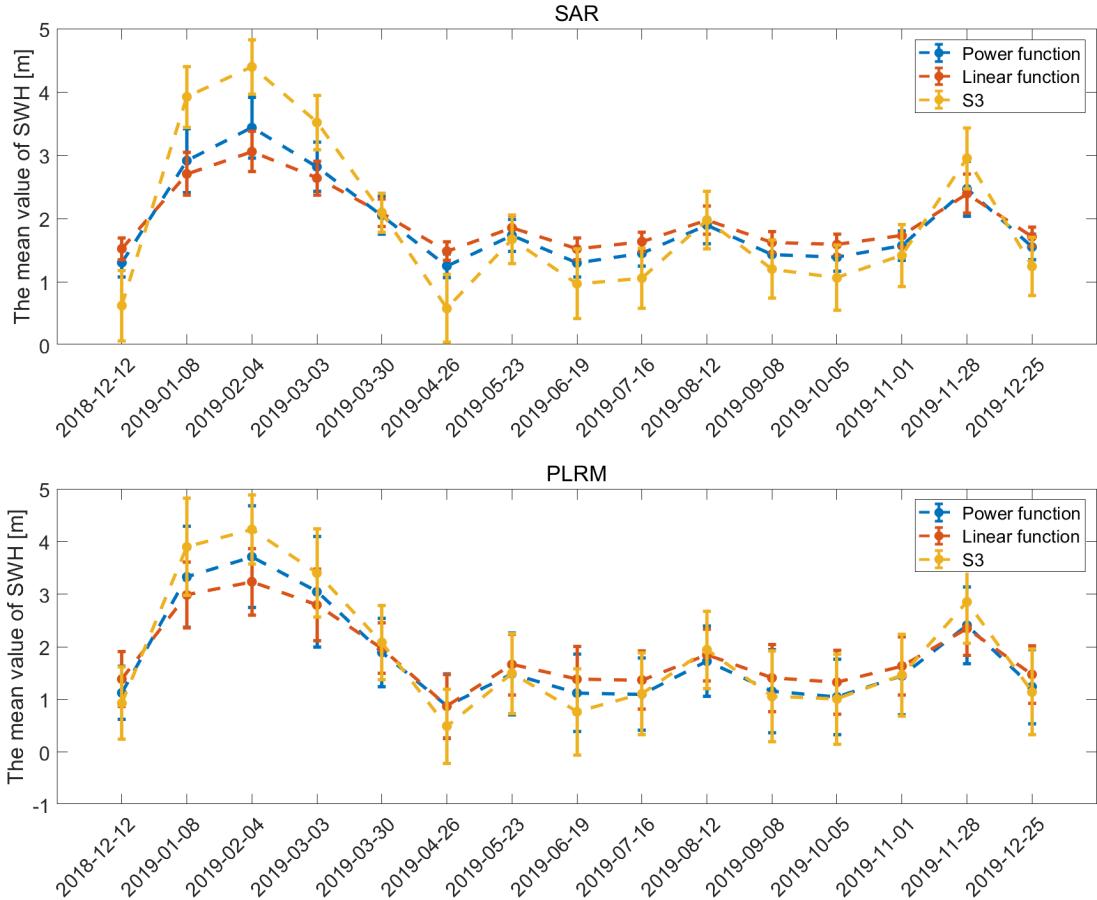
#### 5.1.1 Map and time series of the SWH

The results of apparently failed fitting ( $\sigma > 0.05$ ) will be deleted. The single SWH (using the power function) depending on the location is shown in Fig. 5.1.



**Figure 5.1:** Map of the significant wave heights in the oceanic area between 2018-12-12 and 2019-12-25. Left: SAR, Right: PLRM

The time series of SWH can be reached by calculating the median values of SWH on different dates as shown in Fig. 5.2. The median values are the better choice than the mean values for our study because they could avoid the influence of the blunders. From the time series we can find that the fitting results using the power function show a better quality because they fit the results provided by Sentinel-3 better. Fig. 5.2 confirms the conclusion in Section 4.8. Therefore, we will choose the fitting model using power function for our further research.



**Figure 5.2:** The median values and the standard deviations of the significant wave heights in the oceanic area between 2018-12-12 and 2019-12-25

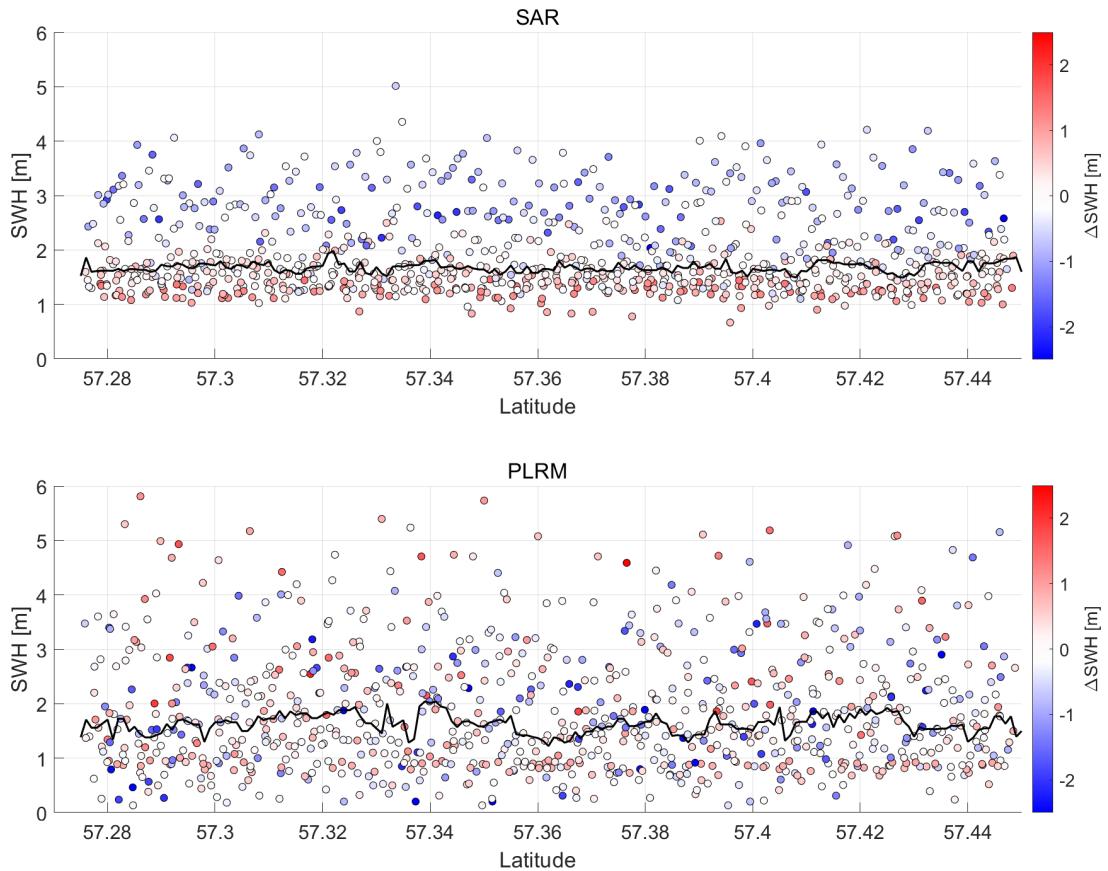
### 5.1.2 Discussion about the quality of the results

Besides, we can also find from Fig. 5.2 that because of the fitting algorithm, the results using the 4- $\beta$  retracker are smoothed. The differences between the estimated SWH and the SWH provided by Sentinel-3 are shown in detail in Fig. 5.3. The  $\Delta$ SWH is defined as the estimated SWH minus the SWH provided by Sentinel-3. The black line shows the median values in a searching window with the width of  $0.005^\circ$  and the step size of  $0.001^\circ$ . Observing the median values in the searching window, we can find that the data are stable between 1.5 m and 2 m, which means that the median values of SWH observed at different times are basically the same in this oceanic area and have not changed significantly with position (latitude). The waveforms are also not obviously affected by the human factors. This phenomenon is more

prominent when we compare the results with that in the coastal area.

By comparing SAR and PLRM, we can see that the stability of SAR SWH is much stronger than that of PLRM SWH, and the variation of the  $\Delta$ SWH (the smooth influence) shows a more systematic trend, i.e. when SWH is larger, estimated SWH is smaller than Sentinel-3 provided SWH, and when SWH is smaller, estimated SWH is higher than Sentinel-3 provided SWH. At the same time, the red and blue points of PLRM are all over the whole interval, which does not show such a trend.

If we look at the performance of the  $4\beta$  retracker, we can find that it performs better on the SAR SWH than the PLRM SWH (see Fig. 5.4). One possible reason is that although the algorithm has tried to avoid the influence of the noise on the trailing edge, the determination of the peak point is sometimes still interfered by that noise. Since the trailing edge of the SAR waveform contains much less noise than the PLRM waveforms, the determination of peak points is more accurate, and the results obtained by the  $4\beta$  retracker are more reliable. This conclusion will be proved again in the coastal area.



**Figure 5.3:** The scatters of the SWH and the median values in the searching windows (oceanic area)

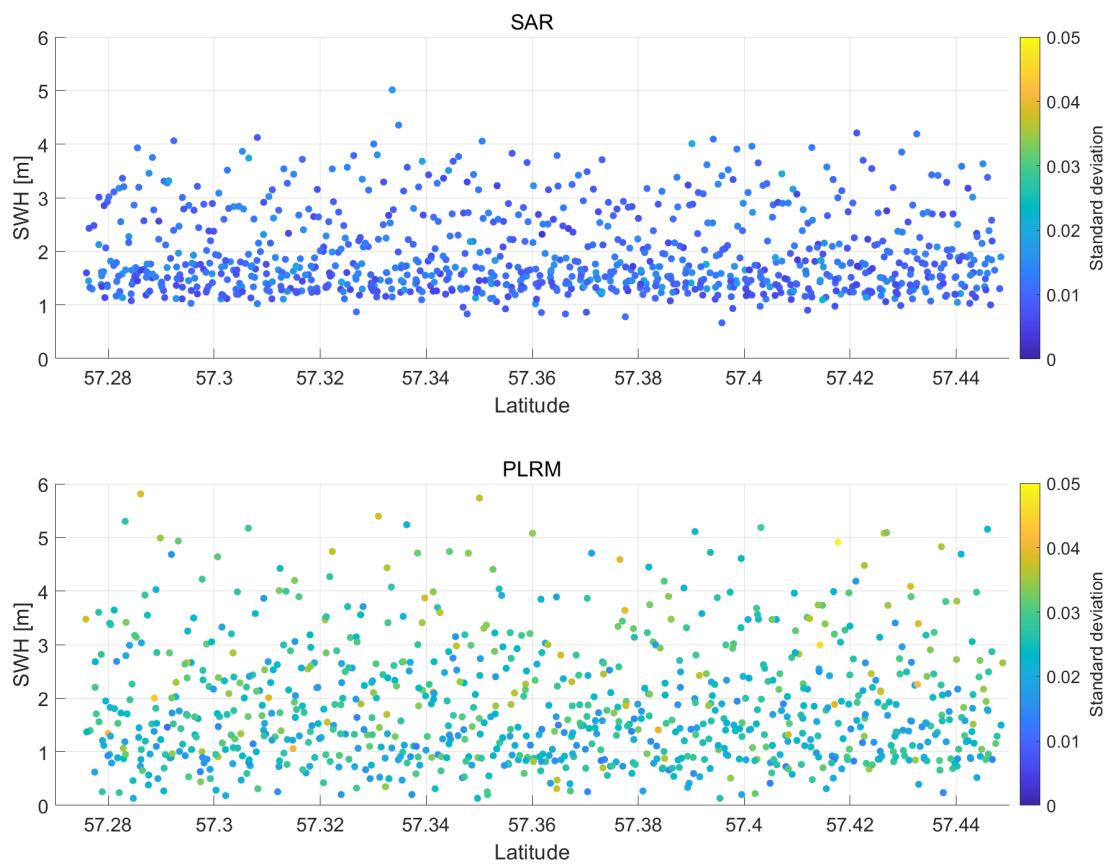


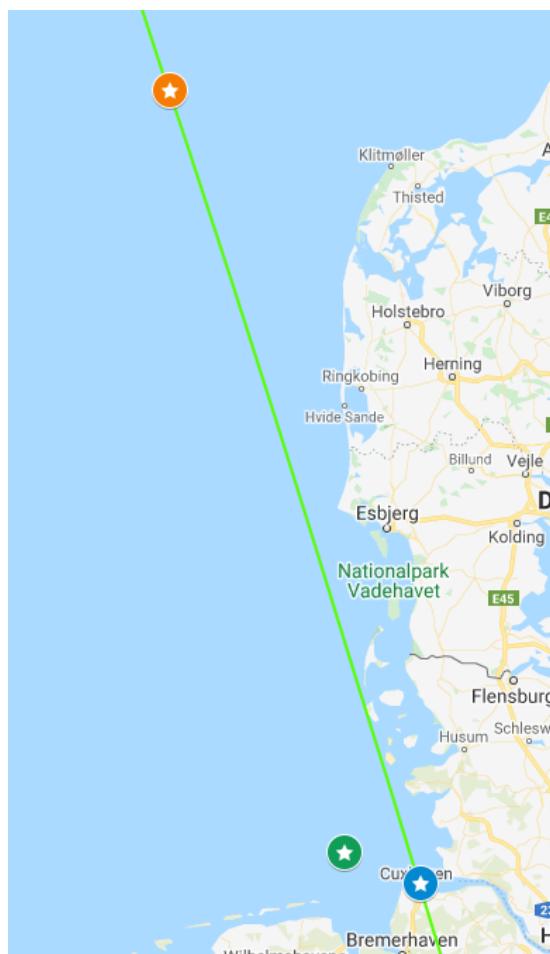
Figure 5.4: The scatters of the SWH and the standard deviations of the 4- $\beta$  retracked waveforms (oceanic area)

## 5.2 Coastal area

### 5.2.1 In-situ data

Since the Sentinel-3 cannot provide meaningful SWH in the coastal area, we need to find another way to verify our results in the coastal area. Except satellite altimetry, measuring stations are still one of the most common way to collect the information of water surfaces.

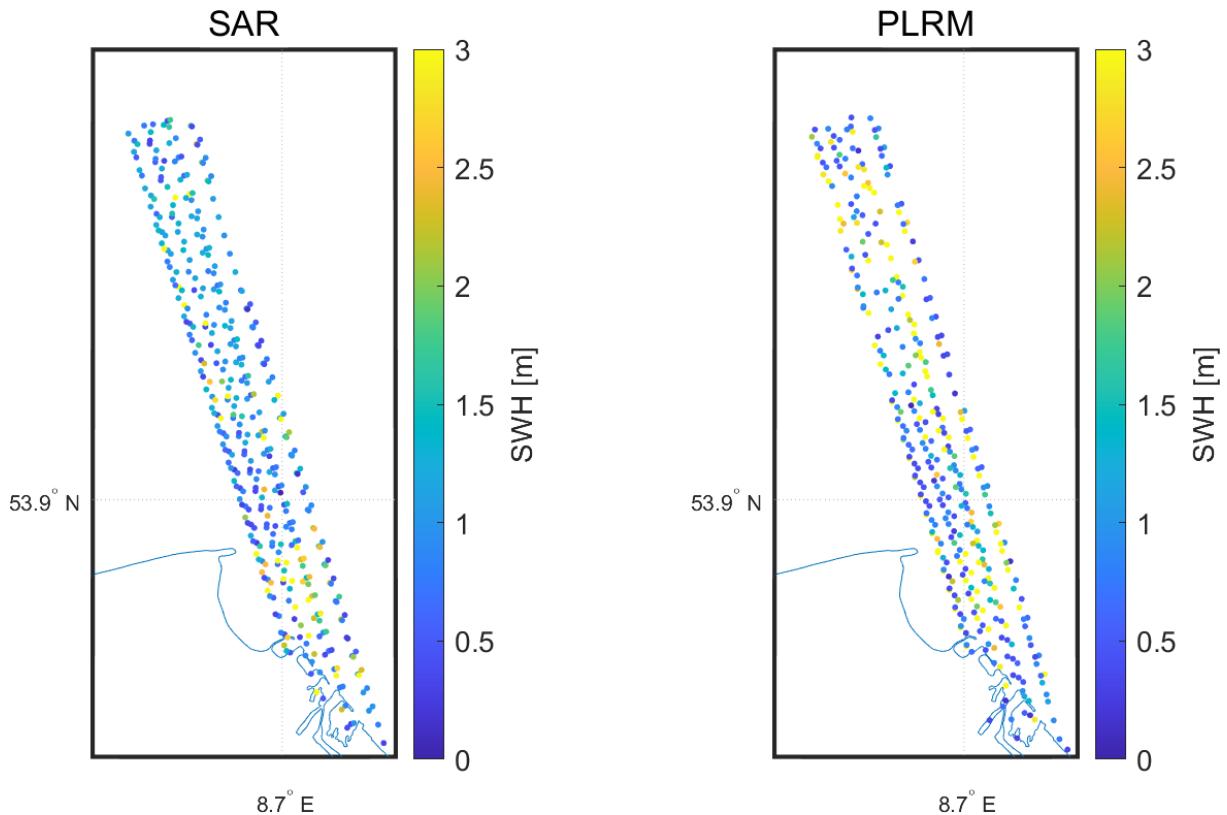
The Federal Maritime and Hydrographic Agency of Germany, in German Bundesamt für Seeschifffahrt und Hydrographie (BSH) operationally collects measurements of the parameters wave height, peak period and wave direction. We can use their measuring station in North Sea to get the in-situ SWH to compare with our results. We have chosen the measuring station *Elbe* (54.0167, 8.1139), which is about 40 km far from the coastal virtual station close to Cuxhaven, see Fig. 5.5. The time resolution of the Elbe provided data is about 30 min. We will take the median values as well to avoid the influence of the blunders.



**Figure 5.5:** The measuring station Elbe (green star), the coastal virtual station (blue star) and the oceanic virtual station (orange star)

### 5.2.2 Map and time series of the SWH

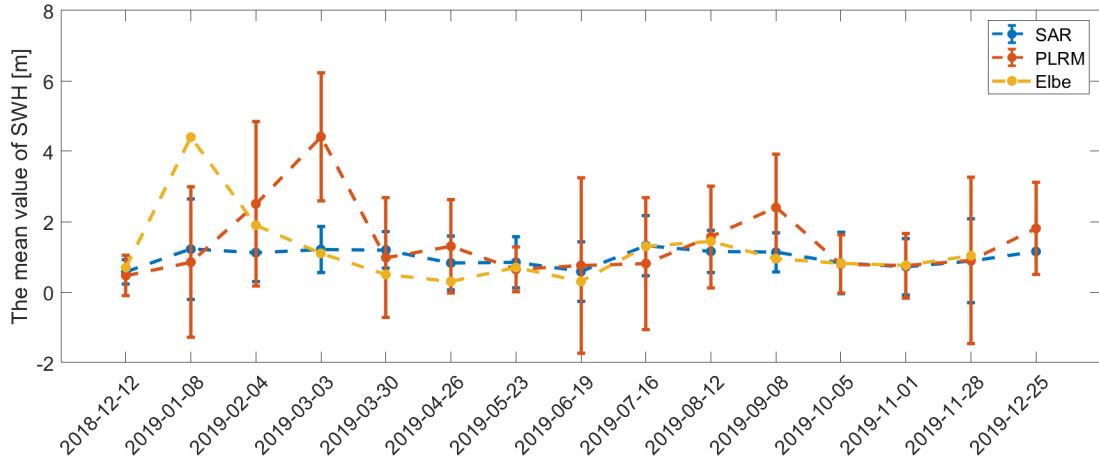
In the coastal area, not only the results of apparently failed fitting but also the results on the land will be deleted. The results are shown in Fig. 5.6. We can find that the results of PLRM are significantly less than SAR, because the PLRM waveforms are more sensitive to the noise, causing more waveform fitting failures. At the same time, comparing Fig. 5.6 and Fig. 5.1, it can be found that the SAR SWH contains more noise in the coastal area than in the oceanic area. Although the SAR waveforms contain much less noise than the PLRM waveforms, the accuracy of the fitted results in the coastal area will be lower than in the oceanic area because the environmental situation in the coastal area (especially the haven) is too complicated.



**Figure 5.6:** Map of the significant wave heights in the coastal area between 2018-12-12 and 2019-12-25. Left: SAR, Right: PLRM

The time series can be generated in the same way as in the oceanic area. The results are shown in Fig. 5.7. The figure shows that the SAR results have a much higher quality than the PLRM results, which fits our expectations. The SAR SWH and the in-situ data provided by the Elbe are in good agreement with each other, except on 2019-01-08. It can be seen that on 2019-01-08, the sea level was relatively unstable with a higher SWH. This is also reflected in the satellite data that the standard deviation of SAR SWH on that day is much higher than other dates, which means the SAR SWH on 2019-01-08 is not that reliable as on the other dates. At the same time, the Elbe data on that date is not that reliable as well. The high wind speed and waves may affect the reading of the buoy. Comparing Fig. 5.2 and Fig. 5.7, we can find that the SWH in the oceanic area on 2019-02-04 is higher than the data on 2019-01-08. In the Elbe data,

the SWH on 2019-01-08 is much higher than the SWH on 2019-02-04 and is even close to the oceanic SWH on the same day, which may not match the real situation.



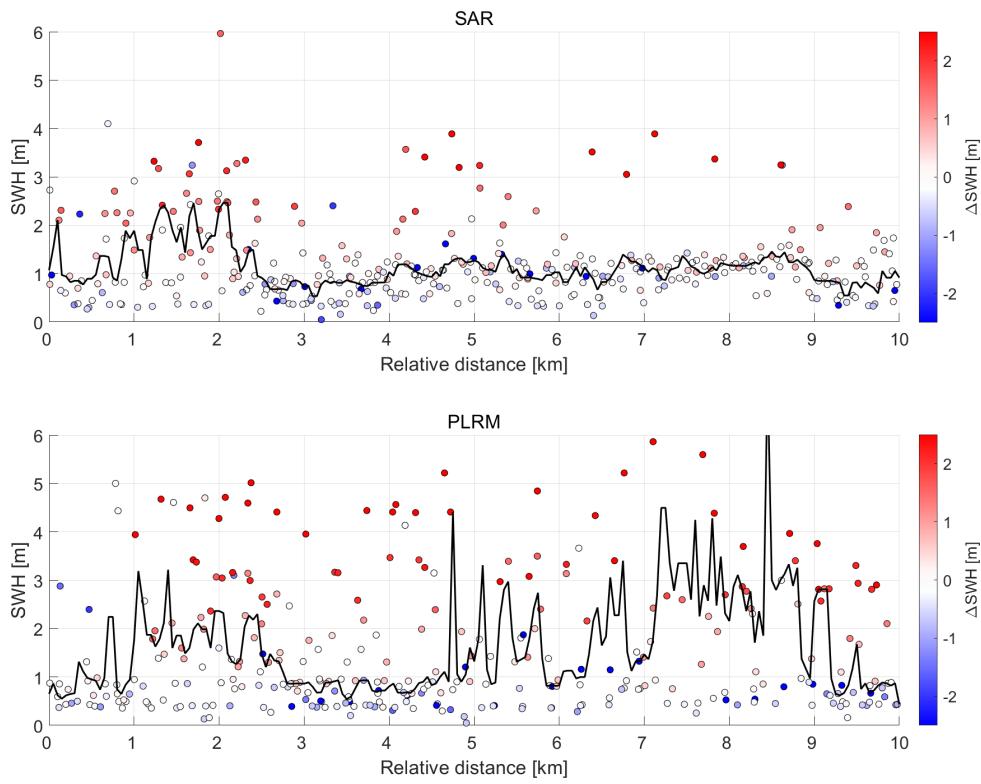
**Figure 5.7:** The median values and the standard deviations of the significant wave heights in the coastal area between 2018-12-12 and 2019-12-25

### 5.2.3 Discussion about the quality of the results

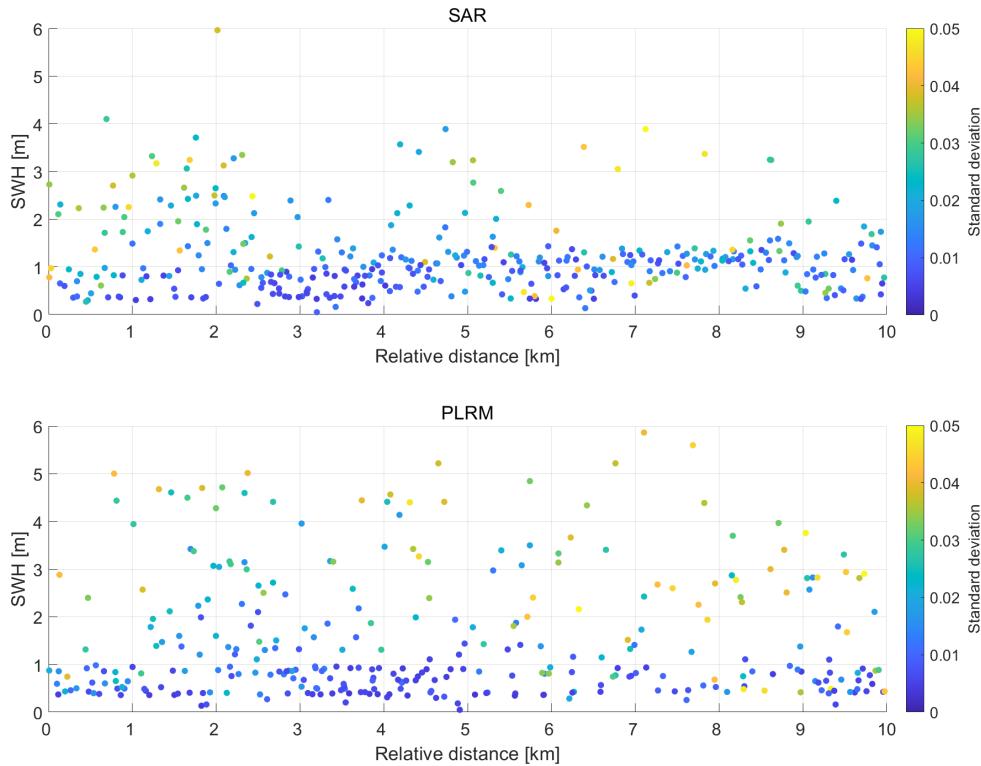
In the coastal area, we can use the same method to evaluate the quality of the estimated SWH. Since we do not have a one-to-one in-situ SWH corresponding to the satellite data, here the  $\Delta$ SWH is defined as the estimated SWH minus the median of the Elbe SWH of the day. The searching window is based on the relative distances between the location of data and the virtual station. The width of the searching window is 250 m, and the step size is 50 m. The result is shown in Fig. 5.8. The y-axis (SWH values) is limited between 0 and 6 m to make the comparison between SAR and PLRM clearer. It should be mentioned that all the SAR data lay in this interval whereas some PLRM data are out of this interval (maximal blunder up to 14 m).

From the figure, we can find that, after 3 km (approximately 1 km offshore), the SAR SWH gradually becomes stable (a clear cluster appears), which means that the quality of the SAR SWH becomes higher. Although there is still some noise, it can already provide more reliable SWH data. At the same time, the PLRM data are still very dispersed and cannot provide reliable data. This is where the advantages of SAR altimeter lay. The higher along-track resolution of SAR significantly reduces the complicated environmental impacts on the leading edge, as described in Chapter 2. Therefore, the 4- $\beta$  retracker can successfully extract the leading edge of the waveform to determine the SWH further.

Fig. 5.9 shows the quality of the 4- $\beta$  retracked waveforms, which proves that the SAR results are much better than the PLRM results in the coastal area again. After 2.5 km the performance of the 4- $\beta$  retracker on the SAR waveforms is better than within 2.5 km. However, the performance of the 4- $\beta$  retracker on the PLRM waveforms does not show an improvement with the distance.



*Figure 5.8: The scatters of the SWH and the median values in the searching windows (coastal area)*



*Figure 5.9: The scatters of the SWH and the standard deviations of the 4- $\beta$  retracked waveforms (coastal area)*

# Chapter 6

## Conclusion and outlook

### 6.1 Summary and Conclusion

Due to the limitation of the number and location, tide gauge stations around the world cannot provide a sufficient amount of in-situ data. Therefore, satellite altimetry plays an increasingly important role in the hydrological analysis. The use of SAR also greatly improves the accuracy of waveform analysis. Nevertheless, due to the complicated situation of the coastal waveforms and excessive interferences, the onboard retracker of Sentinel-3 cannot provide SWH data in the coastal area. Therefore, based on the already mature  $5\beta$  retracker and the Brown model, this thesis developed a retracker ( $4\beta$  retracker), which avoids the effect of trailing edge, making it useful for analyzing the coastal waveforms, thereby determining the coastal SWH.

All in all, the main contributions of this thesis can be summarized as follows:

- Found the limitation of Gauss-Markov model for the waveform adjustment, and chose Levenberg-Marquardt method to adjust the waveform.
- Developed the algorithm to determine the peak of the leading edge, which can avoid noise interferences and extract the leading edge more accurately.
- Due to the different character of the SAR waveforms and the PLRM waveforms,  $Dwf > 0.03$  is selected as the condition to determine the starting point of the leading edge for the SAR waveforms, and for the PLRM waveforms is  $Dwf > 0.01$ .
- Developed the  $4\beta$  retracker and verified the quality of the  $4\beta$  retracker in the oceanic area.
- Found the relationship between the raising time ( $\Delta\text{Bins}$ ) and the  $\beta_4(\sigma_c)$ . Made it possible to determine the SWH using  $\Delta\text{Bins}$ .
- Applied the  $4\beta$  retracker in the coastal area (Cuxhaven) to generate the time series of the significant wave heights in the coastal area successfully.
- Proved that the performance of the SAR altimeter is much better than the PLRM altimeter in the coastal area.

We have employed the developed methodology to determine the significant wave height in the coastal area near the Cuxhaven. The validation showed that the proposed method can determine reliable SWH from approximately 1 km offshore, which is an improvement of earlier results.

## 6.2 Outlook

Although the 4- $\beta$  retracker can already provide some reliable SWH in the coastal area, there is still something that can be improved. For example, the difference between the 4- $\beta$  and the 5- $\beta$  fitted results in the oceanic area shows that the performance of the 4- $\beta$  retracker is affected by environmental factors. In the future, other specific data can be combined with the 4- $\beta$  retracker to determine the correlation between the performance of the 4- $\beta$  retracker and the environmental conditions, thereby further improving the accuracy of the 4- $\beta$  retracker.

Besides, due to the complexity of the coastal waveforms, the performance of the 4- $\beta$  retracker in the coastal area is still worse than in the oceanic area. In future work, we can consider combining with the coastal waveform classification, adjusting the algorithms of extracting the thermal noise and the leading edge for different types of the waveform, such as adding some algorithms to reduce the abrupt noise on the leading edge, thereby reducing biases in the fitted coastal waveforms.

The is still one problem left in the relationship between  $\Delta\text{Bins}$  and  $\beta_4$ . When  $\Delta\text{Bins}$  is smaller than 3.3369,  $\beta_4$  is smaller than  $\sigma_p$  (Eq.4.44). Some further researches are necessary to solve this issue.

The whole algorithm is still based on single-waveform retracking. Therefore, the spatio-temporal information has been ignored. In future research, we can also analyse the radargram stacks to get the spatio-temporal information. It may further improve the accuracy of the waveform retracking and reduce the impact of blunders.

# Bibliography

- Brown, G. (1977), 'The average impulse response of a rough surface and its applications', *IEEE transactions on antennas and propagation* **25**(1), 67–74.
- Deng, X. and Featherstone, W. (2006), 'A coastal retracking system for satellite radar altimeter waveforms: Application to ers-2 around australia', *Journal of Geophysical Research: Oceans* **111**(C6).
- Dinardo, S. (2020), Techniques and Applications for Satellite SAR Altimetry over water, land and ice, PhD thesis, Technical University of Darmstadt.
- EUMETSAT (2017), 'Sentinel-3 SRAL Marine User Handbook'.
- Fu, L.-L. and Cazenave, A. (2001), *Satellite altimetry and earth sciences: a handbook of techniques and applications*, Elsevier.
- Levenberg, K. (1944), 'A method for the solution of certain non-linear problems in least squares', *Quarterly of applied mathematics* **2**(2), 164–168.
- Marquardt, D. W. (1963), 'An algorithm for least-squares estimation of nonlinear parameters', *Journal of the society for Industrial and Applied Mathematics* **11**(2), 431–441.
- Mathworks (2020), 'Least-squares (model fitting) algorithms', *Matlab help center* .  
**URL:** <https://www.mathworks.com/help/optim/ug/least-squares-model-fitting-algorithms.html>
- Passaro, M., Cipollini, P., Vignudelli, S., Quartly, G. D. and Snaith, H. M. (2014), 'ALES: A multi-mission adaptive subwaveform retracker for coastal and open ocean altimetry', *Remote Sensing of Environment* **145**, 173–189.
- Quartly, G., Srokosz, M. and McMillan, A. (2001), 'Analyzing altimeter artifacts: Statistical properties of ocean waveforms', *Journal of Atmospheric and Oceanic Technology* **18**(12), 2074–2091.
- Raney, R. K. (1998), 'The delay/doppler radar altimeter', *IEEE Transactions on Geoscience and Remote Sensing* **36**(5), 1578–1588.
- Rosmorduc, V., Benveniste, J., Lauret, O., Maheu, C., Milagro, M. and Picot, N. (2011), 'Radar altimetry tutorial', *ESA, Europe* .
- Sneeuw, N., Krumm, F. and Roth, M. (2008), 'Adjustment theory', *University of Stuttgart: Stuttgart, Germany* .
- Thibaut, P., Amarouche, L., Zanife, O., Steunou, N., Vincent, P. and Raizonville, P. (2004), 'Jason-1 altimeter ground processing look-up correction tables', *Marine Geodesy* **27**(3-4), 409–431.
- Tourian, M. J. (2012), Controls on satellite altimetry over inland water surfaces for hydrological purposes, Master's thesis.

Vignudelli, S., Kostianoy, A. G., Cipollini, P. and Benveniste, J. (2011), *Coastal altimetry*, Springer Science & Business Media.

Zwally, H. J. and Brenner, A. C. (2001), Ice sheet dynamics and mass balance, in 'International Geophysics', Vol. 69, Elsevier, pp. 351–xxvi.