	ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**CryoSat-2 sUccess over Inland water And Land (CRUCIAL) Contract  
1/6287/11/I-NB**

**D4100 Algorithm Theoretical Basis Document**

Document No: **NCL\_CRUCIAL\_D4100**

Issue: **3**

Issue Date: **4 October 2016**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Author (Chapters 1-4): Philip Moore

Signature: 

Author (Chapter 5): Steve Birkinshaw

Signature: 

Author (Chapter 6): Peter Bauer-Gottwein



Signature: 

Author (Chapter 6): Raphael Schneider

Signature: 



Authorised by: Philip Moore

Signature: 

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

### Document Change Record

Issue	Date	Modified by	Description
1	30.04.16	PM	Created
2	25.08.16	PM/PB	Amendments after ESA review of Issue 1
3	04.10.16	PM	Amendments after ESA review of Issue 2

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



## Abstract

This Algorithm Theoretical Basis Document (ATBD) describes the algorithms used to estimate inland water heights from CryoSat-2 SAR and SARin FBR data and their utilization in river modelling. The report outlines the principles and mathematical background of SAR and SARin FBR processing, justifies the algorithmic approach and considers its limitations and underlying assumptions.

This document outlines the following algorithms:



- CryoSat-2 SAR FBR data processing
- Cryosat-2 SARin FBR data processing
- Geophysical correction data
- River mask data
- Data Assimilation in hydrodynamic modelling using synthetic data.

This document essentially supersedes the Algorithm Interim Report (D4050). Differences between this report and D4050 include improvements, corrections and additional details of the algorithmic approach.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## Contents

<b>Abstract.....</b>	<b>4</b>
<b>Figures.....</b>	<b>8</b>
<b>Tables.....</b>	<b>12</b>
<b>1. Introduction .....</b>	<b>13</b>
<b>1.1 Scope .....</b>	<b>13</b>
<b>1.2 Abbreviations and Acronyms .....</b>	<b>14</b>
<b>2 SAR and SARin FBR Parameters (Baseline B).....</b>	<b>17</b>
<b>2.1 SAR (Baseline B) .....</b>	<b>17</b>
<b>2.2 SARin (Baseline B) .....</b>	<b>18</b>
<b>2.3 Baseline configuration differences .....</b>	<b>18</b>
<b>2.3 Calibration applied. ....</b>	<b>20</b>
<b>3 SAR FBR Theory.....</b>	<b>20</b>
<b>3.1 Beam Formation.....</b>	<b>20</b>
<b>3.2 Geophysical Corrections and Calibrations.....</b>	<b>25</b>
<b>3.3 Beam Steering .....</b>	<b>25</b>
<b>3.4 Burst Centre Ground Points .....</b>	<b>28</b>
<b>3.5 Multi-Look Ground Points.....</b>	<b>36</b>
<b>3.6 Stacking .....</b>	<b>36</b>
<b>3.7 Stack Beam Alignment.....</b>	<b>37</b>
3.7.1 Tracker Bin Correction to Burst Echoes.....	37
3.7.2 Slant Range Correction.....	38
<b>3.8 Multi-look Waveforms.....</b>	<b>38</b>
<b>3.9 Retracking.....</b>	<b>41</b>
3.9.1 Tracker Waveform Descriptions:.....	43
<b>3.9 SARin Analysis .....</b>	<b>59</b>

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



<b>4 SAR FBR Analysis: Summary</b> .....	<b>64</b>
<b>5 Mekong River Masks</b> .....	<b>74</b>
<b>5.1 Mekong: Landsat river mask</b> .....	<b>74</b>
<b>5.2 Landsat river mask summary</b> .....	<b>82</b>
<b>6 Assimilation of CryoSat-2 data to 1-dimensional hydrodynamic models</b> .....	<b>83</b>
<b>6.1 Introduction and background</b> .....	<b>83</b>
<b>6.2 Data and methods</b> .....	<b>85</b>
6.2.1 CryoSat-2 data .....	85
6.2.2 Preprocessing of the data .....	85
6.2.3 Brahmaputra: Landsat river mask .....	87
6.2.4 Filtering and projecting the CryoSat-2 data .....	88
<b>6.3 Hydrologic hydrodynamic model</b> .....	<b>89</b>
<b>6.4 Cross section calibration</b> .....	<b>93</b>
<b>6.5 Data Assimilation</b> .....	<b>95</b>
6.5.1 Data Assimilation in general.....	95
6.5.2 Ensemble Transform Kalman Filter .....	96
6.5.3 Setup with DHI Data Assimilation Framework .....	98
6.5.4 Processing of CryoSat-2 observations .....	100
6.5.5 Observation error .....	101
6.5.6 Model error .....	102
6.5.7 Localization.....	102
6.5.8 Virtual window .....	103
6.5.9 Data assimilation experiment with synthetic CryoSat-2 data.....	104
<b>6.6 Results and Discussion</b> .....	<b>105</b>
6.6.1 Processing of CryoSat-2 data.....	105

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

6.6.2 Synthetic data assimilation experiment.....106

**6.7 Conclusion ..... 109**

**7. References ..... 111**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## Figures

Figure 1: SIRAL transmission and reception timing in the LRM, SAR, and SARIn modes with transmission timing and changes in the sampling of the received echo (Wingham et al., 2006).....	24
Figure 2: Altimetric heights and datums: the index 'i' refers to burst position along the orbit (page 29) and index 'k' to points along the ground (page 36).....	26
Figure 3: Schematic of rock angle.....	27
Figure 4: Schematic of geoidal, ellipsoidal and orthometric heights. ....	30
Figure 5: Ground track across Mekong (2 Nov 2012). Descending pass: first point lat 12.28615°, lon 105.99150°; last point lat 12.24248°, lon 105.98690°.....	31
Figure 6: Burst echo waveforms from a descending pass across the Mekong (#1340-1332), lat 12.2518°-12.2575°. Green line OCOG retracked height. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. ....	32
Figure 7: Burst echo waveforms from a descending pass across the Mekong (#1331-1323), lat 12.2582°-12.2640°. Green line OCOG retracked height. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. ....	33
Figure 8: Burst echo waveforms from a descending pass across the Mekong (#1322-1314), lat 12.2647°-12.2704°. Green line OCOG retracked height. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. ....	34
Figure 9: Burst echo heights across the Mekong. X-axis latitude (deg); y axis (m): Orthometric height (blue) from burst echoes. ....	35
Figure 10: Figure 9 zoomed. ....	35
Figure 11: Schematic of bursts, the fan of Doppler beams, ground points and multi-looks .....	41
Figure 12: SAR empirical retrackers. X-axis is gate number (1-256), Y-axis is power amplitude. .	46
Figure 13: Google Earth image of 19 Apr 2011 Mekong crossing. Satellite ground track N-S.....	49
Figure 14: Waveforms (blue curve) across the Mekong (N=110) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. ....	50
Figure 15: NRE for waveforms across the Mekong .....	51





		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Figure 16: Orthometric heights from empirical retracker (Figure 12) across the Mekong for 19 Apr 2011 (day 109 in year; seconds after 59084) for various number of echoes in multi-look. Upper plot full profile; lower plot a zoomed version of the river surface..... 52

Figure 17: Orthometric heights from OCOG/Threshold retracker across the Mekong for 19 Apr 2011 (day 109 in year; seconds after 59084) for various number of echoes in multi-look. Upper plot full profile; lower plot a zoomed version of the river surface. .... 53

Figure 18: Seasonal variation in extent of Tonle Sap ..... 54

Figure 19: Points across Tonle Sap (3 Dec 2011)..... 55

Figure 20: Geoid height measurements across Tonle Sap (3 Dec 2011) with various multi-look combinations. Upper: Retracked using empirical waveforms (Figure 12). Lower: OCOG/Threshold retracker used. .... 56

Figure 21: Ground track across Amazonas 5 May 2012 (SARin mode) ..... 61

Figure 22: (Upper) Difference between heights from two antennae; (lower) ground slope across. .... 62

Figure 23: SARin waveforms (upper), coherence (middle) and ground slope in degrees (lower) across Amazonas. X-axis is bin number (1-1024 in SARin mode). In the upper plot the left antennae is coloured blue and the right antennae is green. Left hand column location #274 (3.930 °S 70.207 °W); right hand column location #275 (3.927 °S 70.207 °W). Date of pass 5 May 2012. .... 63

Figure 24: Waveforms (i=86, t=18645.69 s, lat= 13.0439°, lon=104.0775°) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. .... 66

Figure 25: Waveforms (i=87, t=18645.73 s, lat= 13.0411°, lon=104.0772°) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. .... 67

Figure 26: Waveforms (i=118, t=18647.12 s, lat= 12.9568°, lon=104.0683°) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. .... 68



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Figure 27: Waveforms (i=119, t=18647.17 s, lat= 12.9541°, lon=104.0680°) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. .... 69

Figure 28: Waveforms (i=120, t=18647.21 s, lat= 12.9514°, lon=104.0677°) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. .... 70

Figure 29: Waveforms (i=121, t=18647.26 s, lat= 12.9486°, lon=104.0674°) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. .... 71

Figure 30: Waveforms (i=122, t=18647.30 s, lat= 12.9459°, lon=104.0671°) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude. .... 72

Figure 31: Geolocation of points i=86 and i=118 corresponding to waveforms of Figure 25 - Figure 30. Other points are omitted for clarity as location of i=87, and i=119-122 are close to points i=86 and i=118 respectively..... 73

Figure 32: Six Landsat images of the Mekong River between 12°N and 18°N ..... 76

Figure 33: Landsat image from 12/11/2014 of the Mekong River between 14.2°N and 14.8°N . 77

Figure 34: Approximate mask of the Mekong River between 12°N and 18°N ..... 77

Figure 35: Approximate mask of the Mekong River between 14.2°N and 14.8°N ..... 78

Figure 36: Water mask of the Mekong River between 12°N and 18°N..... 79

Figure 37: Water mask of the Mekong River between 14.2°N and 14.8°N ..... 80

Figure 38: Final mask of the Mekong River between 12°N and 18°N..... 81

Figure 39: Final mask of the Mekong River between 14.2°N and 14.8°N..... 82

Figure 40: Map of the three main rivers draining into the Gulf of Bengal through Bangladesh .. 84



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Figure 41: Probability distributions of backscatter values of CryoSat-2 datapoints over the Landsat river mask (blue) compared to datapoints outside the river mask, but inside a 20km wide buffer around the river line (blue) ..... 86

Figure 42: Landsat 7 images of the same part of the Brahmaputra river in the Assam valley showing the dynamic changes in river morphology. Left: 2010. Right: 2011. .... 87

Figure 43: Section of the Brahmaputra in the Assam valley showing the Landsat river mask, the CryoSat-2 observations and their mapping to the 1D river model, all for 2013. .... 89

Figure 44: The MIKE 1D computational grid. From MIKE by DHI (2009)..... 90

Figure 45: The Brahmaputra basin model. The main calibration station, Bahadurabad in Bangladesh is shown. The water level calibrated part, i.e. the Assam valley together with the Envisat virtual stations is displayed (see next section). .... 92

Figure 46: Sketch of the two-step cross section calibration. .... 94

Figure 47: Overview over the DHI Data Assimilation Framework. Source: Marc-Etienne Ridler, DHI. .... 99

Figure 48: Development of CPRS (i.e. model error) with ensemble size for a synthetic data assimilation setup of the model as described in section 6.5.9. The chosen ensemble size of 40 is not optimal yet, but considered a reasonable trade-off..... 100



Figure 49: Grouping of individual CryoSat-2 data points from one river crossing based on k-means clustering. The left panel shows a crossing from 09-04-2011, and the right panel from 30-10-2010. Note that the panels are showing different parts of the river, with the 2010 and 2011 river mask respectively. .... 101

Figure 50: Differences in simulated water levels along the Brahmaputra River over different timesteps, i.e. the differences in water levels at each single model grid point within 5, 10, 20, 60, 120, 240, 480, and 960 minutes. .... 103

Figure 51: Comparison of water level (upper panel) and discharge (lower panel) at Bahadurabad station for the original model and the hidden truth run, from which the synthetic CryoSat-2 observations were extracted from. .... 105



Figure 52: CryoSat-2 data 2010 to 2013 after filtering and mapping to the model's river line in comparison with SRTM data. The Assam valley starts around chainage km 2050. .... 106

Figure 53: Results of data assimilation of synthetic CryoSat-2 water level observations in comparison to the synthetic truth model results at Bahadurabad station. The inlet shows April to December 2011 in more detail. .... 107

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## Tables

Table 1: Differences in configurations between Baseline B and Baseline C. ....	19
Table 2: Statistics of fit for pass across Tonle Sap.....	57
Table 3: Tracker types (Figure 12) for the Tonle Sap points. The single peak column is the sum of columns 1-3 and the dual peak column that of 4-5. A total of 68 waveforms were retracked.....	58
Table 4: Landsat 8; Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) bands .	75
Table 5: Decision variables, constraints and objective functions of the genetic algorithm used for the two-step water level calibration .....	95
Table 6: Performance indicators for the synthetic DA experiment compared to an open loop run of the same model. Discharge at Bahadurabad. The continuous rank probability score (CRPS) is a performance indicator commonly used for ensemble forecast. Sharpness is given as the width of the 90% confidence interval. ....	109

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 1. Introduction



### 1.1 Scope

CryoSat-2 was launched on 8 April 2010. It follows on from previous ESA Earth orbiting satellite radar altimeters (e.g. ERS2 and ENVISAT) that have been used for land surface applications including mapping (Berry et al, 2010a,b,c ; Smith & Berry 2011) and measurement of river and lake systems (Berry, 2009; Wheeler et al., 2010, Berry et al., 2012a,b). CryoSat-2's primary instrument is SIRAL (SAR Interferometric Radar Altimeter), which uses radar to determine and monitor the spacecraft's altitude. Although the CryoSat-2 primary aim is to measure sea ice and ice sheets it can provide valuable data over the rest of the Earth surface. SIRAL operates in one of three modes, depending on where (above the Earth's surface) CryoSat-2 is flying. The three modes are: the conventional altimeter mode or Low resolution Mode (LRM), Synthetic Aperture Radar (SAR) and Interferometric Synthetic Aperture Radar (SARin). CryoSat-2 has a low-Earth orbit and is not Sun-synchronous; it has a period of 100 minutes. The CryoSat-2 mission is the first to operate a SAR mode altimeter.

This study is investigating innovative inland water applications from CryoSat-2 with a forward-look component to the Sentinel-3 mission. This ATBD presents the approach developed to process and utilize SAR FBR and SARin data over inland waters. The report does not consider LRM data as that is relatively well understood from previous missions and is not relevant for Sentinel 3 processing. This report thus focuses on SAR FBR and SARin FBR data and hydrological modelling.



Deliverable D4100 presents theoretical and computational aspects of the SAR FBR and SARin FBR algorithms in particular, including comments to high-light the capability of the methodology as well as to identify limitations. D4100 does not contain full results as those will be in the Product Validation Report (D4200).

This document supersedes the Algorithm Interim Report (D4050). Differences between this report and D4050 include improvements, corrections and additional details of the algorithmic approach. D4100 contains the same topics as D4050 but with the addition of SARin data analysis. Furthermore, this report uses synthetic Cryosat-2 data in section 6 'Assimilation of CryoSat-2 data to 1-dimensional hydrodynamic models' to illustrate the theoretical development. The use of real data in section 6, although early results were contained in D4050, is now presented in the Product Validation Report (D4200).



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 1.2 Abbreviations and Acronyms

Abbreviation	Meaning
AIR	Algorithm Interim Report
AR1	Autoregression Model (order 1)
ATBD	Algorithm Theoretical Basis Document
CF	Carrier Frequency
CRPS	Continuous Ranked Probability Score
CRUCIAL	CryoSat-2sUCcess over Inland water And Land
DA	Data Assimilation
DEM	Digital Elevation Model
DHI	Danish Hydraulic Institute
DORIS	Doppler Orbitography and Radiopositioning Integrated by Satellite
DTU	Danish Technical University
ECMWF	European Centre for Medium-Range Weather Forecasts
EGM96/08	Earth Gravitational Model 1996 (2008)
EnKF	Ensemble Kalman filter
ETKF	Ensemble Transform Kalman Filter
ERS2	European Remote Sensing satellite 2
Envisat	Environmental Satellite
ESA	European Space Agency
FBR	Full Bit Rate
FFT	Fast Fourier Transform
HPF	High Pulse Frequency



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

GMES	Global Monitoring for Environment and Security programme (renamed as <i>Copernicus</i> )
GRACE	Gravity Recovery And Climate Experiment
GOCE	Gravity field and steady state Ocean Circulation Explorer
L1A	Level 1A
L1B	Level 1B
LOTUS	Preparing Land and Ocean Take Up from Sentinel-3 (project supporting the development of GMES)
LRM	Low Resolution Mode
JASON-2	US/French Altimeter Satellite
MAE	Mean Absolute Error
MAMSL	Metres Above Mean Sea Level
MATLAB	Matrix Laboratory (High-Level Technical Supporting Language for Algorithm Development)
MIKE	Generalized River Modeling Package
MODIS	Moderate-Resolution Imaging Spectroradiometer
NAM	North American Mesoscale Forecast System
NASA	National Aeronautics and Space Administration (USA)
NCL	Newcastle University
NDVI	Normalised Difference Vegetation Index
NDWI	Normalised Difference Water Index
NIR	Near Infrared
NRE	Normalised Residual Error
NSE	Nash-Sutcliffe efficiency
OCOG	Offset Centre of Gravity

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

OLI	Operational Land Imager
POCA	Point of Closest Approach
PRF	Pulse Repetition Frequency
PVR	Product Validation Report
RA	Radar Altimeter
RMSE	Root Mean Square Error
R&L	River and Lake
SAR	Synthetic Aperture Radar mode of CryoSat-2 SIRAL
SARin	Interferometric Synthetic Aperture Radar mode of CryoSat-2 SIRAL
Sentinel-3	ESA Earth Observation Satellite Mission
SIRAL	SAR Interferometric Radar Altimeter
SRTM	Shuttle Radar Topography Mission
SRC	Slant Range Correction
SWIR	Short Wave Infrared
SWOT	Surface Water Ocean Topography mission
TIRS	Thermal Infrared Sensor
TRMM	Tropical Rainfall measurement mission
UTC	Coordinated Universal Time
VS	Virtual station
.NET/C#	C# programs run on the .NET Framework



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 2 SAR and SARin FBR Parameters (Baseline B)

This section details some of the fundamental parameters of CryoSat-2 used in the subsequent FBR processing. These parameters relate to the Baseline B release. Relevant differences in parameters for SAR/SARin FBR Baseline B to Baseline C are noted in section 2.3.

Data used in this Study is from **Baseline B**, except for data in March 2015.

### 2.1 SAR (Baseline B)

Burst repetition interval = 11.7 ms

Echoes per burst = 64

Tracking Samples per echo = 128

Measurement range window = 30 m

Range bin samples ~ 0.23 m

Pulse Repetition Frequency (PRF) = 18181.82 Hz

Speed of light (c) = 299792458 m/s

Carrier Frequency (CF) = 13.575e9 Hz



Wavelength:  $\lambda = c/CF$

Carrier wave-number:  $k_0 = 2\pi/\lambda$

Along-track antennae 3 dB width 1.0766° (full beam width)

Cross-track antennae 3 dB width 1.2016° (full beam width)

Antenna: Left of the flight direction

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 2.2 SARin (Baseline B)

As for SAR with the exceptions

Burst repetition interval = 46.7 ms

Measurement range window = 120 m



Interferometer baseline:  $\beta = 1.172$  m

Tracking Samples per echo = 512

Antennae: Left and right of the flight direction



## 2.3 Baseline configuration differences

Pertinent differences taken from ESA-ESRIN (2015) available at <https://wiki.services.eoportal.org/wiki/index.php?page=CryoSat+Technical+Notes>. Note that Baseline C Level-1B products are distributed in an updated format, including attitude information (roll, pitch, yaw) and, for SAR/SARin, the waveform length doubled with respect to Baseline B. Further details of corrections applied are given in section 3.2.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Description	Baseline B	Baseline C	Comments
Range window samples	SAR 128 SARin 512	SAR 256 SARin 1024	Loss of trailing edge in Baseline B. No impact on quasi-specular waveforms
Range window size	SAR~ 30m SARin ~ 120m	SAR~ 60m SARin ~ 240m	As above
Window delay reference sample (range bin starting from zero)	SAR 64 SARin 256	SAR 128 SARin 512	
Attitude bias	Pitch = 0.0° Roll = 0.0°	Pitch = 0.0550° Roll = 0.1062°	
Datation bias	SAR/SARin = -0.5195 ms	SAR/SARin = 0	
Range bias	SAR/SARin = 0.673 m	SAR/SARin = 0	
External phase correction	SARin = -0.612 rad	SARin = 0	

**Table 1: Differences in configurations between Baseline B and Baseline C.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 2.3 Calibration applied.

Details of the calibrations needed to be applied can be found in ESA-ESRIN (2016) available from <https://wiki.services.eoportal.org/tiki-index.php?page=CryoSat+Technical+Notes>. Calibrations and corrections applied include



- Instrument range bias
- USO correction
- Datation correction
- Range Bias
- Roll correction

## 3 SAR FBR Theory

This section, a simplified version of Wingham et al (2006), describes the background theory of the SAR FBR processing algorithm.

### 3.1 Beam Formation

The SAR mode operates by transmitting a burst of pulses at a HPF and pausing to receive the returned echoes before transmitting the next pulse (see Figure 1). Only the antenna to the left of the flight direction is used. In SAR mode the FBR I (in-phase carrier) and Q (quadrature-phase carrier) components of the individual echoes of the pulses are telemetered to the ground. This data, along with positioning, geophysical, orbital and other associated altimetric corrections, are provided on the SAR FBR level 1A (L1A) data files. It is noted that CryoSat-2 has two antennae and receiving chains for SARin with the left antenna used during the SAR mode of operation (Wingham et al., 2006). In SARin the tracking measurement window is 240 m compared with 60 m in SAR and the samples per echo are 512 compared with the 128 in SAR. Furthermore, the burst repetition interval is 46.7 ms against 11.7 ms in SAR. Thus, the ground track spacing of about 80

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

m for SAR is near 310 m in SARin. In the theoretical development of this section  $\pm$  denote the two antennae. A SAR only development can be inferred by taking the + antennae with the baseline distance between the antennae  $\beta = 0$ .

The first step from a theoretical viewpoint is to form the complex waveform

$$\phi_b^\pm(m) = I_b^\pm(m) + iQ_b^\pm(m) \quad m=0, 1, \dots, 63 \quad (1)$$

where  $i = \sqrt{-1}$  is the imaginary number,  $\phi_b^\pm$ , represent the complex valued echoes from the two antennae,  $m=0, 1, \dots, 63$  denotes the echo number in the burst centred at time  $t$  and based at orbital position  $\underline{x}_b$  and  $I$  is the in-phase component of the waveform and  $Q$  the quadrature component. From Wingham et al. (2006),  $\underline{x}_b$  lies mid-way between the 32<sup>nd</sup> and 33<sup>rd</sup> individual start-stop locations; each start-stop location being related to an individual echo in the burst of 64 echoes. The high repetition rate of the burst, means that the reflectors,  $k$ , on the ground are the same for each pulse. During a burst the carrier phase is locked to the transmission timing so that the transmission within the burst is phase coherent. After allowance for the phase shift (due to the Doppler effect from the satellite motion) the returned waveforms themselves exhibit phase coherence. Under this assumption we can represent



$$\phi_b^\pm(m) = \sum_k \tilde{\phi}_b(x_k) e^{\pm i k_0 \beta \hat{n}_k \cdot \hat{m}_b / 2} e^{2i(m-63/2)\Delta k_0 \hat{n}_k \cdot \hat{n}_v} \quad m=0, 1, \dots, 63 \quad (2)$$

In Eq. (2),  $k$  denotes the elemental scatterers on the ground at location  $x_k$ ,  $\tilde{\phi}_b(x_k)$  the reflectance from the  $k$ th scatterer,  $\Delta$  the along-track sampling interval,  $\hat{n}_k$  the unit vector from burst centre to elemental scatterer,  $\hat{n}_v$  the direction of the velocity vector,  $\beta$  the antennae interferometer baseline and  $\hat{m}_b$  a unit vector in the direction of the interferometer baseline. Since the satellite moves only 60m in SAR mode (310 m in SARin) during the duration of the burst the velocity vector is considered as a straight line.

Equation (2) is useful as it underpins the mathematics behind the “beam formation” stage where we form

$$\psi_b^\pm(\theta_1) = \sum_{m=0}^{63} \phi_b^\pm(m) e^{-2i k_0 \Delta \theta_1 (m-63/2)} \quad (3)$$

where  $\theta_1$  is called the “look-angle” of the beam, taken as

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

$$\theta_1 = \theta_R^b + \frac{\pi n}{64k_0\Delta} \quad n = 0, 1, \dots, 63 \quad (4)$$

where we can identify the beam inclination =  $\frac{\pi}{64k_0\Delta}$ .

In Eq. (4) the so-called rock-angle,  $\theta_R^b$ , is used to direct the beam to either to

1. the nadir pointing direction

**OR**

2. to pre-defined equi-angular ground points.

By substituting Eq. (4) into Eq. (3) and re-arranging gives

$$\psi_b^\pm(n) = \sum_{m=0}^{63} \phi_b^\pm(m) e^{-2ik_0\Delta\theta_R^b(m-63/2)} e^{-2i\pi n(m-63/2)/64} \quad (5)$$

The dependence of  $\psi^\pm$  has changed from the look-angle  $\theta_1$  of Eq. (3) to the integer n where  $\theta_1$  and n are linked by Eq. (4).

On substituting Eq. (2) into Eq. (3) it can be shown that



$$\psi_b^\pm(n) = \sum_k \tilde{\phi}_b^\pm(x_k) e^{\pm ik_0\beta \hat{n}_k \cdot \hat{m}_b / 2} d(\hat{n}_k \cdot \hat{n}_v - \theta_R^b - \frac{\pi n}{64k_0\Delta}) \quad (6)$$

where

$$d(x) = \frac{\sin(64\Delta k_0 x)}{\sin(\Delta k_0 x)} \quad (7)$$

Equations (6) and (7) show that the echoes are a product of the echo illumination with the beam pattern. This is a function of direction only. Thus, the burst beam is steered into a fan of equi-angular directions for  $n=0, 1, \dots, 63$ . With the satellite near 717 km in altitude and using the 3 dB antenna width the altimetric footprint is about 13.47 km along-track and 15.04 km cross-track

We may consider beam steering as slicing the ground footprint into 64 sections along-track leaving the full cross-track illumination.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Now by denoting

$$\hat{\phi}_b^\pm(m) = \phi_b^\pm(m) e^{-2ik_0 \Delta \theta_R^b m} \quad (8)$$

in Eq. (5) we have

$$\psi_b^\pm(n) = e^{ik_0 63 \Delta \theta_R^b} e^{\frac{i63\pi n}{64}} \sum_{m=0}^{63} \hat{\phi}_b^\pm(m) e^{\frac{-2i\pi nm}{64}} \quad n=0, 1, \dots, 63 \quad (9)$$



Hence, beam formation is undertaken through multiplication (phase shift) and a FFT as the summation

$$\sum_{m=0}^{63} \hat{\phi}_b^\pm(m) e^{\frac{-2i\pi nm}{64}}$$

can be simply computed via the Fast Fourier Transform. The maximum angle of the beam is such that the maximum of the (n+1) beam coincides with the first zero of the nth beam, this equates to 0.0239°, or equivalently 300 m along-track.

The SARin mode is designed for recovering altimetric heights along ice-cap margins where the slope of the terrain can lead to incorrect point of closest approach (POCA). For a cross-track sloping surface the POCA will not be at the nadir point but at an angle,  $\alpha$ , to the nadir pointing direction. With the antennae baseline across-track the transmitted signal from the left antennae will be received at different times at the two antennae as the transmission travels slightly different distances. Alternatively, this can be viewed as a phase shift between the two received signals. For inland waters the across-track slope of the river will be negligible.

For both SAR and SARin, the inland water is a better reflector than land so that the waveform will be dominated by water returns.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

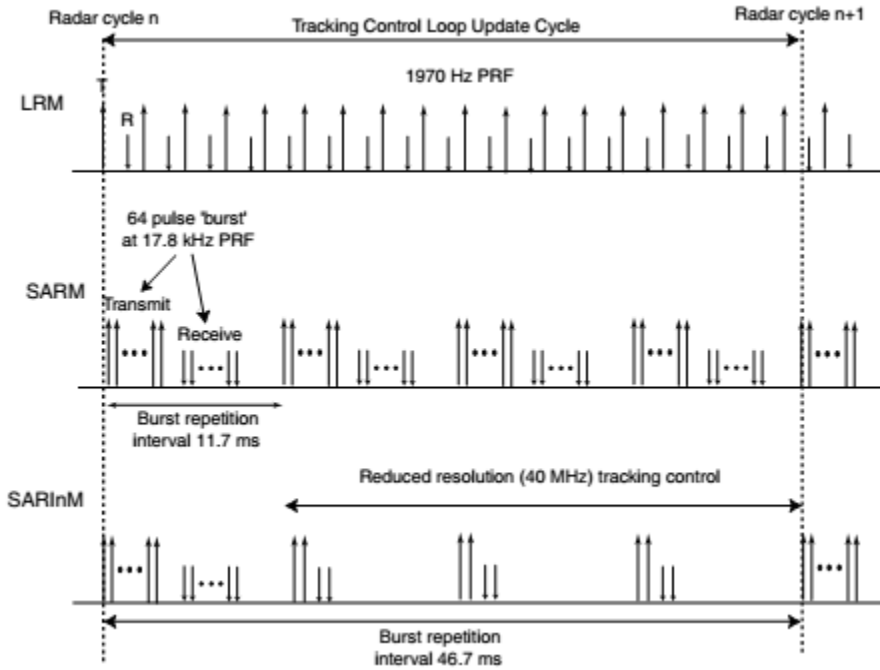




Figure 1: SIRAL transmission and reception timing in the LRM, SAR, and SARIn modes with transmission timing and changes in the sampling of the received echo (Wingham et al., 2006).



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 3.2 Geophysical Corrections and Calibrations

The geophysical corrections, to be applied to the altimetric distance to account for media delays and other known effects, are added at an early stage to use the true geometric (i.e. straight-line) distances. Thus, the corrected altimetric distance,  $AltitudeN_i$ , where  $i$  denotes the burst is given by

$$AltitudeN_i = Alt_i * (1 + USO_{corr,i}) + GEO_{corr,i} + INS_{range\ corr,i}^{\pm} \quad (10)$$

where  $Alt_i$  is the datum of bin 63 in the data file, and the  $USO_{corr}$ ,  $GEO_{corr}$  and  $INS_{range\ corr,i}^{\pm}$  are the DORIS Ultra-Stable Oscillator correction, the geophysical corrections (troposphere, ionosphere, tides) and the instrument range correction respectively for the burst (ESA-ESRIN, 2013).

The height  $H_{surface}$  of Figure 2 is derived as the geometric height of the altimetric datum for the burst echo retracking above the reference ellipsoid, i.e.



$$H_{surface_i} = H_{sat_i} - AltitudeN_i \quad (11)$$

The retracked height,  $retrackN_i$ , (i.e. waveform leading edge departure from central bin) has to be added to  $H_{surface}$  to derive the ellipsoidal inland water surface height.

The epoch of the burst centre is corrected to UTC by allowing for the leap seconds in the time stamp.

## 3.3 Beam Steering

Equation (9) describes the beam formation over 64 equi-angles from the burst centred at time  $t_b$  and orbital position  $\underline{x}_b$ . The rock-angle,  $\theta_R^b$ , is used to steer the beam either precisely in the nadir direction as the Doppler beam axis is perpendicular to the velocity vector or to further direct to predefined ground points.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

In detail

$$\alpha = -\arcsin(\dot{h} / v)$$

$$\beta = -\arccos\left(\frac{(\underline{X}_{ground} - \underline{X}_{sat}) \cdot (\underline{X}_g - \underline{X}_{sat})}{|\underline{X}_{ground} - \underline{X}_{sat}| \cdot |\underline{X}_g - \underline{X}_{sat}|}\right)$$

(12)

whence

$$\theta_R^b = \begin{cases} \alpha \\ \alpha + \beta \end{cases}$$

where the upper definition is for the nadir direction and the lower for a ground point (see Figure 3). In Eq. (12)  $h$  is the satellite height,  $v$  the satellite velocity,  $\underline{X}_{ground}$  the Cartesian coordinates at the first ground point on the ground track behind the satellite position,  $\underline{X}_{sat}$  the Cartesian coordinates of the satellite burst centre location  $x_b$ , and  $\underline{X}_g$  the Cartesian coordinates of the burst location on the ground beneath  $\underline{X}_{sat}$ . The sign of the first term in Eq. (12) changes with a change in altitude rate from positive to negative. The sign of the rock angle component involving the ground point is consistent if the ground point is constrained to lie on the ground track and is the first backward looking ground point from the current satellite position.

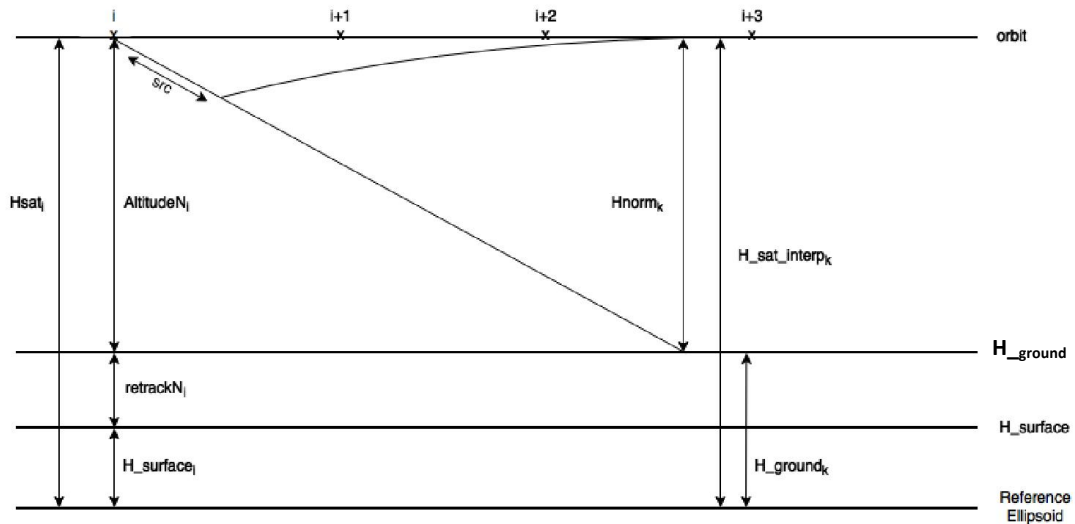




Figure 2: Altimetric heights and datums: the index 'i' refers to burst position along the orbit (page 29) and index 'k' to points along the ground (page 36).

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

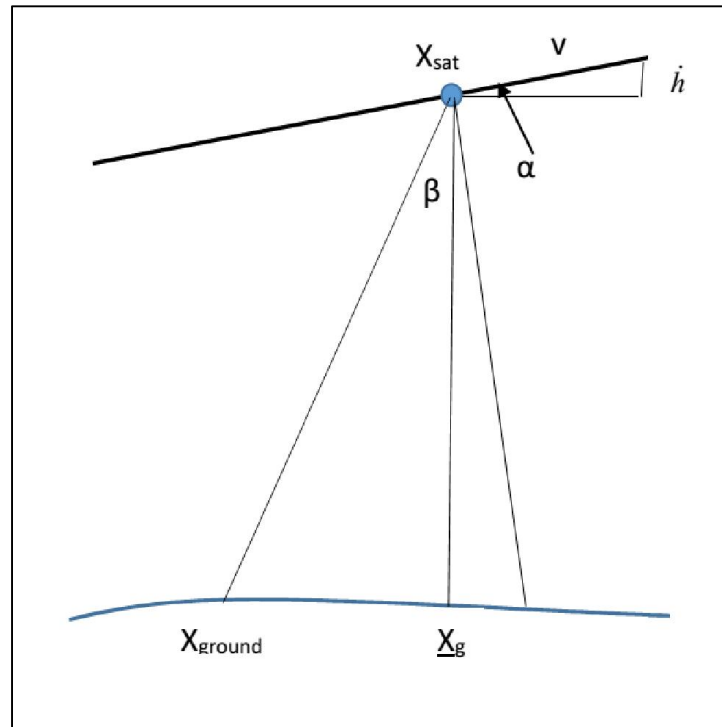




Figure 3: Schematic of rock angle

With the rock angle directed towards the nadir in Eq. (12), the beam formation component  $n=32$  in Eq. (9) gives the nadir pointing beam with  $n=0, 1, \dots, 31$  backward looking from the satellite and  $n=33, 34, \dots, 63$  forward looking. With the forward and backward looking beams directed at ground points sooner or later the beam with number  $n=63$  will be directed at an undefined ground point. In this case a new ground point is estimated preserving the equi-angular spacing of the beam inclination  $\pi/64k_0\Delta$  of Eq. (4). Computationally, we can identify the necessity to define a new ground point when the rock angle for the ground point designated as the first backward looking exceeds the beam inclination. The record count for the ground points is then incremented by unity as a new ground point will be added and the beam counting over  $n$  reduced by unity. A new first ground point on the ground track behind the current position is identified. At the same time the ground point furthest behind the current satellite position (now with  $n=-1$ ) is excluded being replaced by a new ground point (for  $n=63$ ). This new point is at the intersection of the ground and a line inclined at the beam inclination to the  $n=62$  beam.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



### 3.4 Burst Centre Ground Points

Equation (9) requires that the ground points are located by lines radiating from the altimetric burst centre at multiples of the beam-angle  $\pi / 64k_0\Delta$ . In this step the burst echoes are utilized to derive the orthometric height (i.e. height above the geoid) of Figure 4. As a precursor to determining these ground points a first-look surface is determined by:

1. Finding the waveform from each burst with redirection towards the nadir direction of the reference ellipsoid. Zero padding employed to increase range sampling from 128 to 256 for SAR and from 512 to 1024 for SARin adopting Baseline B data.
2. Retrack the burst waveform using an Offset Centre of Gravity (OCOG)/Threshold (Gommenginger et al., 2011) retracker to derive first estimate of ellipsoidal heights. Derive orthometric heights by consideration of the geoid height from a high resolution geoid model (EGM96 in our case)
3. Use a coarse river mask to group orthometric heights to an inland water crossing if appropriate.
4. Filter orthometric heights across the water body and reinstate geoidal height.

This first-look provides approximate geodetic (ellipsoidal) coordinates of the sub-surface ground track of the SAR burst centres at approximately 80m (310 m SARin) along the ground track. Later this will facilitate the recovery of the equi-angular ground points as required for the multi-looking stage.

For illustration the descending pass in Figure 5 across the Mekong (SAR mode) on 2 Nov 2012 is used. Figure 6 - Figure 8 show burst echo waveforms from South to North while Figure 9 plots the orthometric height of the retracked heights across land and inland water surfaces. There are a number of heights near 18 m as derived for example from the first 3 waveforms in Figure 6. These are multi-peaked land/inland water waveforms with strong returns from the land as well as from off-nadir ranging to the tributary flowing into the Mekong. Subsequent waveforms are in general more specular with an orthometric height of about 9.98m (Figure 10). The filter process will give the lower value for the crossing based on a  $3\sigma$  rejection criterion.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



The above procedure only modifies the geodetic heights over water bodies. Coordinates over land are not critical but good estimates of inland water surface coordinates will mitigate against the introduction of fuzziness in the multi-look waveforms. The use of the orthometric heights in the above is particularly pertinent to large lakes where the lake surface will closely approximate an equipotential surface of the Earth's gravity field, that is a level surface of constant height above mean sea-level. Note that the use of the ellipsoidal height to approximate a water surface is not advised as water can flow from low to high ellipsoidal heights such as for lakes sited over areas of high geoidal spatial variability. The recovery of a first-look orthometric surface thus reduces uncertainty in the ground points across inland waters but is also used later to normalize the burst waveforms to a common datum. EGM96 can be replaced by a higher resolution model such as EGM08 or a GRACE/GOCE model but at additional computational cost. Trials with EGM08 had negligible impact and hence the simpler lower order model was retained here although EGM08 was utilized in a later stage to derive the final orthometric heights.

At this stage the algorithm has derived the geodetic surface  $H_{\_ground}$  of Figure 2, where for the  $i^{\text{th}}$  burst centre

$$H_{\_ground_i} = H_{\_surface_i} + retrackN_i \quad (13)$$

where  $H_{\_surface_i}$  is given in Eq. (11) and  $retrackN_i$  is the OCOG/Threshold retracked height.

The retracked burst echoes in SAR mode provide height measurements every 80 m across the water body and land. Figure 10 shows a segment of heights from Figure 9, zooming in on the river. The heights are at low precision with values between 9.68 m and 9.80 m after removing the obvious outliers.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

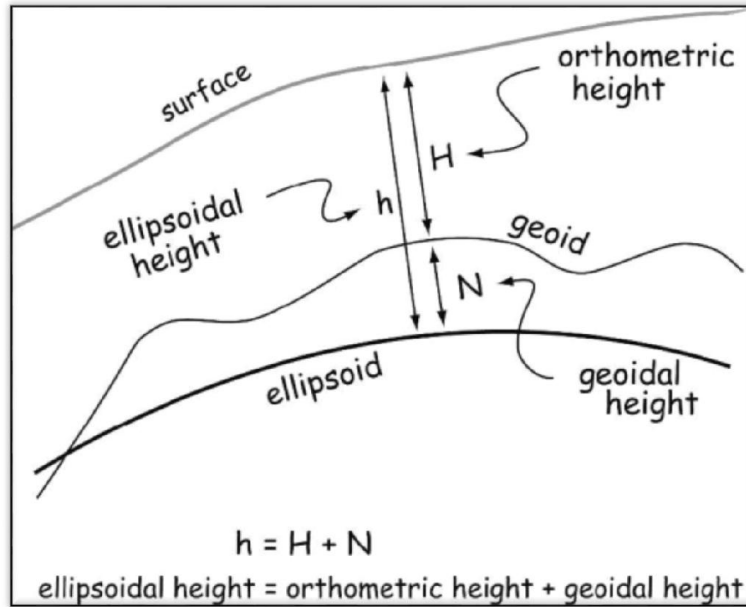


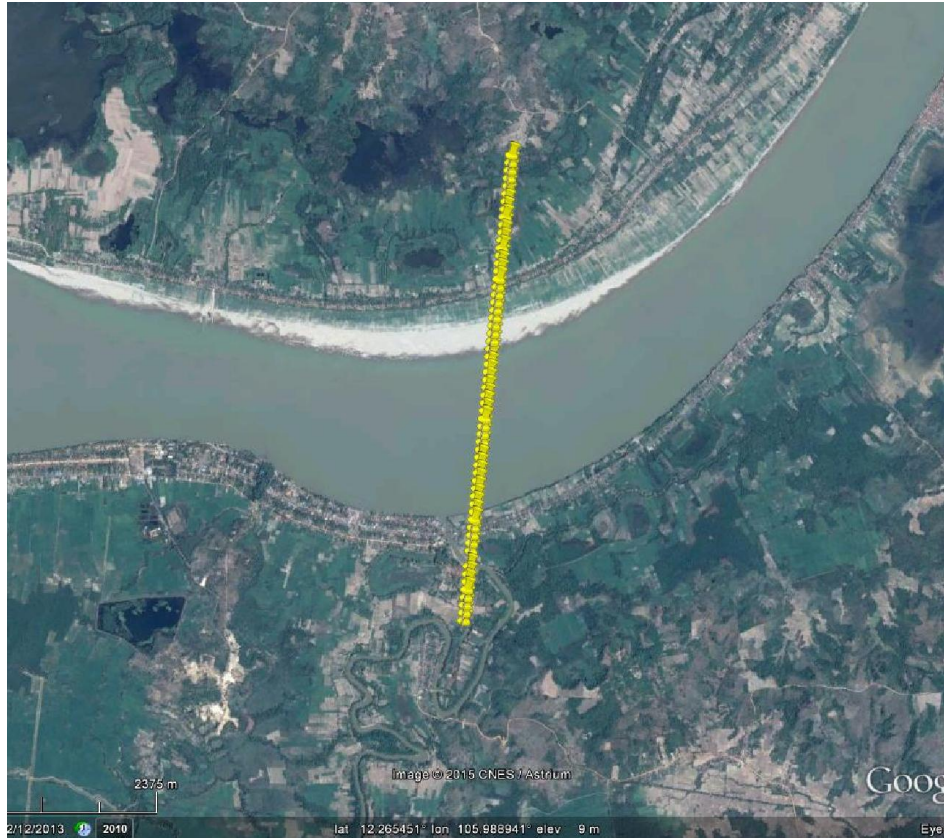




Figure 4: Schematic of geoidal, ellipsoidal and orthometric heights.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 5: Ground track across Mekong (2 Nov 2012). Descending pass: first point lat 12.28615°, lon 105.99150°; last point lat 12.24248°, lon 105.98690°.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

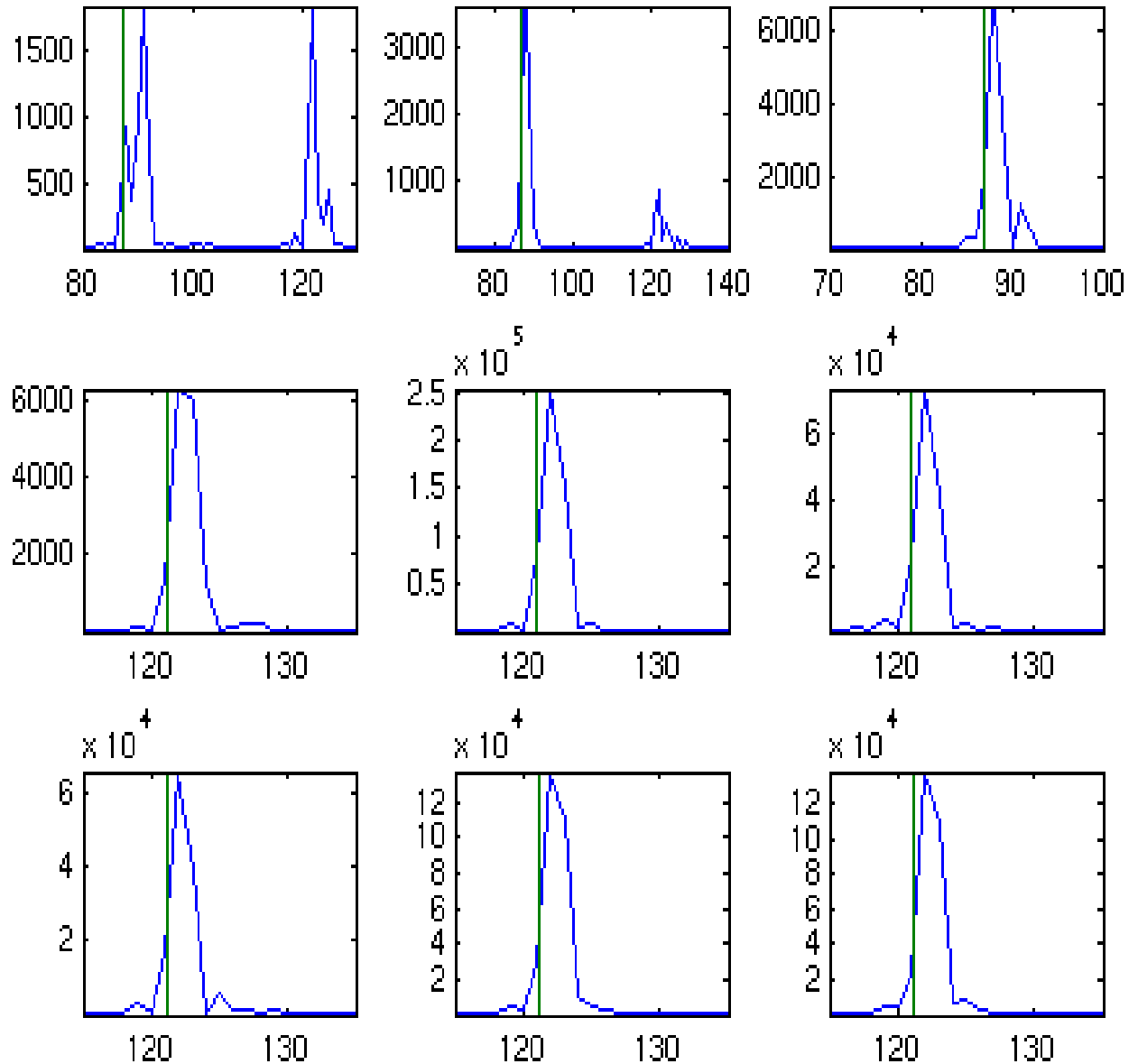




Figure 6: Burst echo waveforms from a descending pass across the Mekong (#1340-1332), lat 12.2518°-12.2575°. Green line OCOG retracked height. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

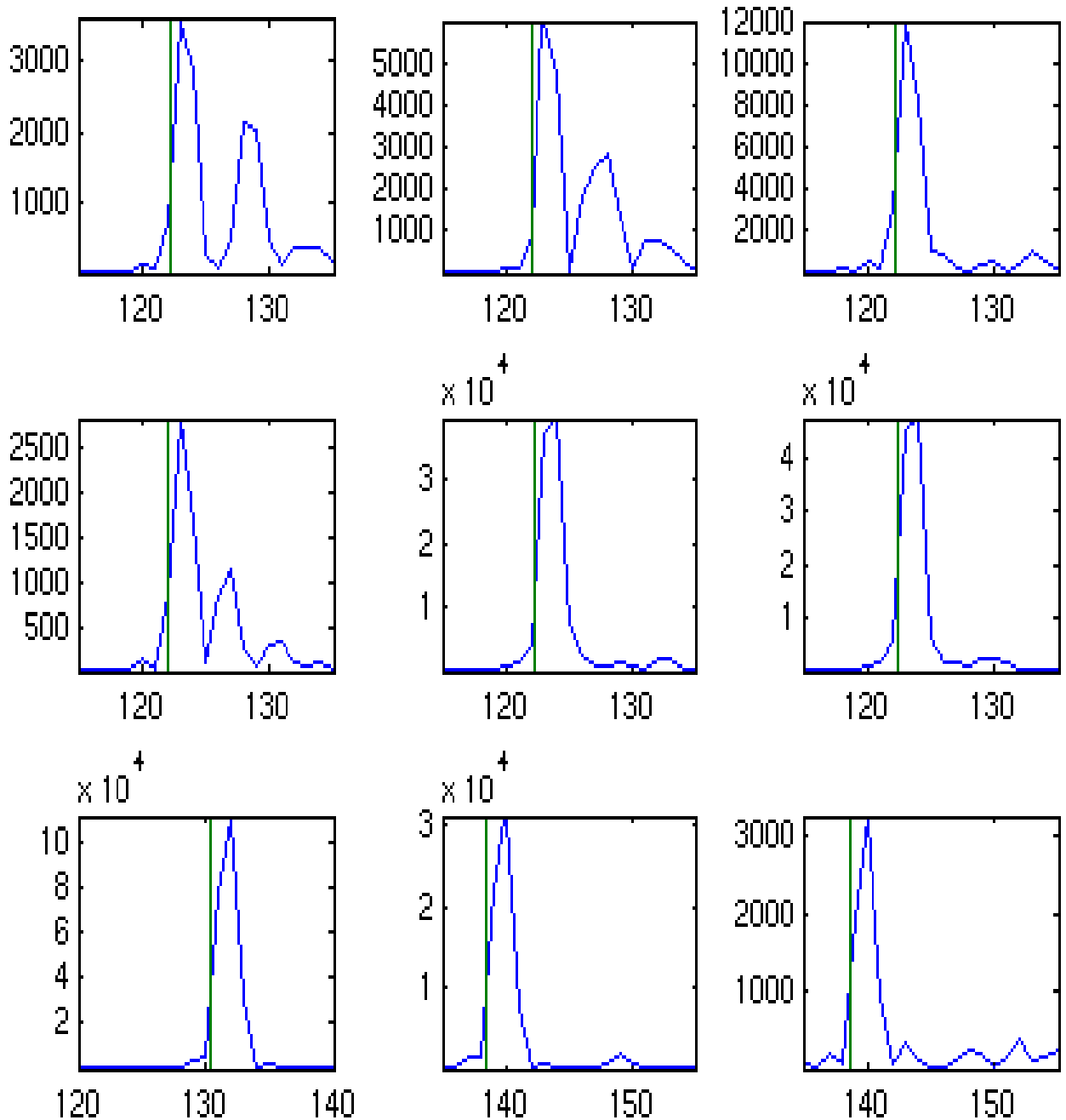




Figure 7: Burst echo waveforms from a descending pass across the Mekong (#1331-1323), lat 12.2582°-12.2640°. Green line OCOG retracked height. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

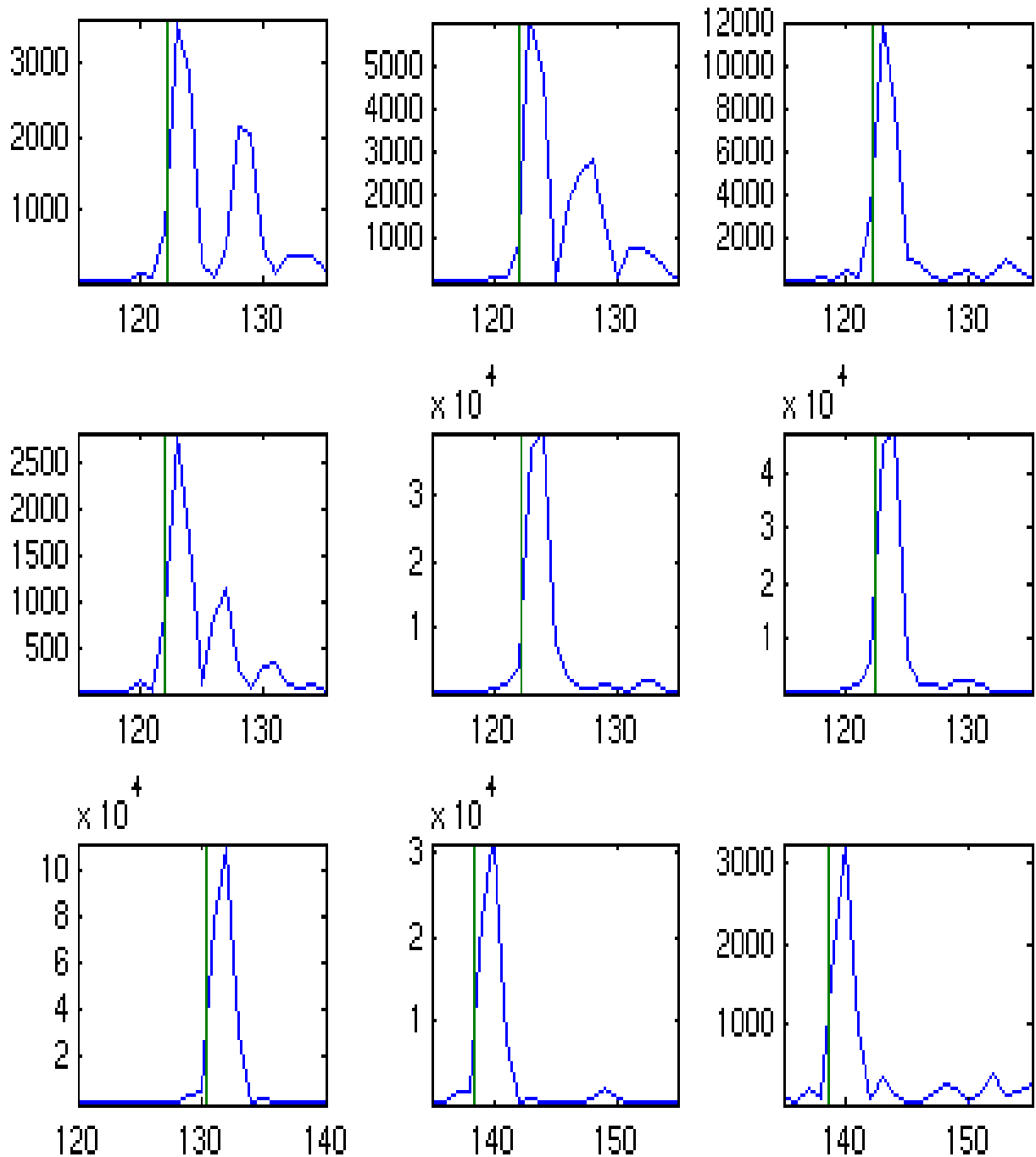




Figure 8: Burst echo waveforms from a descending pass across the Mekong (#1322-1314), lat 12.2647°-12.2704°. Green line OCOG retracked height. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

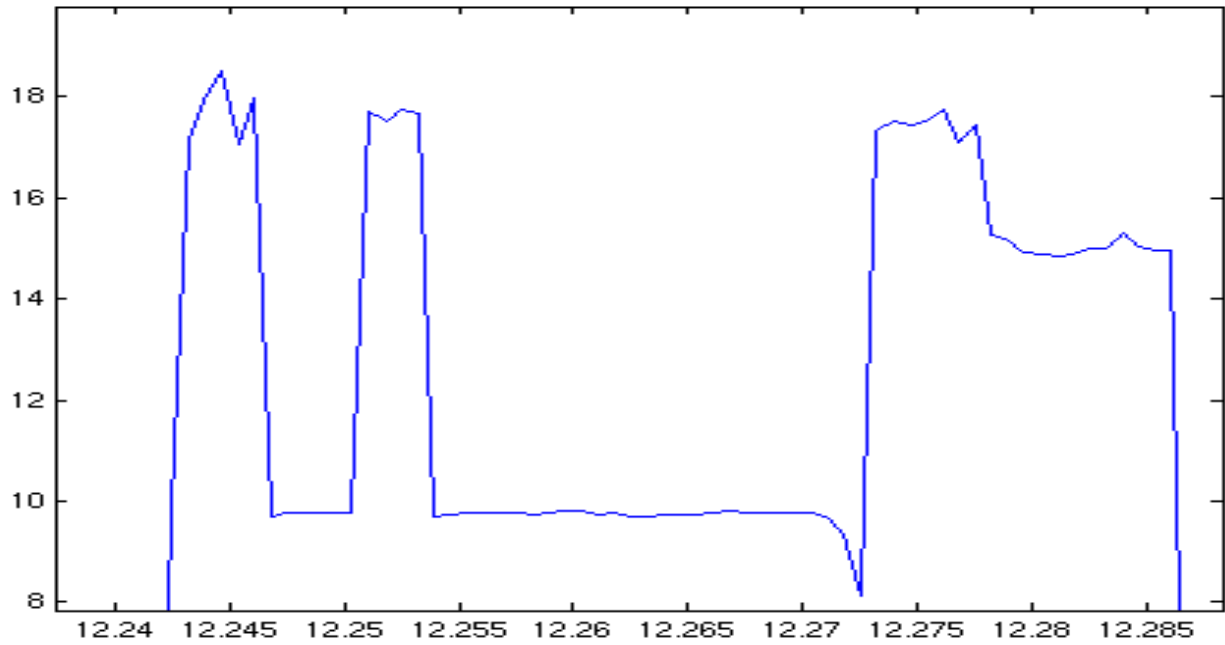


Figure 9: Burst echo heights across the Mekong. X-axis latitude (deg); y axis (m): Orthometric height (blue) from burst echoes.

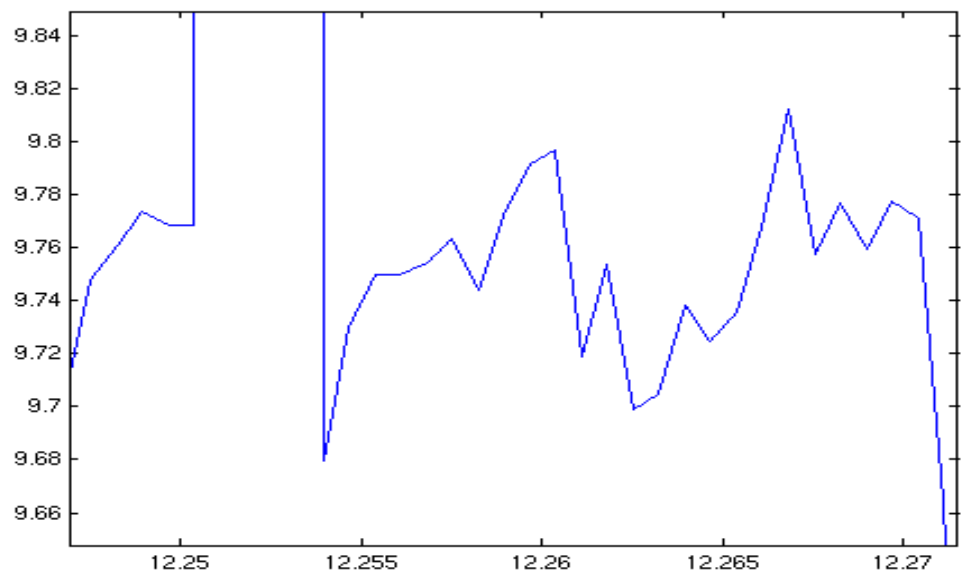




Figure 10: Figure 9 zoomed.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



### 3.5 Multi-Look Ground Points

The equi-angular spacing of  $\pi/64k_0\Delta$  for the multi-look directions with the first-look sub-satellite ground track enables the computations of the multi-look ground points at the intersection of the surface and the beam direction. This is undertaken by first initializing the zero ground point to lie along the satellite nadir of the first burst centre with subsequent ground points defined by the equi-angular spacing. New ground points are introduced as described in section 3.3. This has been called approximate ground point spacing but is adequate for inland water heights where multiple ground points across a water body are essentially at the same orthometric height. At each burst centre the beam is “rocked” by angle  $\theta_R^b$  such that the n=32 beam is steered towards the ground point closest to the nadir. The algorithm considers absence of burst echoes and interpolates across small gaps in the data.

The use of the first-look surface from OCOG retracking gives heights  $H_{-surface_i}$  every 80 m and 310 m along-track for SAR and SARin respectively which can be interpolated to derive the geodetic height at the ground points, namely  $H_{-ground}$ . The height  $H_{-ground_k}$  corresponds to the height of the kth ground point above the reference ellipsoid (see Figure 2).

### 3.6 Stacking

The waveforms  $\psi_b^\pm(n)$  of Eq. (5) depend on the burst centre and the multi-look number. For each multi-look ground point a stack of waveforms directed at the ground location is formed based on the burst and the multi-look number, that is the burst location  $\underline{x}_b$  and the number, n, in the beam steering (see Eq. (9)). Computational evidence has shown that this stack can comprise up to 240 multi-looks in SAR mode and 60 in SARin (Cullen and Wingham, 2002). Each waveform from a single burst is subject to radar speckle which limits the precision of the retracked heights from the burst echoes. By multi-looking, the random nature of the speckle is reduced. The phase is locked only during a single burst to the extent that it changes from burst to burst and the echoes of the stack, acquired from different bursts, are no longer coherent.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

### 3.7 Stack Beam Alignment

The stack comprises a single beam from each burst beam fan that has been steered to the same ground location. The beams in the same stack are views of the identical ground footprint but need alignment in range. Two corrections are applied, namely

- Tracker Bin correction to burst echoes
- The slant range correction to beam steered echoes.

A third correction namely the Doppler range correction to align the Doppler shifts in each beam can be applied but as the beams in the stack are later incoherently averaged in power has no effect on the result. In SARin mode stacks are formed for each antennae.

#### 3.7.1 Tracker Bin Correction to Burst Echoes

This is performed immediately before the 64 individual echoes in a burst are combined using the Doppler shift into a single echo beam. The centre of the  $i^{\text{th}}$  burst echo corresponds to the height  $AltitudeN_i$  from the L1A data (see Figure2 and section 3.2). This height parameter exhibits step changes with discontinuous first derivative. To correct all bin centres to a common height the waveforms are phase and range shifted to the height  $H_{surfaceN}$  of Eq. (13). In this way all bin centres correspond to the inland water orthometric height with all echoes over the inland water corrected to the inland water height datum.



The height correction is thus

$$\Delta h_1 = H_{sat} - AltitudeN_i - H_{ground_i} \quad (14)$$

where, repeating Eq. (13)

$$H_{ground_i} = H_{surface_i} + retrackN_i \quad (15)$$

The algorithm performs this shift by utilizing the FFT to the required phase adjustment followed by a shift of the waveform with respect to the window centre by  $\Delta h_1/hwin$  where  $hwin$  is the

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

tracking window divided by the tracking samples per echo. Note that  $h_{win}$  is the same for SAR and SARin due to the quadrupling of both the range window and the tracking samples per echo, i.e.  $h_{win} = 60/256 \text{ m} = 0.2344 \text{ m}$  after doubling the tracking samples due to the applied zero padding (Table 1 and ESA-ESRIN, 2015).

### 3.7.2 Slant Range Correction

This correction is applied after the beam steering stage. The multi-looks at a given groundpoint location,  $k$ , are components of the beam fans that are identified within the algorithm. Thus, the beam angle will differ from each burst centre to the groundpoint giving rise to a different slant range. The goal is to correct the slant range to the vertical height of the satellite above the ground point. This correction is denoted as slant range correction ( $src$ ) in Figure 2. Thus, the range correction is



$$\Delta h_c = src .$$

This is calculated as the difference between the burst centre and the  $k$  th ground point (using Cartesian coordinates) and distance over the ellipsoid normal from the ground point to the orbit. The Cartesian coordinates of the orbital position above the ground point are derived by interpolating the Cartesian orbital positioning of the burst centres with respect to the latitude of the ground point. Linear interpolation is adequate as over a short distance of 80 (SAR) or 310 m (SARin) the orbit can be assumed to be a straight line.

As in the previous sub-section, modification for the slant range will affect the centre of the range window and the bin to which the retracked height is referenced. Thus the window is shifted by a number of bins corresponding to  $src/h_{win}$ .

### 3.8 Multi-look Waveforms

There is now the possibility to form a weighted averaging of the power of the waveforms directed from the burst echo centres to the ground points (Figure 11). This is performed to derive the waveform for final retracking. The rationale is that the multi-looks from the different burst centres will reduce the speckle in the waveform. Over oceans this has proved to be a useful operation.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Although to some extent location dependent about 240 (60) bursts, 120 (30) either side of the SAR (SARin) beam direction closest to the nadir, can be used in a multi-look procedure for a particular ground point. The actual number to be used and whether some weighting strategy is used to reduce the impact of beams with a large beam angle is a subject warranting study with a clear distinction between inland and ocean waters.



Inland water waveforms are potentially different to oceanic waveforms as the waveforms generally exhibit near specular reflections off the river or lake surface. Except for large lakes where wind driven waves can occur the water surface is relatively flat to the extent that the waveforms are non-ocean like. In addition, the scatterers for the radar echoes are dominated by those over the water as land is a poor reflector. Thus, within the SAR or SARin waveform slice (broad across track, but short along track) the reflectors can be assumed to be dominated by those from the inland water. Thus, specular waveforms are common for inland waters whilst non-existent over the oceans.

Another difference between inland and ocean water is that over the oceans all echoes are from the ocean surface, a surface of near constant height above a reference ellipsoid. Over land the river banks or lake edges can be metres above the water level. For example, rivers such as the Mekong have a 10-16m range between high and low flow. Assuming that the river is contained within its banks at high flow regimes the river level can be over 10m below bank level. Any area of high reflectance on the banks (e.g. irrigated fields) will cause the burst echo height to refer to a different level to the actual river surface. Thus, the tracker bin correction for the burst echoes can refer to different reflectors. Accordingly, the number of multi-looks that contribute to a sharply defined waveform could be much lower over inland waters. It was this rationale that first suggested that correcting all bursts to an inland water height was the way forward.

From a practical viewpoint we are interested in the power,  $P\psi_{ml}^{\pm}(\underline{x}_g)$ , of the multi-look waveforms, namely

$$P\psi_{ml}^{\pm}(\underline{x}_g) = \frac{\sum_{i=1}^{N_{ml}} w_{b_i} \psi_{b_i}^{\pm}(\underline{x}_g, \Delta h_{c_i}) \bar{\psi}_{b_i}^{\pm}(\underline{x}_g, \Delta h_{c_i})}{\sum_{i=1}^{N_{ml}} w_{b_i}} \quad (16)$$

where the summation is over the multi-look waveforms  $i$  at the ground point  $\underline{x}_g$ ,  $N_{ml}$  the total number of multi-looks,  $\Delta h_{c_i}$  the slant range correction for beam centre  $i$ , and  $w_{b_i}$  the associated weight for the  $i^{\text{th}}$  waveform; the bar over  $\psi$  denotes the complex conjugate.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

For SARin the power is recovered for the right and left antennae separately. The power of the multi-look waveforms is retracked to provide the height from the two antennae. Other parameters useful for SARin include the cross-power

$$\psi_{ml}^* = \sum_{i=1}^{N_{ml}} w_{b_i} \psi_{b_i}^+ \bar{\psi}_{b_i}^- / \sum_{i=1}^{N_{ml}} w_{b_i} ; \quad (17)$$

the multi-looked argument or phase,  $A\psi_{ml}$ ,

$$A\psi_{ml} = Arg(\psi_{ml}^*); \quad (18)$$

and the multi-look coherence,

$$C\psi_{ml} = \frac{|\psi_{ml}^*|}{(P\psi_{ml}^+ P\psi_{ml}^-)^{1/2}} \quad (19)$$

Eq. (18) and Eq. (19) are essentially those in Wingham et al. (2006) except that the full weighting in the multi-look waveforms is included in Eq. (17) rather than unit weighting of Wingham et al. The denominator in Eq. (19) utilizes the power of both antennae so that  $C\psi_{ml} = 1$  if  $\psi_{b_i}^+ = \psi_{b_i}^-$  for all waveforms  $i$  in the stack.

Similar to Wingham et al. (2006) we propose determining the phase difference by minimizing the functional

$$\frac{C\psi_{ml}^3(\tau)}{1 - C\psi_{ml}(\tau)} (A\psi_{ml}(\tau) - g(\tau, \chi))^2 \quad (20)$$



with respect to parameters  $\chi = \{\chi_1, \chi_2\}$  where

$$g(\tau, \chi) = \begin{cases} \chi_1 & \tau < \tau_0 \\ \chi_1 + \chi_2(\tau - \tau_0) & \tau > \tau_0 \end{cases} \quad (21)$$

In Eq. (21)  $\tau_0$  is the time determined from the range retracker. After determining  $\chi_1$  the surface slope is recovered from

$$\theta = \sin^{-1}(\chi_1 / (k_0 \beta)) + \theta_{roll} \quad (22)$$



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

where  $k_0$  is the carrier wave-number,  $\beta$  the interferometer baseline and  $\theta_{roll}$  is the satellite roll angle as read from the SARin FBR data file and corrected for any bias.

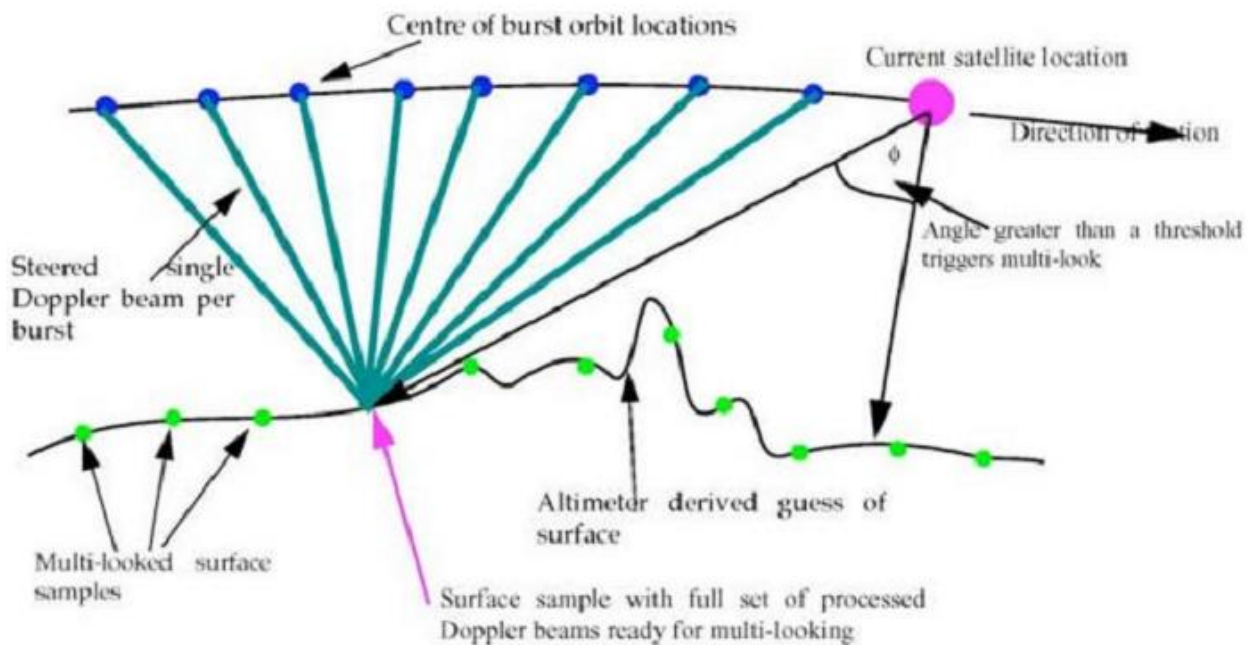






Figure 11: Schematic of bursts, the fan of Doppler beams, ground points and multi-looks

### 3.9 Retracking

Retracking of the multi-looked waveforms provides the geodetic height above the reference ellipsoid, or, with a geoid model, the orthometric height. For simple waveforms such as specular shapes it is possible that the OCOG/Threshold retracker could perform as well as more sophisticated retrackers. However, we have incorporated a number of simple empirical retrackers (see Figure 12) based on an exponential rise of the leading edge followed by an exponential decay at a different rate. In addition, trackers with two peaks were considered. The latter can yield erroneous water heights if the strongest return does not correspond to the required water level but to a secondary water body. In practice, we preferred the earliest returned peak unless the

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

subsequent peak was substantially (factor of 2) larger in amplitude. Of course, in retrospect, selection of the wrong waveform peak can be inferred from the height data across a wide body of water by comparison against adjacent heights. However, any procedure based on comparison with adjacent values is inapplicable for narrow rivers.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

### 3.9.1 Tracker Waveform Descriptions

The set of empirical retrackerers below involves a number of constant parameters estimated by a least squares approach by fitting the each tracker shape to the observed waveform.

During extensive recomputations after submission of D4050 a computing error was identified. This error affected the construction of the multi-look waveforms with the consequence that there is a significantly reduced number of double peaked waveforms compared to D4050. As a result the number of empirical waveforms has been reduced from 6 in D4050 to 5 in this issue as waveform Type 6 of D4050 is never used.

#### Type 1 (Specular):

Waveform given by positive and negative exponential branches with continuity at branch  $t=t_0$ .

Parameters:  $a, t_0, k_1, k_2$



$$f = \begin{cases} (-k_1(t-t_0))^{1/2} & t \leq t_0 \\ (k_2(t-t_0))^{1/2} & t \geq t_0 \end{cases}$$

$$P = a e^{-f^2}$$

This waveform is intended to capture quasi-specular waveforms.

#### Type 2 (Ocean like – calm water)

Waveform has three sections: A positive exponential function, a negative exponential joined by a cubic polynomial to ensure continuity of function and first derivative. This waveform allows for a wider peak than Type 1 but reduces to Type 1 as  $t_1 \rightarrow 0$ .

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Parameters:  $a, t_0, t_1, k_1, k_2$

$$t_a = t_1^{-3/4}$$

$$c_2 = (2.75k_2t_a - 2k_1) / t_1$$

$$c_3 = (k_1 - 1.75k_2t_a) / t_1^2$$

$$f = \begin{cases} k_1(t-t_0) & t \leq t_0 \\ k_1(t-t_0) + c_2(t-t_0)^2 + c_3(t-t_0)^3 & t_0 \leq t \leq t_0 + t_1 \\ k_2(t-t_0)^{1/4} & t \geq t_0 + t_1 \end{cases}$$

$$P = a e^{-f^2}$$

### Type 3 (Ocean like – ruffled water)

Waveform has 4 branches. The first three being same as in Type 2 but with a fourth branch enabling a more gradual decay at the tail.

Parameters:  $a, t_0, t_1, t_2, k_1, k_2, k_3$



$$t_a = t_1^{-3/4}$$

$$c_2 = (2.75k_1t_a - 2k_2) / t_1$$

$$c_3 = (k_1 - 1.75k_2t_a) / t_1^2$$

$$f = \begin{cases} k_1(t-t_0) & t \leq t_0 \\ k_1(t-t_0) + c_2(t-t_0)^2 + c_3(t-t_0)^3 & t_0 \leq t \leq t_0 + t_1 \\ k_2(t-t_0)^{1/4} & t_0 + t_1 \leq t_0 + t_2 \\ k_2t_2^{1/4} + k_3(t-t_0-t_2) & t \geq t_0 + t_2 \end{cases}$$

$$P = a e^{-f^2}$$

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

#### Type 4

Waveform is a Type 2 with additional Type 1 peak

Parameters:  $a_1, a_2, t_0, t_1, t_2, k_1, k_2, k_3, k_4$

$$t_a = t_1^{-3/4}$$

$$c_2 = (2.75k_2t_a - 2k_1) / t_1$$

$$c_3 = (k_1 - 1.75k_2t_a) / t_1^2$$

$$f_1 = \begin{cases} k_1(t-t_0) & t \leq t_0 \\ k_1(t-t_0) + c_2(t-t_0)^2 + c_3(t-t_0)^3 & t_0 \leq t \leq t_0 + t_1 \\ k_2(t-t_0)^{1/4} & t \geq t_0 + t_1 \end{cases}$$

$$f_2 = \begin{cases} (-k_3(t-t_0-t_2))^{1/2} & t \leq t_0 \\ (k_4(t-t_0-t_2))^{1/2} & t \geq t_0 \end{cases}$$

$$P = a_2 e^{-f_2^2} + a_1 e^{-f_1^2}$$

#### Type 5



Waveform a composite of two Type 1 peaks.

Parameters:  $a_1, a_2, t_1, t_2, k_1, k_2, k_3, k_4$

$$f_1 = \begin{cases} (-k_1(t-t_1))^{1/2} & t \leq t_1 \\ (k_2(t-t_1))^{1/2} & t \geq t_1 \end{cases}$$

$$f_2 = \begin{cases} (-k_3(t-t_1))^{1/2} & t \leq t_2 \\ (k_4(t-t_1))^{1/2} & t \geq t_2 \end{cases}$$

$$P = a_1 e^{-f_1^2} + a_2 e^{-f_2^2}$$

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

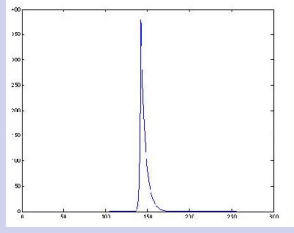
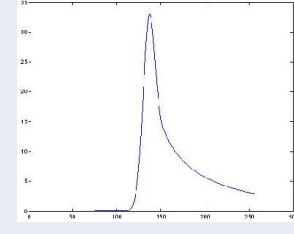
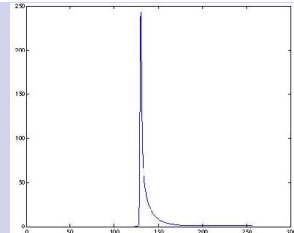
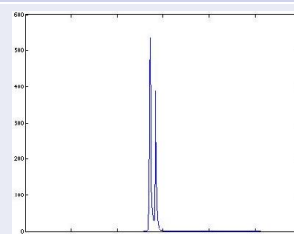
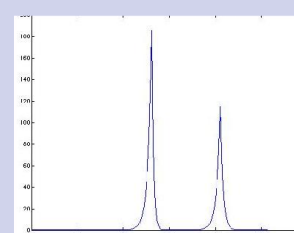


Retracker #	Description	Waveform Shape
1	Specular (still water)	
2	Ocean like (ruffled water)	
3	Ocean like with fall away at high # bins (ruffled waters)	
4	Two specular peaks (strong returns off two patches of still water)	
5	Retracker Type 2 with additional specular peak (ruffled and still water)	

Figure 12: SAR empirical retracker. X-axis is gate number (1-256), Y-axis is power amplitude

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

The empirical waveform was fitted to the observed data using least squares solving for the location of the peak(s), the peak amplitude and the rate of exponential rise and decay of the peak. A total of 4-7 parameters were estimated with the larger numbers for the twin peak forms.

Each waveform was fitted and the best fitted waveform in the sense of the lowest Normalized Residual Error as defined by (SAR)

$$NRE = \frac{\sum_{i=1}^{N_{pad}} (P_i^{obs} - P_i^{mod})^2}{\sum_{i=1}^{N_{pad}} (P_i^{obs})^2} \quad (23)$$

was chosen. In Eq. (23)  $P_i^{obs}$ ,  $P_i^{mod}$  are the observed and modelled power within bin  $i$  respectively. In practice the single peak waveforms were preferred unless the NRE reduced by 10% or more to avoid over parameterization with twin peaks. NRE in Eq. (23) is the sum of the residual power squared normalized with respect to the total waveform power squared. Further, for Baseline B,  $N_{pad} = 256(1024)$  for SAR(SARin) denotes the number of samples after zero padding the original samples as in section 3.4.



All heights from the empirical retrackers and OCOG/Threshold (threshold=0.75) were stored. The value of the threshold was chosen to give comparability between the two height values.

The PVR (D4200) will contain details of the results of the multi-look analysis including NRE performance, consistency in derived height values and comparison against in situ and Jason-2 data where possible. The multi-look waveforms are weighted using a ‘‘Hamming window’’ with unit weight for the central burst beam with minimum angle to the ground-point decreasing as the beam angle increases for bursts before and after this central beam, via

$$w(n) = 0.54 + 0.46 \cos(\pi n / N) \quad (24)$$

where  $N$  is the number of beams before or after the burst of minimum beam angle and  $n = 0, \pm 1, \pm 2, \dots, \pm N$ . Thus, if  $N=2$ ,  $w = 0.31, 0.77, 1.0, 0.77, 0.31$  for the 5 beams centred on the burst with minimum beam angle. We note that Eq. (24) is equivalent to the standard expression for a Hamming window as the usual definition has  $2\pi n / 2N$  for the cosine argument.



Analyses over the Mekong in D4050 erroneously suggested that a very low number of multi-looks may be preferred, as little as  $N=3$ , that is 2 either side of the burst echo beam direction closest to the nadir. However, after identifying a computing error multi-looking over a large number of waveforms in the stack is now seen to be preferable. As an example consider the 20 points of the pass on 19 April 2011 across the Mekong where the satellite is travelling from North to South

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

(Figure 13). Figure 14 shows typical specular waveforms from points over the Mekong and the very close agreement between the empirical and OCOG/Threshold retrackerers. The NRE values for the 21 points are plotted in Figure 15 for sample values of  $N=10$  to  $N=110$ . High values of NRE at several epochs implies a poor fit of the empirical curves to the observed waveform; at the third epoch (0.2 sec) the NRE for  $N=10$ , 20 and 40 exceed the NRE cutoff limit of 0.4 utilised in this study. This limit rejects the empirical fit where the observed and fitted waveforms differ in relative power by 40% or more with the inference that the waveform is complex and not from an inland water body. However, with the exception of  $N=10$  all other  $N$  values are acceptable in terms of NRE.

Figures 16 and 17 plot the derived orthometric heights where the selection of multi-look constructions have been retracked using the empirical and OCOG/Threshold retrackerers respectively. This pass crosses the Mekong in low flow where the heights of the river banks are clearly evident some 10m above the water level. The first three points are over land with various degrees of off-pointing to the Mekong with the next 8 points showing predominance of reflection off the water. The final 10 points are over the southern bank. At the scale in the upper plots of Figures 16 and 17 there is little to distinguish between the derived heights apart from the unrealistic values of  $N=10$ . For the larger values of  $N$ , the number of acceptable heights is 7 or 8 which decreases to 4 for the lower values of  $N = 10$ . The orthometric heights for  $N=110$ ,  $N=90$  and  $N=70$  are indistinguishable. The lower plots of Figures 16 and 17 are zoomed onto the river surface. Given the low number of points statistical analysis is not viable and only qualitative inferences are possible including the near equivalence of the values for  $N \geq 40$  and between the empirical and OCOG/Threshold retrackerers.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

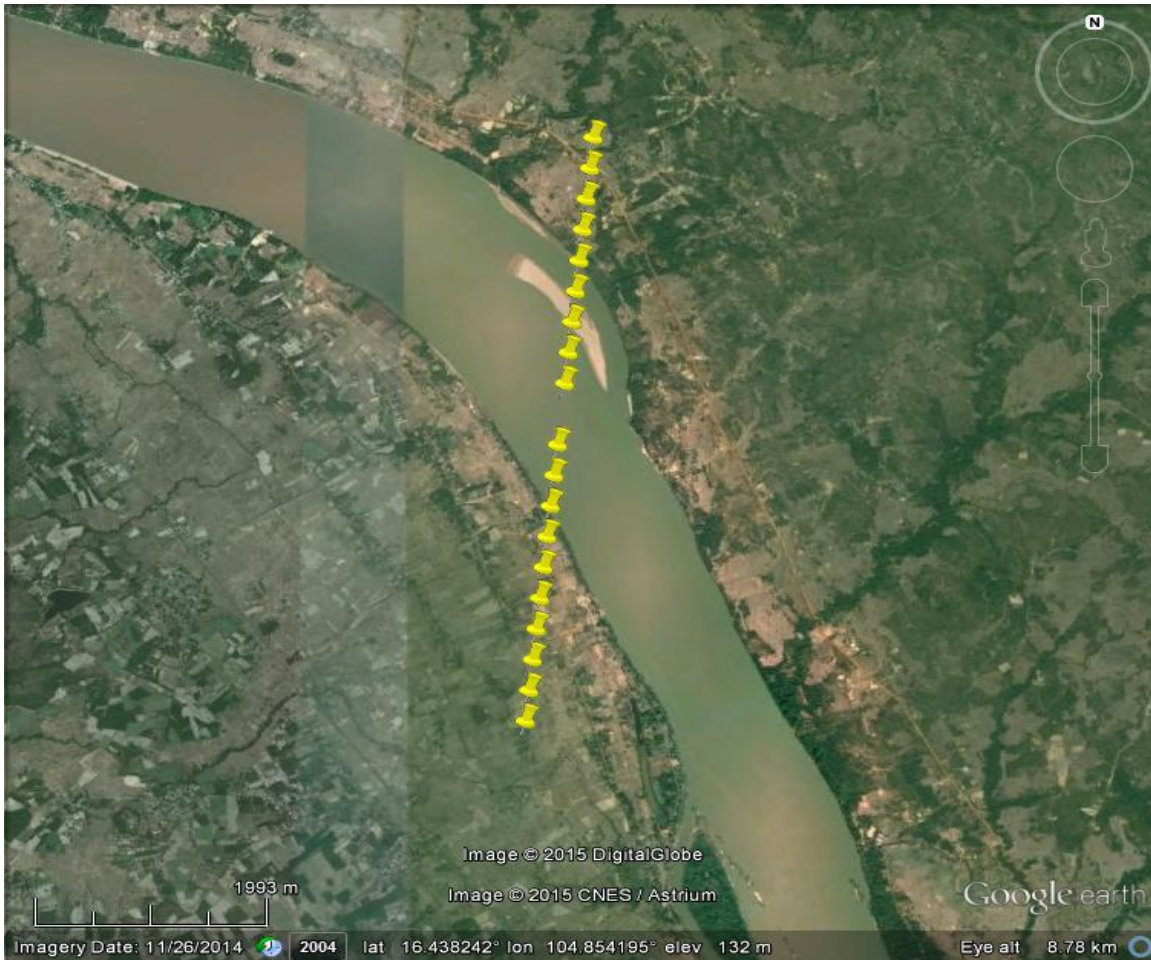




Figure 13: Google Earth image of 19 Apr 2011 Mekong crossing. Satellite ground track N-S.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

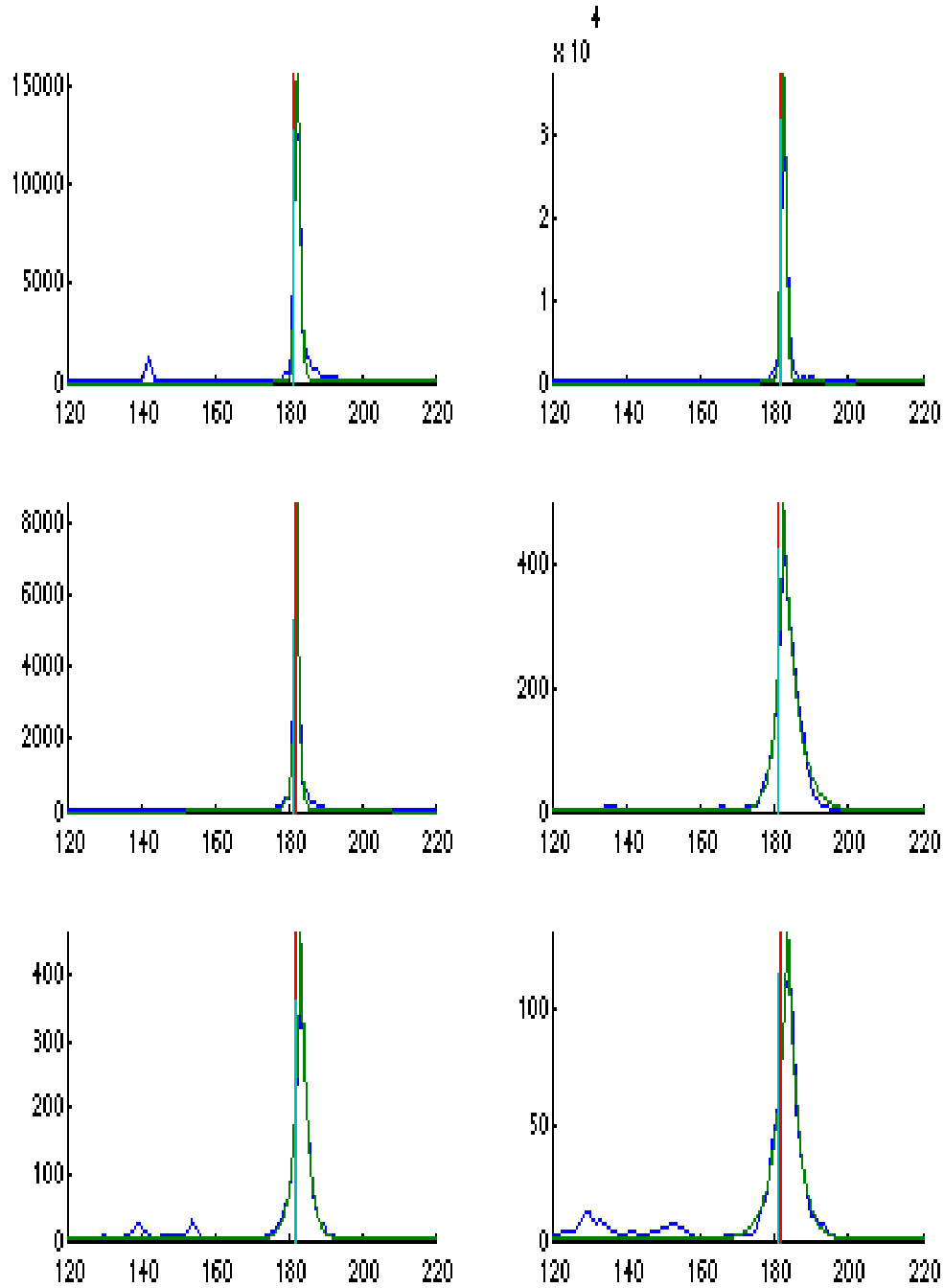




Figure 14: Waveforms (blue curve) across the Mekong (N=110) with empirical retracker (green curve) with retracted bin given by red line; cyan line shows OCOG/Threshold retracted bin.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

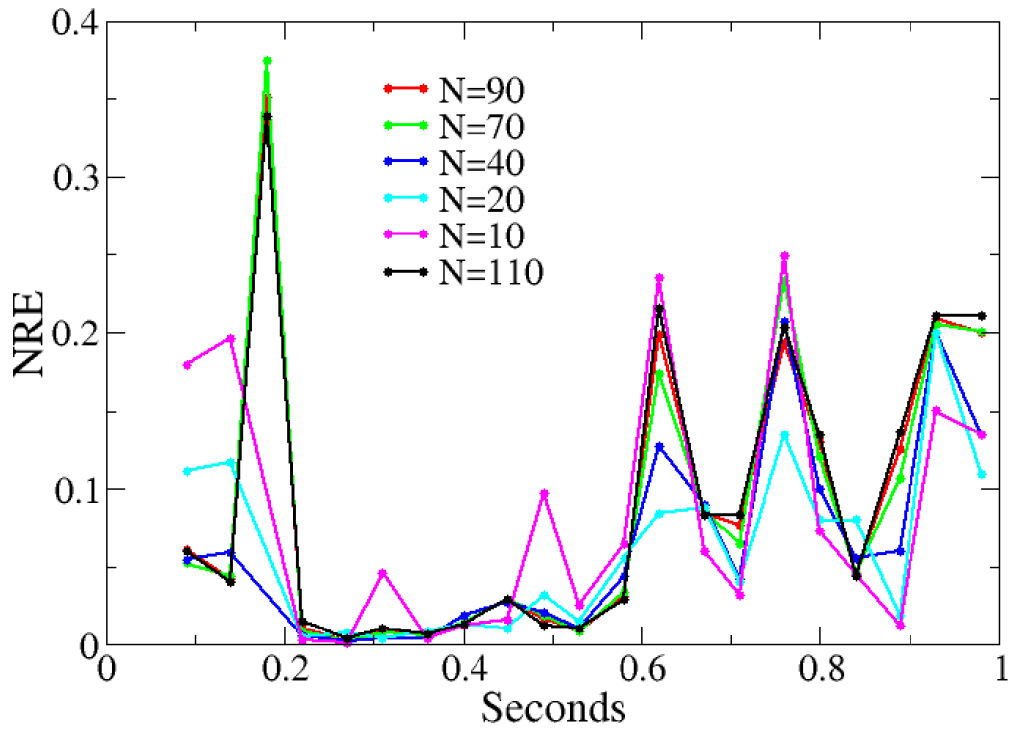


Figure 15: NRE for waveforms across the Mekong

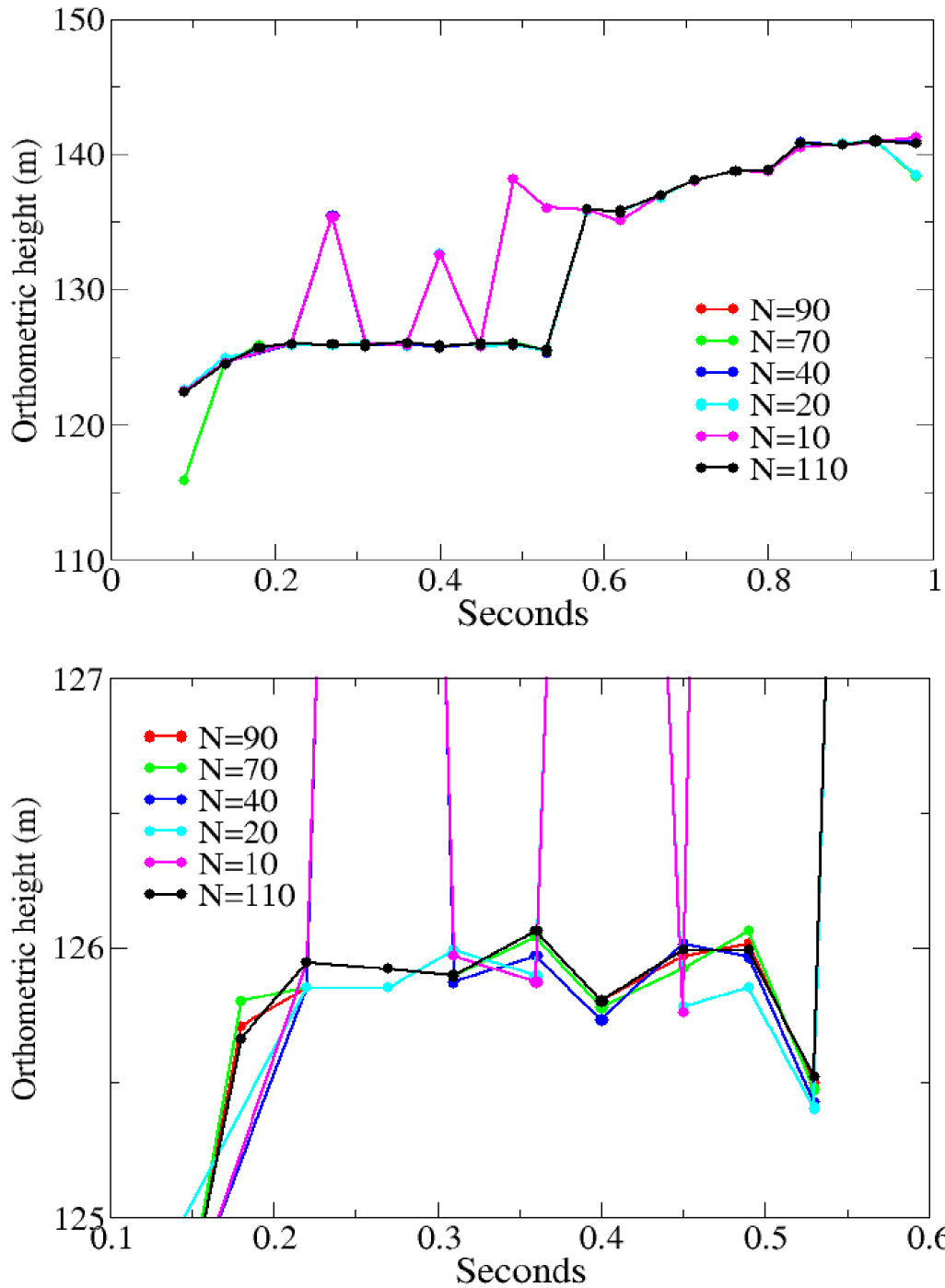


Figure 16: Orthometric heights from empirical retracker (Figure 12) across the Mekong for 19 Apr 2011 (day 109 in year; seconds after 59084) for various number of echoes in multi-look. Upper plot full profile; lower plot a zoomed version of the river surface.

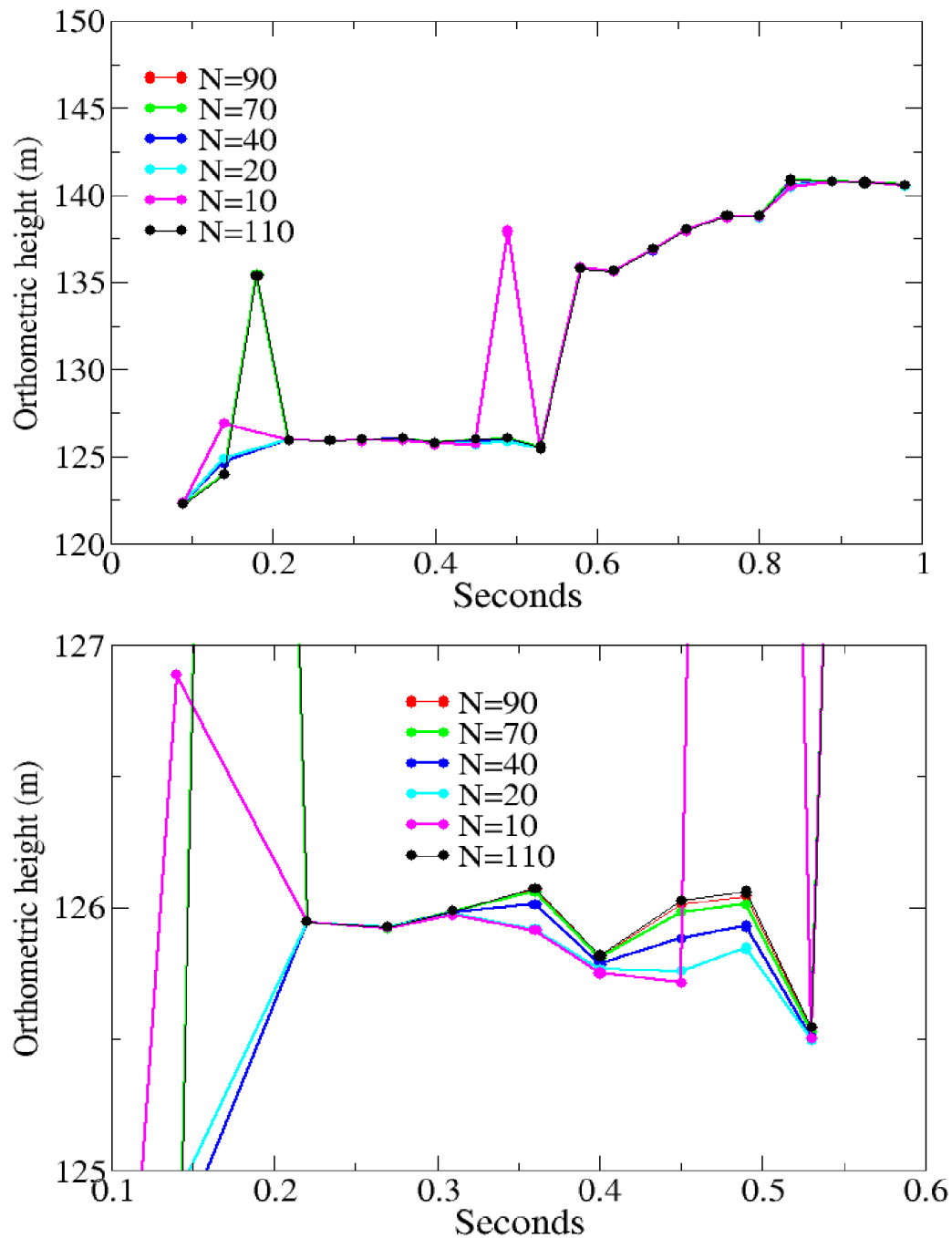


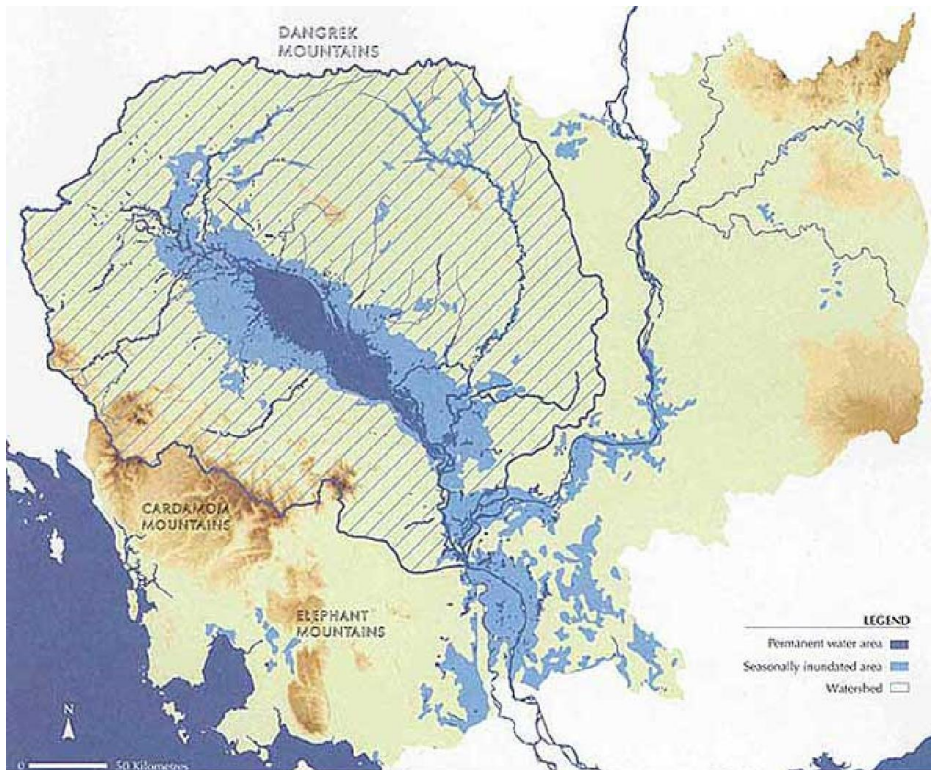


Figure 17: Orthometric heights from OCOG/Threshold retracker across the Mekong for 19 Apr 2011 (day 109 in year; seconds after 59084) for various number of echoes in multi-look. Upper plot full profile; lower plot a zoomed version of the river surface.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



To allow statistical inference analysis was also made across the Tonle Sap lake near its full extent (Figure 18).



**Figure 18: Seasonal variation in extent of Tonle Sap**

For the pass on 3 Dec 2011 a total of 68 ground points (Figure 19) were identified as over the lake. The analysis followed a similar approach to the river crossing but with more ground points a statistical study was possible. The derived orthometric heights were fitted with a quadratic to remove any residual geoid. Figure 20 plot the orthometric heights derived using the empirical retracker and the OCOG Threshold retracker (threshold =0.75) for the pass. The heights for the selected N show systematic offsets from each other due to the differing slope of the leading edge but this is not considered important as long as a consistent choice of N is used through the SAR or SARin processing.



The residuals of fit in Figure 20 were taken to represent the scatter in the estimated water heights. The degree of fit is summarized in Table 2. That table shows that for the empirical and the OCOG/Threshold retracker the minimum sum of the squares of the errors is for N = 40 and N=20

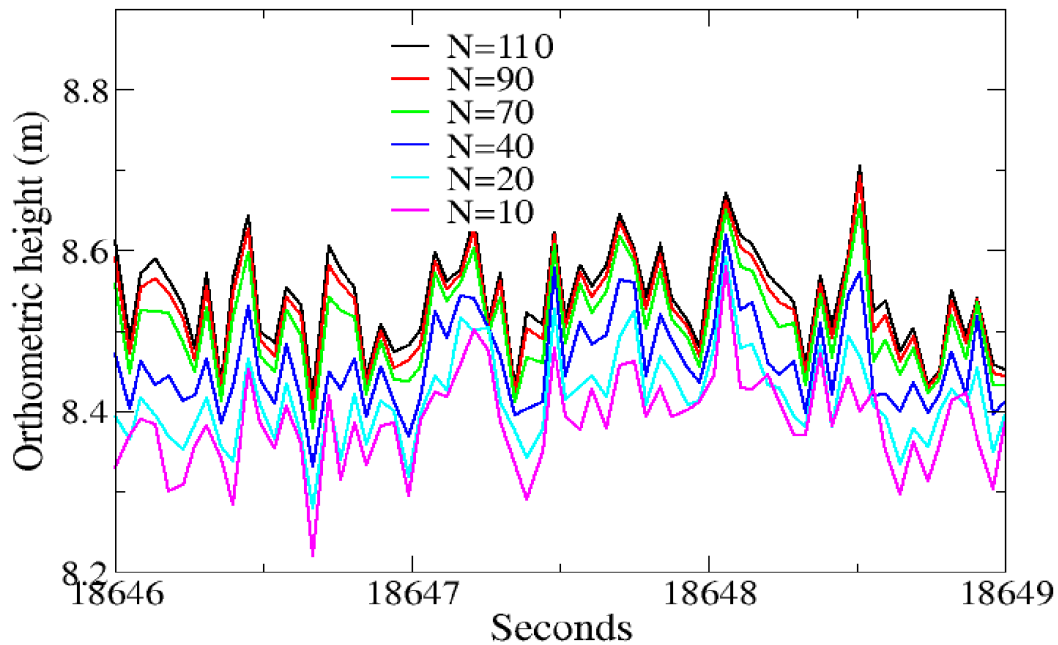
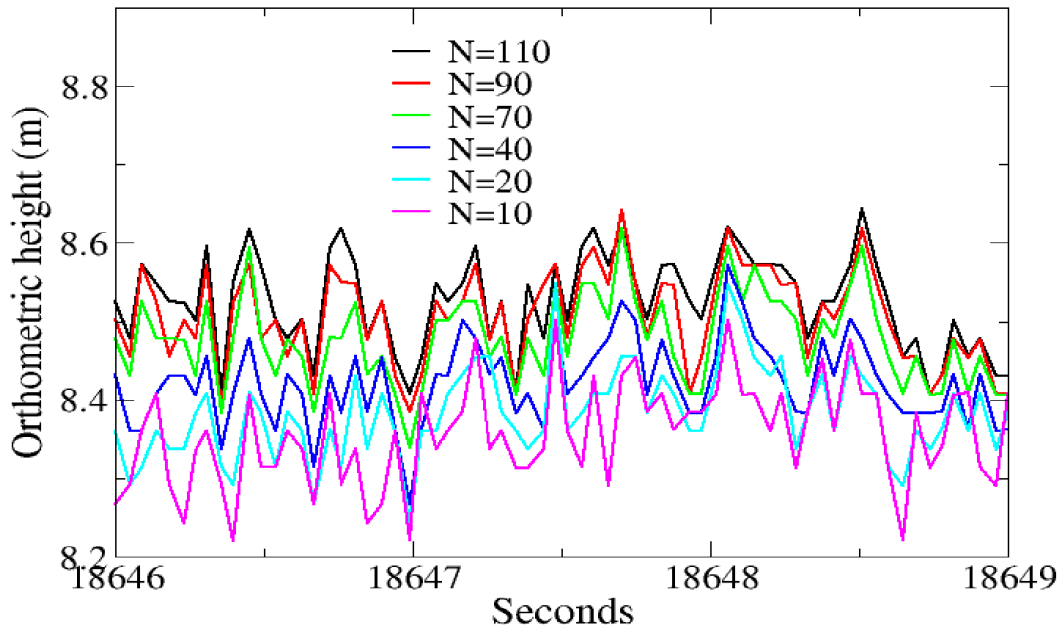
		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

respectively. All retrackerers appear somewhat sensitive to N but the best fit values are near identical. Table 3 gives the number of waveforms retracked with the single peak retrackerers (1-3) and the dual peak retrackerers (4-6). The total numbers of single and dual peaks is also given. It is noted that the number of dual peaks increases as the number N decreases; that is as fewer looks are included in the multi-look waveform. Further, the number of waveforms that were retracked best with retracker type 3 (ocean-like form) increased for N=40 and N=20. This reflects that the tail of the waveform has been modified by inclusion of a larger number of looks and that the retracker of Type 1 can adapt to best-fit the waveform. It does not mean that the waveforms are becoming more specular but rather that the exponential decrease in the tail is modelled by the simulated waveform. Based on Table 2 we decided to use N=40 as our preferred value giving multi-looks over 79 burst echoes.





**Figure 19: Points across Tonle Sap (3 Dec 2011)**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





**Figure 20: Geoid height measurements across Tonle Sap (3 Dec 2011) with various multi-look combinations. Upper: Retracked using empirical waveforms (Figure 12). Lower: OCOG/Threshold retracker used.**



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Multi-look N	Sum of Square of the errors: Empirical retrackers (m <sup>2</sup> )	Sum of Square of the errors: OCOG/Threshold (m <sup>2</sup> )
110	0.222	0.251
90	0.225	0.245
70	0.212	0.230
40	0.171	0.196
20	0.184	0.176
10	0.240	0.194



**Table 2: Statistics of fit for pass across Tonle Sap.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

	Empirical tracker type			
	2	3	4	5
<b>N=110</b>	30	37	1	0
<b>N=90</b>	26	40	2	0
<b>N=70</b>	18	49	1	0
<b>N=40</b>	5	57	3	3
<b>N=20</b>	5	57	3	3
<b>N=10</b>	8	34	11	15

	Empirical tracker type						
	1	2	3	4	5	Single Peak	Dual Peak
<b>N=110</b>	0	30	37	1	0	67	1
<b>N=90</b>	0	26	40	2	0	66	2
<b>N=70</b>	0	18	49	1	0	67	1
<b>N=40</b>	0	5	57	3	3	62	6
<b>N=20</b>	0	5	57	3	3	62	6
<b>N=10</b>	0	8	34	11	15	42	26

**Table 3: Tracker types (Figure 12) for the Tonle Sap points. The single peak column is the sum of columns 1-3 and the dual peak column that of 4-5. A total of 68 waveforms were retracked.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

### 3.9 SARin Analysis



(This section contains material not present in D4050)

The SARin mode is particularly pertinent over polar ice margins where the ground slope can affect the point of closest approach (POCA). As the first radar return is along the direction perpendicular to the ground a sloping surface will have a POCA that is different from the nadir point. Adjustment over sloping surfaces for the POCA is necessary for correct interpretation of the waveform return and height recovery. Analysis of the waveforms from the two antennae in SARin mode follows that for SAR under the assumption that the two waveforms illuminate the same surface area. If the satellite roll is accounted for, heights from the two antennae should be near identical over flat terrain and inland waters. Disparity between the measurements can then be used as a quality control with any difference identifying inconsistencies.

Use can also be made of the phase difference between the waveforms (Eq 18) and the coherence (Eq. 19) from the two antennae to investigate the ground slope. Such an approach requires least squares analyses of the weighted phase difference between the two antennae waveforms. Over inland waters reflectance is often near specular with the waveform power limited to a very small number of bins in the waveform. Outside these bins the coherence decreases as the signal to noise ratio reduces. Recovery of the ground slope from the waveform phase will also be somewhat ill-conditioned from a few bins leading to spurious results. Thus, after experimentation in varying the value, we used 20 bins either side of the retracked height as a compromise.



As an exemplar of SARin we consider the South-North pass on 5 May 2012 across the Amazonas near the Tabatinga gauge (Figure 21). Figure 22 plots the difference between the heights from the two antennae and the derived ground slope. The Amazonas flood plain is visible as the latitudinal extent of low variability centered on  $4.2^{\circ}\text{S}$ . The difference in height measurements between the two antennae has a mean near zero and standard deviation near 1 cm. Note this difference uses the convention of height from left antennae along flight direction minus height from the second (right-side) antenna. Figure 23 shows the waveforms from the two antennae, the coherence and the ground slope for two locations over the Amazon. The waveforms from the left and right antennae are near identical while the coherence near the retrack point is near unity in both waveforms. The least squares estimated ground slope angle is  $0.027^{\circ}$  for location #274 and  $0.030^{\circ}$  for location #275.

The Amazon near Tabatinga has passes that are about 2 days apart, but differ in longitude by  $1.2^{\circ}$ . Thus, given the small time difference we can assume that these measurements record the same river flows. The difference between the river heights at the two locations provides an estimate of

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



the slope of the river over a stretch of the Amazon some 150 km in length. Taking the average of these values gives a river slope of 3.95 cm per km along the river. This is equivalent to a ground slope of  $-0.002^\circ$ . This illustrates the accuracy needed to be achieved by ground slope deviation from SARin to be meaningful in the context of river hydrology particularly taking into account the noise of the SARin instrument and the uncertainty in the roll angle.

Thus, it is probably unrealistic to attempt to determine the river slope from SARin unless all systematic errors in the slope equation are less than this value. In practice, the analyses in D4200 will show both a positive ground-slope for North-South and South-North passes across the Amazon near the Tabatinga gauge and across the lower Brahmaputra. For confidence in the solution, the river slope should change sign with the pass direction. Results in D4200 will show that the mean slope across both rivers is near  $+0.02^\circ$ . One possibility for this positive value is an over-estimation of the roll-bias to be applied to BaselineB and applied at source in BaselineC. A reduction of  $0.02^\circ$  in the roll-bias to about  $0.086^\circ$  is almost identical to that found by Smith and Scharroo (2011) from subcycle 014 (19 Apr 2011 – 18 May 2011) and as used in Garcia et al. (2014).

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 21: Ground track across Amazonas 5 May 2012 (SARin mode)**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

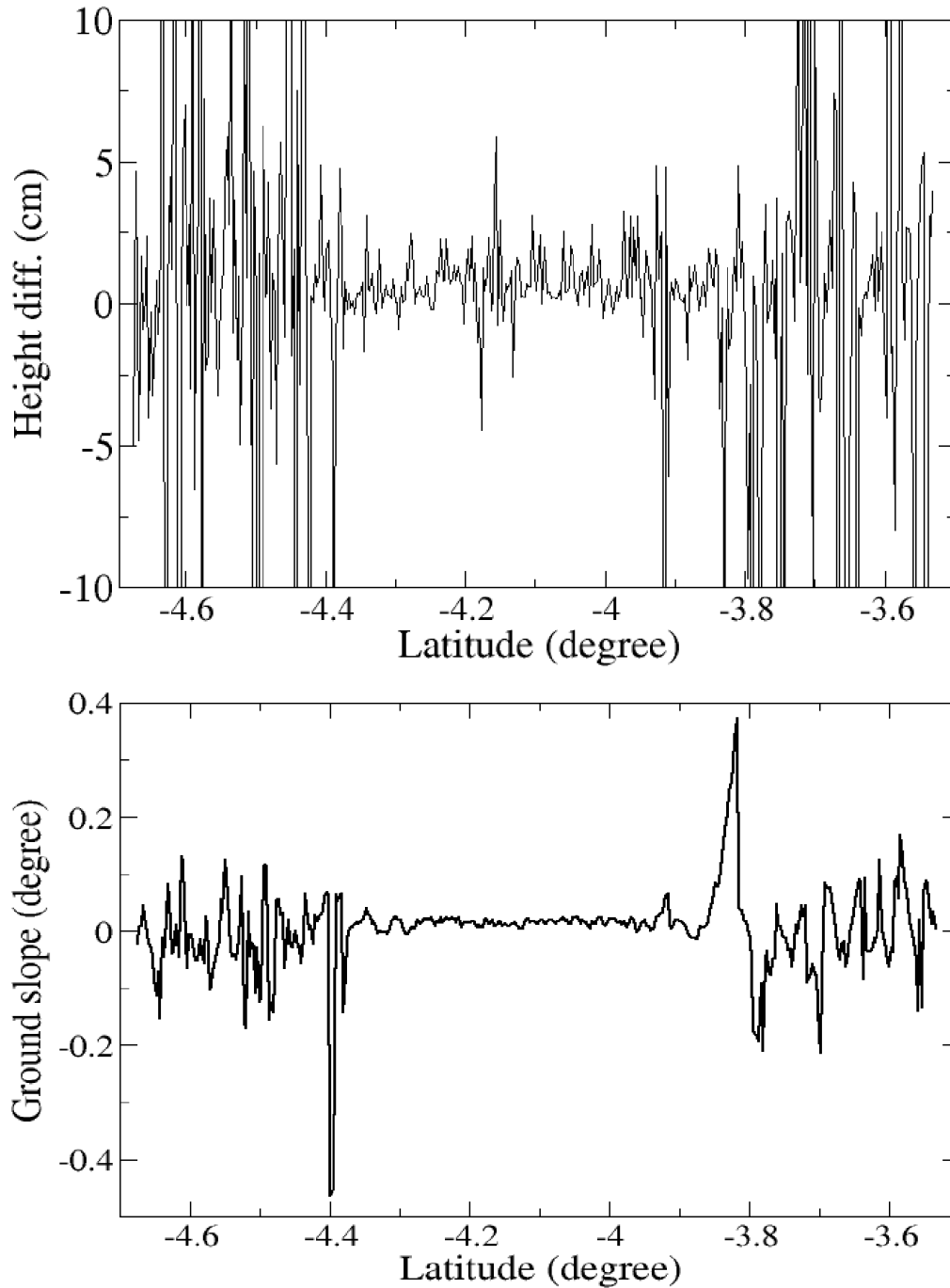


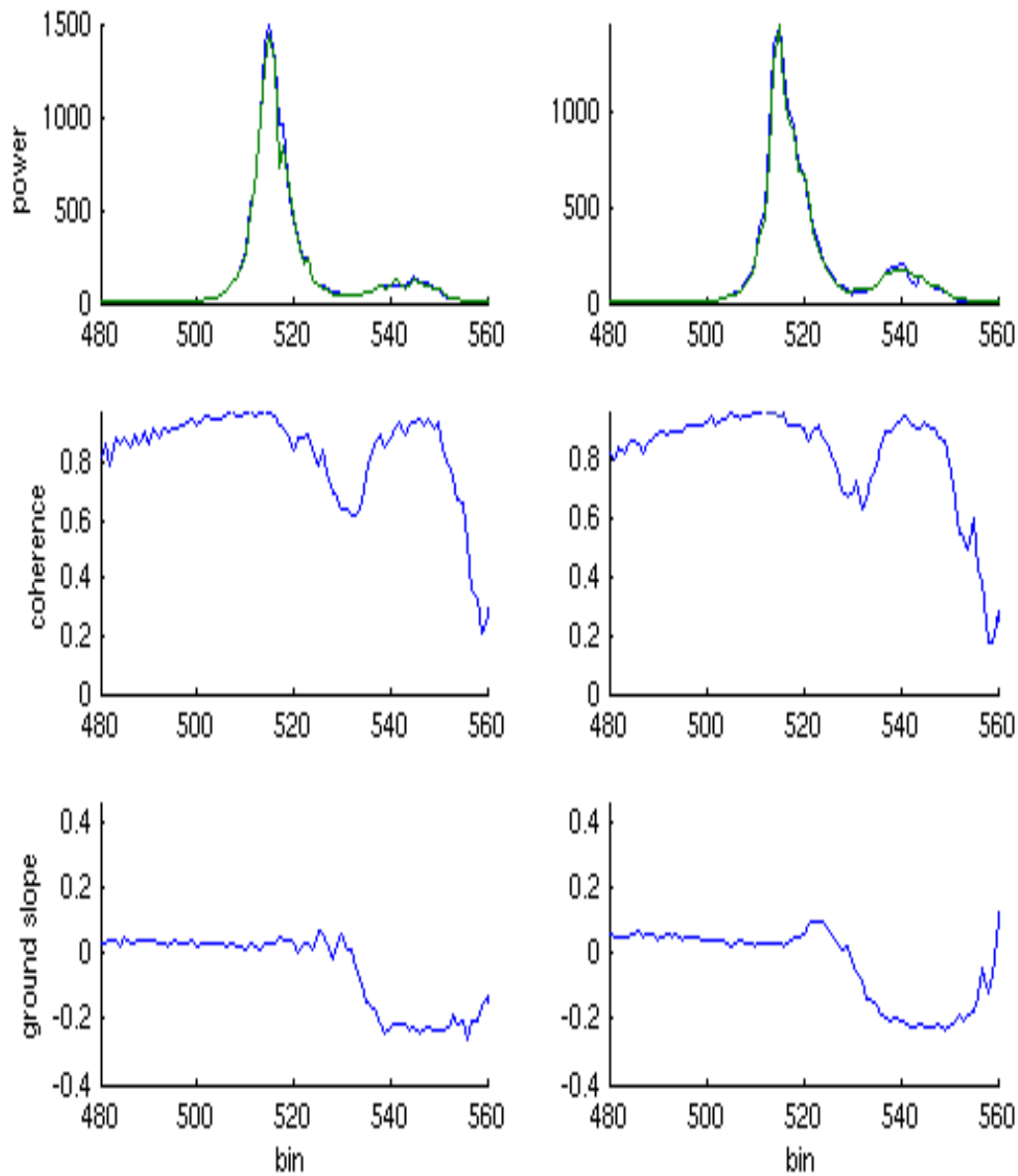




Figure 22: (Upper) Difference between heights from two antennae; (lower) ground slope across. SARin Amazonas 5 May 2012. (Latitude -4.2° corresponds to longitude 70.179°W)

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 23: SARin waveforms (upper), coherence (middle) and ground slope in degrees (lower) across Amazonas. X-axis is bin number (1-1024 in SARin mode). In the upper plot the left antennae is coloured blue and the right antennae is green. Left hand column location #274 (3.930 °S 70.207 °W); right hand column location #275 (3.927 °S 70.207 °W). Date of pass 5 May 2012.**



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 4 SAR FBR Analysis: Summary



Sections 2-4 have summarized the algorithmic development undertaken to process FBR SAR and SARin L1a data for inland waters. There are a number of pertinent points.

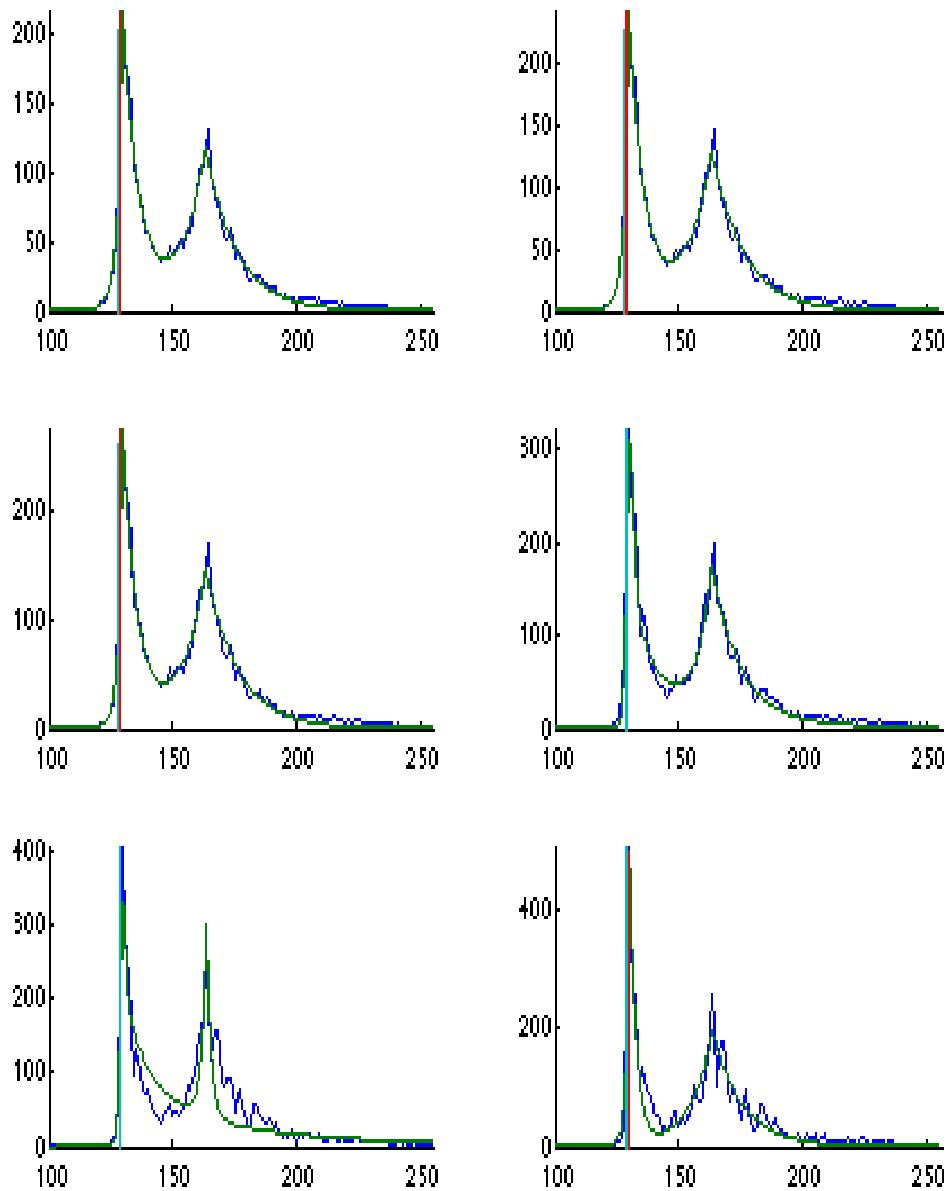
- Inland waters cannot be simply compared with the ocean surface where the distribution of reflectors in the bursts are all oceanic and will tend to be homogeneous given the stronger influence of waves on the reflectance. Further, the reflection characteristics from consecutive bursts are likely to be very similar in the open ocean as the illuminated surface is entirely ocean. In other words, the oceanic reflectors in one burst are near indistinguishable from previous and subsequent bursts. Over non-oceanic surfaces the forward motion of the satellite means that small bodies of water will not fill the illuminated surface; the slices of the illuminated surface will contain varying proportions of inland water reflectance. There is contamination from the land, potentially other water or high reflectance surfaces close to the water body as well as highly reduced surface ruffling with small waves or ripples. Thus, the steering of beams to the ground point within a stack will be less homogeneous in water reflections than over oceans. Given the reduced homogeneity it is not unsurprising that a value of N significantly less than the maximum possible number in the stack gives slightly better results. For rivers a value of N=40 is sufficient to provide a well-defined waveform.
- That noise increases in the waveform is evident for  $N \leq 40$  in Figure 24 – Figure 30 for the pass across Tonle Sap. These points are plotted on a Google Earth image as Figure 31. The sequence in Figure 25 and Figure 26 show that waveforms exhibit a dual peak and that the peaks become less smooth and the empirical tracker a poorer fit for  $N \leq 40$  and particularly N=20 and N=10. These waveforms are close to the edge of Tonle Sap with dual reflections. For locations closer to the centre of Tonle Sap, i.e. Figure 26 - Figure 30, the waveforms are more typical of that seen over the oceans.
- The algorithmic approach is to use N=40 for SAR multi-looked waveforms. This provides sufficient waveforms to reduce the speckle in the individual waveforms.
- For SARin we similarly adopt N=30. This therefore includes all but a few of the steered waveforms in the multi-look.
- In terms of the retracker the OCOG retracker with threshold = 0.75 has proven to be as precise as the empirical retracker employed. Furthermore, the simplicity of the OCOG/Threshold retracker is also attractive. Due to the high reflectance of water compared with land many waveforms are quasi-specular (c.f. Figure. 14) which are easily retracked by OCOG/Threshold.





		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

- The more specular nature of inland water waveforms is shown in Figure 14 compared to the more ocean-like waveforms over lakes as in Figures 27-31.
- We are not excluding the development of more advanced retrackerers to enhance the accuracy but the expectation is that such a refinement is likely to be small given the other uncertainties and potential systematic errors in the retracking process. For example, errors in the wet tropospheric correction will give rise to a systematic error that will not be picked up by a precision statistic such as the standard deviation or root mean square error. Hence the need for external validation using gauge or other satellite heights (e.g. Jason -2).
- For SARin evidence suggests that the two antennae give near identical waveforms and height estimates.
- The river slope of the Amazon near Tabatinga is  $0.002^\circ$  which therefore illustrates the level accuracy needed to be obtained from SARin for any slope information to be of use in hydrology.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 24: Waveforms ( $i=86$ ,  $t=18645.69$  s,  $lat= 13.0439^\circ$ ,  $lon=104.0775^\circ$ ) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1  $N=110$ ,  $N=90$ ; row2  $N=70$ ,  $N=40$ ; row 3  $N=20$ ,  $N=10$ . X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

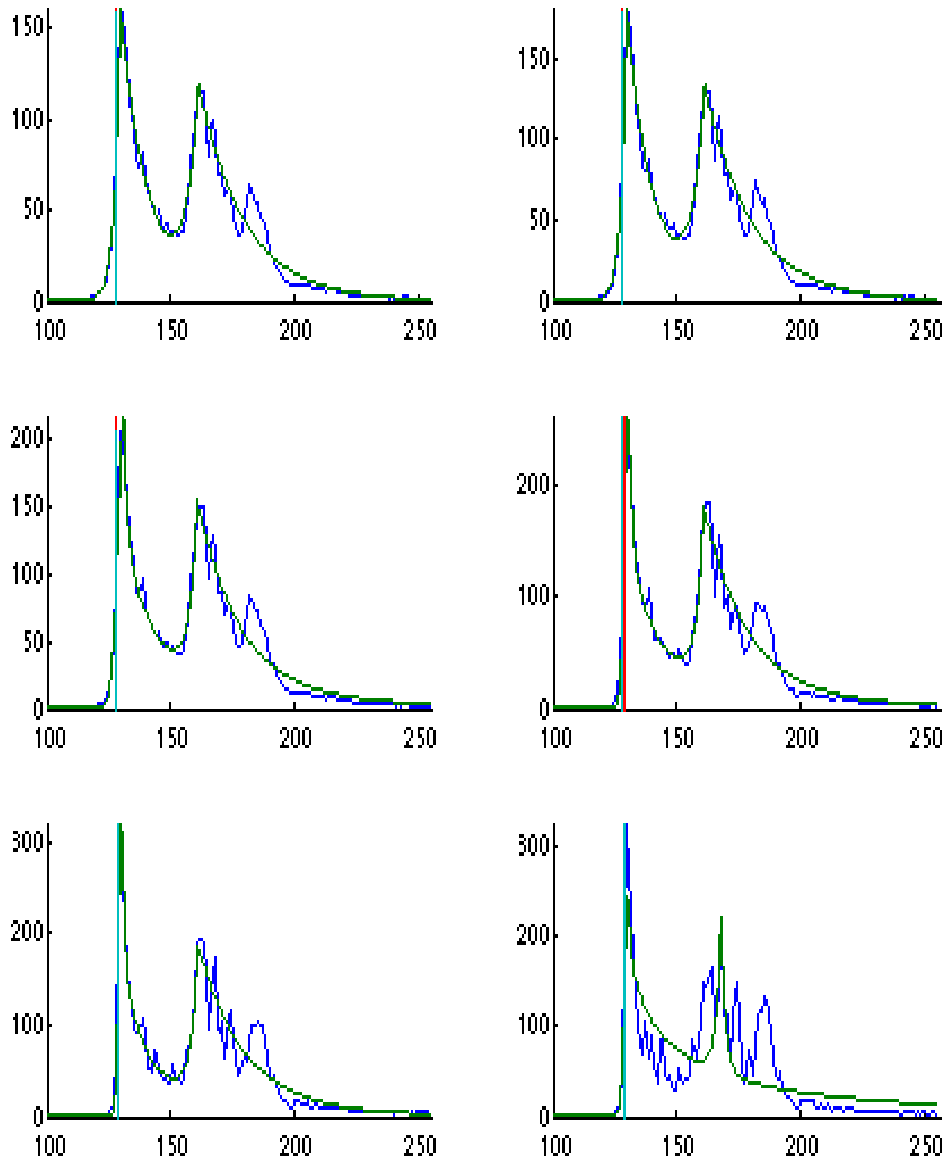


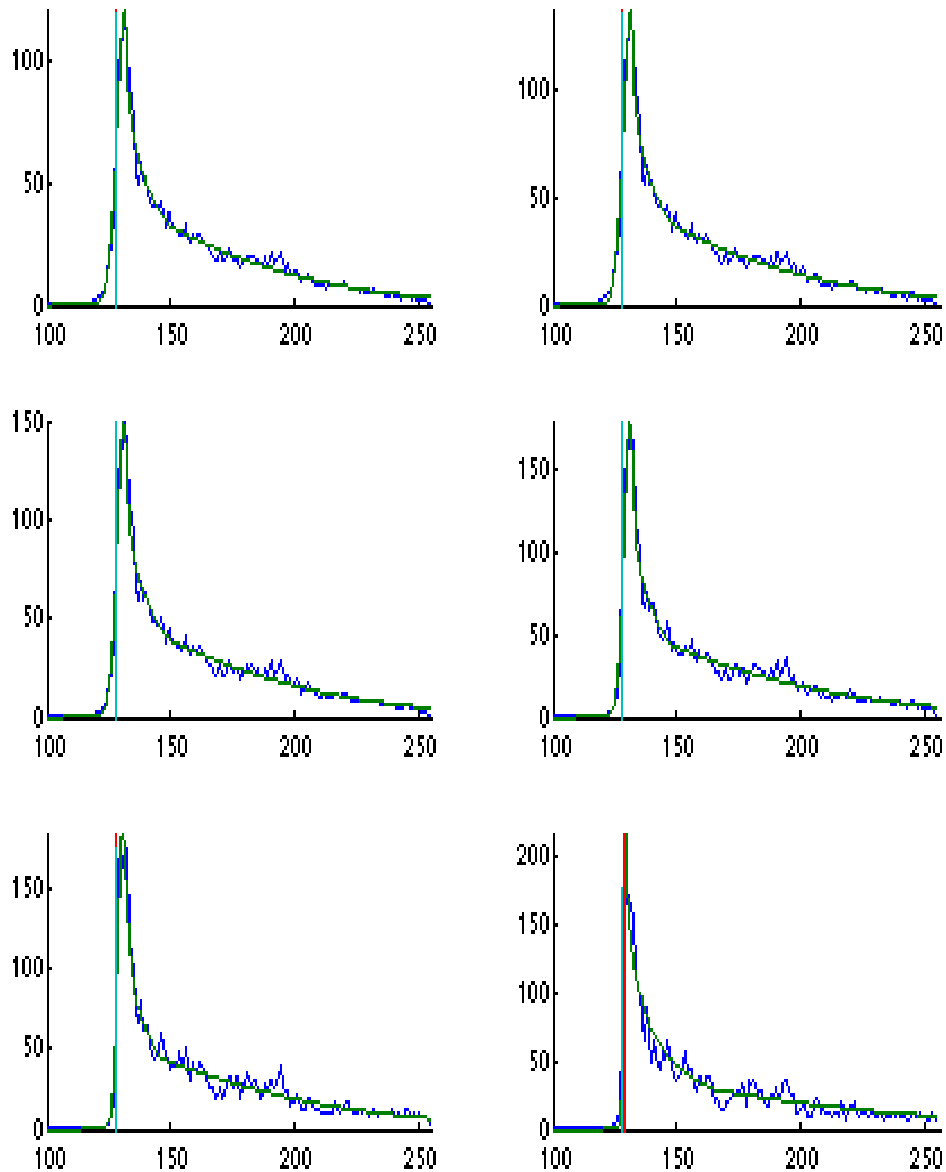




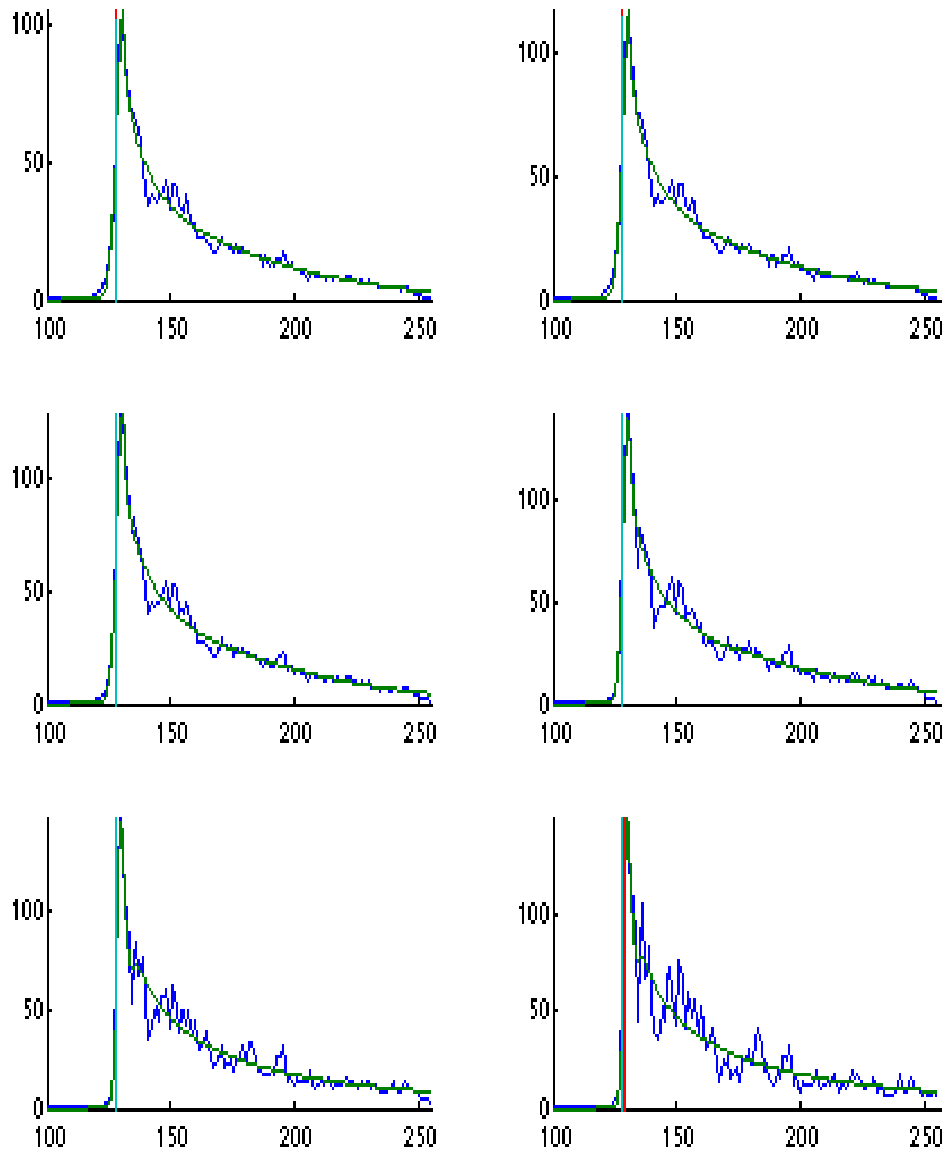
Figure 25: Waveforms ( $i=87$ ,  $t=18645.73$  s,  $\text{lat}=13.0411^\circ$ ,  $\text{lon}=104.0772^\circ$ ) with empirical retracker (green curve) with retracted bin given by red line; cyan line shows OCOG/Threshold retracker bin. From left to right: row 1  $N=110$ ,  $N=90$ ; row 2  $N=70$ ,  $N=40$ ; row 3  $N=20$ ,  $N=10$ . X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





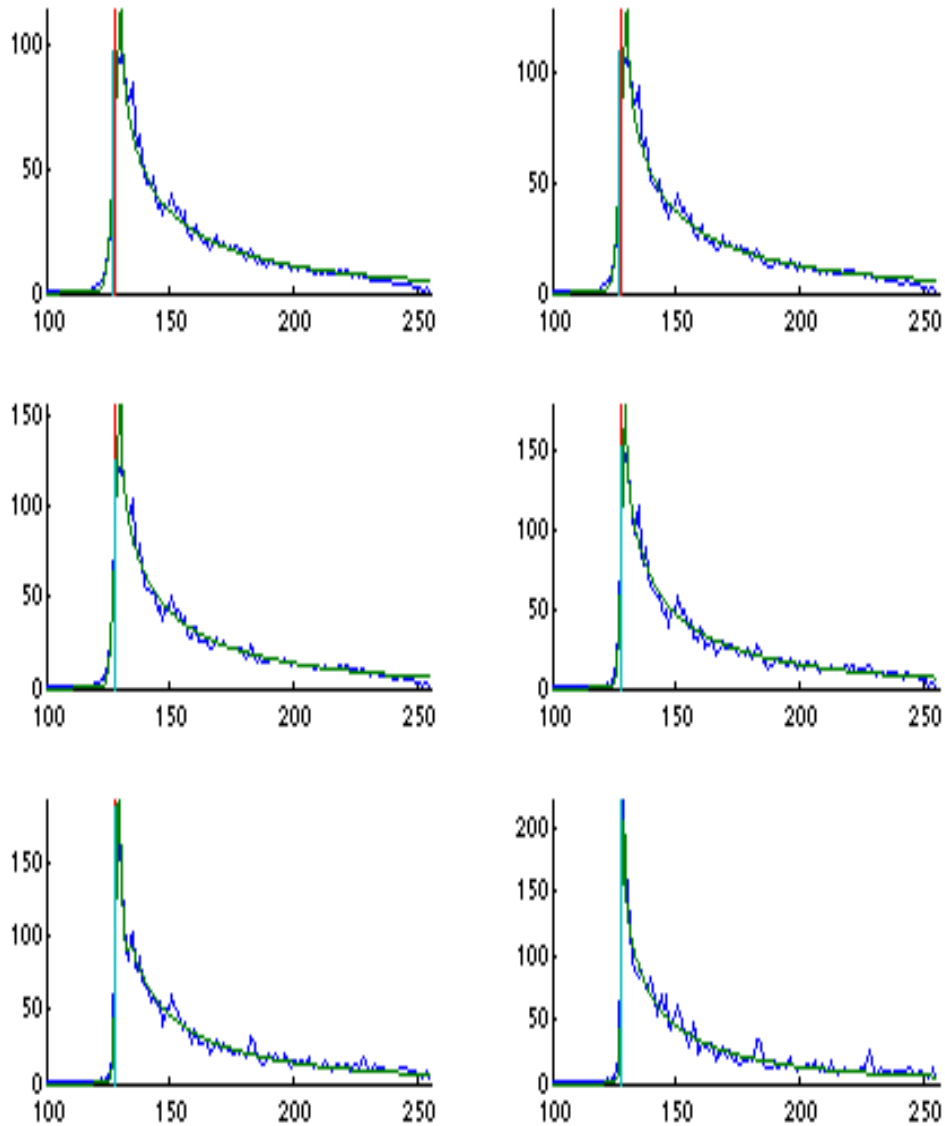
**Figure 26: Waveforms ( $i=118$ ,  $t=18647.12$  s,  $lat= 12.9568^\circ$ ,  $lon=104.0683^\circ$ ) with empirical retracker (green curve) with retracted bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1  $N=110$ ,  $N=90$ ; row2  $N=70$ ,  $N=40$ ; row 3  $N=20$ ,  $N=10$ . X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





**Figure 27: Waveforms ( $i=119$ ,  $t=18647.17$  s,  $\text{lat}=12.9541^\circ$ ,  $\text{lon}=104.0680^\circ$ ) with empirical retracker (green curve) with retraced bin given by red line; cyan line shows OCOG/Threshold retraced bin. From left to right: row 1  $N=110$ ,  $N=90$ ; row2  $N=70$ ,  $N=40$ ; row 3  $N=20$ ,  $N=10$ . X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 28: Waveforms ( $i=120$ ,  $t=18647.21$  s,  $lat= 12.9514^\circ$ ,  $lon=104.0677^\circ$ ) with empirical retracker (green curve) with retracted bin given by red line; cyan line shows OCOG/Threshold retracted bin. From left to right: row 1  $N=110$ ,  $N=90$ ; row2  $N=70$ ,  $N=40$ ; row 3  $N=20$ ,  $N=10$ . X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

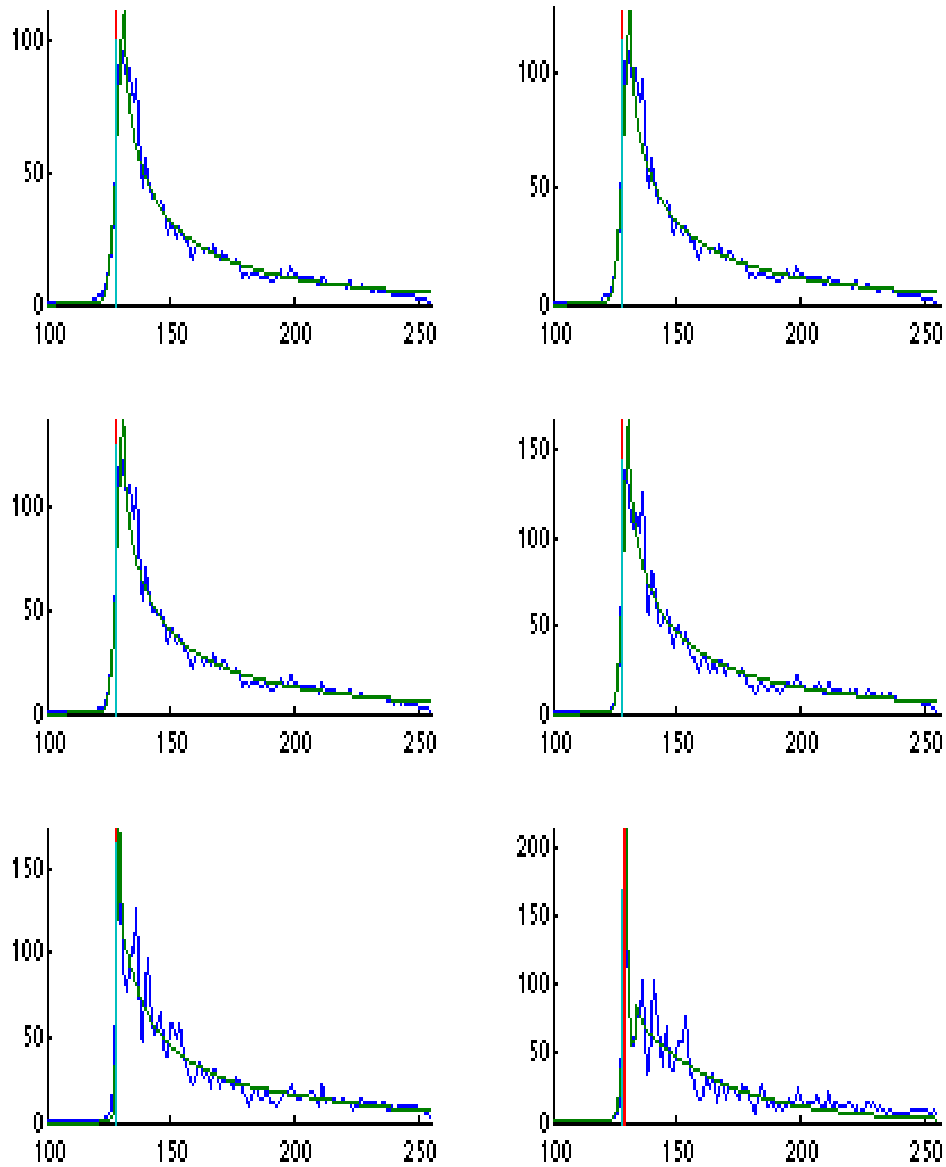




Figure 29: Waveforms ( $i=121$ ,  $t=18647.26$  s,  $lat=12.9486^\circ$ ,  $lon=104.0674^\circ$ ) with empirical retracker (green curve) with retracted bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1  $N=110$ ,  $N=90$ ; row 2  $N=70$ ,  $N=40$ ; row 3  $N=20$ ,  $N=10$ . X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

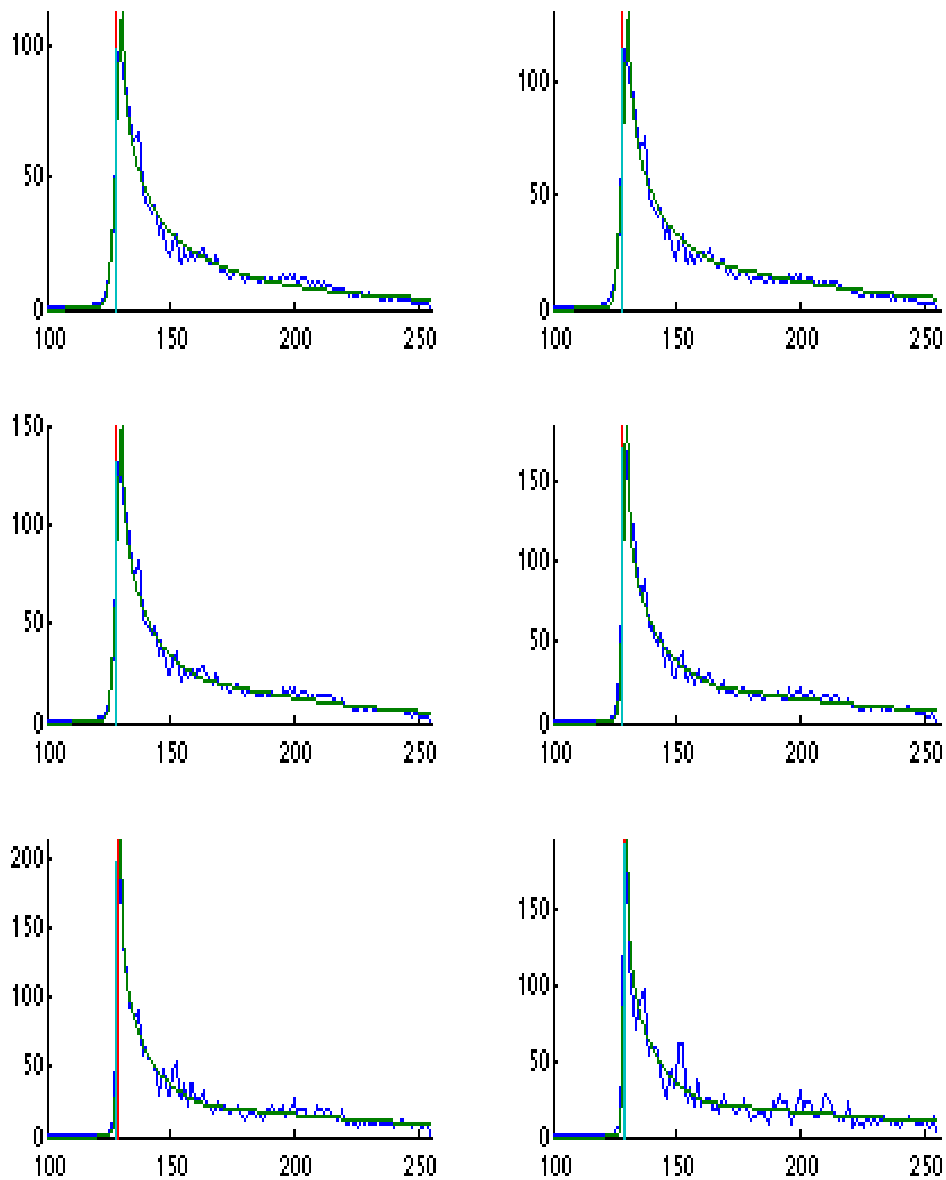




Figure 30: Waveforms ( $i=122$ ,  $t=18647.30$  s,  $lat=12.9459^\circ$ ,  $lon=104.0671^\circ$ ) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin.





		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

From left to right: row 1 N=110, N=90; row2 N=70, N=40; row 3 N=20, N=10. X-axis is gate number (1-256 in SAR mode), Y-axis is power amplitude.



**Figure 31: Geolocation of points i=86 and i=118 corresponding to waveforms of Figure 25 - Figure 30. Other points are omitted for clarity as location of i=87, and i=119-122 are close to points i=86 and i=118 respectively.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 5 Mekong River Masks

Existing river masks for complex braided rivers such as the Mekong derived as part of the ESA/ESRIN R&L contract were reassessed for accuracy and compared against Google earth imagery. The conclusion, as well as from previous experience with R&L Mekong locations, highlighted deficiencies with some so-called “river” locations being outside the expected river banks. Within CRUCIAL we constructed masks from Landsat optical data although SAR images could be used. Landsat was preferred as the discrimination between land and water is more distinct.



### 5.1 Mekong: Landsat river mask

The algorithm to generate Landsat river masks is general but is illustrated by reference to the Mekong River. The algorithm follows the following steps:

1. Download Landsat imagery

For example, six Landsat 8 images were obtained covering the area upstream of the tidal limit for which SAR data is available for the Mekong river. This is between latitudes 12°N and 18°N. The images are nearly cloud free and were obtained on the following dates: 12/11/2014, 19/11/2014 and 21/11/2014. November typically has mid-season water levels. The levels are decreasing but do not reach their minimum until February or March (Figure 32 and Figure 33).

Table 4 shows the Landsat 8 bands that are available ([http://landsat.usgs.gov/best\\_spectral\\_bands\\_to\\_use.php](http://landsat.usgs.gov/best_spectral_bands_to_use.php)). Most bands have a resolution of 30m but band 8 has a resolution of 15m and bands 10 and 11 a resolution of 100m.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

**Table 4: Landsat 8; Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) bands**

Band	Wavelength	Useful for mapping
Band 1 – coastal aerosol	0.43 - 0.45	coastal and aerosol studies
Band 2 – blue	0.45 - 0.51	Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation
Band 3 - green	0.53 - 0.59	Emphasizes peak vegetation, which is useful for assessing plant vigor
Band 4 - red	0.64 - 0.67	Discriminates vegetation slopes
Band 5 - Near Infrared (NIR)	0.88 - 0.85	Emphasizes biomass content and shorelines
Band 6 - Short-wave Infrared (SWIR) 1	1.57 - 1.65	Discriminates moisture content of soil and vegetation; penetrates thin clouds
Band 7 - Short-wave Infrared (SWIR) 2	2.11 - 2.29	Improved moisture content of soil and vegetation and thin cloud penetration
Band 8 - Panchromatic	0.50 - 0.68	15 meter resolution, sharper image definition
Band 9 – Cirrus	1.36 - 1.38	Improved detection of cirrus cloud contamination
Band 10 – TIRS 1	10.60 – 11.19	100 meter resolution, thermal mapping and estimated soil moisture
Band 11 – TIRS 2	11.5 - 12.51	100 meter resolution, Improved thermal mapping and estimated soil moisture

## 2. Merge Landsat images into a single mosaic

The individual Landsat images were merged into a single image with an approximate mask manually drawn extending the river width by approximately 1km either each of the river (Figure 34 and Figure 35)

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

### 3. Use of Normalised Difference Water Index to derive high resolution mask

To obtain a high resolution mask of the surface water the standard technique is to use the Normalized Difference Water Index (NDWI) (McFeeters, 1996). This uses the reflectance of the green band 3 and near-infrared band 5 of Landsat as  $NDWI = (Green - NIR) / (Green + NIR)$ .



A variety of other techniques have been developed more recently (e.g. Feyisa et al. 2014) to improve the classification in built up areas and where there are clouds. In this case (as there is very little built up area and clouds) a simple technique was all that was necessary. Water was defined if the Digital Number of the red band was less than 60 (the range is 0-255). This picked up all surface water (including any lakes such as Tonle Sap) as shown in Figure 36 and Figure 37.

### 4. Combined high resolution and approximate mask

A final mask of the Mekong River (Figure 38 and Figure 39) was obtained by combining the approximate mask (Figure 34 and Figure 35) with the water mask (Figure 36 and Figure 37). This was considered desirable as the high resolution mask identified the large spatial extent of surface water in the region apart from the Mekong.



**Figure 32: Six Landsat images of the Mekong River between 12°N and 18°N**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

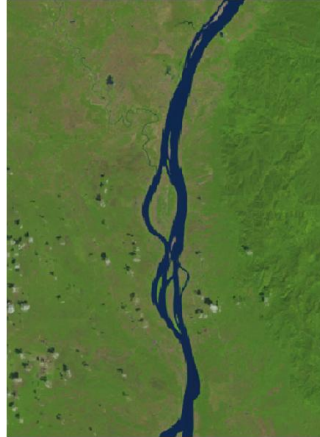




Figure 33: Landsat image from 12/11/2014 of the Mekong River between 14.2°N and 14.8°N



Figure 34: Approximate mask of the Mekong River between 12°N and 18°N

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
Version No	3	Date	04.10.16



**Figure 35: Approximate mask of the Mekong River between 14.2°N and 14.8°N**



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



Figure 36: Water mask of the Mekong River between 12°N and 18°N.





		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





Figure 37: Water mask of the Mekong River between 14.2°N and 14.8°N



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 38: Final mask of the Mekong River between 12°N and 18°N.**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





**Figure 39: Final mask of the Mekong River between 14.2°N and 14.8°N.**

## 5.2 Landsat river mask summary

This section has detailed the approach to generate a river mask for the Mekong. Pertinent points include

- Landsat imagery is straightforward to use but may not be applicable in all locations due to cloud cover. For such areas SAR imagery can be used.
- The mask for the Mekong is not time dependent but can be temporal if imagery is used in near real-time.
- River masks over areas with large inland water distributions need to be cleared to remove spurious off-river points.
- DTU has adopted a similar approach (see section 6).



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

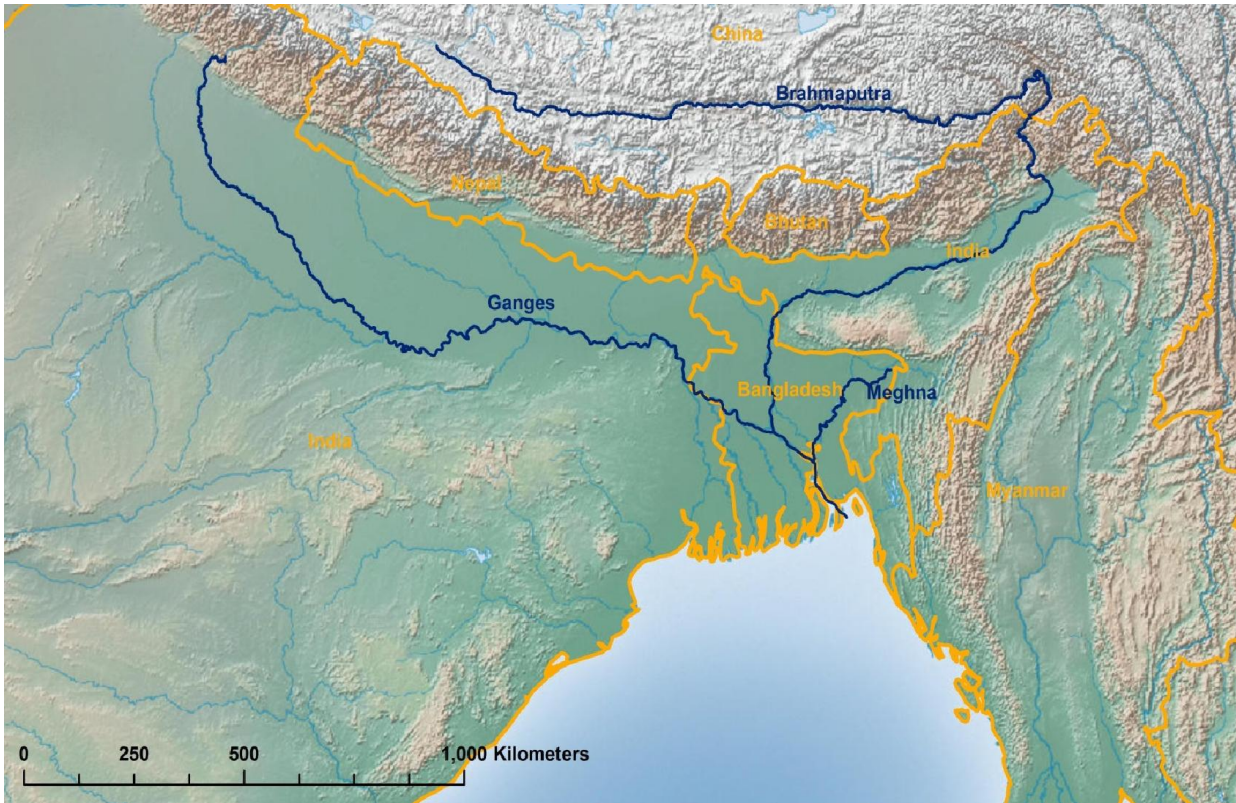
## 6 Assimilation of CryoSat-2 data to 1-dimensional hydrodynamic models

### 6.1 Introduction and background

A number of recent studies combine satellite altimetry with hydrologic river models using data from repeat orbit satellites such as Envisat, ERS or Jason. One popular river is the Amazon due to its large size and favourable direction of flow in relation to most satellites' orbits, for example Yamazaki et al. (2012). Other examples include other big rivers, such as the Mekong and Ob in the work of Birkinshaw et al. (2014) where daily discharge data was estimated from Envisat and ERS-2 altimetry. The discharge [ $\text{m}^3/\text{s}$ ] represents the volume rate of water flow transported through a given cross-sectional area. A combination of MODIS data of river velocity and Envisat water levels was used by Tarpanelli et al. (2014) to estimate the discharge in the Po River. Using satellite altimetry data is particularly attractive over poorly gauged basins where in-situ data is lacking. Becker et al. (2014) have used Envisat altimetry data in the Congo basin. Moreover, application of data from the wide-swath drifting type orbit mission SWOT has been considered (for example Biancamaria et al. (2011a) or Yoon et al. (2012)), however only with synthetically generated data: the SWOT mission is expected to be launched in 2020 (NASA, 2016).

For the Brahmaputra basin in South Asia, previous work has shown the value of Envisat altimetry data (Michailovsky et al., 2013). The Brahmaputra river drains a basin of over 580.000  $\text{km}^2$  and flows along 2880 km from its source north of the Himalayas to the Bay of Bengal. (Michailovsky et al., 2013). The basin and its main river, the Brahmaputra, are being monitored closely by India and China, however almost none of this in-situ hydrologic data is publically available. The Brahmaputra catchment for example is considered a classified basin by the Indian government (Central Water Commission, 2009). Water level and river discharge data are not publicly available for the Indian portion of the basin, which highlights the importance of remote sensing data for independent hydrologic analysis and modelling in the basin, for instance for flood forecasting in downstream Bangladesh.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 40: Map of the three main rivers draining into the Gulf of Bengal through Bangladesh**

Bangladesh, a low lying country at the Gulf of Bengal in the estuary region of the three large rivers Ganges, Brahmaputra and Meghna (Figure 40) is often hit by devastating floods. More than 90% of its surface waters are imported from outside of the country, i.e. mainly India, but still little data is shared between Bangladesh and India to aid water resource management or flood forecasting (Biancamaria et al., 2011b).

A hydrologic-hydrodynamic model for the entire Brahmaputra basin was set up. The course of the Brahmaputra River can be roughly divided into two parts: The upstream part in the Tibetan Plateau draining through the Himalaya into India, and the downstream part draining through the approximately 800 km long Assam valley into Bangladesh where it merges into the Ganges-Brahmaputra-Meghna delta region and finally flows into the Gulf of Bengal. In the downstream part, the Brahmaputra is a wide braided river. In the upstream part however the Brahmaputra's width usually is below 500 metres, and is often surrounded by steep slopes (Jain et al., 2007). This makes it hard to extract satellite altimetry data, and therefore the focus of this study was on the Assam valley part of the river. However it is noteworthy that with the methodology presented in

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

the following sections it still was possible to extract a relevant amount of CryoSat-2 altimetry in this stretch of the river (see section 6.6.1). In the Assam valley, the river is flowing in a wide, flat valley. The total width of the water body during the high flow season is typically around 10 km, exceeding 2 km in most places, making extraction of satellite altimetry data easier.

The CryoSat-2 data used in this work have already been shown by Villadsen et al. (2015) to be useful to monitor water levels in the Ganges and Brahmaputra. There also exists a number of studies using CryoSat-2 altimetry to extract water levels in lakes, for example Kleinherenbrink et al. (2014), Nielsen et al. (2015), or Song et al. (2015). Even though the potential of the special drifting orbit of CryoSat-2 over rivers has been considered (Bercher et al., 2013), this work is one of the first actual applications of CryoSat-2 altimetry over rivers.



## 6.2 Data and methods

### 6.2.1 CryoSat-2 data

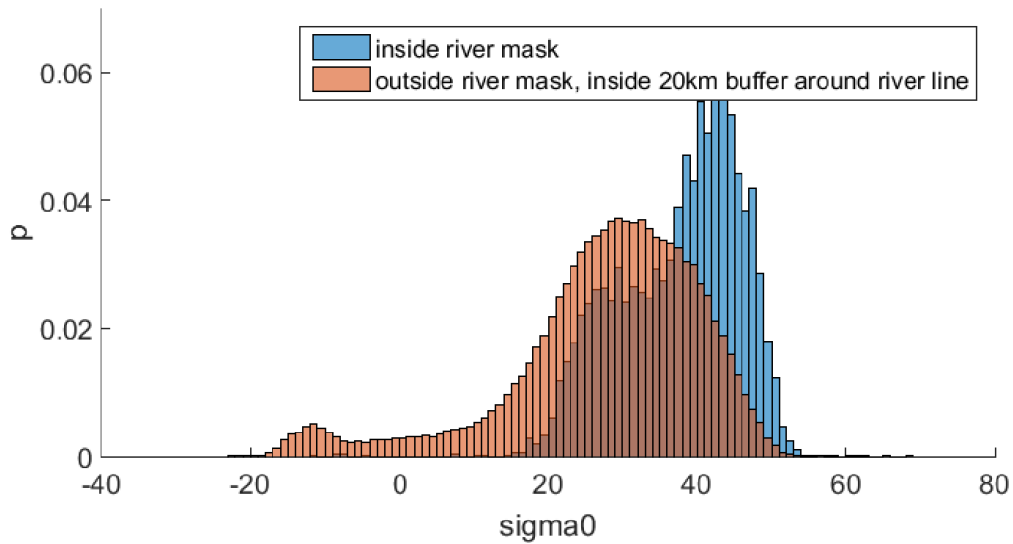
We used CryoSat-2 level 2 altimetry data that was processed and provided by DTU Space in the framework of the LOTUS project (<http://www.fp7-lotus.eu/>). The basis for the data is the ESA baseline-b L1b 20 Hz product. This was retracked by Villadsen et al. (2015) using a primary peak threshold retracker. Most of the study area is covered in the Synthetic Aperture Radar Interferometer (SARin) mode of CryoSat-2. SAR mode of CryoSat-2 offers an along track resolution of 300m, whereas the SARin mode is adopted to deal with areas presenting a high surface slope by additionally using a second across-track antenna. This allows to determine off-nadir locations of the altimetric observations (Armitage and Davidson, 2014). The data is available since July 2010. For this report, data until the end of 2013 have been used.

### 6.2.2 Preprocessing of the data



The CryoSat-2 data itself does not currently deliver reliable information on whether it was acquired over land surface or over (small inland) water surfaces. Often, the backscatter values ( $\sigma_0$ ) can give an indication of the surface it was acquired over. Over the Brahmaputra, however, the backscatter could not be used, which might be due to the fact that the river is relatively narrow and its waters often turbulent or turbid and consequently cannot be distinguished from land surface. See Figure 41 for an illustration of this: the backscatter values of CryoSat-2 datapoints over a Landsat river mask (details are explained in the following section) cannot be clearly distinguished from the backscatter values of points outside the river mask. Especially backscatter values around 30 can both be found in relevant amounts over water and

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

over land surface. Hence, the distinction between CryoSat-2 data points representing river water surface and land surface has to be based on independent data – in this case a water mask from satellite imagery. Furthermore, the Brahmaputra in the Assam valley is a very dynamic braided river and experiences significant changes to its course from one year to another (see Figure 42), which requires dynamic river masks.



**Figure 41: Probability distributions of backscatter values of CryoSat-2 datapoints over the Landsat river mask (blue) compared to datapoints outside the river mask, but inside a 20km wide buffer around the river line (blue)**

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





**Figure 42: Landsat 7 images of the same part of the Brahmaputra river in the Assam valley showing the dynamic changes in river morphology. Left: 2010. Right: 2011.**

### 6.2.3 Brahmaputra: Landsat river mask

For mapping the dynamic braided river system, high resolution data with at least a seasonal time step is needed. Because no SAR data covering the whole time frame is freely available, it was decided to use optical imagery from the Landsat program. Optical imagery from the Landsat programme has been used repeatedly to create high resolution river masks, see for example Michailovsky et al. (2012) using it over the Zambezi River with Envisat altimetry data or Birkinshaw et al. (2014) using it over the Mekong and Ob rivers with Envisat and ERS-2 data.

Water masking was performed based on Normalized Difference Vegetation Index (NDVI) images acquired by the Landsat multi-spectral imager as in section 5. Here, NDVI is an index computed from the spectral reflectances in the red (Band 4 of Table 4) and near-infrared (Band 5 of Table 4) spectral regions as  $NDVI = (NIR-RED)/(NIR+RED)$  where NIR is the near infrared reflectance and RED is the reflectance in the red range.

Landsat 7 and 8 NDVI imagery is available every 8 days, however due to issues such as cloud cover and sensor failures only 32-day composites give a reasonable result (Google, 2015). All areas with an NDVI below 0 were considered water, everything else was considered land. Because of the above mentioned morphological dynamics, an individual mask had to be created for each year



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

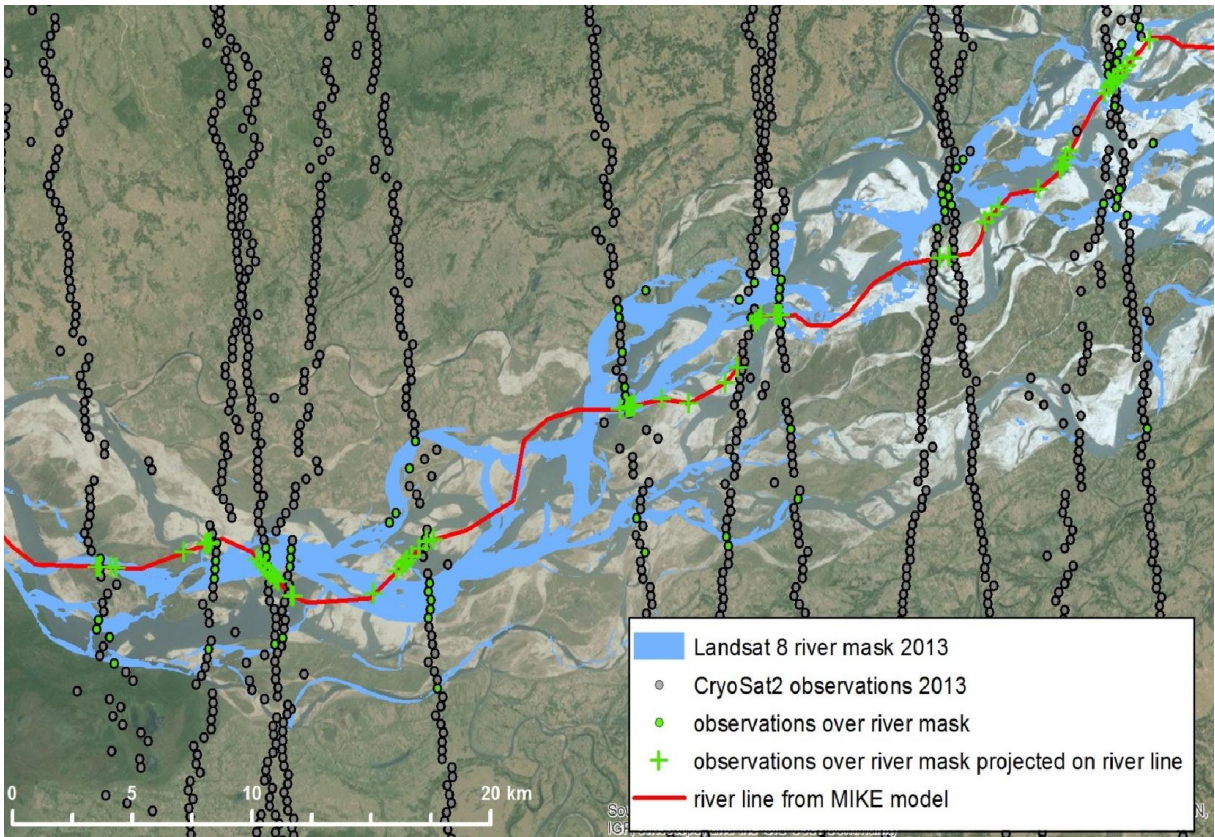
from 2010 to 2013. However, even using the 32-day composites no sufficient coverage could be achieved during the high-flow monsoon season in summer due to cloud cover. Thus, the river masks created by merging all available images from one year represent a minimum- or low-flow mask. The resulting river masks have a width of approximately 1 km to 20 km for the entire multi-channel river bed, with single channels reaching up to 2 km in width.

#### 6.2.4 Filtering and projecting the CryoSat-2 data

Figure 43 shows the filtering and projecting processes applied to the CryoSat-2 level 2 data points. Only data points above the Landsat minimum river mask for the respective year are used, i.e. considered to represent the river water surface. After filtering those points they have to be projected onto the river line of the 1D hydrodynamic model to determine the correspondence between CryoSat-2 observation location and model state space. The model state consists of water level (measured in metres above mean sea level, MAMSL) at each  $h$  discretization point and discharge ( $m^3/s$ ) at each  $Q$  discretization point of the model.





		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



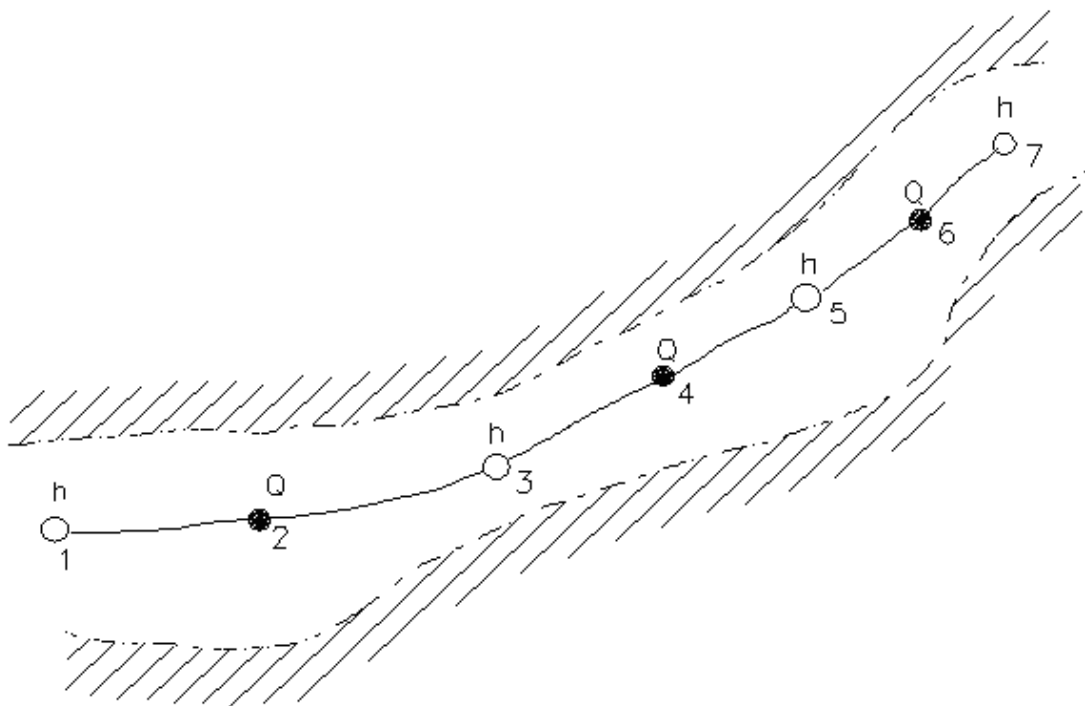
**Figure 43:** Section of the Brahmaputra in the Assam valley showing the Landsat river mask, the CryoSat-2 observations and their mapping to the 1D river model, all for 2013.

### 6.3 Hydrologic hydrodynamic model

In cooperation with DHI (Danish Hydraulic Institute, [www.dhi.dk](http://www.dhi.dk)), a model of the entire Brahmaputra basin was set up in the DHI MIKE software. The model consists of a hydrologic and a hydrodynamic part: 33 lumped NAM rainfall-runoff subcatchments (Nielsen and Hansen, 1973) and a river network draining all the catchments. The NAM subcatchments included the modelling of snow melt processes where applicable. The river flow in MIKE 1D is modelled using a 1D dynamic wave routing based on the Saint Venant equations for unsteady flow (MIKE by DHI, 2009). The one-dimensional model space is discretized into finite intervals, which are separated by alternating Q and h points (Figure 44). In this specific model, the default distance between these calculation points was set to 5 km. A smaller spacing is applied where needed, for example at connections between river branches and at oddly spaced cross sections (as cross sections are



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

always placed on an h-point). A discrete version of the continuity equation is formulated for each h-point and a discrete version of the momentum equation is formulated for each Q-point. The resulting equation system is solved using an implicit Abbott-Ionesco 6-point finite difference scheme. Water level in meters above mean sea level is thus simulated at every h-point of the model domain. The measurement operator interpolates simulated water levels at the h-points to the locations of the individual CryoSat-2 height readings along the model river line.





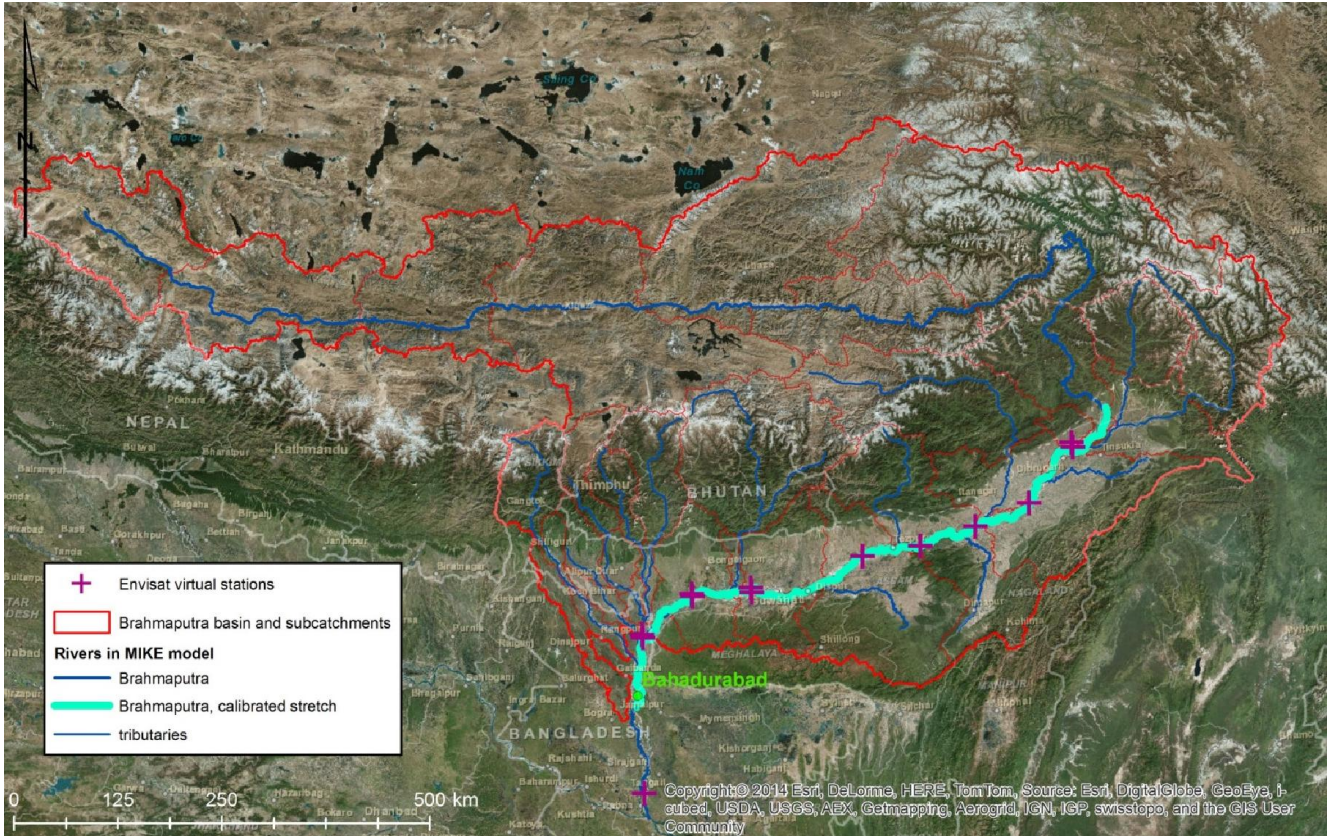
**Figure 44: The MIKE 1D computational grid. From MIKE by DHI (2009).**

Figure 45 provides an overview of the Brahmaputra basin model setup. The lumped subcatchments of the Brahmaputra for which the NAM rainfall-runoff models have been set up are shown as the red-bordered shapes in Figure 45. Subcatchment discretization was chosen taking into account variations in land surface elevation, climate, land cover and soil types, and maintaining a manageable total number of subcatchments. The Brahmaputra model is part of a larger model of the Ganges Brahmaputra region (Figure 40). For some of the subcatchments of this larger model in-situ discharge observations were available. Calibration catchments were mostly located in the Nepalese headwaters of the Ganges due to an open data policy in Nepal. Only one calibration subcatchment is located in the Brahmaputra basin itself. Furthermore,

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



discharge observations from the station Bahadurabad on the Brahmaputra (see Figure 45), close to its confluence with the Ganges River, were available. Besides these few in-situ observations the entire model was based on remote sensing data: For the precipitation forcing, Tropical Rainfall Measurement Mission Project (TRMM) v7 3B42 data was used (TRMM, 2011). Temperature and evapotranspiration data were derived from the ERA-interim reanalysis products from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al., 2011). Note that this changed since D4050 was submitted, hence there are some differences in the results, too. The SRTM DEM (Jarvis et al., 2008) was used to delineate the subcatchments and the drainage network. The calibration period included the years 2002 to 2007. As it was impossible to generate enough runoff from the rainfall-runoff models especially in the Himalaya it was assumed that there is a bias in the TRMM precipitation data: For all Himalaya subcatchments the precipitation forcing was scaled with a factor of 1.4. Such a bias in the TRMM precipitation product has been experienced before, for example by Michailovsky et al. (2013). The runoff from the NAM subcatchments with available in situ discharge observations was calibrated individually, and the resulting parameters transferred to the remaining catchments.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 45: The Brahmaputra basin model. The main calibration station, Bahadurabad in Bangladesh is shown. The water level calibrated part, i.e. the Assam valley together with the Envisat virtual stations is displayed (see next section).**

The hydrodynamic model was subsequently calibrated to the discharge at Bahadurabad by adjusting the Manning's number. River bed friction in the hydrodynamic model is parameterized with the empirical Manning equation:  $S_f = (n^2 Q^2) / (A^2 R^{4/3})$  where  $S_f$  is the friction slope (-),  $n$  is the Manning number ( $\text{sm}^{-1/3}$ ),  $Q$  is the river discharge ( $\text{m}^3/\text{s}$ ),  $A$  is the flow cross-sectional area ( $\text{m}^2$ ) and  $R$  is the hydraulic radius (m). The Manning number is a key parameter affecting the balance between friction and gravity and thus the kinematic wave velocity.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

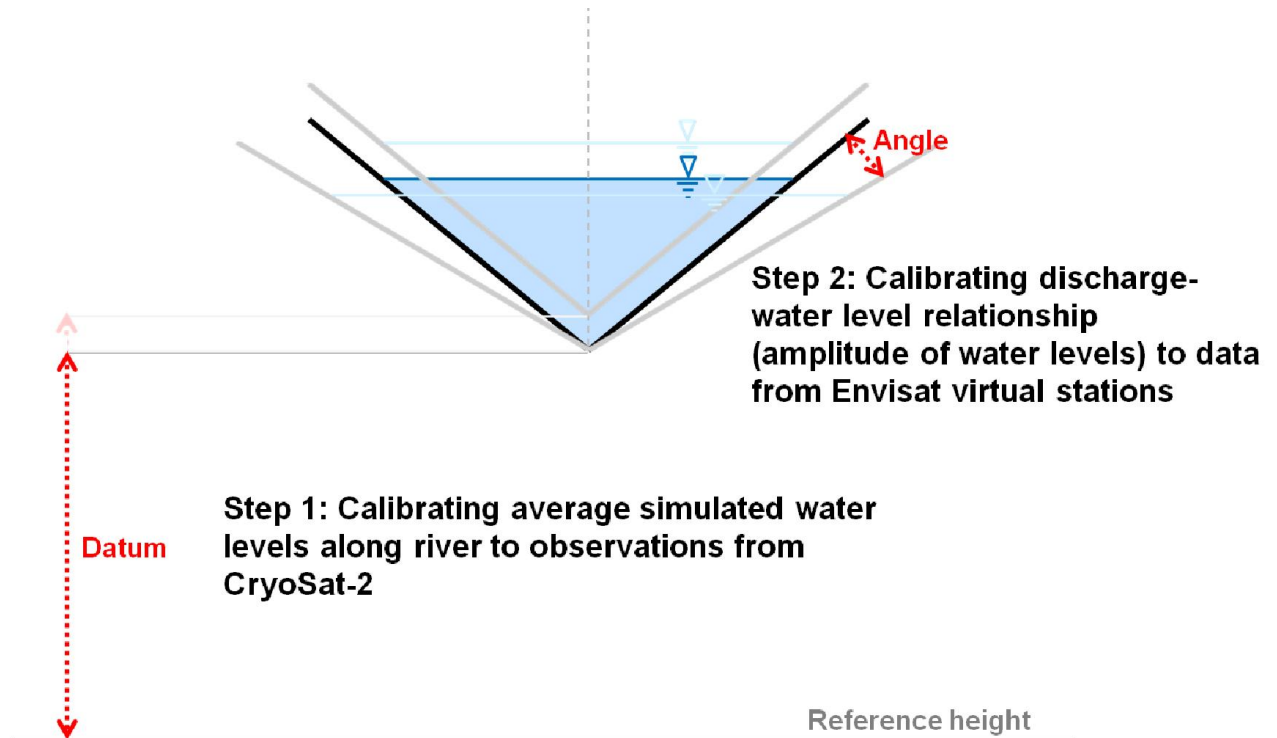
## 6.4 Cross section calibration

The calibration procedure requires cross section data to be inserted into the Manning Equation (see the 'A' term). As no in-situ cross section data was available, the SRTM DEM was used to derive the rivers' course and a first guess of the cross section datums and widths. However, this can only be seen as a first guess, as the SRTM DEM has a horizontal resolution of 90 meters, and the vertical standard error is in the range of a few meters (Rodríguez et al., 2006). Furthermore the SRTM data cannot represent river bathymetry. An approach based on altimetry data was chosen to ensure the model accurately reproduces water levels along the river: Cross sections with a simple triangular shape were placed every 50 km along the entire Brahmaputra River and then calibrated using the above described CryoSat-2 altimetry data along the river in combination with Envisat virtual station data.

The cross section calibration was performed after the discharge calibration presented in the previous section. The calibration was based on a combination of data from the Envisat mission and the CryoSat-2 data: The Envisat mission with its 35-day repeat orbit (lowered to 30 days in the last part of the mission) provides virtual station time series. These show water level time series with a 35-day timestep at distinct points in the Brahmaputra River. 13 Envisat virtual stations along the Brahmaputra in the Assam valley (see Figure 45) covering the years 2002 to 2010 were used. Those stations have an average distance along the river of 75 km between each other. The data originates from the ESA River&Lake project (Berry, 2009). The Envisat observations will only be used for calibration in this study and will not be assimilated. However, the assimilation scheme presented here is perfectly able to also ingest data from repeat orbit radar altimetry mission.



CryoSat-2 has a repeat orbit of 369 days with subcycles of 30 days. Subcycles are related to the off nadir pointing capability of the antenna. CryoSat-2 observations cannot be used directly to extract water level time series. However, due to the drifting orbit, when several years are taken into account, they show the average longitudinal water level profile along the entire river. Both datasets combined provide the necessary information to fully calibrate the model's water levels, in a two-step process illustrated in Figure 46:

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



**Figure 46: Sketch of the two-step cross section calibration.**

Both calibration steps were performed by coupling the MIKE 1D model to MATLAB and using MATLAB's genetic algorithms to find the optimal parameter combination. Table 5 displays the decision variables and objective functions, based on the Root Mean Square Error (RMSE), for the two calibration steps. In Step 1, the parameters to be optimized (decision variables) are the cross section datums for 24 model cross sections in the Assam Valley, i.e. 24 parameters were fitted in total in this calibration run. In Step 2, the parameters to be optimized are the opening angles of 27 triangular cross sections, i.e. 27 parameters were fitted in total in this calibration run. The angles of cross sections without neighbouring Envisat virtual station only show little impact on the calibration's objective, a fit to the observed water level amplitudes at the virtual stations. Hence such cross section angles were linearly interpolated between their neighbouring cross sections, effectively reducing the number of decision variables. The RMSE between CryoSat-2 water levels and average simulated water levels was weighted according to the assumed standard deviation of the respective CryoSat-2 measurements. See section 6.5.4 for a description of the procedure.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

**Table 5: Decision variables, constraints and objective functions of the genetic algorithm used for the two-step water level calibration**

	Decision variables	Constraints	Objective function
<b>Step 1: Calibrating absolute average water levels to CryoSat-2 observations</b>	24 cross section datums	Cross section datums continuously decreasing from upstream to downstream	Standard error-weighted RMSE between CryoSat-2 observations and average simulated water levels from 2010 to 2013
<b>Step 2: Calibrating amplitude of water levels to Envisat VS data</b>	27 cross section angle	Cross section datums without neighbouring virtual stations linearly interpolated from their neighbours	RMSE between yearly amplitudes of Envisat VS data and of simulated water levels from 2002 to 2010

## 6.5 Data Assimilation

### 6.5.1 Data Assimilation in general

In general, a (hydrologic) model at time  $t$  can be written as (Bauer-Gottwein et al., 2014)



$$x_t = M(x_{t-1}, p, f_t) + v_t \quad (24)$$

where

- $x$  model state vector
- $p$  vector of model parameters
- $f$  vector of model forcings
- $v$  model noise
- $M$  model operator

which means that a state vector, displaying the model's states of interest is propagated from timestep to timestep by the parameterized model (model operator), using the prior model's state and the forcings as an input. Furthermore, a model error is considered.

Now we consider some observations at time  $t$

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

$$y_t = H(x_t) + u_t \quad (24)$$

where  $y$  vector of observations  
 $H$  observation model or measurement operator  
 $u$  observation noise

The observations have to be mapped into model space by the observation model, and also have a related uncertainty. Trying to incorporate the information from the observations into the model, i.e. updating the model is referred to as data assimilation (DA). The actual update can be written as

$$x_t^a = x_t^f + K_t(y_t - H(x_t^f)) \quad (24)$$

where  $x^a$  updated forecast of the model states (analysis)  
 $x^f$  model forecast  
 $K$  gain operator

This means the model is propagated from timestep to timestep until an observation is available. Whenever observations are available, an analysis is performed, i.e. the model's forecast is updated using the information from the observation. The gain operator basically weighs between the model's forecast and the observed value, depending on their individual uncertainties. The forecast is an optimal estimate of the model's state with both information sources combined.

### 6.5.2 Ensemble Transform Kalman Filter

Commonly, some implementation of the Kalman Filter is used in Data Assimilation to perform the analysis as presented in equation (24) and the propagation of the model's uncertainties. The Kalman Gain  $K$  is determined based on the state error covariance matrix



$$K = P^f H^T (H P^f H^T + R)^{-1} \quad (24)$$

where  $R$  is the observation error covariance and superscript T denotes matrix transposition. The state error covariance matrix  $P$  is propagated by

$$P^a = (I - KH)P^f \quad (24)$$

where  $I$  is the identity matrix.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

As seen above, the classical Kalman Filter requires the calculation of the covariance matrices between the model states. This excludes its use for high order or non-linear models and requires a linear measurement operator, i.e. a linear relationship between the model states and the simulated observations. The hydrodynamic model used here is non-linear. Consequently for our case an ensemble based method, the Ensemble Transform Kalman Filter (ETKF) has been used. For ensemble based methods the covariance is computed from an ensemble of model realizations accounting for the related model uncertainty. Those ensemble based methods are often also based on the Kalman Filter theory.

The Ensemble Kalman Filter (EnKF) is one of the most commonly used ensemble based filters. The covariance matrix  $P$  is usually propagated by the ensemble by

$$P = \frac{1}{m-1} \sum_{i=1}^m (X_i - x)(X_i - x)^T = \frac{1}{m-1} AA^T \quad (24)$$

where

- $X$  ensemble of model states,  $X = [X_1, \dots, X_m]$
- $m$  ensemble size
- $x$  ensemble mean,  $x = \frac{1}{m} \sum_{i=1}^m X_i$
- $A$  ensemble of anomalies,  $A = [A_1, \dots, A_m]$ , and  $A_i = X_i - x$

The update then is

$$X_i^a = X_i^f + K(d - HX_i^f) \quad (24)$$

If the system state is represented by the ensemble mean, the average of (24) is the analysis equation (24). This does however not apply to the covariance equation (24), and can lead to an ensemble collapse, i.e. too little ensemble spread. To circumvent this, the EnKF often is used with perturbed observations.



Another possibility to circumvent this problem is the ETKF. In contrast to the EnKF, in the ETKF the ensemble anomalies are explicitly updated via a transformation matrix  $T$

$$A^a = A^f T \quad (24)$$

where  $T$  must satisfy

$$T = T^S U \quad (24)$$

and  $U$  is an arbitrary orthonormal matrix, and

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



$$T^S = [I + \frac{1}{m-1} (HA^f)^T R^{-1} HA^f]^{-1/2} \quad (24)$$

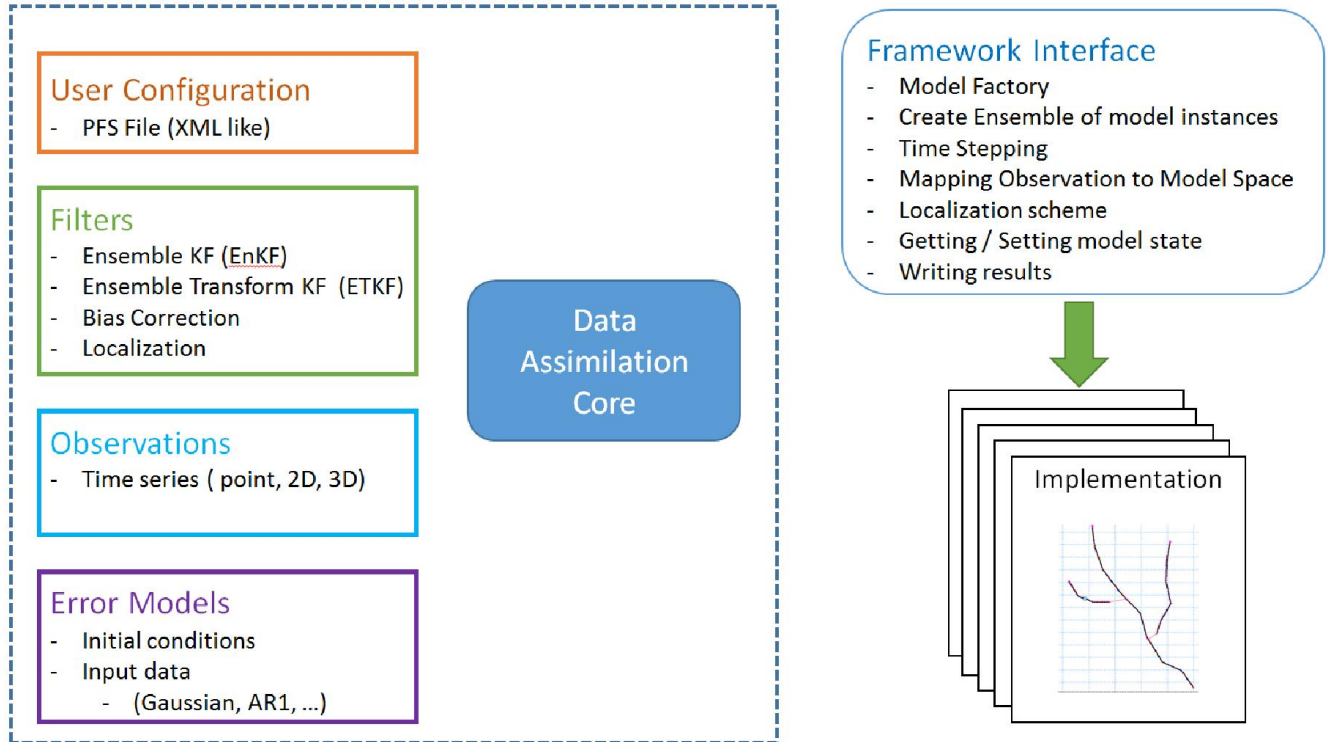
where  $R$  observation error covariance matrix  
 $I$  identity matrix

For details of the ETKF please refer to Sakov and Oke (2008).

### 6.5.3 Setup with DHI Data Assimilation Framework

For the data assimilation, the DHI Data Assimilation Framework was used. The DHI Data Assimilation Framework provides various filters, error models and observation mapping methods to assimilate various datasets to MIKE models. It is written in .NET/C# and communicates directly with the MIKE models. An overview over the framework can be seen in Figure 47. Configuration details are provided through an ASCII configuration file (.altpfs file). The framework provides different filtering algorithms such as EnKF and ETKF, or also a defined gain method. It also allows for state augmentation with a time-constant bias term, which is then estimated by the filter. Localization approaches are also provided, i.e. the state updating is restricted to model states in the neighbourhood of the measurement location. This is useful to avoid artefacts created by spurious correlations due to limited ensemble size. The framework also provides templates for the parameterization of model errors, including perturbation of the initial conditions and autocorrelated error representations for the model forcings. Autocorrelated model error is integrated in the model state using state augmentation.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



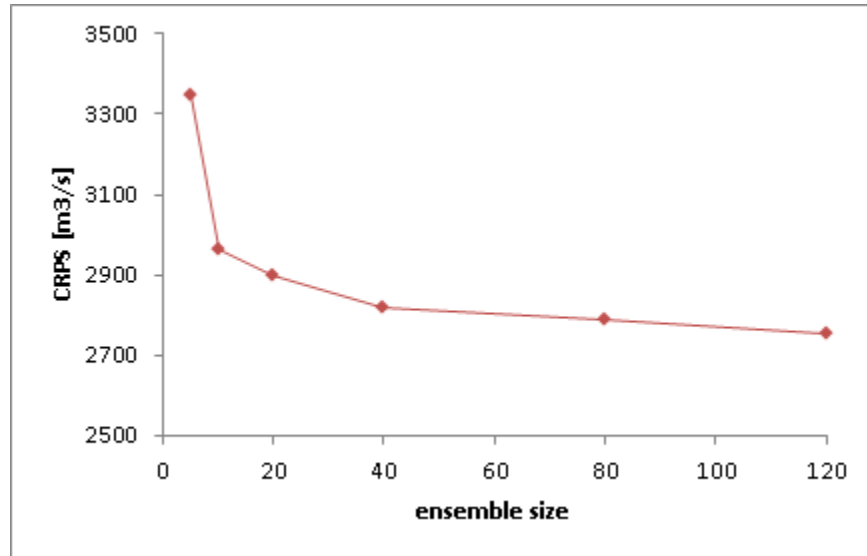
**Figure 47: Overview over the DHI Data Assimilation Framework. Source: Marc-Etienne Ridler, DHI.**

For our case, a version that ingests altimetry data as for example delivered by CryoSat-2 was used. The challenge with assimilating CryoSat-2 data is its spatial and temporal distribution. Commonly in hydrologic modelling assimilated data consists of time series of some value at certain points in the model space. CryoSat-2 however provides observations distributed over the entire model space at various points in time. Hence, the above described ETKF was chosen, and the modelled water levels were calibrated across the entire model as described in section 6.4.

To fully represent the model uncertainty across space, the ensemble size of Ensemble filters should be chosen large enough. However, computational limitations will restrict ensemble sizes, as each ensemble member represents an individual realisation of the model. Often an ensemble size below 100 is considered sufficient. For the Brahmaputra model, synthetic experiments were conducted to get an idea of the ideal ensemble size: the same synthetic setup (described in section 6.5.9) of the Brahmaputra model was run repeatedly with different ensemble sizes. Then their errors in terms of continuous ranked probability score (CRPS) (more details on the CRPS in section 6.6.2) were evaluated. The average CRPS for 3 evaluations of the model for each 5, 10, 20, 40, 80, and 120 ensemble members is reported in Figure 48. For the experiments described in the following an ensemble size of 40 was chosen. It can be seen that this is not the optimal ensemble size yet, as the error still decreases towards larger ensemble sizes, however this is considered a

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



reasonable trade-off between computational demand and model uncertainty representation for the ETKF.

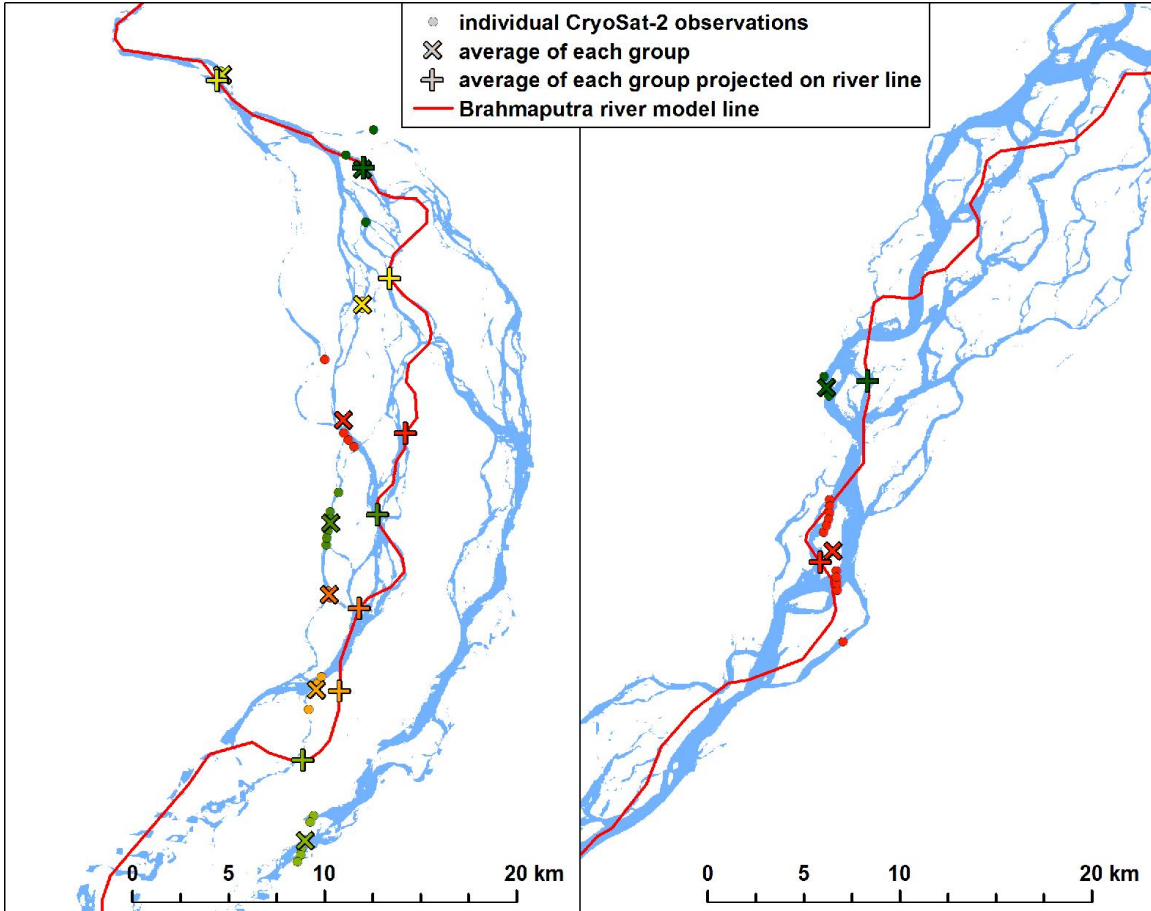


**Figure 48: Development of CPRS (i.e. model error) with ensemble size for a synthetic data assimilation setup of the model as described in section 6.5.9. The chosen ensemble size of 40 is not optimal yet, but considered a reasonable trade-off.**

#### 6.5.4 Processing of CryoSat-2 observations

After filtering over the river mask as described previously (see Figure 43), a total of 2714 CryoSat-2 observation points over the Assam valley from 2010 to 2013 remained. Those individual observations originate from 471 river crossings of CryoSat-2. If the satellite ground track crosses the river more or less perpendicular, as mostly the case in Figure 43, all observations from one specific crossing can be aggregated into one single observation to be assimilated to the model. Often however, as for example displayed in Figure 49, the individual CryoSat-2 data points from one river crossing are spread out over considerable length along the river. Hence, a grouping method was applied: First, the total distance along the river spanned by the individual data points (dots in Figure 49) of the current transect is determined. Then, the number of cluster groups is determined as the total distance divided by some threshold distance. For the example in Figure 49 this threshold distance was set to 5 km, for the actual data assimilation also a threshold value of 3 km was considered. The individual data points then are grouped into the number of clusters determined above by k-means clustering of their location (different colours represent different clusters). Their location and observed elevation was averaged (tilted crosses in Figure 49), and then projected onto the model river line (crosses on the red river line in Figure 49).

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





**Figure 49: Grouping of individual CryoSat-2 data points from one river crossing based on k-means clustering. The left panel shows a crossing from 09-04-2011, and the right panel from 30-10-2010. Note that the panels are showing different parts of the river, with the 2010 and 2011 river mask respectively.**

The standard deviation of the individual observations' observed elevations can be used as a (at least relative) indicator of the related observations uncertainty, assuming that optimally all observations in one group observe the same river water level.

### 6.5.5 Observation error

As described above, the standard deviation of individual crossings can be considered as the related measurement uncertainty. Hence, this was assumed to be the standard error of the observations. For single measurements or groups consisting of less than three individual observations, the standard error was set to a defined default (maximum) value; in our case 1.5 metres. However, these observations are considered to still contain some information useful in the DA, hence we did not want to exclude them completely. The mean observation error from this method for all CryoSat-2 observations used is 0.45 metres, and the median is 0.22 metres. If

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

the cases are omitted where the standard error has to be set to a default value as described above, the mean standard error is 0.29 metres and the median is 0.18 metres.

### 6.5.6 Model error

Initially, the main source of model error was assumed to stem from the runoff generated in the NAM subcatchments. This is mainly because of the large uncertainty of the remote sensing precipitation product. Hence, the ensemble was generated by perturbing the catchments' runoffs. Due to the size of the model and the number of subcatchments, the perturbations of the single catchments have to be correlated in both space and time. Otherwise the perturbations of the single catchments cancel each other out when aggregated in the full model.

As in Michailovsky et al. (2013), the perturbation was assumed to be fully spatially correlated, i.e. it is the same across all subcatchments at a specific timestep. The temporal correlation was modelled by a first-order autoregressive model (AR1):

$$w_t = \delta w_{t-1} + \varepsilon_t \quad (25)$$



where

- $w_t$  runoff perturbation from catchment:  $w_t = \frac{Q_{sim,t} - Q_{obs,t}}{Q_{obs,t}}$
- $Q$  river discharge (simulated and observed)
- $\delta$  AR1 parameter
- $\varepsilon_t$  white Gaussian noise

It was assumed that the autocorrelation of the relative runoff error is the same as the autocorrelation of the relative model error at Bahadurabad. The AR1 parameter determined based on the residuals at Bahadurabad station was 0.9606, and was used for the perturbation in the data assimilation. The magnitude of the relative runoff error at Bahadurabad station however was found to be too small to result in a large enough ensemble spread, hence the Gaussian noise was assumed to have a standard deviation of 0.3.

### 6.5.7 Localization

Because of the previously mentioned spurious correlations of model states across the entire model space, updates often have to be limited in their effect to their closer surroundings. This was also the case for the Brahmaputra model, where in data assimilation experiments unphysical updates across the entire river could be seen in some cases. Hence, localisation was introduced limiting the effect of the updates to 200 km upstream and downstream of each observation in the approximately 1000 km long river model.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

### 6.5.8 Virtual window

With some gaps, CryoSat-2 measurements over the Brahmaputra river occur approximately once per day. The simulation time step of a hydrodynamic model however is usually much smaller, for the Brahmaputra basin model it is five minutes. Still, each CryoSat-2 measurement only coincides with one simulation time step, and only would be used to update the model in said time step. It is however reasonable to assume that a CryoSat-2 measurement delivers information about a longer time frame than just this specific simulation time step: the typical dynamics of discharge and water level changes in such a hydrodynamic model are slower than its simulation time step. This can for example be seen in Figure 50: here, the differences in simulated water levels over different timesteps, starting from 5 minutes, are shown. It can be seen that within a window of 120 minutes water levels in the model change by a maximum of 2.5 mm, hence the error made by assuming using a virtual window of 120 minutes is maximum 2.5 mm. For a 240 minute-window however, the maximum water level change, rises to 11 mm. Hence, each measurement was spread out over 120 minutes and consequently assimilated at all simulation time steps during this so called virtual window.

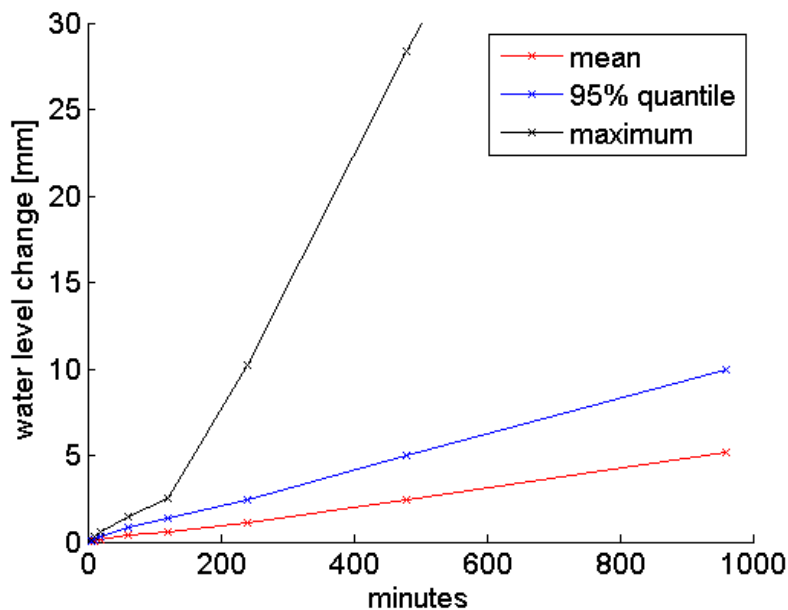




Figure 50: Differences in simulated water levels along the Brahmaputra River over different timesteps, i.e. the differences in water levels at each single model grid point within 5, 10, 20, 60, 120, 240, 480, and 960 minutes.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

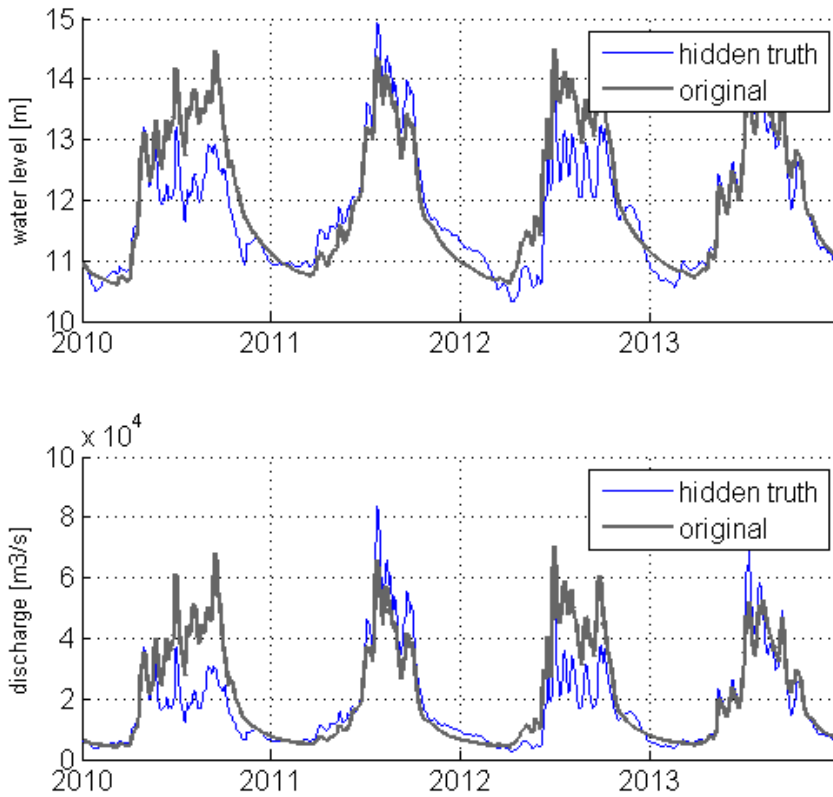
### 6.5.9 Data assimilation experiment with synthetic CryoSat-2 data

To test the DHI Data Assimilation Framework and its capabilities under realistic conditions, a synthetic data assimilation was performed. This was done with the actual Brahmaputra basin model and synthetically extracted CryoSat-2 measurements, in a controlled hidden truth experiment.

The synthetic observations were derived from a hidden truth run of the Brahmaputra basin model. The hidden model is a run of the model with one random realization of the forcing perturbation that also was used for the data assimilation. I.e. the model was run with the model error description as described in section 6.5.6. This hidden truth run then has significantly different discharge and water levels than the original model with unperturbed forcings, as can be seen in Figure 51, where discharge and water level of the hidden truth model are shown in comparison to the original model at Bahadurabad station. The synthetic observations were extracted from the simulated water levels in the hidden truth model. To be able to judge the value of CryoSat-2 observations, the actual sampling pattern of CryoSat-2 was kept, i.e. the synthetic observations were extracted at the exact same locations and times as real CryoSat-2 observations occur. To account for measurement error, the synthetic observations were perturbed with a Gaussian white noise error with a standard deviation of 0.2 metres. This means that the synthetic observations exhibit the same sampling pattern of the Brahmaputra water levels as the real CryoSat-2 observations, but are replaced with the simulated water levels from the hidden truth model (with added uncertainty). When assimilating those synthetic measurements from the hidden truth model to the original model, one expects the assimilation to move the model closer to the hidden truth model.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16





**Figure 51: Comparison of water level (upper panel) and discharge (lower panel) at Bahadurabad station for the original model and the hidden truth run, from which the synthetic CryoSat-2 observations were extracted from.**

## 6.6 Results and Discussion

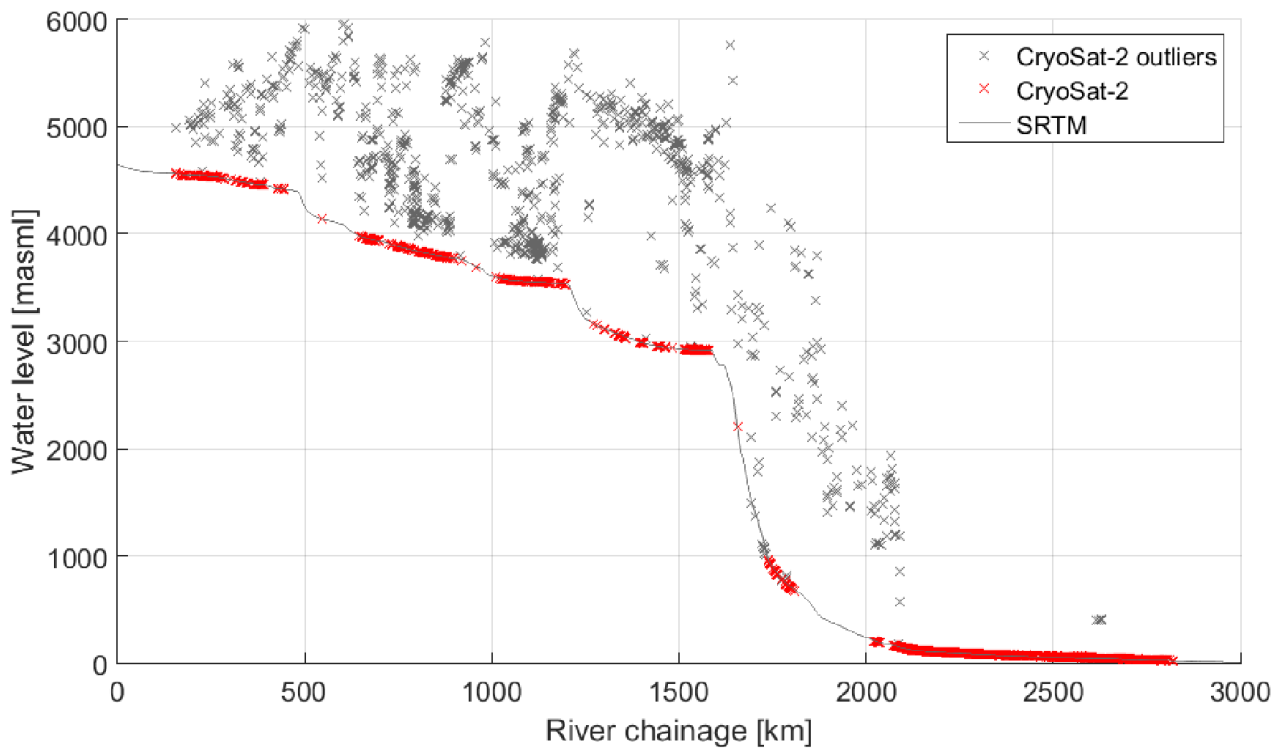
### 6.6.1 Processing of CryoSat-2 data

Figure 52 shows the CryoSat-2 data after filtering over the Brahmaputra river mask and projecting it onto the model's river line (Figure 43). The CryoSat-2 data shows many outliers, mainly in the upstream part before river km 2000. In this area, the river is narrow and the terrain surrounding the river is very steep, explaining some of the outliers. In this portion, the results are very sensitive to the roll bias correction applied to the CryoSat-2 data. Still, meaningful data can be gathered in many places, as can be seen when compared to SRTM elevations along the same river line. If outliers in the CryoSat-2 data are defined as showing an elevation difference of more than 20

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

meters from the SRTM data, about 20% of the data are discarded. For the Assam Valley there however are only 1% outliers, whilst in the upstream Brahmaputra 46% of the CryoSat-2 observations are outliers.



As mentioned, calibration and data assimilation were limited to the Assam valley, i.e. the river after km 2050.



**Figure 52: CryoSat-2 data 2010 to 2013 after filtering and mapping to the model's river line in comparison with SRTM data. The Assam valley starts around chainage km 2050.**

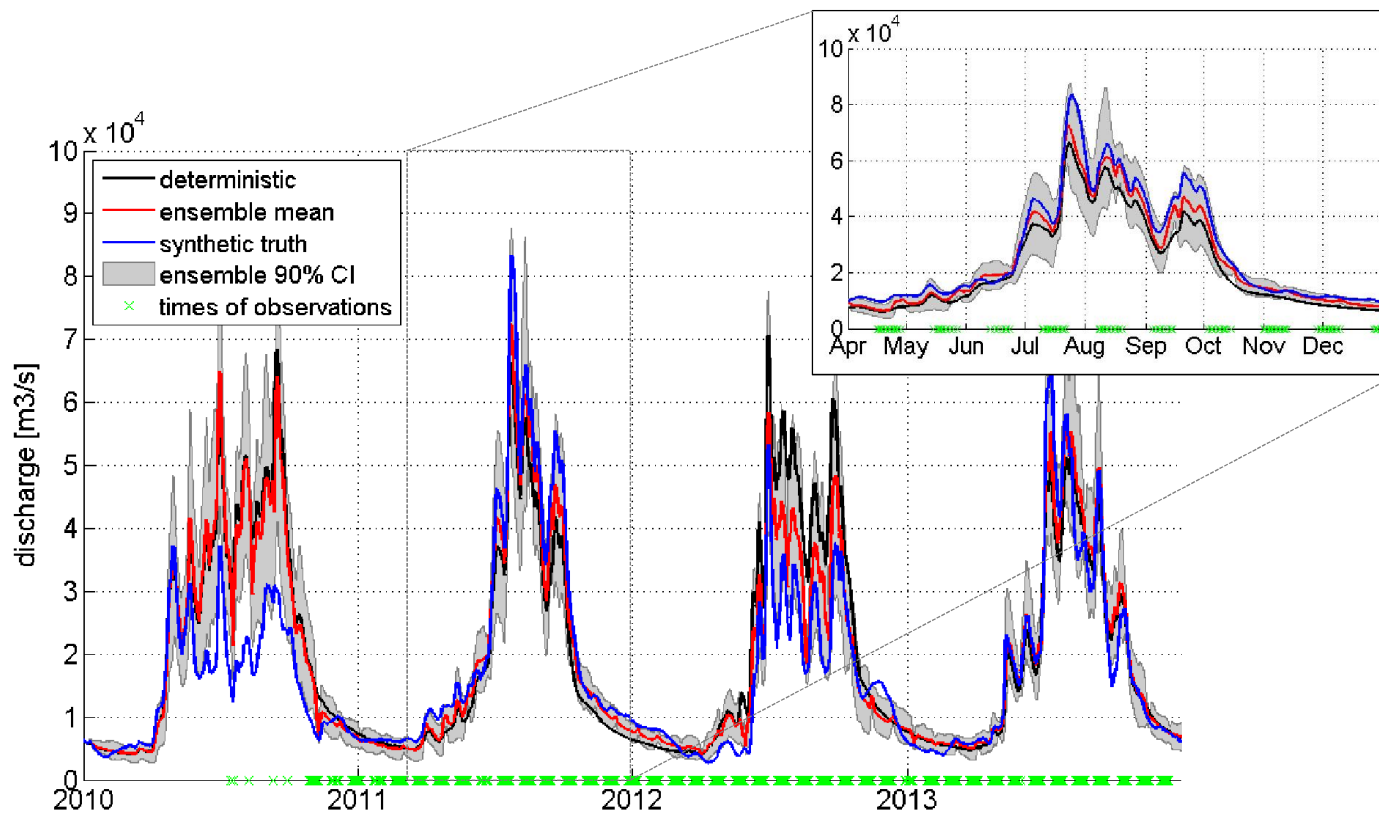
### 6.6.2 Synthetic data assimilation experiment

For the following section please note that between the submission of D4050 and this document the forcings of the model have been changed: the temperature and evapotranspiration products used are now ECMWF ERA Interim (see also section 6.3), hence precise results differ. The general conclusion that can be drawn from the results however remains unchanged.



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

The synthetic data assimilation setup as described above is a way to test whether the suggested data assimilation setup can work as expected. The model used is the actual Brahmaputra basin model, the assimilated observations occur at the points in space and time of the original CryoSat-2 data. Only the elevations of the synthetic data are extracted from a hidden truth model.

The results of such a run, in terms of discharge at Bahadurabad station, i.e. close to the downstream end of the model, can be seen in Figure 53. The deterministic model's discharge is displayed in black. The deterministic model is the "original" model as referred to in section 6.5.9, i.e. the model with the unperturbed, original forcing. The discharge of the hidden truth (or: synthetic truth) model, from which the elevations of the assimilated synthetic observations were extracted, is displayed in blue. The main model from the data assimilation run (i.e. the ensemble mean) is displayed in red, together with the ensemble's 90% confidence intervals in grey. As the observations occur across the entire model domain, only their times can be indicated in Figure 53 as green crosses along the x-axis.



**Figure 53: Results of data assimilation of synthetic CryoSat-2 water level observations in comparison to the synthetic truth model results at Bahadurabad station. The inlet shows April to December 2011 in more detail.**



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

It can be seen that the effect of the assimilation varies a lot over time. This is for example affected by the location of the observations – the further upstream in the model the observations are, the smaller the impact on the discharge at Bahadurabad station is expected to be: the update is localized, and furthermore upstream updates always will be diluted in their effect by the (randomly perturbed) lateral inflows along the river. In general however, it can be seen that the data assimilation is working as expected. Remember also that the synthetic observations occur at the locations of the real CryoSat-2 observations, i.e. they are spread out across the entire Assam Valley. At most times the data assimilation is correcting the model’s discharge towards the synthetic truth, and also the model uncertainty represented by the ensemble confidence intervals is reduced by data assimilation. This can also be seen in Table 6 showing that the data assimilation run displayed in Figure 53 provides a better simulation of discharge compared to the hidden truth than an open loop run of the exact same model, improving its continuous ranked probability score (CRPS) by 24%.

An open loop run refers to an ensemble run of the model setup, as it would be used in data assimilation, just without assimilating any data. This means that an ensemble of model realizations is created in the same way as for the data assimilation (described in section 6.5.6), and also otherwise ran with the same setup and parameters. The deterministic model run (displayed also in Figure 53) serves as a basis for both the data assimilation and the open loop run ensemble. Such an open loop run can be seen as an ensemble forecast of the system, accounting for model uncertainty. Furthermore it serves as a baseline the data assimilation run of the model can be compared to. The deterministic model run could also be chosen as a baseline, however it is harder to compare to the data assimilation results as it only provides one deterministic solution instead of an ensemble.

The CRPS is a tool to evaluate probabilistic forecasts, like for example the ensembles representing the data assimilation and open loop runs described in this section. Its value can be interpreted as the over all timesteps accumulated area between the cumulative distribution function of the forecast and the observation. The observation is considered being deterministic, i.e. it is a Heaviside step function. Consequently, the smaller the CRPS value the better the forecast, with a CRPS of 0 indicating a perfect forecast. For details on the CRPS please refer to Gneiting et al., 2005, and its application in discharge forecasting with hydrologic models is discussed by Bauer-Gottwein et al., 2015.

The sharpness is included in Table 6 as it indicates the model uncertainty. A prime assumption of data assimilation is that it reduces uncertainty, hence the sharpness should be reduced. However, for example by assuming to low observation uncertainty, the model also can be overcorrected, reducing sharpness but also the reliability. This is however not the case here, as the CRPS value

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

does improve. Again, see Bauer-Gottwein et al., 2015 or Michailovsky and Bauer-Gottwein, 2013 for details.

To conclude, it could be shown (see Table 6) that assimilating the data (with the ETKF) improves the performance of the model compared to the open-loop run. This means that it can be expected that the data assimilation setup described here also will provide good results when real altimetry data from CryoSat-2 or any other source will be assimilated.

**Table 6: Performance indicators for the synthetic DA experiment compared to an open loop run of the same model. Discharge at Bahadurabad. The continuous rank probability score (CRPS) is a performance indicator commonly used for ensemble forecast. Sharpness is given as the width of the 90% confidence interval.**



	CRPS [m3/s]	sharpness [m3/s]
open loop run	3688	15647
DA of synthetic data with ETKF	2805	11291

## 6.7 Conclusion

A method to i) set up large-scale 1D hydrologic-hydrodynamic models based entirely (or mainly) on publically available remote sensing data, ii) calibrate its water levels, and iii) assimilate CryoSat-2 altimetry data has been developed. The validity of this method was shown in a data assimilation experiment using synthetic CryoSat-2 data.



The 1D hydrologic-hydrodynamic model used in this study is the Brahmaputra river basin in South Asia. After discharge calibration, also water levels in the model's downstream part, the Assam valley, are calibrated. The calibration procedure suggested relies only on remote sensing data and uses a combination of the repeat cycle orbit altimetry data from the Envisat mission with altimetry data from the drifting orbit mission CryoSat-2 to fit synthetical cross sections datums and shapes to match the observed water level amplitudes and absolute water level profile along the river. This method allows the calibration of water levels in a 1D hydrodynamic model without the knowledge of any bathymetry or cross sections.

A model representing precise discharges and water levels at any point in time and model space is also the precondition to be used in a data assimilation setup with the drifting orbit altimetry data from CryoSat-2. The CryoSat-2 level 2 data over the Assam valley have been filtered over river masks from Landsat imagery, and projected onto the model's river line. The data can be assimilated to the hydrodynamic part of the Brahmaputra model by the help of the DHI Data Assimilation Framework, using an ETKF. This DA scheme is operational and working as expected

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16



with a synthetic experiment. For the synthetic experiment, the Brahmaputra basin model was updated by synthetic observations at the same points in time and space as the original CryoSat-2 observations, only that the water levels were replaced with values extracted from the water levels of a perturbed run of the model.

This is the first time that a DA scheme was developed for CryoSat-2-type altimetry data, i.e. distributed in space and time and not occurring at virtual stations only. In a next step, real CryoSat-2 altimetry data can be assimilated to the model. If other errors such as errors in the in-situ data considered the truth, errors in the filtering and projection of the CryoSat-2 data onto the river model, errors in the retracked heights of the CryoSat-2 data as such etc. are of controllable magnitude, it is expected that the assimilation of CryoSat-2 altimetry data will improve the model's forecast.

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

## 7. References

- Armitage, T. W. K. and Davidson, M. W. J. 2014. Using the interferometric capabilities of the ESA CryoSat-2 mission to improve the accuracy of sea ice freeboard retrievals, *IEEE Trans. Geosci. Remote Sens.*, Vol. 52, 529–536.
- Bauer-Gottwein, P., Bates, P., Getirana, A., Andreadis, K., Matgen, P., Neal, J., Biancamaria, S., Michailovsky, C.I., 2016. Altimetry Data Assimilation in Hydrological Modeling, in: Benveniste, J., Vignudelli, S., Kostianoy, A. (Eds.), *Inland Water Altimetry*. Springer.
- Bauer-Gottwein, P., Jensen, I.H., Guzinski, R., Bredtoft, G.K.T., Hansen, S., Michailovsky, C.I., 2015. Operational river discharge forecasting in poorly gauged basins: the Kavango River basin case study. *Hydrol. Earth Syst. Sci.* 19, 1469–1485. doi:10.5194/hess-19-1469-2015
- Becker, M., da Silva, J., Calmant, S., Robinet, V., Linguet, L., Seyler, F., 2014. Water Level Fluctuations in the Congo Basin Derived from ENVISAT Satellite Altimetry. *Remote Sens.* 6, 9340–9358. doi:10.3390/rs6109340
- Berry, P., 2009. *River and Lake Product Handbook v3.5*.
- Berry, P.A.M. et al. 2010a, An enhanced ocean and coastal zone retracking technique for gravity field computation, *Gravity, Geoid and Earth Observation, International Association of Geodesy Symposia 2010*, 135,(3), pp. 213-220
- Berry P.A.M., Smith R.G., Witheridge S., Wheeler J., 2010b; *Global Inland water monitoring from Satellite Radar Altimetry - a glimpse into the future*. ESA Living Planet Symposium, Bergen, Norway, 27th June - 2nd July, ESA SP-686, ISBN 978-92-9221-250-6
- Berry P.A.M., Smith R.G., 2010c; *A Global Assessment of the Envisat RA-2 Performance over Non-Ocean Surfaces*. ESA Living Planet Symposium, Bergen, Norway, 27th June - 2nd July, ESA SP-686, ISBN 978-92-9221-250-6
- Berry, P.A.M., Smith, R.G. and Benveniste, J., 2012a. *EnviSat altimetry for river and lakes monitoring*. *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1844-1847

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical Basis Document
		Doc. No	NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Berry, P.A.M., Smith, R.G., Salloway, M.K. and Benveniste, J., 2012b. Global Analysis of EnviSat Burst Echoes Over Inland Water. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (5), pp 1980-1985

Biancamaria, S., Durand, M., Andreadis, K.M., Bates, P.D., Boone, A, Mognard, N.M., Rodríguez, E., Alsdorf, D.E., Lettenmaier, D.P., Clark, E. A., 2011a. Assimilation of virtual wide swath altimetry to improve Arctic river modeling. *Remote Sens. Environ.* 115, 373–381. doi:10.1016/j.rse.2010.09.008

Biancamaria, S., Hossain, F., Lettenmaier, D.P., 2011b. Forecasting transboundary river water elevations from space. *Geophys. Res. Lett.* 38, L11401. doi:10.1029/2011GL047290

Birkinshaw, S.J., Moore, P., Kilsby, C.G., O'Donnell, G.M., Hardy, A.J., Berry, P.A.M., 2014. Daily discharge estimation at ungauged river sites using remote sensing. *Hydrol. Process.* 28, 1043–1054. doi:10.1002/hyp.9647

Central Water Commission, 2009. *Integrated Hydrological Data Book (non-classified river basins)*. New Delhi.

Cullen, R.A. and Wingham, D.J., 2002. CryoSat Level 1b Processing Algorithms and Simulation Results, *Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International*, 3, 1762-1764, DOI 10.1109/IGARSS.2002.1026246



Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J. N., and Vitart, F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteorol. Soc.*, 137, 553–597, doi:10.1002/Qj.828, 2011

ESA-ESRIN. (2013) *Cryosat Product Handbook*

ESA-ESRIN, (2015) Main evolutions and expected quality improvements in BaselineC Level1b products, DOC C2-TN-ARS-GS-5154, <https://wiki.services.eoportal.org/tiki-index.php?page=CryoSat+Technical+Notes>

ESA-ESRIN (2016) *Cryosat characterisation for FBR users*, DOC C2-TN-ARS-GS-5179, <https://wiki.services.eoportal.org/tiki-index.php?page=CryoSat+Technical+Notes>



		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Feyisa G.L., Meilby H., Fensholt R., Proud S.R. 2014. Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* 140:23–35. doi: 10.1016/j.rse.2013.08.029.

Garcia, E.S., Sandwell, D.T., Smith, W.H.F., 2014, Retracking CryoSat-2, Envisat and Jason-1 radar altimetry waveforms for improved gravity field recovery. *Geophys. J. Int.*, doi: 10.1093/gji/ggt469.

Gneiting, T., Raftery, A.E., Westveld, A.H., Goldman, T., 2005. Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *Mon. Weather Rev.* 133, 1098–1118. doi:10.1175/MWR2904.1

Gommenginger, C., Thibaut, P., Fenoglio-Marc, L., Quartly, G., Deng, X., Gomez-Enri, J., Challenor, P. and Gao, Y. Retracking Altimeter Waveforms Near the Coasts, in S. Vignudelli, A.G. Kostianoy, P. Cipollini, J. Benveniste (eds.), *Coastal Altimetry*, Springer-Verlag Berlin Heidelberg, 2011. DOI: 10.1007/978-3-642-12796-0\_

Google, 2015. Google Earth Engine - Landsat 7 32-Day NDVI Composite [WWW Document]. URL [https://earthengine.google.org/#detail/LANDSAT%2FLE7\\_L1T\\_32DAY\\_NDVI](https://earthengine.google.org/#detail/LANDSAT%2FLE7_L1T_32DAY_NDVI)

Jain, S.K., Agarwal, P.K. and Singh, V.P. (2007) *Hydrology and Water Resources of India*, Netherlands: Springer.



Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database.

Kleinherenbrink, M., Ditmar, P.G., Lindenbergh, R.C., 2014. Retracking CryoSat data in the SARin mode and robust lake level extraction. *Remote Sens. Environ.* 152, 38–50. doi:10.1016/j.rse.2014.05.014

McFeeters, S.K. 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.*, 17, 1425–1432.

Michailovsky, C.I., Bauer-Gottwein, P., 2013. Operational reservoir inflow forecasting with radar altimetry: the Zambezi case study. *Hydrol. Earth Syst. Sci. Discuss.* 10, 9615–9644. doi:10.5194/hessd-10-9615-2013

Michailovsky, C. I., McEnnis, S., Berry, P. A. M., Smith, R., and Bauer-Gottwein, P.:2012. River monitoring from satellite radar altimetry in the Zambezi River basin, *Hydrol. Earth Syst. Sci.*, 16, 2181-2192, doi:10.5194/hess-16-2181-2012

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

Michailovsky, C.I., Milzow, C., Bauer-Gottwein, P., 2013. Assimilation of radar altimetry to a routing model of the Brahmaputra River. *Water Resour. Res.* 49, 4807–4816. doi:10.1002/wrcr.20345

MIKE by DHI, 2009. MIKE 11 - A Modelling System for Rivers and Channels - Reference Manual.

NASA, 2015. NASA Science Missions - SWOT [WWW Document].

Nielsen, S.A., Hansen, E., 1973. Numerical simulation of the rainfall runoff process on a daily basis. *Nord. Hydrol.* 4, 171–190.

Nielsen, K., Stenseng, L., Andersen, O.B., Villadsen, H., Knudsen, P., 2015. Validation of CryoSat-2 SAR mode based lake levels. *Remote Sens. Environ.* 171, 162–170. doi:10.1016/j.rse.2015.10.023

Rodríguez, E., Morris, C.S., Beiz, E.J., 2006. A Global Assessment of the SRTM Performance. *Photogramm. Eng. Remote Sens.* 72, 249 – 260.

Sakov, P., Oke, P.R. 2008. Objective Array Design: Application to the Tropical Indian Ocean. *Journal of Atmospheric and Oceanic Technology*, 25, 794–807, doi: 10.1175/2007JTECHO553.1.

Smith, R.G. and Berry, P.A.M., 2011. Evaluation of the differences between the SRTM and satellite radar altimetry height measurements and the approach taken for the ACE2 GDEM in areas of large disagreement. *Journal of Environmental Monitoring*, 13 (6), pp. 1646-1652



Smith, W.H.F. & Scharroo, R., 2011. Retracking range, SWH, sigma-naught, and attitude in CryoSat conventional ocean data. In proceedings of Ocean Surface Topography Science Team Meeting. San Diego, October 19-21, 2011.

Song, C., Ye, Q., Cheng, X., 2015. Shifts in water-level variation of Namco in the central Tibetan Plateau from ICESat and CryoSat-2 altimetry and station observations. *Sci. Bull.* 60, 1287–1297. doi:10.1007/s11434-015-0826-8

Tarpanelli, A., Brocca, L., Barbetta, S., Faruolo, M., Lacava, T., Moramarco, T., 2014. Coupling MODIS and Radar Altimetry Data for Discharge Estimation in Poorly Gauged River Basins. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8, 1–8. doi:10.1109/JSTARS.2014.2320582

Tropical Rainfall Measurement Mission Project (TRMM) (2011), TRMM/TMPA 3B42 TRMM and Others Rainfall Estimate Data V7, version 7, Greenbelt, MD: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), Accessed 16/02/2016 at [http://disc.sci.gsfc.nasa.gov/datacollection/TRMM\\_3B42\\_V7.html](http://disc.sci.gsfc.nasa.gov/datacollection/TRMM_3B42_V7.html)

Villadsen, H., Andersen, O.B., Stenseng, L., Nielsen, K., Knudsen, P., 2015. CryoSat-2 altimetry for

		ESA Contract:	1/6287/11/I-NB
		Doc. Title	D4100 Algorithm Theoretical
		Doc. No	Basis Document NCL_CRUCIAL_D4100
		Version No	3
		Date	04.10.16

river level monitoring - Evaluation in the Ganges-Brahmaputra basin. Remote Sens. Environ. 168, 80–89. doi:10.1016/j.rse.2015.05.025

Wheeler J., Berry P.A.M., Smith R.G., Benveniste J., 2010; The ESA Near-Real-Time River & Lake Processor. ESA Living Planet Symposium, Bergen, Norway, 27th June - 2nd July, ESA SP-686, ISBN 978-92-9221-250-6

Wingham D. J., Francis C.R., Baker S., Bouzinac C., Cullen R., de Chateau-Thierry P., Laxon S.W., Mallow U., Mavrocordatos C., Phalippou L., Ratier G., Rey L., Rostan F., Viau P., and Wallis D., CryoSat: A Mission to Determine the Fluctuations in Earth's Land and Marine Ice Fields. Advances in Space Research 37 (2006) 841-871

Yamazaki, D., Lee, H., Alsdorf, D.E., Dutra, E., Kim, H., Kanae, S., Oki, T., 2012. Analysis of the water level dynamics simulated by a global river model: A case study in the Amazon River. Water Resour. Res. 48, n/a–n/a. doi:10.1029/2012WR011869

Yoon, Y., Durand, M., Merry, C.J., Clark, E. a., Andreadis, K.M., Alsdorf, D.E., 2012. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. J. Hydrol. 464-465, 363–375. doi:10.1016/j.jhydrol.2012.07.028