| | | | |
|---|---|---|---|
| Newcastle University School of Civil Engineering &Geosciences | DTU | ESA Contract: | 1/6287/11/I-NB |
| | | Doc. Title<br>Doc. No | D4050 Algorithm Interim Report<br>NCL_CRUCIAL_D4050 |
| | | Version No<br>Date | 4<br>16.02.16 |

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

# D4050 Algorithm Interim Report

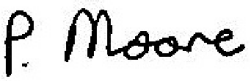
Document No: **NCL_CRUCIAL_D4050**

Issue: **4**

Issue Date: **16 February 2016**

| | | | |
|---|---|---|---|
| Newcastle University School of Civil Engineering &Geosciences | DTU | ESA Contract: | 1/6287/11/I-NB |
| | | Doc. Title<br>Doc. No | D4050 Algorithm Interim Report<br>NCL_CRUCIAL_D4050 |
| | | Version No<br>Date | 4<br>16.02.16 |

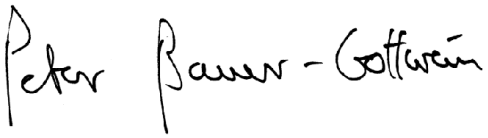
Author (Chapters 1-4): Philip Moore

Signature: 

Author (Chapter 5): Steve Birkinshaw

Signature:

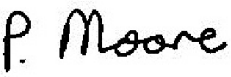
Author (Chapter 6): Peter Bauer-Gottwein

Signature: 

Author (Chapter 6): Raphael Schneider

Signature:

Authorised by: Philip Moore

Signature: 

**Document Change Record**

| Version | Date | Modified by | Description |
|---|---|---|---|
| 1 | 20.10.15 | PM | Created |
| 2 | 07.12.15 | PM, SJB, PB-G | Revisions in response to ESA/ESRIN review |
| 3 | 03.02.16 | PM, PB-G | Revisions in response to ESA/ESRIN review |
| 4 | 16.02.16 | PM, PB-G | Revisions in response to ESA/ESRIN review |
| | | | |

# Abstract

This Algorithm Interim Report (AIR) describes the algorithms used to estimate inland water heights from CryoSat-2SAR FBR data and their utilization in river modelling. The report identifies the source of the input data, outlines the principles and mathematical background of SAR 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
- Geophysical correction data
- River mask data
- River modelling

The report, as the title states, is interim and not intended to give all algorithms in full. These will be included in the Algorithm Theoretical basis Documents (ATBD) at Milestone 4.

# Contents


**Abstract** ..... 4

**Figures** ..... 7

**Tables** ..... 11

**1. Introduction** ..... 12

**1.1 Scope** ..... 12

**1.2 Abbreviations and Acronyms** ..... 13

**2 SAR FBR Parameters** ..... 16

**3 SAR FBR Theory** ..... 17

**3.1 Beam Formation** ..... 17

**3.2 Geophysical Corrections** ..... 20

**3.3 Beam Steering** ..... 21

**3.4 Burst Centre Ground Points** ..... 23

**3.5 Multi-Look Ground Points** ..... 31

**3.6 Stacking** ..... 31

**3.7 Stack Beam Alignment** ..... 32

3.7.1 Tracker Bin Correction to Burst Echoes ..... 32

3.7.2 Slant Range Correction ..... 33

**3.8 Multi-look Waveforms** ..... 33

**3.9 Retracking** ..... 35

**4 SAR FBR Analysis: Summary** ..... 45

**5 Mekong River Masks** ..... 54

**5.1 Mekong: Landsat river mask** ..... 54

**5.2 Landsat river mask summary** ..... 62

**6 Assimilation of CryoSat-2 data to 1-dimensional hydrodynamic models** ..... 63

**6.1 Introduction and background** ..... 63

**6.2 Data and methods** ..... 65

6.2.1 CryoSat-2 data ..... 65

6.2.2 Preprocessing of the data ..... 65

6.2.3 Brahmaputra: Landsat river mask ..... 66

6.2.4 Filtering and projecting the CryoSat-2 data ..... 67

**6.3 Hydrologic hydrodynamic model ..... 68**

**6.4 Cross section calibration ..... 71**

**6.5 Data Assimilation ..... 73**

6.5.1 Data Assimilation in general ..... 73

6.5.2 Ensemble Transform Kalman Filter ..... 74

6.5.3 Setup with DHI Data Assimilation Framework ..... 76

6.5.4 Processing of CryoSat-2 observations ..... 77

6.5.5 Observation error ..... 77

6.5.6 Model error ..... 77

**6.6 Results and Discussion ..... 78**

6.6.1 Processing of CryoSat-2 data ..... 78

6.6.2 Hydrologic model calibration ..... 79

6.6.3 Cross section calibration ..... 82

6.6.4 Brahmaputra Data Assimilation ..... 84

**6.7 Conclusion ..... 86**

**7. References ..... 87**

# Figures

Figure 1: SAR mode bursts: transmit and receive ..... 20

Figure 2: Altimetric heights and datums: the index 'i' refers to burst position along the orbit (page 24) and index 'k' to points along the ground (page 31). ..... 22

Figure 3: Schematic of rock angle ..... 22

Figure 4: Schematic of geoidal, ellipsoidal and orthometric heights. ..... 25

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˚. ..... 26

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), Y-axis is power amplitude. ..... 27

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), Y-axis is power amplitude. ..... 28

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), Y-axis is power amplitude. ..... 29

Figure 9: Burst echo heights across the Mekong. X-axis latitude (deg); y axis (m): Orthometric height (blue) from burst echoes. ..... 30

Figure 10: Figure 9 zoomed. ..... 30

Figure 11: Schematic of bursts, the fan of Doppler beams, ground points and multi-looks ..... 34

Figure 12: Empirical retrackers. X-axis is gate number (1-256), Y-axis is power amplitude ..... 37

Figure 13: Google Earth image of 19 Apr 2011 Mekong crossing. Satellite ground track N-S. ..... 39

Figure 14: Geoid heights across the Mekong for 19 Apr 2011 (day 109 in year) for various number of echoes in multi-look. Upper plot full profile; lower plot a zoomed version of the river surface ..... 40

Figure 15: Seasonal variation in extent of Tonle Sap ..... 41

Figure 16: Points across Tonle Sap (3 Dec 2011) ..... 42

Figure 17: Geoid height measurements across Tonle Sap (3 Dec 2011) with various multi-look combinations. Empirical retrackers used (Figure 12). ..... 43

Figure 18: Geoid height measurements across Tonle Sap (3 Dec 2011) with various multi-look combinations. OCOG/Threshold retracker used. .....43

Figure 19: 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=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. . X-axis is gate number (1-256), Y-axis is power amplitude.....46

Figure 20: 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=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.....47

Figure 21: 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=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.....48

Figure 22: 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=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude. .....49

Figure 23: 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=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.....50

Figure 24: 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=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude. .....51

Figure 25: 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=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.....52

Figure 26: Geolocation of points i=86 and i=118 corresponding to waveforms of Figure 19 - Figure 25. 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. .....53

Figure 27: Six Landsat images of the Mekong River between 12°N and 18°N.....56

Figure 28: Landsat image from 12/11/2014 of the Mekong River between 14.2°N and 14.8°N..57

Figure 29: Approximate mask of the Mekong River between 12°N and 18°N.....57

Figure 30: Approximate mask of the Mekong River between 14.2°N and 14.8°N.....58

Figure 31: Water mask of the Mekong River between 12°N and 18°N. .....59

Figure 32: Water mask of the Mekong River between 14.2°N and 14.8°N .....60

Figure 33: Final mask of the Mekong River between 12°N and 18°N. .....61

Figure 34: Final mask of the Mekong River between 14.2°N and 14.8°N.....62

Figure 35: Map of the three main rivers draining into the Gulf of Bengal through Bangladesh ...64

Figure 36: Landsat 7 image of the same part of the Brahmaputra river in the Assam valley showing the dynamic changes in river morphology. Left: 2010. Right: 2011. .....66

Figure 37: 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. .....67

Figure 38: The Mike-11 computational grid .....68

Figure 39: 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). .....70

Figure 40: Sketch of the two-step cross section calibration. MIKE 11 assumes vertical walls for cross sections outside their defined limits. .....72

Figure 41: Overview over the DHI Data Assimilation Framework. Source: Marc Ridler, DHI.....76

Figure 42: 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 2000. .....79

Figure 43: Observed vs. simulated discharge for the Brahmaputra at Bahadurabad station for the calibration period 2002 - 2007 .....80

Figure 44: Observed vs. simulated discharge for the Brahmaputra at Bahadurabad station for the data assimilation period 2010 – 2013. Note that observed data only exists for the high-flow periods. .....80

Figure 45: Result of water level calibration step 1 for the Assam valley for the period 2010 to 2013. All levels are shown as elevations relative to the reference model's cross section datums based on the SRTM DEM. .....82

Figure 46: Water levels after step 2 of the cross section calibration for one virtual station. All levels relative to the water levels at the time of the first Envisat observation. .....83

Figure 47: Comparison of probabilistic discharge simulations produced with and without assimilation of CryoSat-2 data. The central model prediction is shown as a solid line, grey shades indicate 95% confidence intervals, the in-situ observations are in red. .....85

# Tables

Table 1: Statistics of fit for pass across Tonle Sap. ..... 44

Table 2: 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 4-6. A total of 68 waveforms were retracked. ..... 44

Table 3: Landsat 8; Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) bands.. 55

Table 4: Decision variables and objective functions of the genetic algorithm used for the two-step water level calibration ..... 73

Table 5: Number of observations and performance indicators for the station Bahadurabad for the full calibration period 2002 - 2007, the high-flow seasons of 2002 - 2007 and the high flow seasons of 2010 - 2013 ..... 81

Table 6: Comparative performance of the deterministic model run, the probabilistic model run without data assimilation (“open loop”) and the run with assimilation of CryoSat-2 data.. 85

# 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 land and water applications from CryoSat-2 with a forward-look component to the future Sentinel-3 mission. This Algorithm Interim Report presents the approach developed to process and utilize SAR FBR 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 is classified as "interim" to D4100 Algorithm Theoretical basis Documents (ATBD). This report focuses on SAR FBR data and hydrological modelling and hence partially satisfies the adjusted objectives of the ATBD in the sense that the ATBD will include further algorithmic developments including a preliminary analysis of SARin data.

Deliverable D4050 presents theoretical and computational aspects of the SAR FBR algorithms in particular including comments to high-light the capability of the methodology as well as to identify limitations. D4050 does not contain full results as those will be in the Product Validation Report (D4200).

## 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 |
| 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 |
| GMES | Global Monitoring for Environment and Security programme (renamed as *Copernicus*) |
| 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 |
| OLI | Operational Land Imager |
| 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 |

# 2 SAR FBR Parameters

This section details some of the fundamental parameters of CryoSat-2 used in the subsequent SAR FBR processing.

Burst repetition interval = 11.7 ms

Echoes per burst = 64

Tracking Samples per echo = 128

Track window = 60 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)

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

$$\phi_b(m) = I_b(m) + iQ_b(m) \qquad \text{m=0, 1, ..., 63} \qquad (1)$$

where $i = \sqrt{-1}$ is the imaginary number, $\phi$, represents the complex valued echo, m =0, 1, ..., 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* quadrature component. From Wingham et al. (2006), $\underline{x}_b$ lies mid-way between the 32nd and 33rd 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

$$\phi_b(m) = \sum_k \tilde{\phi}_b(\mathrm{x}_k)\, \mathrm{e}^{2i(m-63/2)\Delta k_0 \hat{n}_k \cdot \hat{n}_y} \qquad \text{m=0, 1, ..., 63} \qquad (2)$$

In Eq. (2), k denotes the elemental scatterers on the ground at location $x_k$, $\tilde{\phi}_b(x_k)$ the reflectance from the kth scatterer, Δ the along-track sampling interval, $\underline{\hat{n}}_k$ the unit vector from

burst centre to elemental scatterer and $\hat{\underline{n}}_v$ the direction of the velocity vector. Since the satellite moves only 60m 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

$$\psi_b(\theta_1) = \sum_{m=0}^{63} \phi_b(m) e^{-2ik_0\Delta\theta_1(m-63/2)} \quad , \tag{3}$$

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

$$\theta_1 = \theta_R^b + \frac{\pi n}{64k_0\Delta} \quad \text{n = 0, 1,....,63} \tag{4}$$

where we can write 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

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

The dependence of $\psi$ 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

$$\psi_b(n) = \sum_k \tilde{\phi}_b(x_k) d(\hat{\underline{n}}_k . \hat{\underline{n}}_v - \theta_R^b - \frac{\pi n}{64k_0\Delta}) \tag{6}$$

where

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

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, …., 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. 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).

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

Now by denoting

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

in Eq. (5) we have

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

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

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

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.

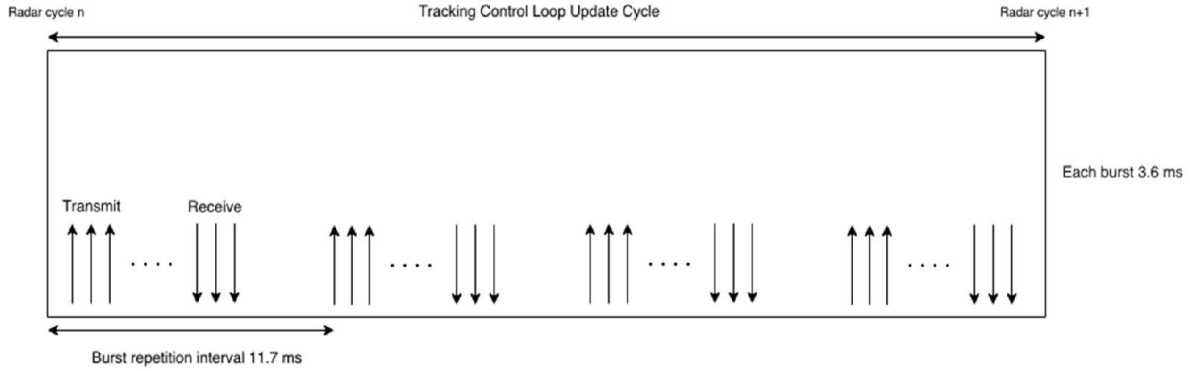



**Figure 1: SAR mode bursts: transmit and receive**

## 3.2 Geophysical Corrections

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

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

where $Alt_i$ is the datum of bin 63 in the data file, and the $USO_{corr}$, $GEO_{corr}$ and $INS_{rangecorr}$ 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.

In detail

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

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

whence (12)

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

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.

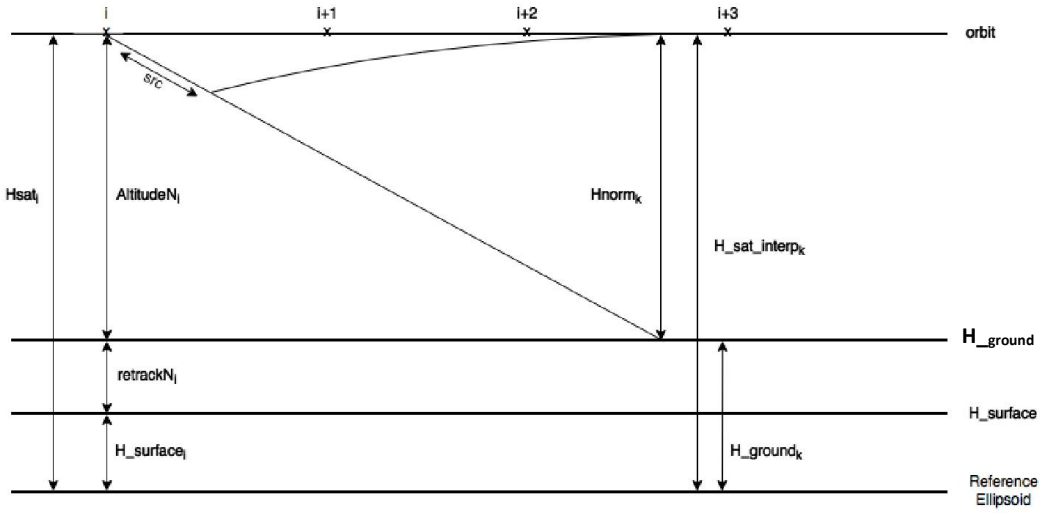



**Figure 2: Altimetric heights and datums: the index 'i' refers to burst position along the orbit (page 24) and index 'k' to points along the ground (page 31).**

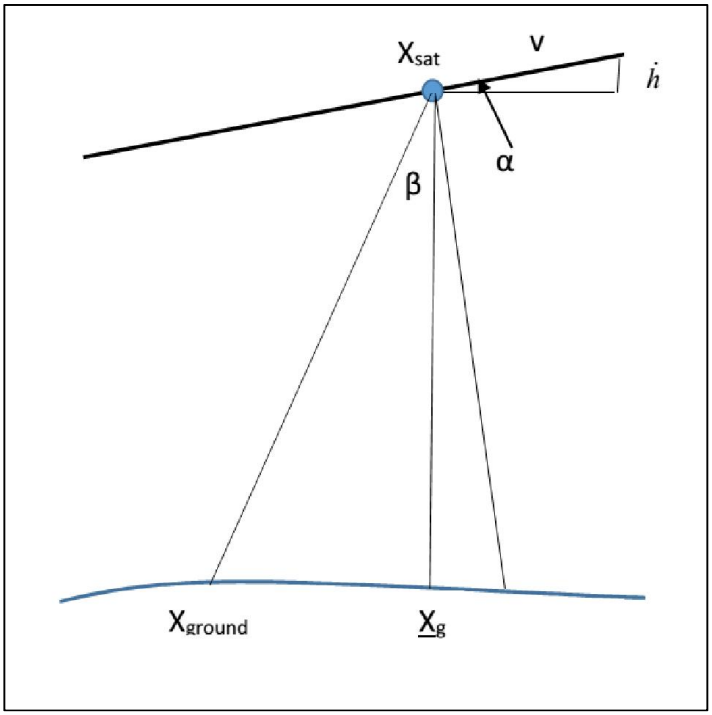



**Figure 3: Schematic of rock angle**

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, …, 31 backward looking from the satellite and n=33, 34, …, 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.

## 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.

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.

4. Filter orthometric heights across the water body and reinstate geoidal height.

5. This first-look provides approximate geodetic (ellipsoidal) coordinates of the sub-surface ground track of the burst centres at approximately 80m 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 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. The filter process will give the lower value for the crossing based on a 3σ rejection criterion.

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

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

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

The retracked burst echoes 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 by removing the obvious outliers. This point will be revisited later after the multi-looked waveforms have been analysed.

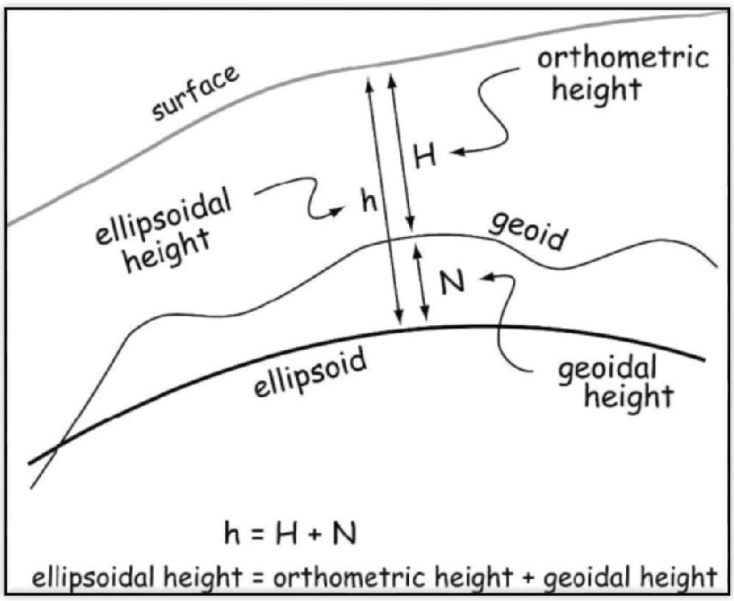



**Figure 4: Schematic of geoidal, ellipsoidal and orthometric heights.**

**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˚.**

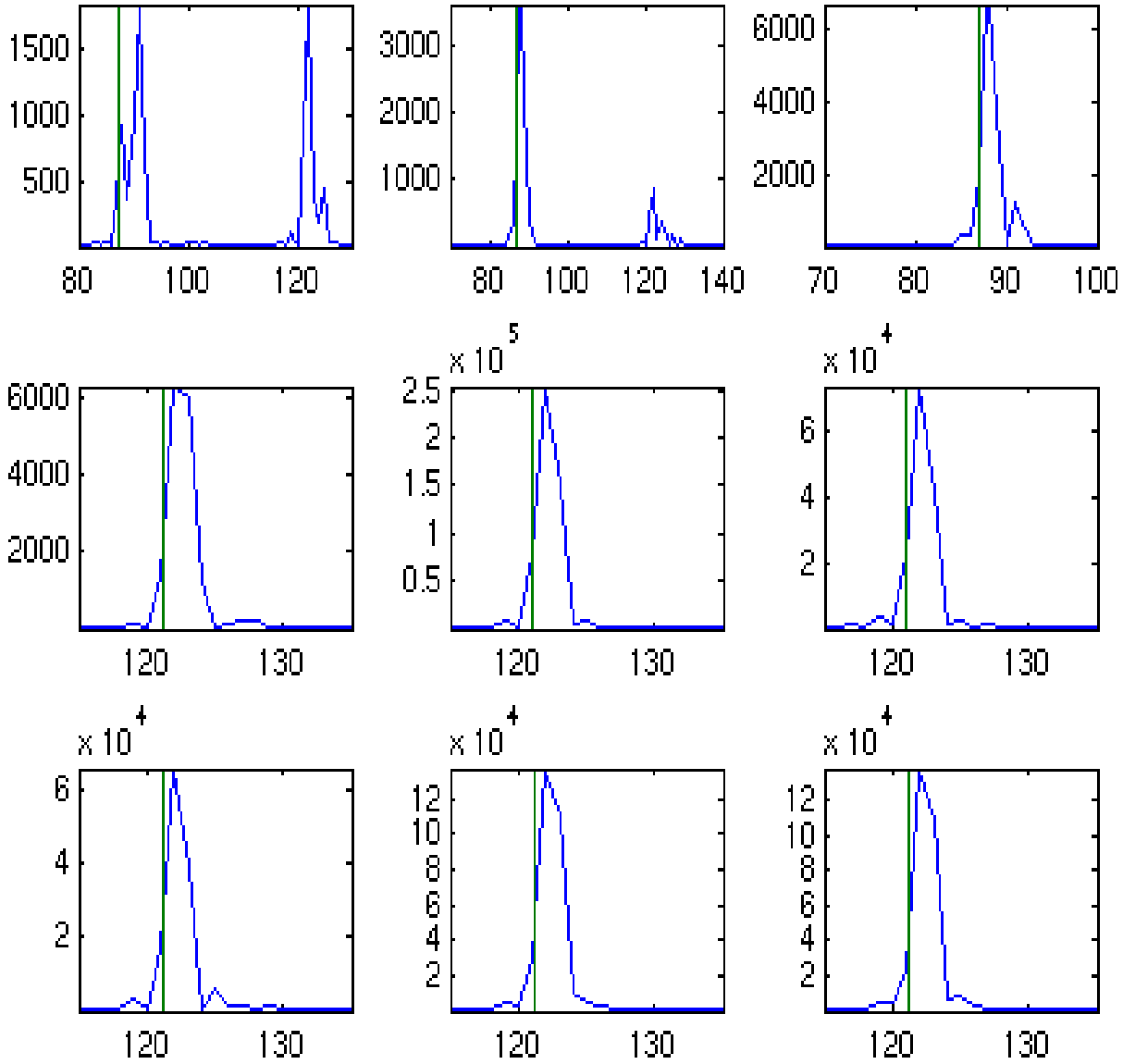



**Figure 6: Burst echo waveforms from a descending pass across the Mekong (#1340-1332), lat 12.2518˚-12.2575˚. Green line OCOG retracked height. X-axis is gate number (1-256), Y-axis is power amplitude.**

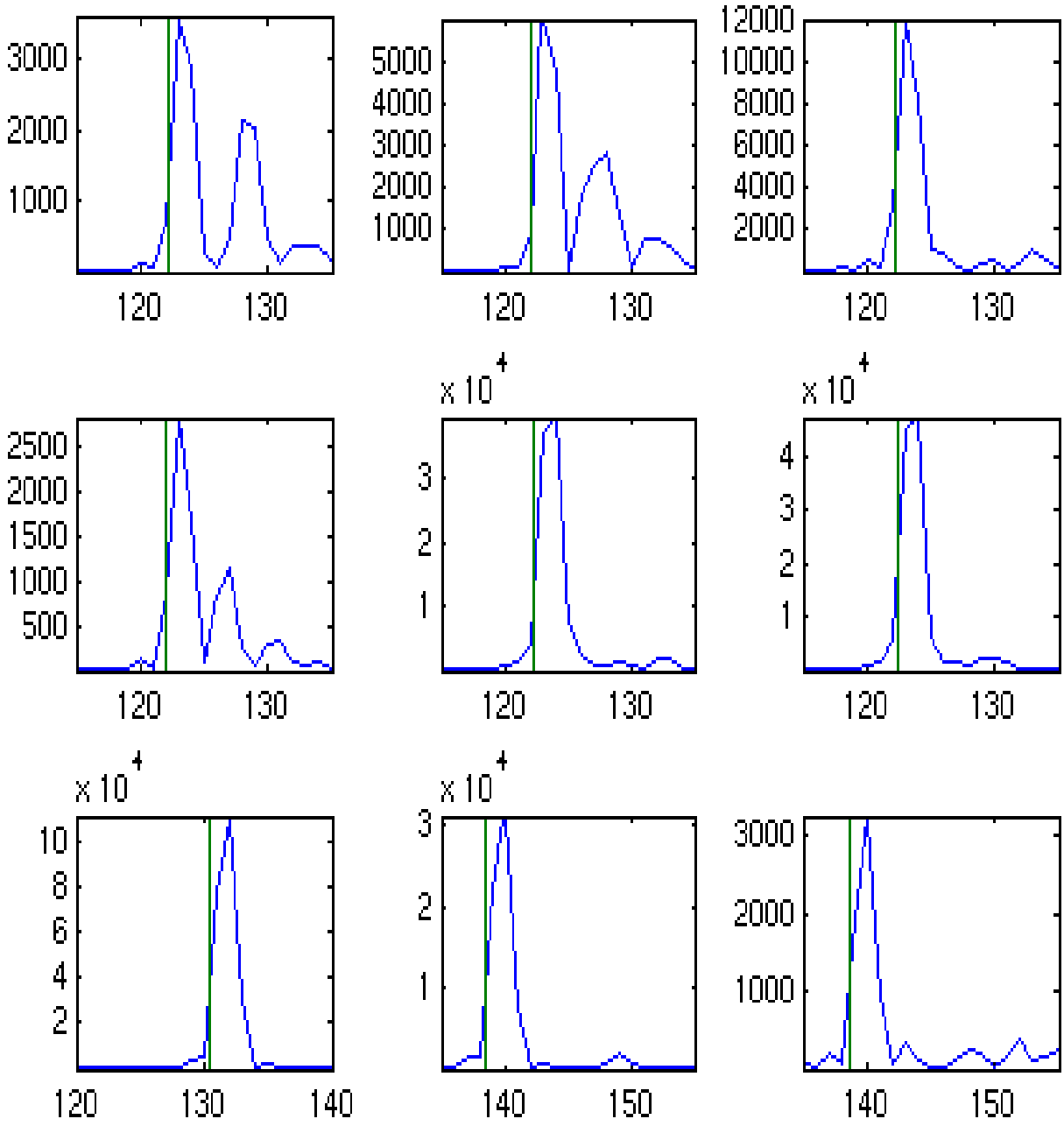



**Figure 7: Burst echo waveforms from a descending pass across the Mekong (#1331-1323), lat 12.2582˚-12.2640˚. Green line OCOG retracked height. X-axis is gate number (1-256), Y-axis is power amplitude.**

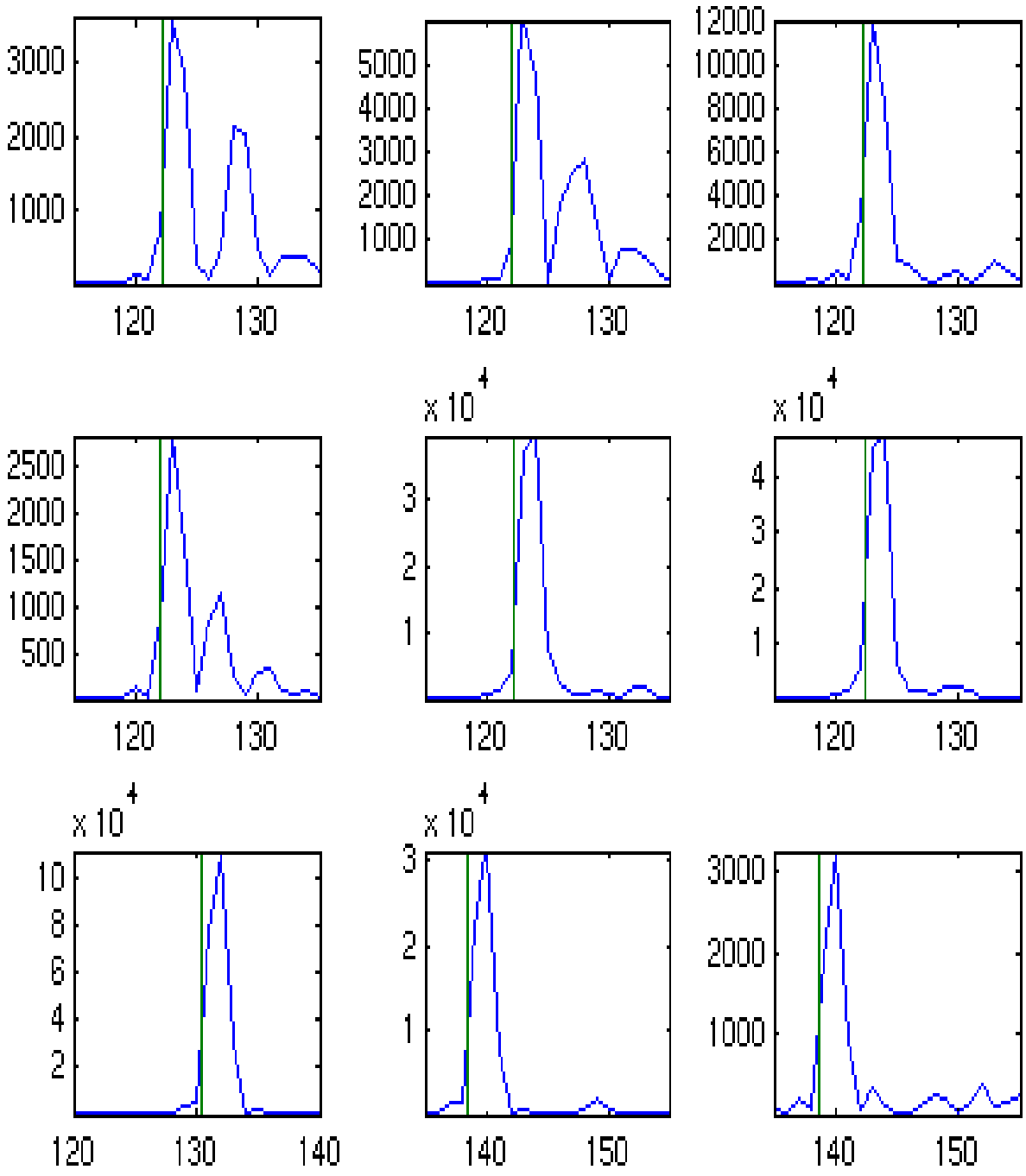



**Figure 8: Burst echo waveforms from a descending pass across the Mekong (#1322-1314), lat 12.2647˚-12.2704˚. Green line OCOG retracked height. X-axis is gate number (1-256), Y-axis is power amplitude.**

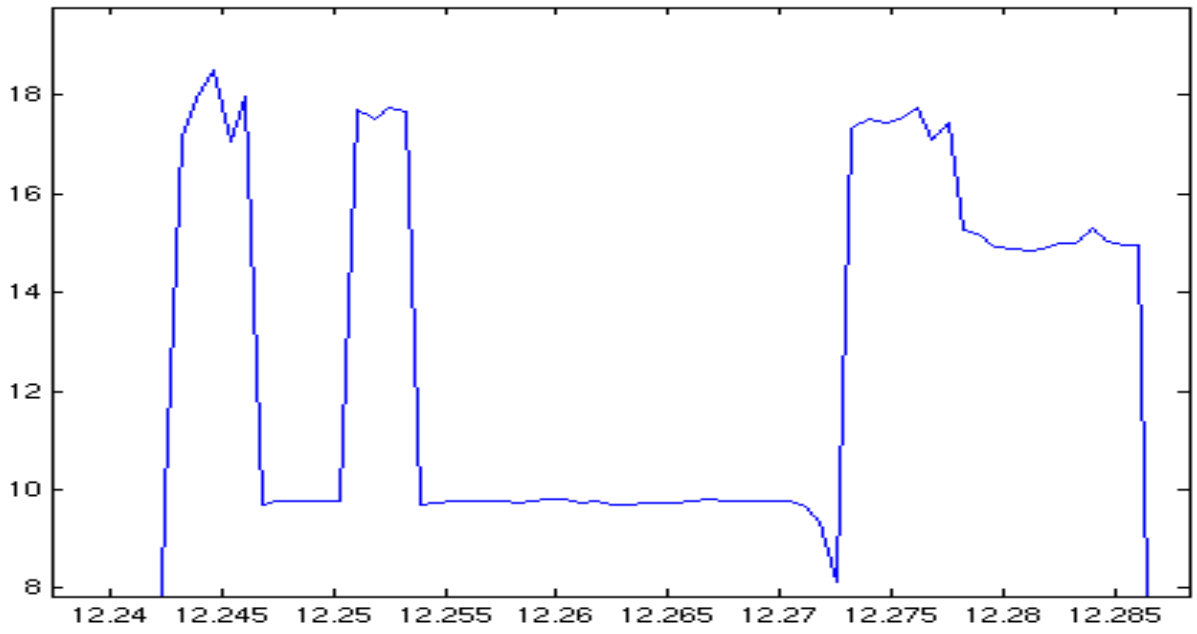



**Figure 9: Burst echo heights across the Mekong. X-axis latitude (deg); y axis (m): Orthometric height (blue) from burst echoes.**

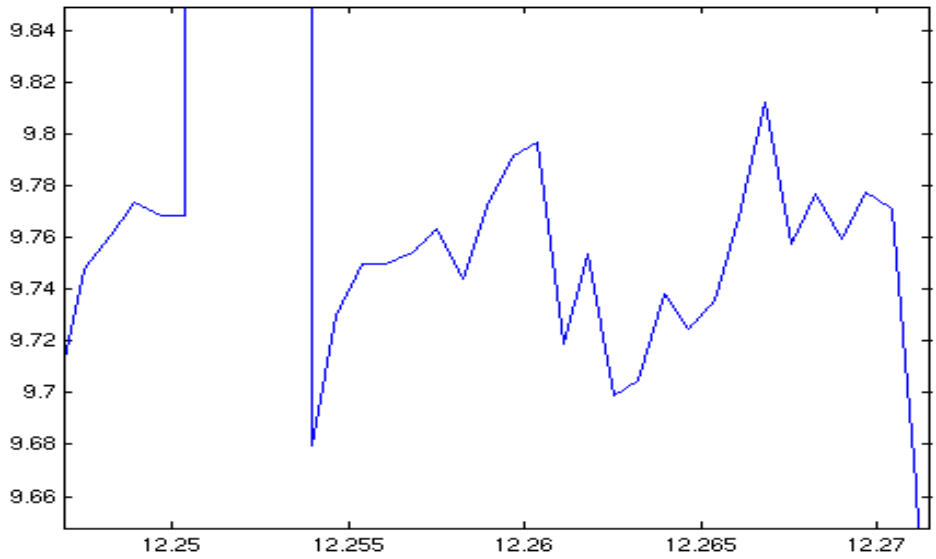



**Figure 10: Figure 9 zoomed.**

## 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 along-track 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 $\psi_b(n)$ 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)). This stack can comprise up to 240 multi-looks. Each waveform from a single burst is subject to radar speckle which limits the precision of the retracked heights from the burst echoes. However, by multi-looking, the random nature of the speckle is reduced. However, given uncertainties in the orbital positioning etc the waveforms in this stack will be incoherent in phase given the wavelength of the altimeter. Further, 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.

## 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.

### 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 $i^{th}$ 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 tracking window (60 m) divided by the tracking samples per echo.

### 3.7.2 Slant Range Correction

This correction is applied after the beam steerage 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 (80 m) 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/hwin.

## 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. Although to some extent location dependent about 240 bursts, 120 either side of the 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 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.

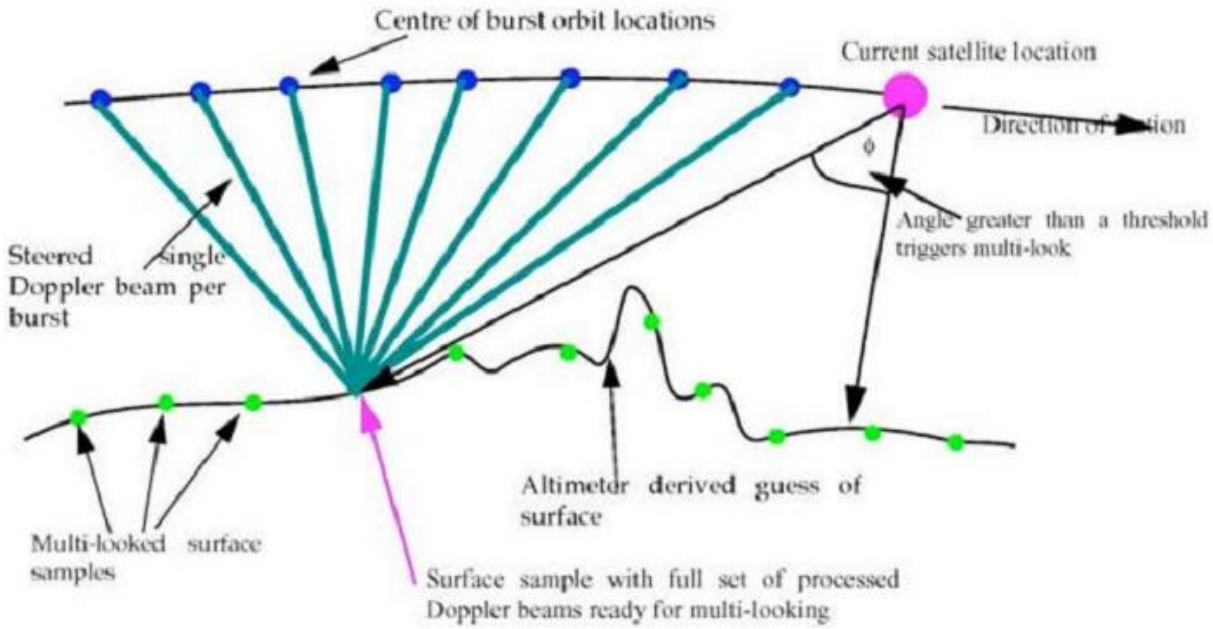
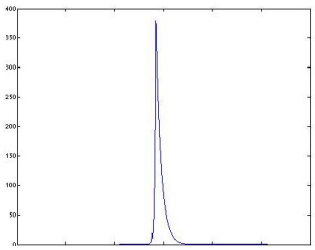
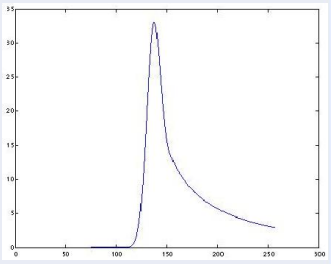
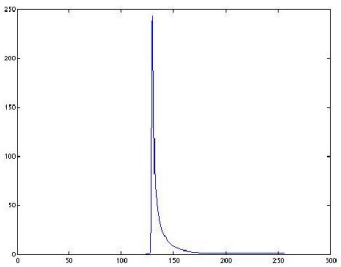



**Figure 11: Schematic of bursts, the fan of Doppler beams, ground points and multi-looks**

## 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 specular shapes it was suggested 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 subsequent peak was substantially (factor of 2) larger in amplitude. Of course, in retrospection, 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.

| Retracker # | Parameters fitted | Description |
|---|---|---|
| Simple waveform shapes | | |
| 1 | | Specular: (still water)  |
| 2 | | "Ocean-like": (ruffled water)  |
| 3 | | "Ocean-like" with fall away in power at high # bins: (ruffled water)  |

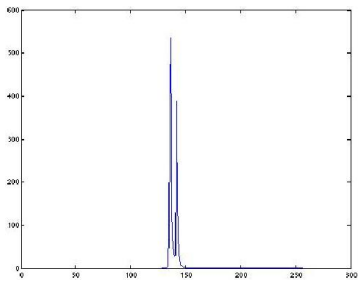
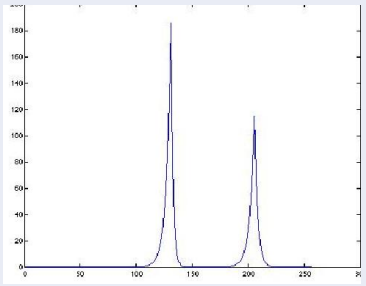
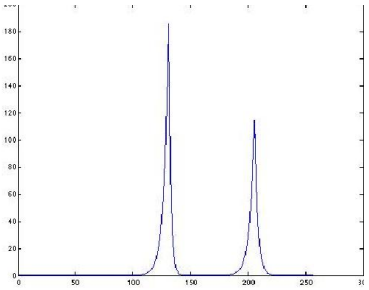
| Retracker # | Parameters Fitted | Description |
|---|---|---|
| Double peak waveform shapes | | |
| 4 | | Twin specular peaks (strong returns off two still water Patches)  |
| 5 | | Retracker type 2 with additional specular peak (ruffled and still water)  |
| 6 | | Retracker type 3 With additional Specular peak (ruffled water)  |

**Figure 12: Empirical retrackers. X-axis is gate number (1-256), Y-axis is power amplitude**

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

$$NRE = \sum_{i=1}^{256} (P_i^{obs} - P_i^{\mathrm{mod}})^2 \Big/ \sum_{i=1}^{256} (P_i^{obs})^2 \qquad (16)$$

was chosen. In Eq. (16) $P_i^{obs}$, $P_i^{\mathrm{mod}}$ 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. NRE in Eq. (16) is the sum of the residual power squared normalized with respect to the total waveform power squared. The 256 samples are derived from zero padding the original 128 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 cosine weighted with unit weight for the central burst beam with minimum angle to the ground-point decreasing as the beam angle increase for bursts before and after this central beam, via

$$wt(n) = 0.54 - 0.46\cos(\pi n / \mathrm{N})$$

where N-1 is the number of beams before or after the burst of minimum beam angle and $n = 0, \pm 1, \pm 2, \ldots . \pm (N-1)$. Thus, if N=3, wt= 0.31, 0.77, 1.0, 0.77, 0.31 for the 5 beams centred on the burst with minimum beam angle.

Preliminary analyses over the Mekong do suggest that a very low number of multi-looks are preferred, as little as N=3, that is 2 either side of the burst echo beam direction closest to the nadir. 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 (Figure 13). Figure 14 shows a pass where a selection of multi-look constructions has been retracked. The thicker green curve corresponding to N=3 is the preferred choice. This pass crosses the Mekong in low flow where the heights of the river banks are clearly evident some 10m above the water level. For the larger values of N, the number of acceptable heights is 3 which increases to 6 or 7 for the low values of N. This is in practice an

extreme case and for most passes the choice of N is not significant. Even here, the heights attributable to the water surface are in excellent agreement. Interestingly, N=2 gave the most river height values close to 101.6 m in geoid height (Figure 14).

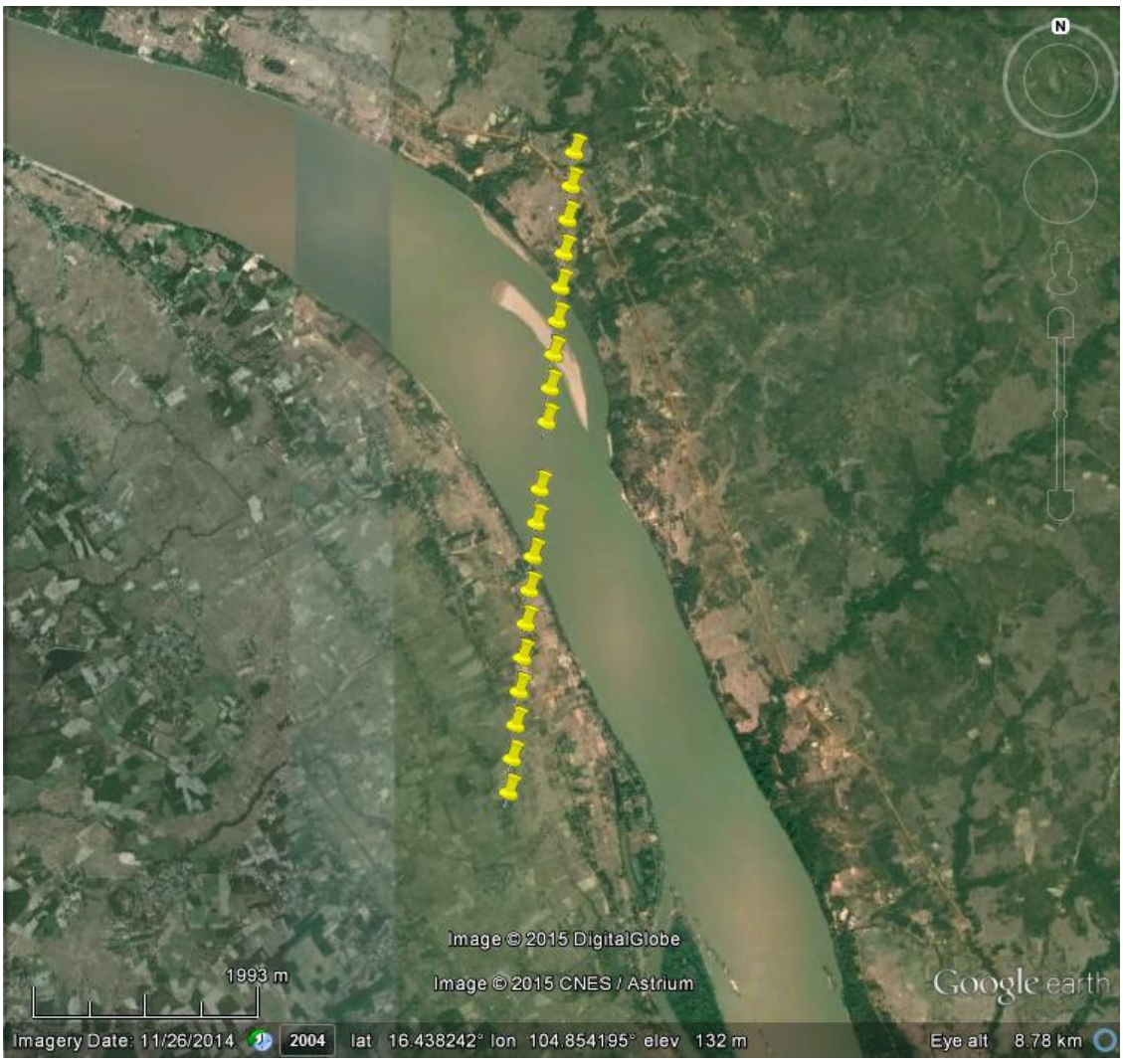



**Figure 13: Google Earth image of 19 Apr 2011 Mekong crossing. Satellite ground track N-S.**

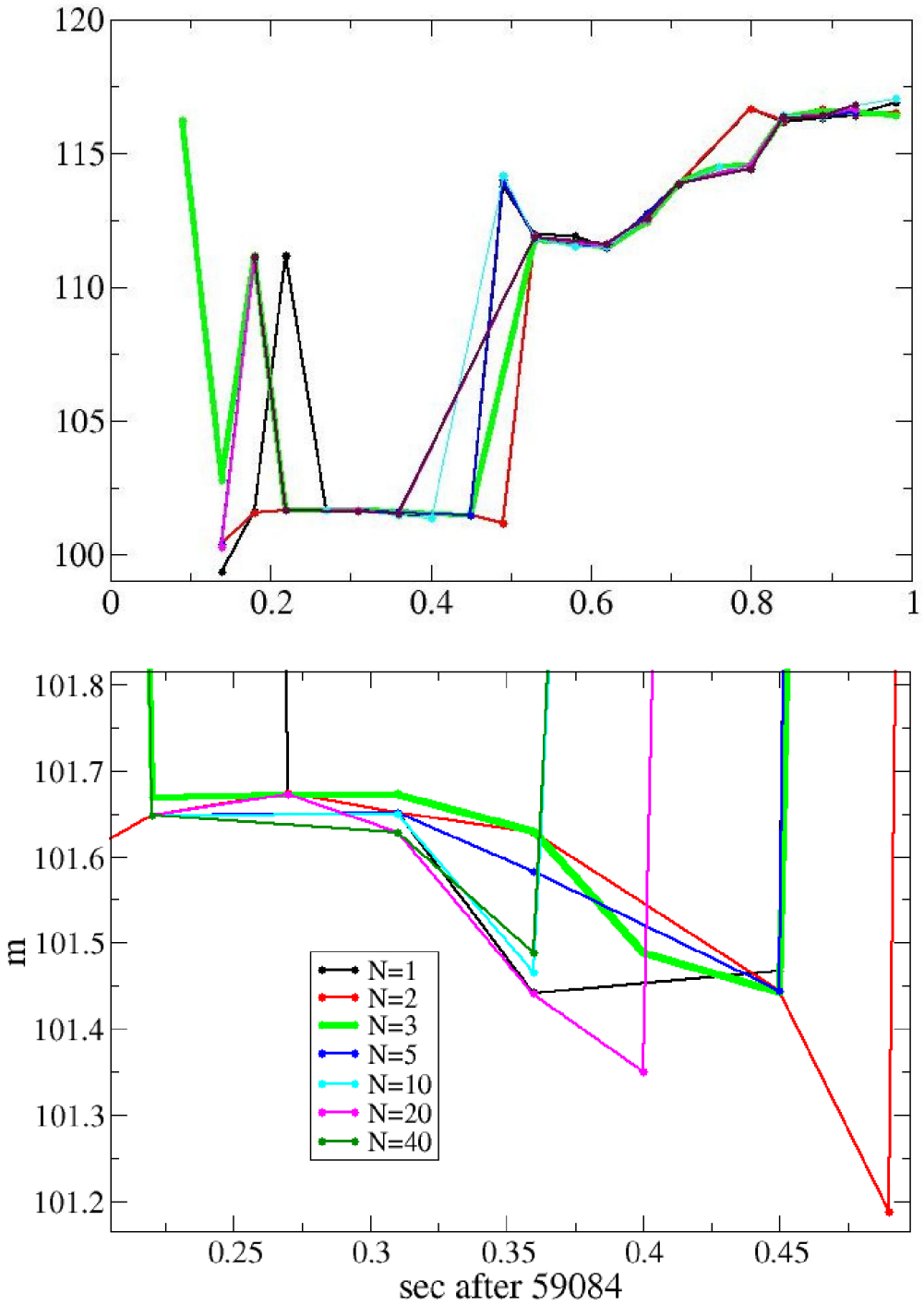



**Figure 14: Geoid heights across the Mekong for 19 Apr 2011 (day 109 in year) for various number of echoes in multi-look. Upper plot full profile; lower plot a zoomed version of the river surface.**

Analysis was also made across the Tonle Sap lake near its full extent (Figure 15) for the pass on 3 Dec 2011 (Figure 16).

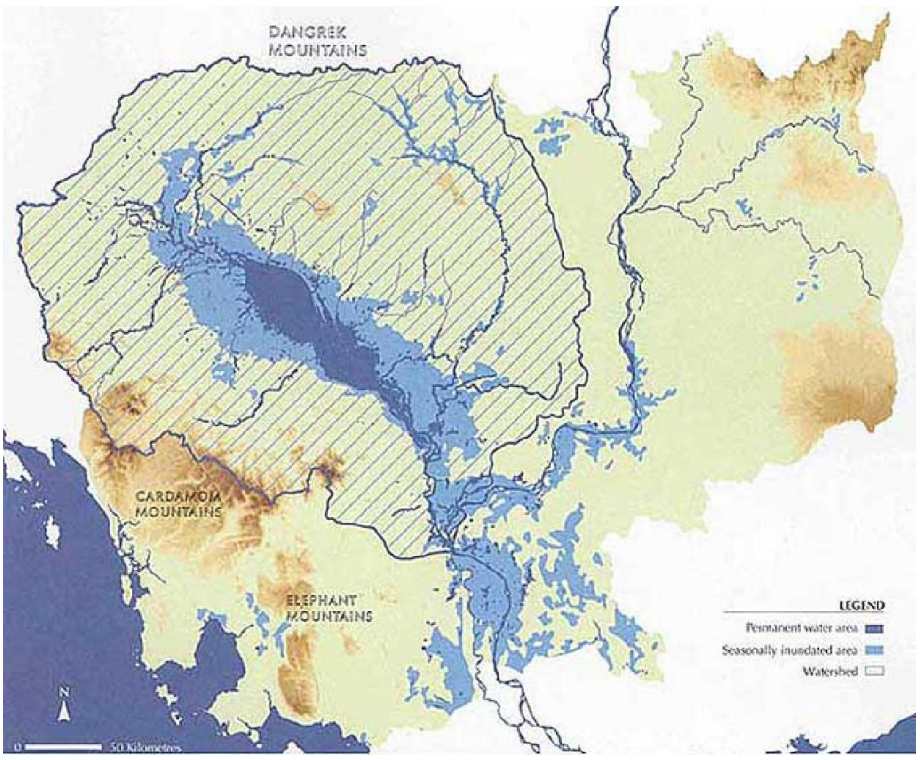



**Figure 15: Seasonal variation in extent of Tonle Sap**

For the pass on 3 Dec 2011 a total of 68 ground points 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 geoid heights were fitted with a quadratic to remove the geoid giving residuals of fit that were taken to represent the scatter in the estimated water heights. Figure 17 and Figure 18 plot the geoid heights derived using the empirical retrackers and the OCOG Threshold retracker (threshold =0.75) for the pass. The degree of fit is summarized in Table 1. That table shows for both the OCOG/Threshold and empirical retrackers the minimum sum of the squares of the errors is for N = 10. The empirical retrackers (Figure 12) appear more sensitive to N but the best fit values are near identical. Table 2 gives the number of waveforms retracked with the single peak retrackers (1-3) and the dual peak retrackers (4-6). The total numbers of single and dual peaks is also given. It is noted that the number of dual peaks reduces as the

number N increases, that is as more looks are included in the multi-look waveform. Further, the number of waveforms that were retracked best with retracker type 1 (single peak retracker) increased for N=40 and N=60. 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.

**Figure 16: Points across Tonle Sap (3 Dec 2011)**

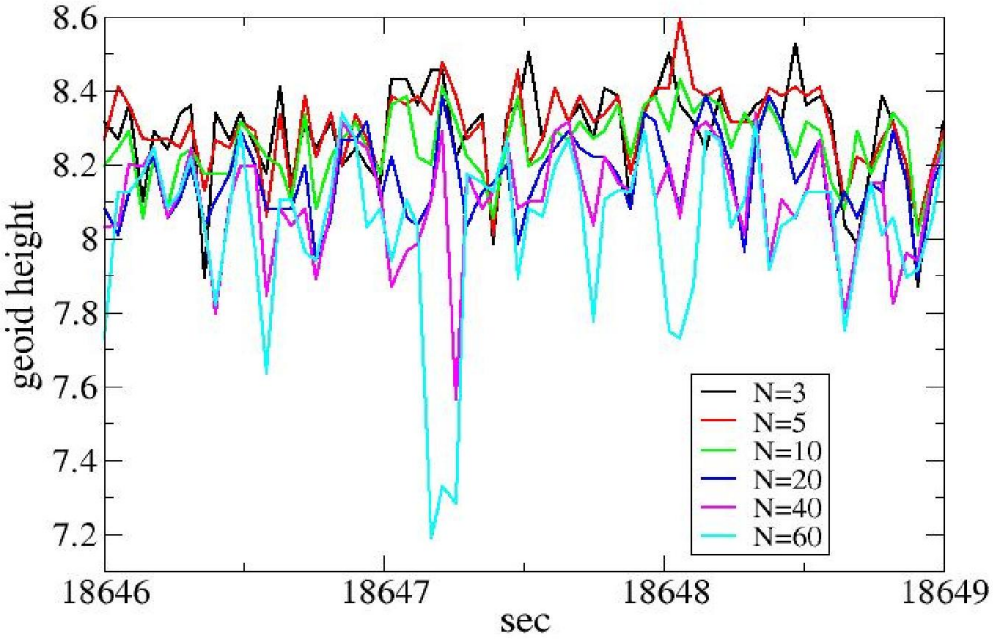



**Figure 17: Geoid height measurements across Tonle Sap (3 Dec 2011) with various multi-look combinations. Empirical retrackers used (Figure 12).**

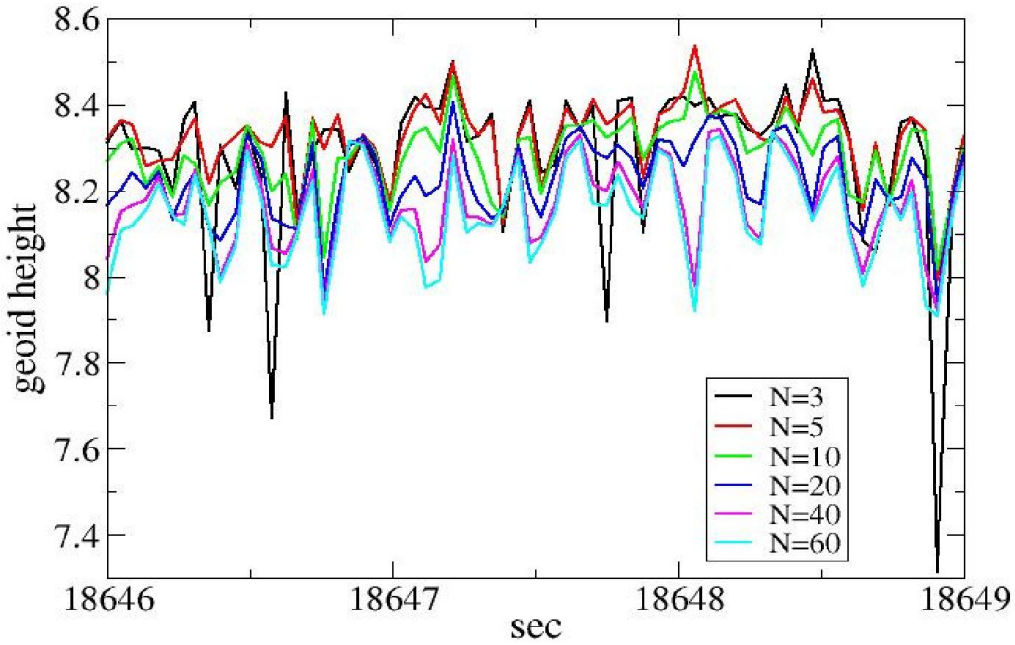



**Figure 18: Geoid height measurements across Tonle Sap (3 Dec 2011) with various multi-look combinations. OCOG/Threshold retracker used.**

| Multi-look N | Sum of Square of the errors: Empirical retrackers ($m^2$) | Sum of Square of the errors: OCOG/Threhold ($m^2$) |
|---|---|---|
| 3 | 1.07 | 2.19 |
| 5 | 0.69 | 0.50 |
| 10 | 0.48 | 0.46 |
| 20 | 0.74 | 0.51 |
| 40 | 1.46 | 0.67 |
| 60 | 3.4 | 0.79 |

**Table 1: Statistics of fit for pass across Tonle Sap.**

| | **Empirical tracker type** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **Single Peak** | **Dual Peak** |
| **N=3** | 2 | 7 | 20 | 17 | 11 | 11 | 29 | 39 |
| **N=5** | 1 | 7 | 33 | 9 | 4 | 7 | 48 | 20 |
| **N=10** | 0 | 8 | 33 | 13 | 6 | 8 | 41 | 27 |
| **N=20** | 0 | 16 | 24 | 20 | 5 | 3 | 40 | 28 |
| **N=40** | 37 | 10 | 9 | 10 | 1 | 1 | 56 | 12 |
| **N=60** | 56 | 5 | 2 | 4 | 1 | 0 | 63 | 5 |

**Table 2: 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 4-6. A total of 68 waveforms were retracked.**

# 4 SAR FBR Analysis: Summary

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

- The failure of the empirical retracker is evident for N≥60 in Figure 19 - Figure 25 for the pass across Tonle Sap. These points are plotted on a Google Earth image as Figure 26. The sequence in Figure 19 and Figure 20 show that the low values of N pick up a clear dual peak but that the peaks become blurred and tend to merge for the higher values N=60 and particularly N=120. The inference is that the reflectors for the multi-looks are not of the same characteristics for all the steered beams. This differs to say 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. In other words, the reflectors in one burst are near indistinguishable from previous and subsequent bursts. Over inland waters 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. The conclusion is that a low value of N is preferred. For narrow rivers a value of N=3, 4 or 5 is sufficient to provide a well-defined waveform for retracking increasing slightly to N=10 over larger water bodies such as small-medium lakes. Of course over the largest lakes a higher value of say N=40 is also acceptable.
- In terms of the retracker the OCOG retracker with threshold = 0.75 has proven to be as precise as the empirical retrackers employed. Furthermore, the simplicity of the Threshold retracker is also attractive. That is not to say that more advanced retrackers would not 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).
- Ongoing studies are extending the analysis to the SAR FBR L1A data for the Ganges delta region in Bangladesh and the Amazon with validation against gauge and Jason-2.
- A preliminary look at SARin data is also underway.
- The EGM96 geoid model used in this report will be replaced in the ATBD by a higher resolution geoid model such as EGM08 or a GRACE/GOCE model.

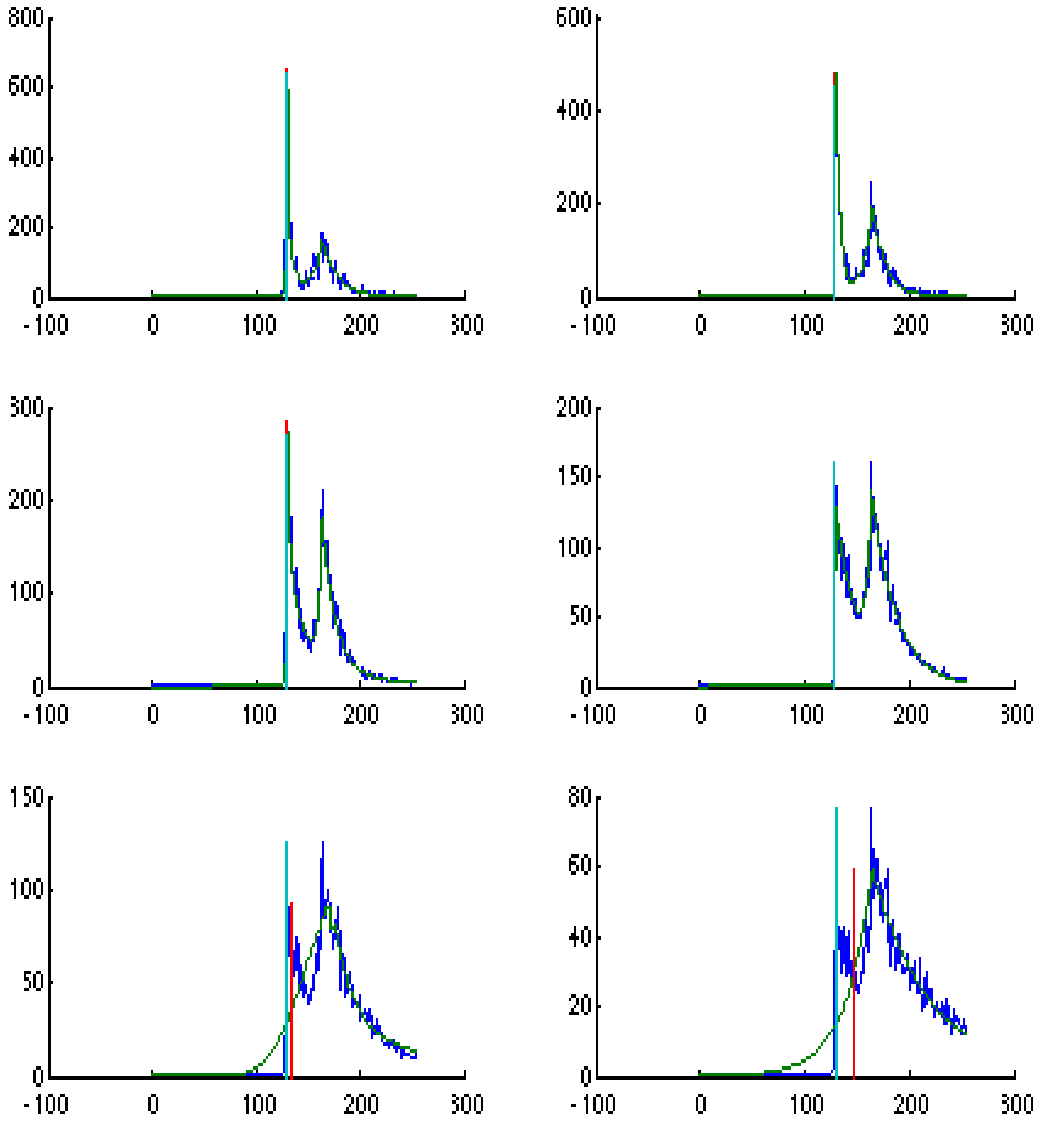



**Figure 19: Waveforms (i=86, t=18645.69 s, lat= 13.0439˚, lon=104.0775˚) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. . X-axis is gate number (1-256), Y-axis is power amplitude.**

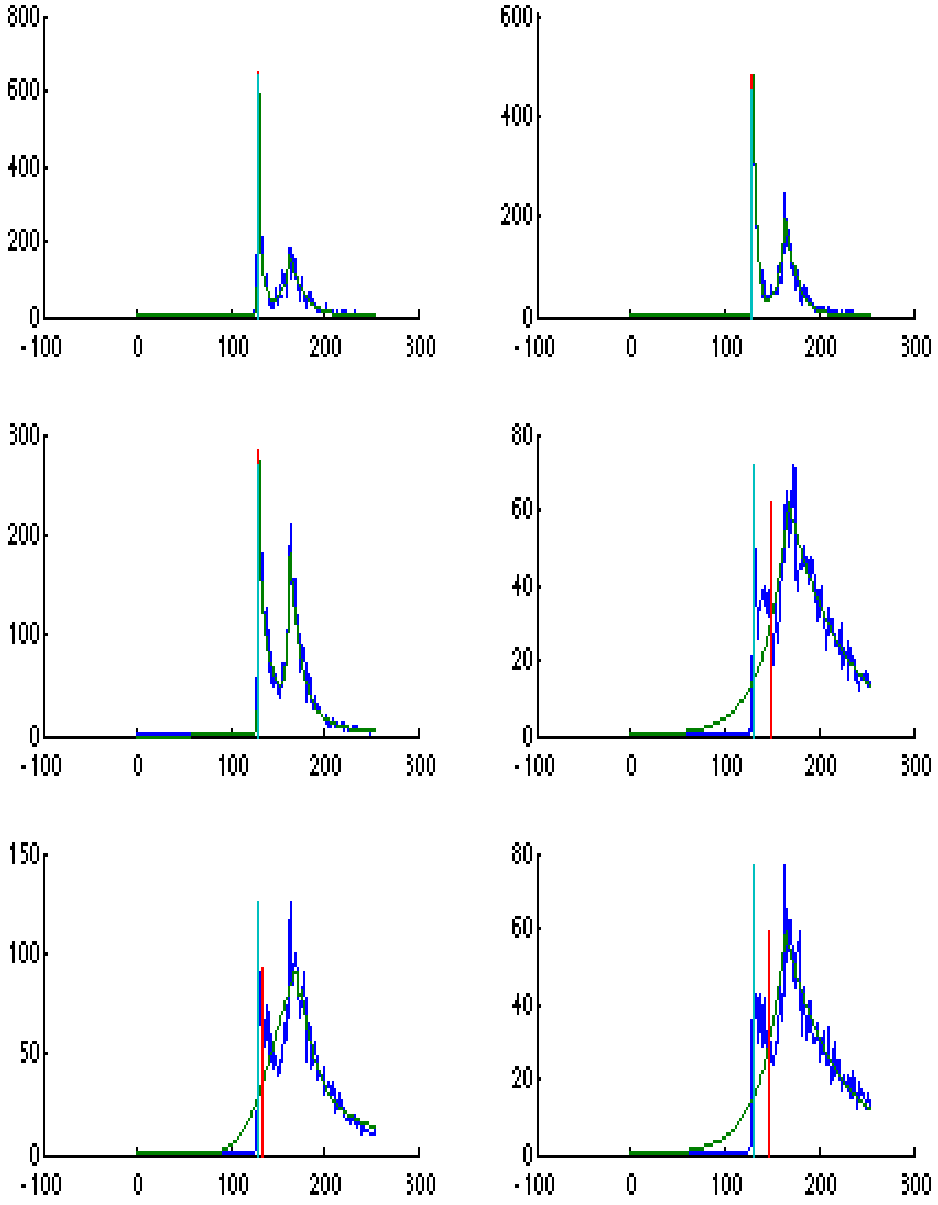


**Figure 20: Waveforms (i=87, t=18645.73 s, lat= 13.0411˚, lon=104.0772˚) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.**

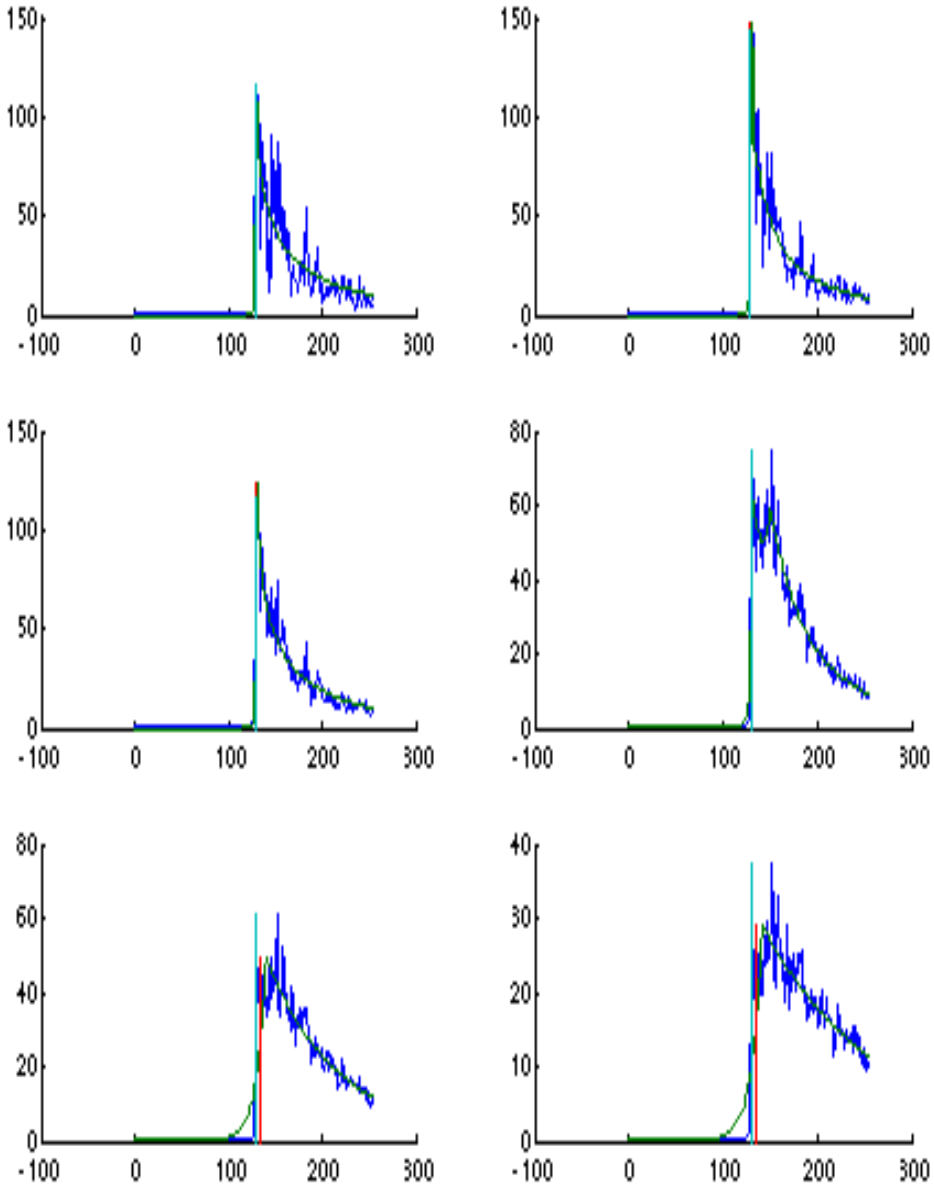


**Figure 21: Waveforms (i=118, t=18647.12 s, lat= 12.9568˚, lon=104.0683˚) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.**

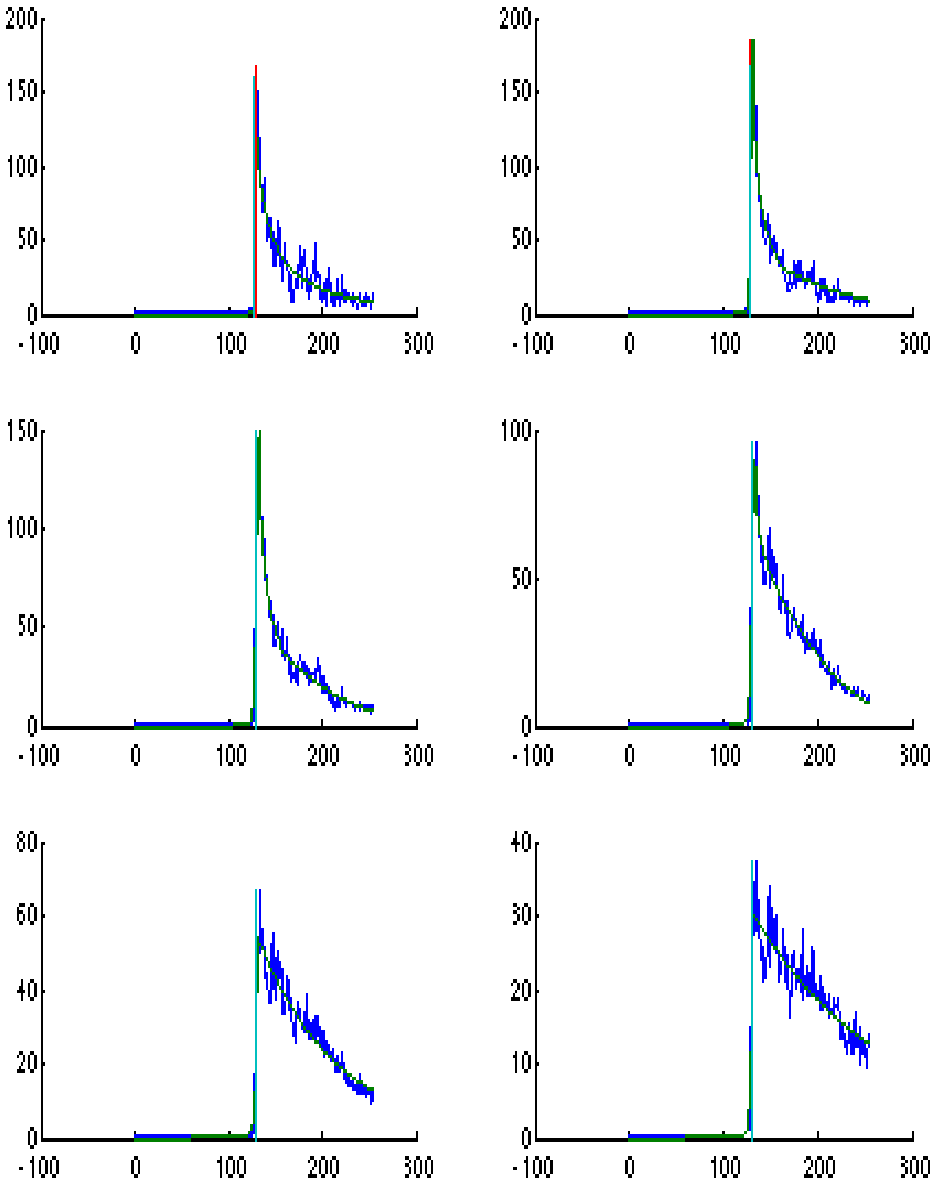


**Figure 22: Waveforms (i=119, t=18647.17 s, lat= 12.9541˚, lon=104.0680˚) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.**

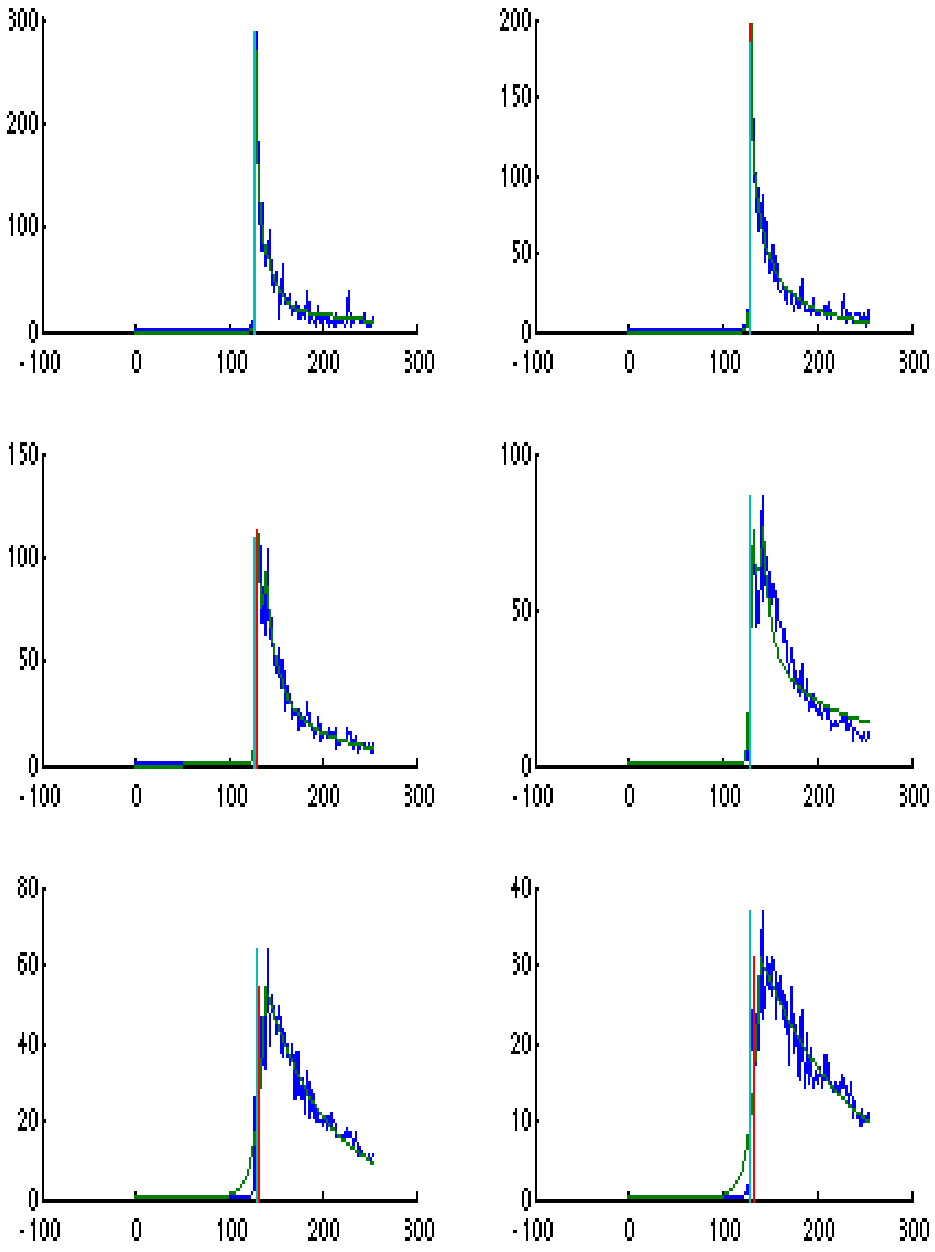



**Figure 23: Waveforms (i=120, t=18647.21 s, lat= 12.9514˚, lon=104.0677˚) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.**

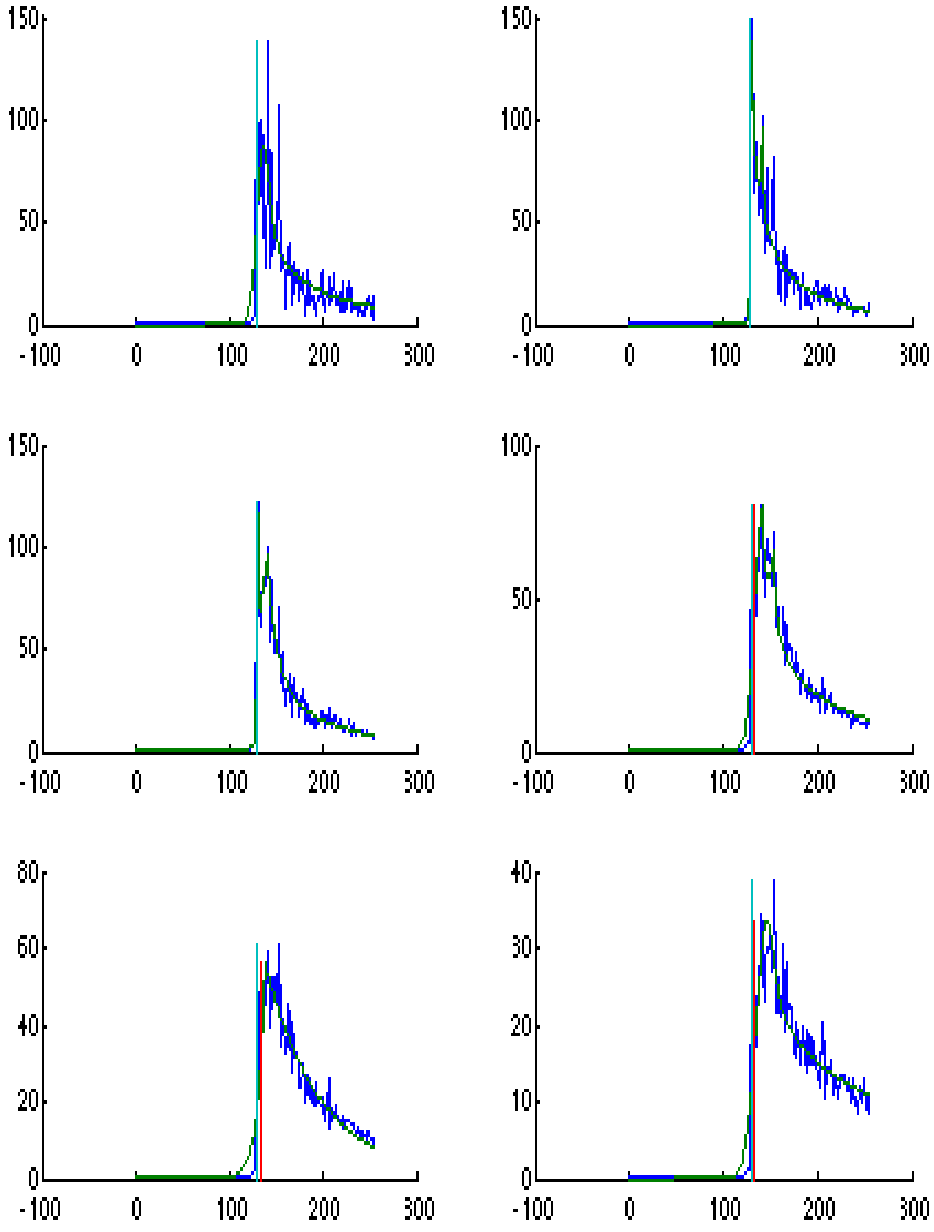


**Figure 24: Waveforms (i=121, t=18647.26 s, lat= 12.9486˚, lon=104.0674˚) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.**

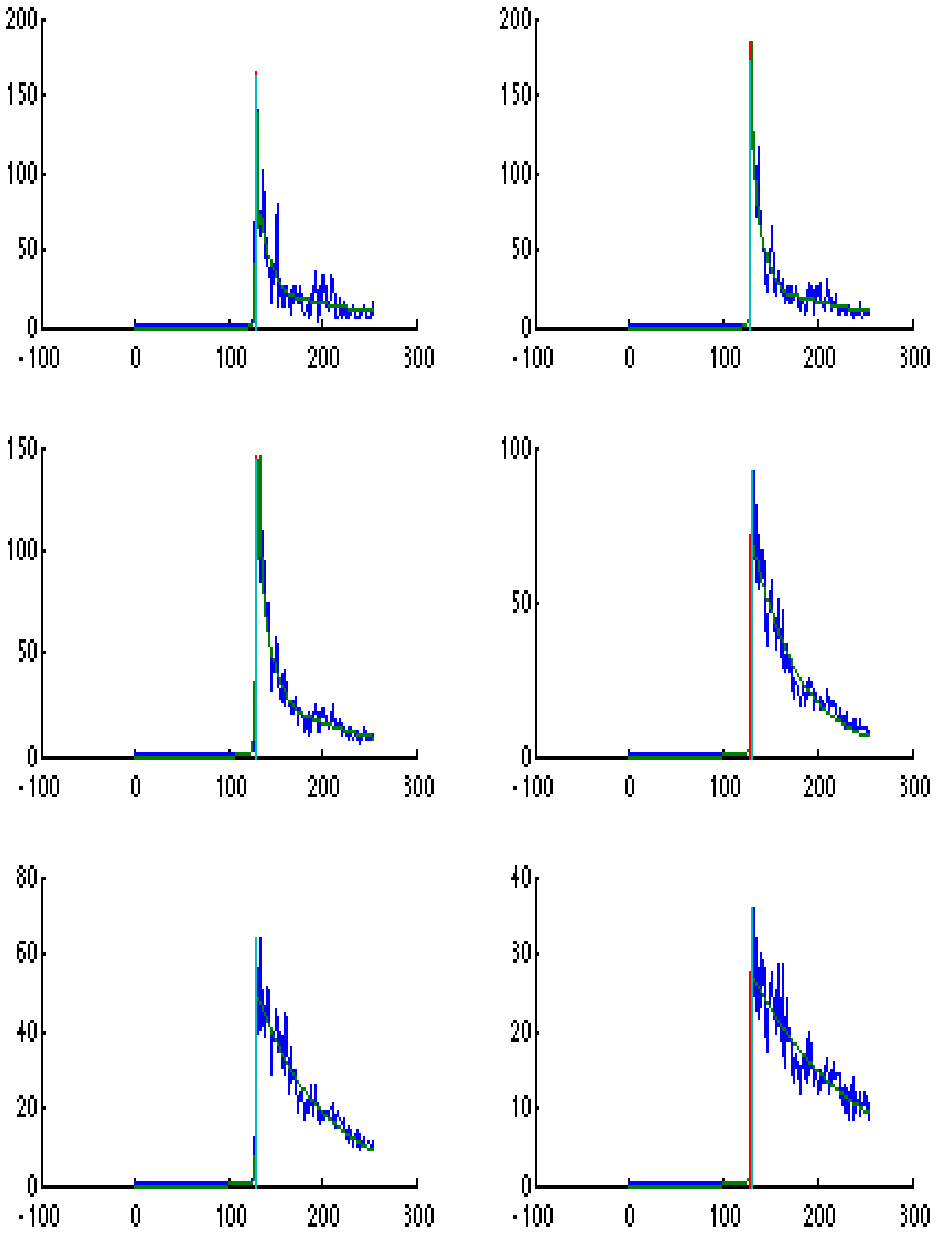


**Figure 25: Waveforms (i=122, t=18647.30 s, lat= 12.9459˚, lon=104.0671˚) with empirical retracker (green curve) with retracked bin given by red line; cyan line shows OCOG/Threshold retracked bin. From left to right: row 1 N=5, N=10; row2 N=20, N=40; row 3 N=60, N=110. X-axis is gate number (1-256), Y-axis is power amplitude.**

**Figure 26: Geolocation of points i=86 and i=118 corresponding to waveforms of Figure 19 - Figure 25. 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.**

# 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 and18°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 27 and Figure 28).

Table 3 shows the Landsat 8 bands that are available (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 bands10 and 11 a resolution of 100m.

**Table 3: 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 29 and Figure 30)

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 31 and Figure 32.

4. Combined high resolution and approximate mask

A final mask of the Mekong River (Figure 33 and Figure 34) was obtained by combining the approximate mask (Figure 29 and Figure 30) with the water mask (Figure 31 and Figure 32). 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 27: Six Landsat images of the Mekong River between 12°N and 18°N**

**Figure 28: Landsat image from 12/11/2014 of the Mekong River between 14.2°N and 14.8°N**

**Figure 29: Approximate mask of the Mekong River between 12°N and 18°N**

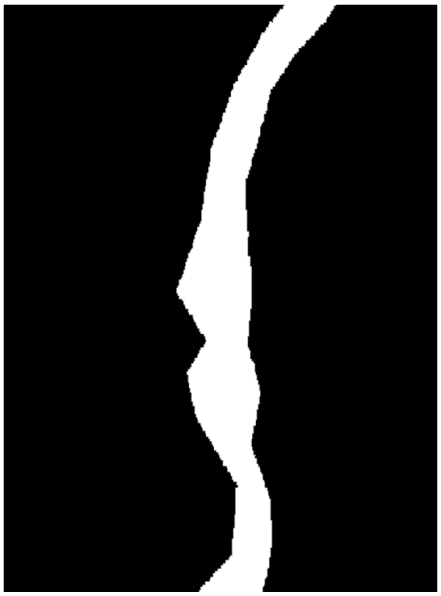


**Figure 30: Approximate mask of the Mekong River between 14.2°N and 14.8°N**

**Figure 31: Water mask of the Mekong River between 12°N and 18°N.**

**Figure 32: Water mask of the Mekong River between 14.2°N and 14.8°N**

**Figure 33: Final mask of the Mekong River between 12°N and 18°N.**

**Figure 34: 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).

# 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 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, 2015).

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 $km^2$ 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 is 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.

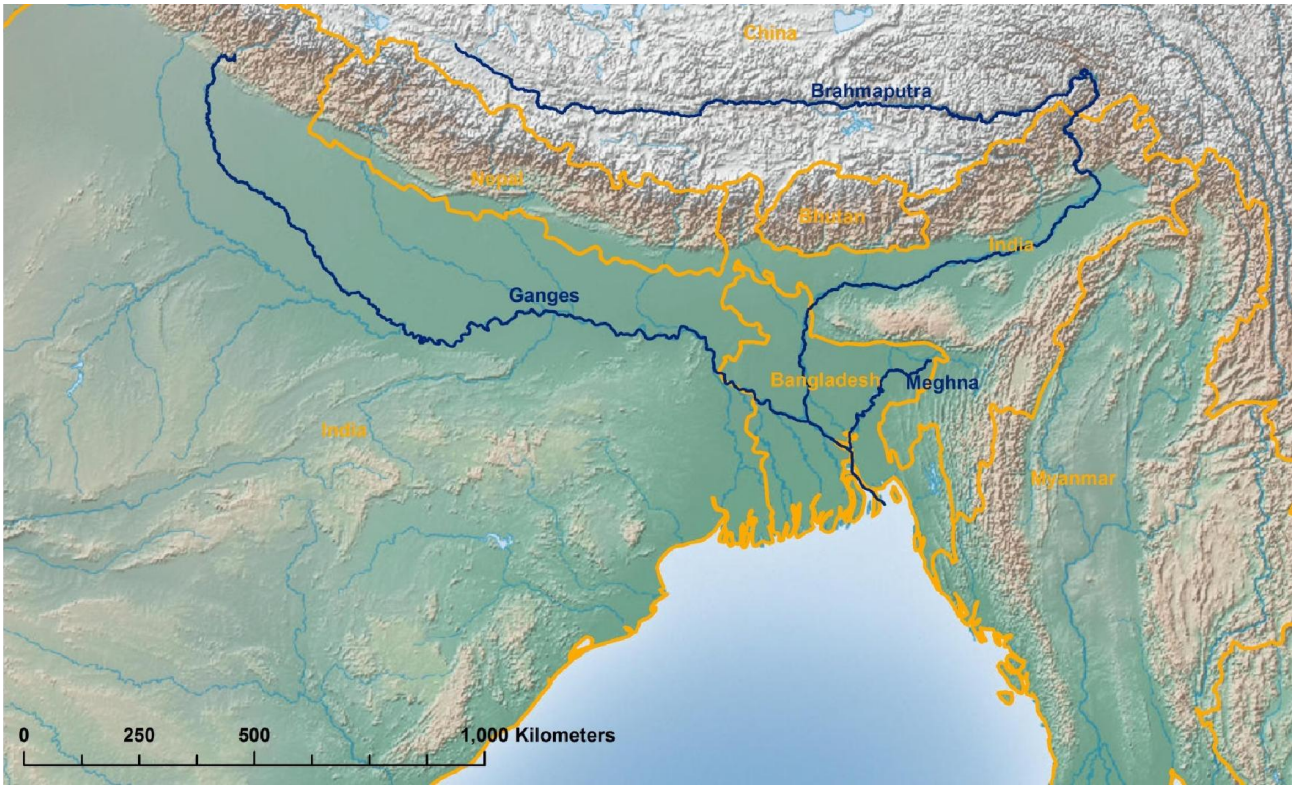



**Figure 35: Map of the three main rivers draining into the Gulf of Bengal through Bangladesh**

Bangladesh, a low lying country at the Gulf of Bengal in the estuary region of the three large rivers Ganges, Brahmaputra and Meghna (Figure 35) 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 (Biancamaria et al., 2011b).

A hydrologic-hydrodynamic model for the entire Brahmaputra basin was set up. The course of the river Brahmaputra 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 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. 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. In the Assam valley, the river is braided and the width of the water body during the high flow season is typically around 10 km, exceeding 2 km in most places.

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) or Song et al. (2015). This work however is one of the first applications of CryoSat-2 altimetry over rivers, showing how its special drifting orbit can be used.

## 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 dense SARin mode of CryoSat-2. SAR mode offers an along track resolution of 300m, whereas the SARin mode is adopted to deal with areas presenting a high surface slope. 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 water (river) or over land surface. Often, the backscatter values (sigma0) 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. Hence, the distinction between CryoSat-2 datapoints 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 36), which requires dynamic river masks.

**Figure 36: Landsat 7 image 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.

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 3) and near-infrared (Band 5 of Table 3) spectral regions as NDNI = (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 from 2010 to 2013. However, even using the 32-day composites no sufficient coverage could be achieved during the high-flow season in summer due to cloud cover. Thus, minimum water extent masks were created by merging all available images from one year.

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

Figure 37 shows the filtering and projecting process applied to the CryoSat-2 level 2 data points. Only datapoints 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 on 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.

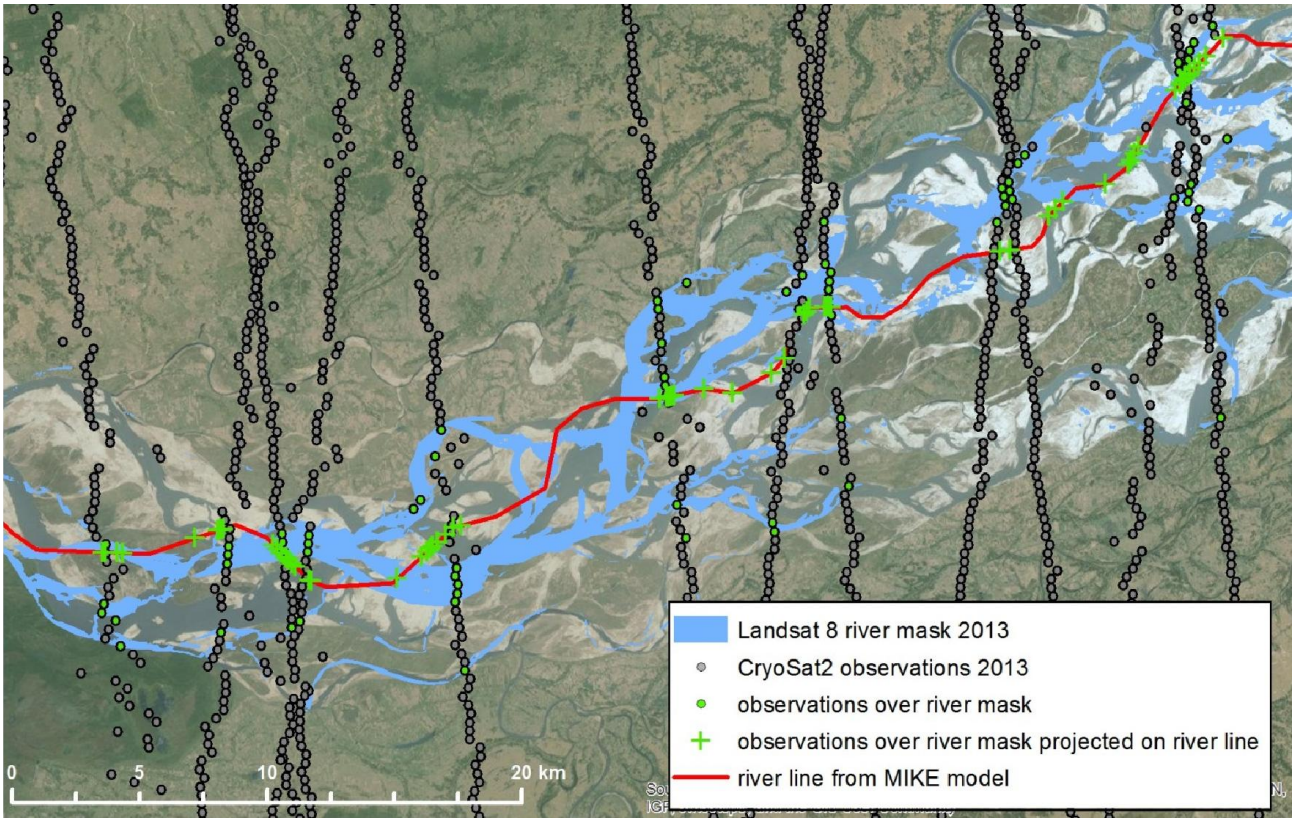



**Figure 37: Section of the Brahmaputra in the Assam valley showing the Landsat river mask, the CryoSat-2 observations and their mapping to the 1D river model, all for 2013.**

## 6.3 Hydrologic hydrodynamic model

In cooperation with DHI (Danish Hydraulic Institute, www.dhi.dk), a model of the entire Brahmaputra basin was set up in the DHI MIKE 11 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 river flow in MIKE 11 is modelled using a 1D dynamic wave routing based on the Saint Venant equations for unsteady flow MIKE (MIKE by DHI, 2009). The one-dimensional model space is discretized into finite intervals, which are separated by alternating Q and h points (Figure 38). 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 exact locations of the individual CryoSat-2 height readings.

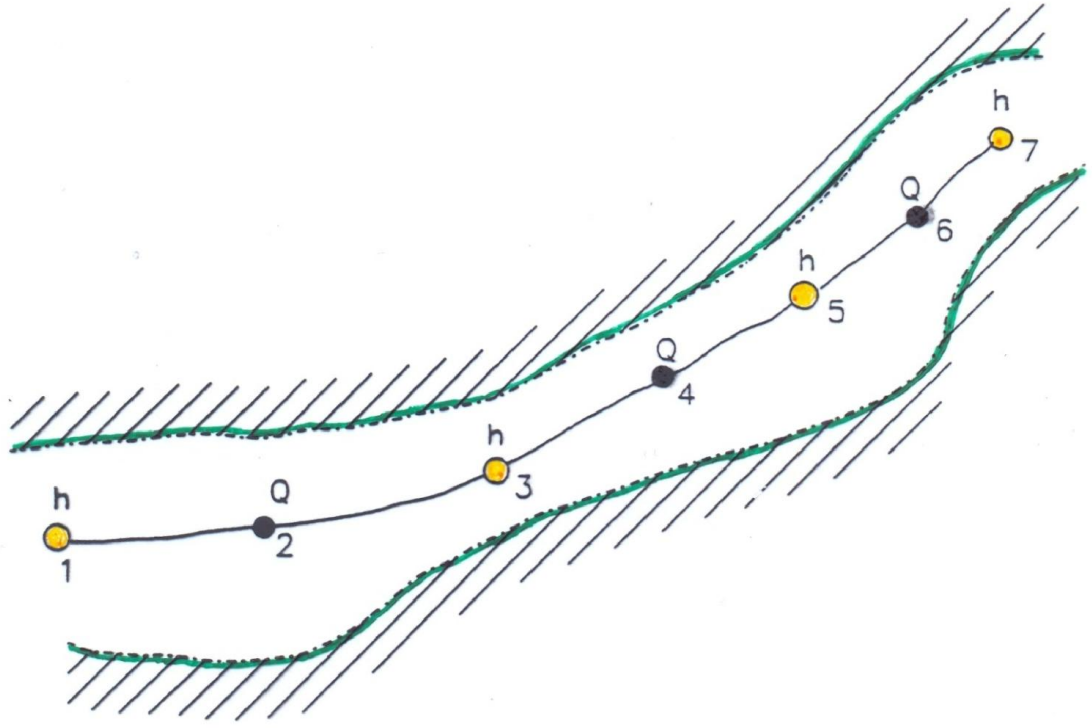



**Figure 38: The Mike-11 computational grid**

Figure 39 provides an overview of the 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 39. 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. For some of the NAM subcatchments in the larger Ganges Brahmaputra model region (Figure 35), 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. No calibration subcatchments are located in the Brahmaputra itself. Furthermore, discharge observations from the station Bahadurabad on the Brahmaputra (see Figure 39), 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 was based on data from the APHRODITE's Water Resources project (APHRODITE's Water Resources, 2014). 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 1.1. Such a bias in the TRMM precipiation 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.

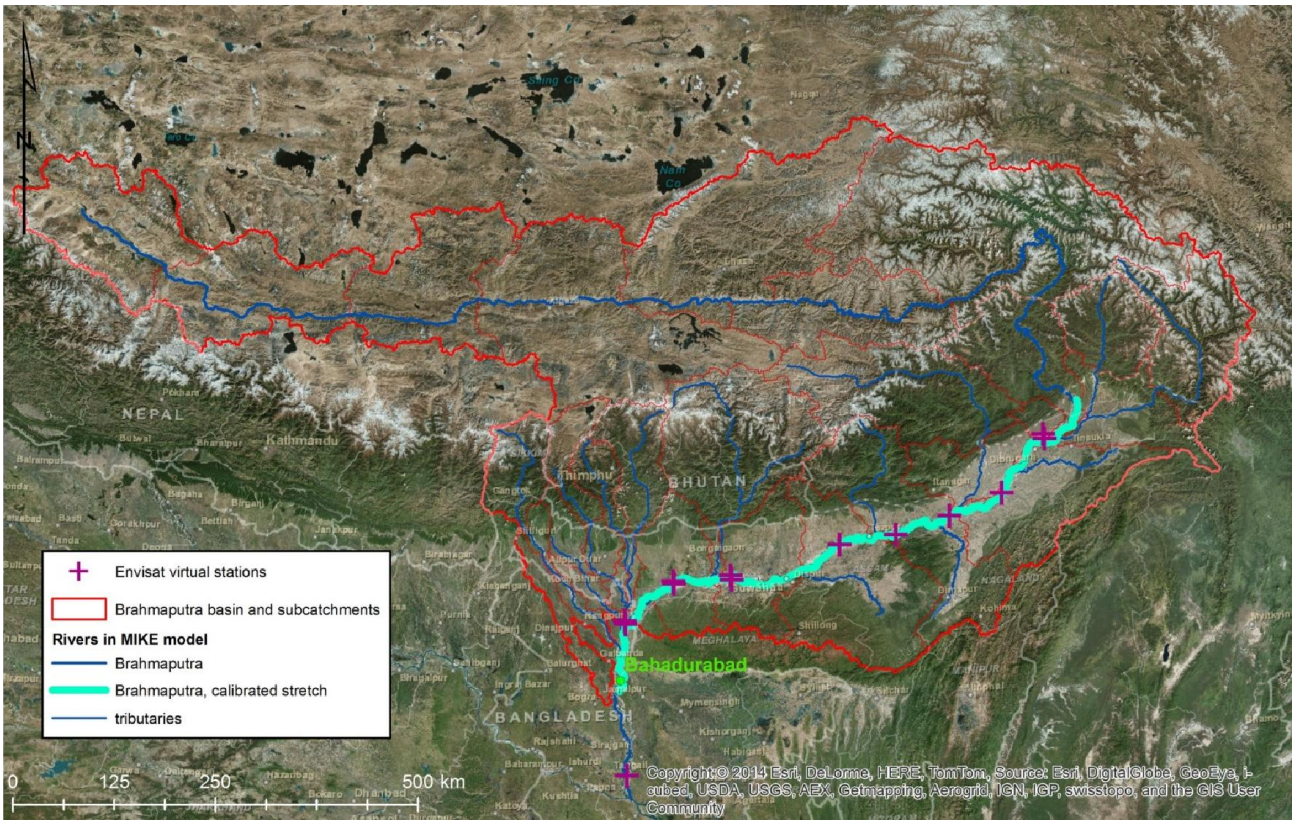



**Figure 39: The Brahmaputra basin model. The main calibration station, Bahadurabad in Bangladesh is shown. The water level calibrated part, i.e. the Assam valley together with the Envisat virtual stations is displayed (see next section).**

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 ($sm^{-1/3}$), Q is the river discharge ($m^3/s$), A is the flow cross-sectional area ($m^2$) 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.

## 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. 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. A different approach 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 elevations extracted from the SRTM DEM and some rough estimates about bathymetry as a starting point for the calibration.

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 39) covering the years 2002 to 2010 were used. 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 40:

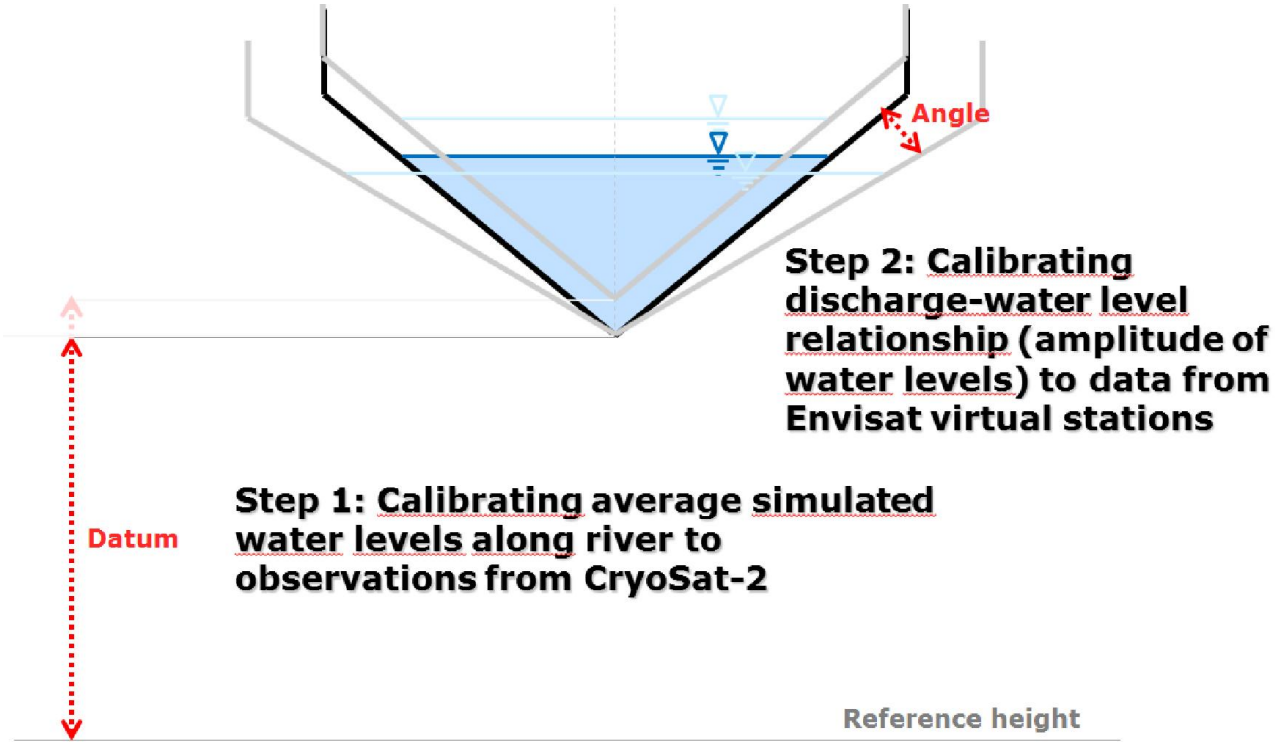



**Figure 40: Sketch of the two-step cross section calibration. MIKE 11 assumes vertical walls for cross sections outside their defined limits.**

Both calibration steps were performed by coupling the MIKE 11 model to MATLAB and using MATLAB's genetic algorithms to find the optimal parameter combination. Table 4 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 18 model cross sections in the Assam Valley, i.e. 18 parameters were fitted in total in this calibration run. In Step 2, the parameters to be optimized are the opening angles of 21 triangular cross sections, i.e. 21 parameters were fitted in total in this calibration run.

**Table 4: Decision variables and objective functions of the genetic algorithm used for the two-step water level calibration**

| | Decision variables | Objective function |
|---|---|---|
| **Step 1: Calibrating absolute water levels to CryoSat-2 observations** | Cross section datums | RMSE between average simulated water levels and CryoSat-2 observations for the period 2010 to 2013 |
| **Step 2: Calibrating amplitude of water levels to Envisat VS data** | Cross sections' triangular shape (determined by angle) | RMSE between time differences of simulated water levels and Envisat VS data for the period 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., 2016)

$$x_t = M(x_{t-1}, p, f_t) + v_t \tag{17}$$

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

$$y_t = H(x_t) + u_t \tag{18}$$

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

$$x_t^a = x_t^f + K_t(y_t - H(x_t^f)) \tag{19}$$

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 (19) and the propagation of the model's uncertainties. The Kalman Gain $K$ is determined based on the state error covariance matrix

$$K = P^f H^T (HP^f H^T + R)^{-1} \tag{20}$$

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

$$P^a = (I - KH)P^f \tag{21}$$

where $I$ is the identity matrix.

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 is usually propagated by the ensemble by

$$P = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - x)(X_i - x)^T = \frac{1}{m-1}AA^T \tag{22}$$

where $X$ ensemble of model states, $X = [X_1, \ldots, X_m]$

$m$ ensemble size

$x$ ensemble mean, $x = \frac{1}{m}\sum_{i=1}^{m} X_i$

$A$ ensemble of anomalies, $A = [A_1, \ldots, A_m]$ , and $A_i = X_i - x$

The update then is

$$X^a{}_i = X^f{}_i + K(d - HX^f{}_i) \tag{23}$$

If the system state is represented by the ensemble mean, the average of (23) is the analysis equation (19). This does however not apply to the covariance equation (21), and can lead to an ensemble collapse, i.e. too small 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 $T$

$$A^a = A^f T \tag{24}$$

where $T$ must satisfy

$$T = T^S U \tag{25}$$

where $U$ is an arbitrary orthonormal matrix, and

$$T^S = [I + \tfrac{1}{m-1}(HA^f)^T R^{-1} HA^f]^{-1/2} \quad (26)$$

### 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 41. Configuration details are provided through a ASCII input file (PFS file). The framework provides different filtering algorithms such as EnKF and Ensemble Transform KF. 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 neighborhood of the measurement location. This is useful to avoid artifacts 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.

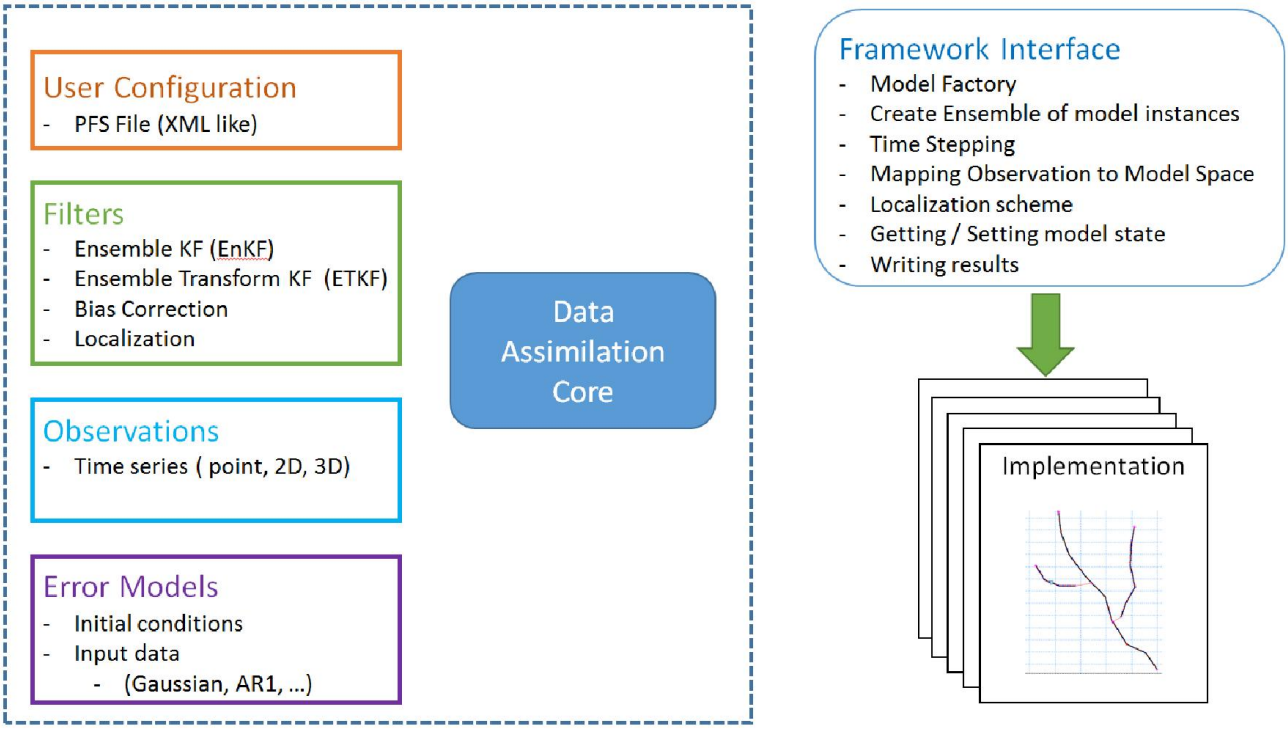



**Figure 41: Overview over the DHI Data Assimilation Framework. Source: Marc Ridler, DHI.**

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.

### 6.5.4 Processing of CryoSat-2 observations

After filtering over the river mask as described previously (see Figure 37), a total of 2794 CryoSat-2 observation points over the Assam valley from 2010 to 2013 remained. Those single observations are grouped within the DA framework to their mean value for each individual crossing of the Brahmaputra, which leaves 471 crossings. Finally, the observations are projected from their actual coordinates onto the 1D river line and assimilated to the model.

### 6.5.5 Observation error

A possible approach to estimate the observation error is to determine the standard deviation of the elevations of all points that are aggregated into one crossing. For our case however, a default standard deviation of the CryoSat-2 observations of 0.5 metres was assumed. The standard deviation of individual elevations in one crossing is in the range of 0.1 to 0.2 meters and seemed too low to give meaningful results in the data assimilation. This can be due to the fact that there may be a general bias for all observation points of an individual crossing which is ignored using this method.

### 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 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):

$$w_t = \delta w_{t-1} + \varepsilon_t$$

where $w_t$ runoff perturbation from catchment: $w_t = \frac{Q_{sim,t} - Q_{obs,t}}{Q_{obs,t}}$

$\delta$ 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 0.25.

## 6.6 Results and Discussion

### 6.6.1 Processing of CryoSat-2 data

Figure 42 shows the CryoSat-2 data after filtering over the Brahmaputra river mask and projecting it onto the model's river line (Figure 37). 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 meters from the SRTM data, about 20% of the data are discarded.

As mentioned, calibration and data assimilation were limited to the Assam valley, i.e. the river approximately after km 2000.

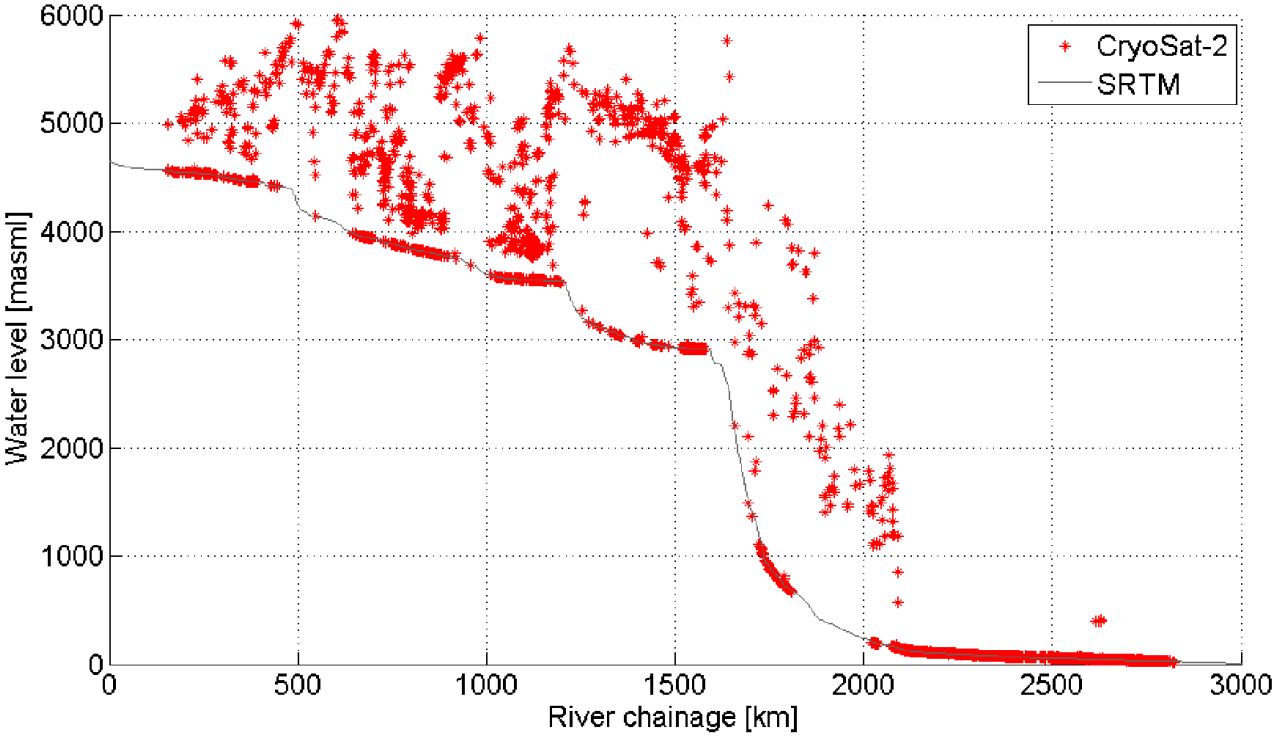



**Figure 42: CryoSat-2 data 2010 to 2013 after filtering and mapping to the model's river line in comparison with SRTM data. The Assam valley starts around chainage km 2000.**

### 6.6.2 Hydrologic model calibration

The main calibration objective was to calibrate the model's discharge at the station Bahadurabad in Bangladesh, close to the confluence of the Brahmaputra river with the Ganges. The results for the calibration period 2002 to 2007 can be seen in Figure 43. A good fit with a NSE of 0.91, a RMSE of 4955 $m^3/s$ and a mean bias of -7.5% could be obtained. The strong seasonality of the discharge, caused by the region's monsoon climate and snow melt in the Himalayas in early summer, is nicely reproduced by the model.

The validation/data assimilation period includes the years 2010 to 2013, as it starts with the availability of CryoSat-2 data. During that period, only discharge data for the high-flow season (April to October) could be obtained. The result for this period can be seen in Figure 44.

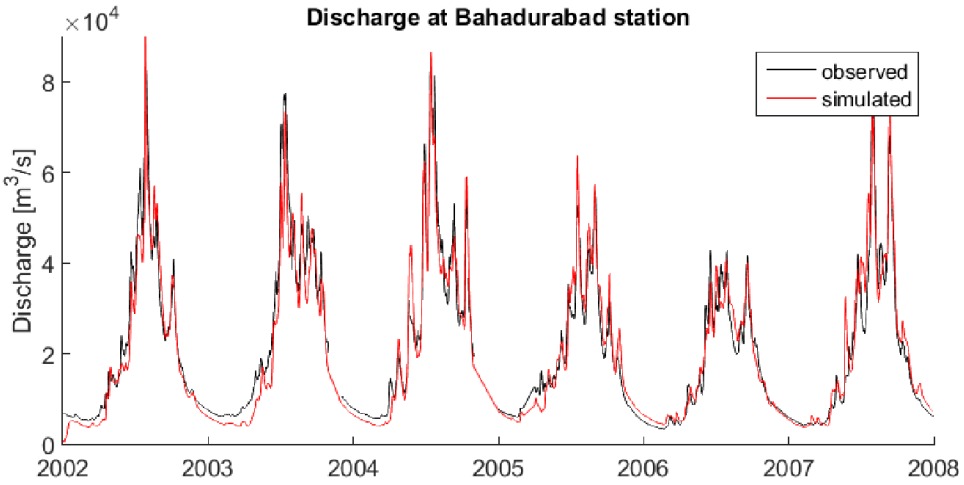



**Figure 43: Observed vs. simulated discharge for the Brahmaputra at Bahadurabad station for the calibration period 2002 - 2007**

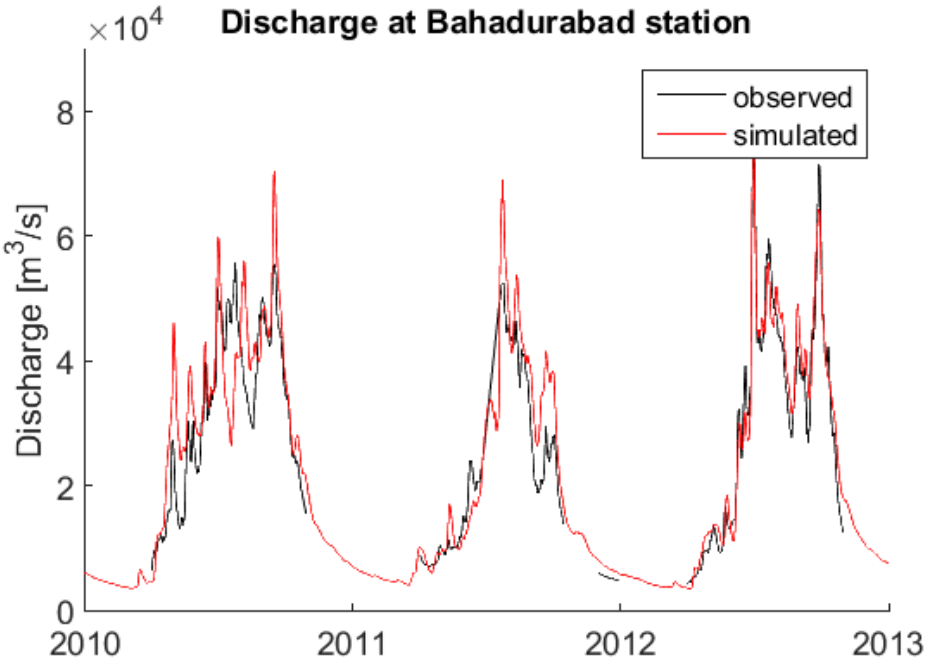



**Figure 44: Observed vs. simulated discharge for the Brahmaputra at Bahadurabad station for the data assimilation period 2010 – 2013. Note that observed data only exists for the high-flow periods.**

A summary of the results for both periods is presented in Table 5, showing minor decreases in the fit of the model's simulated discharge for the data assimilation period.

**Table 5: Number of observations and performance indicators for the station Bahadurabad for the full calibration period 2002 - 2007, the high-flow seasons of 2002 - 2007 and the high flow seasons of 2010 - 2013**

| | # observations | RMSE [m3/s] | Bias ($Q_{sim}$ - $Q_{obs}$)[-] | NSE [-] |
|---|---|---|---|---|
| 2002 – 2007 | 5914 | 4955 | -0.075 | 0.91 |
| 2002 – 2007 Apr – Oct | 3792 | 5471 | -0.109 | 0.90 |
| 2010 – 2013 Apr - Oct | 5156 | 6744 | 0.048 | 0.81 |

The results of the hydrological model for the calibration period 2002 to 2007 with a Nash–Sutcliffe coefficient of 0.91 are good, given the size of the model and the availability of forcing data, using freely available remote sensing only. The NSE quantifies model performance relative to a benchmark, which is the average of all available historical observations. An NSE value of 1 indicates the perfect model, an NSE value of 0 indicates performance equivalent to the benchmark and negative NSE indicates worse performance than the benchmark. For the data assimilation period 2010 to 2013 discharge observations at Bahadurabad are only available for the high-flow season. Comparing the model's performance in the two periods therefore has to be done based on data from April to October. The performance of the model in the calibration period stays almost the same – the NSE is only reduced to 0.90. In the validation/data assimilation period the NSE is decreased to 0.81, which is still a very good performance. To some extent, the reduced performance in the validation period could be due issues in the precipitation forcing or changes in the climatic conditions during those two periods. The RMSE increased by 23% between the calibration and the assimilation periods.

### 6.6.3 Cross section calibration

- Step 1:

  Figure 45 shows the results of step 1 of the water level calibration (Figure 40). For better visibility, the results are all shown in elevations relative to the reference model's cross section datums instead of absolute elevations. The reference model was run with cross section datums derived from the SRTM DEM. It can be seen that the average simulated water levels from the reference model do not accurately represent the CryoSat-2 observations. After calibrating the cross section datums – which resulted in datum adjustments of up to 4 meters – the simulated average water level follows the CryoSat-2 observations more closely. The calibration reduced the RMSE between average simulated water levels and CryoSat-2 observations from 3.04 metres for the reference model to 2.38 meters. The remaining deviation can mainly be explained by seasonal water level variations.

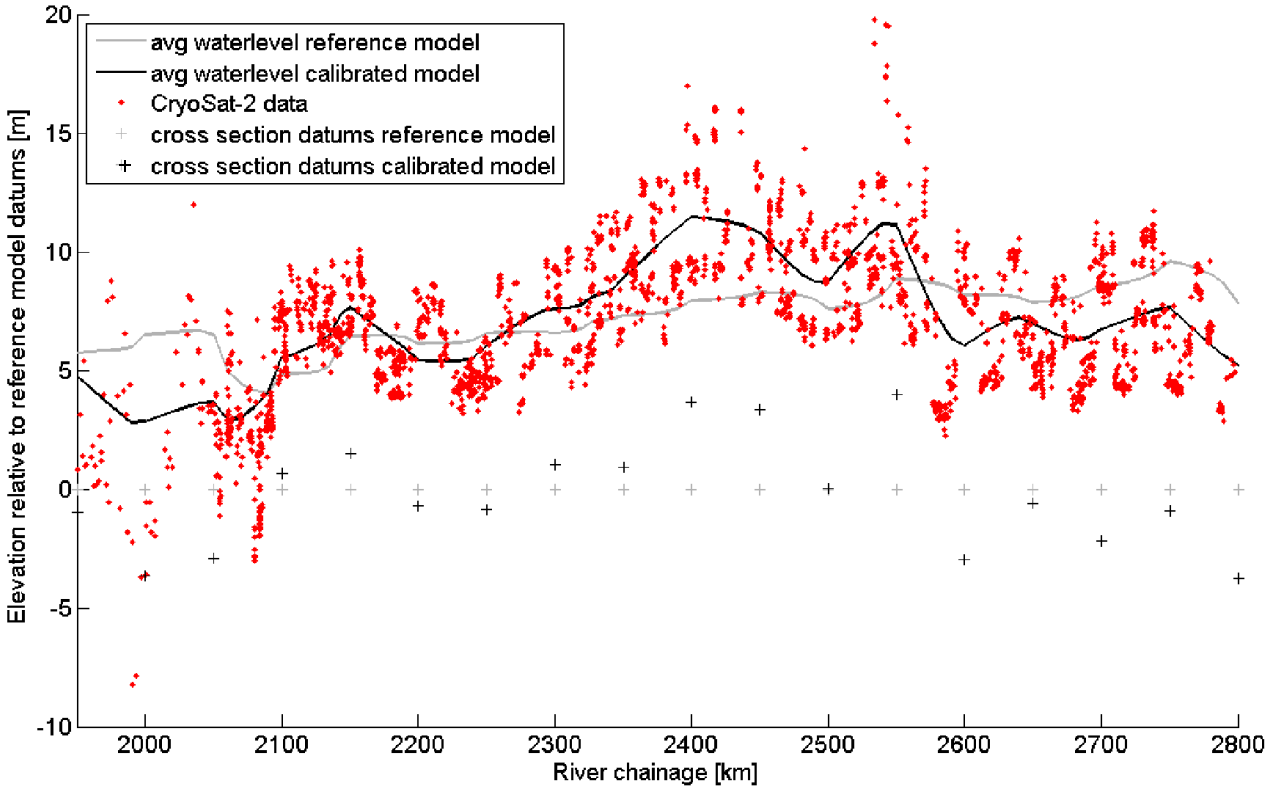



**Figure 45: Result of water level calibration step 1 for the Assam valley for the period 2010 to 2013. All levels are shown as elevations relative to the reference model's cross section datums based on the SRTM DEM.**

- **Step 2**

  The results of the second step of the cross section calibration (Figure 40), adjusting the cross section angles to fit the simulated water level amplitudes to the Envisat observations can be seen in Figure 46 for one of the 13 virtual stations. During the calibration, no absolute water levels were used, but time differences. Consequently, Figure 46 shows water levels relative to the time series value at the time of the first Envisat observation.

  The change of the cross section shapes to calibrate for water level amplitudes showed to have a relevant effect on the absolute water levels. The mean absolute deviation between the simulated mean water level before and after step 2 was 2.6 m. Therefore, step 1 of the cross section calibration had to be repeated using the calibrated opening angles from step 2.

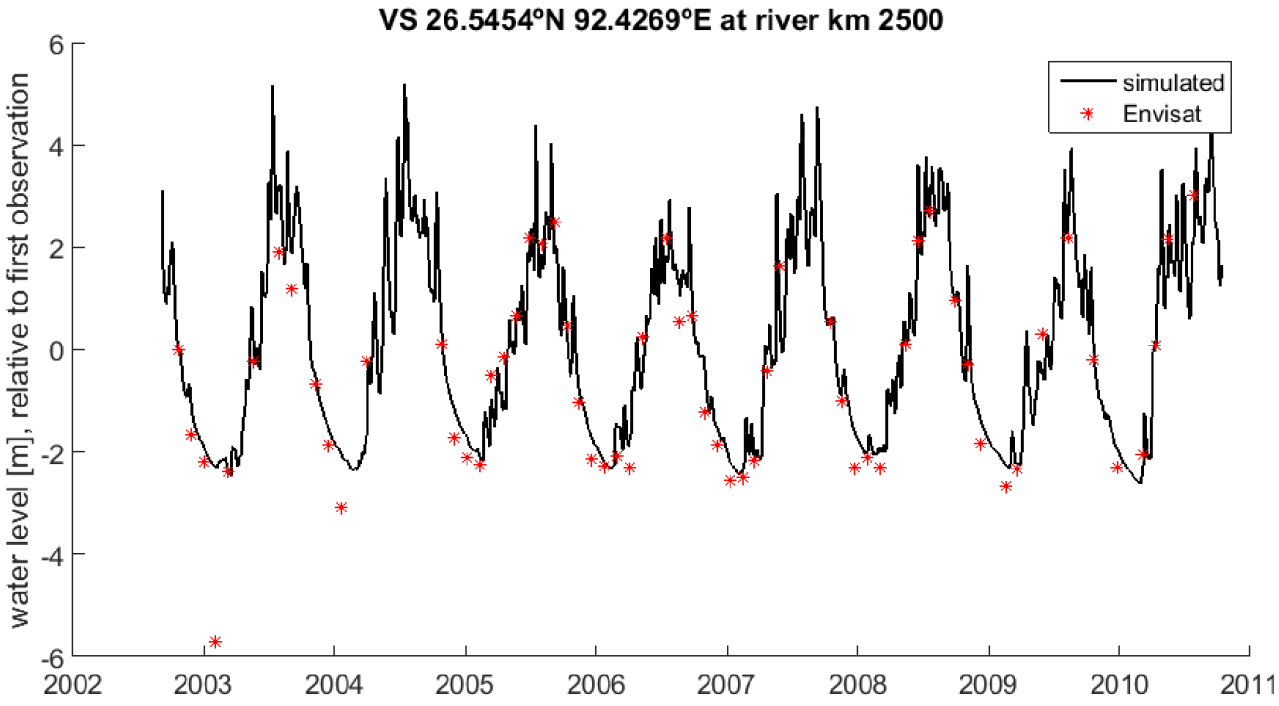



**Figure 46: Water levels after step 2 of the cross section calibration for one virtual station. All levels relative to the water levels at the time of the first Envisat observation.**

The changes to the cross section were found to have a negligible effect on the discharge routing. On average, the simulated discharge changed by less than 0.2% (equivalent to 33 $m^3/s$) after the

cross section calibration. Hence no discharge recalibration of the model was performed after the water level calibration finished.

The presented cross section calibration offers a way to calibrate water levels in the entire model space without precise knowledge of topography and bathymetry. Synthetic cross sections allow the use of practically any shape, however for the sake of reducing the number of decision variables a simple triangular shape has been chosen. The whole process is still computationally expensive; step 2 takes approximately 1 week on a 20 core calculation server.

This study uses a very simplified cross sectional geometry which only has two free parameters: the datum and the opening angle. As shown above, we can reproduce both the average water level and the dynamics of the water level well using this simple model. More highly parameterized cross section geometries could be investigated in future work. However, no major improvement is expected.

### 6.6.4 Brahmaputra Data Assimilation

Figure 47 and Table 6 present initial results of the data assimilation experiments. Figure 47 compares the performance of probabilistic discharge predictions at Bahadurabad, produced with and without assimilation of CryoSat-2 data. Table 6 presents performance statistics and contains mean absolute error (MAE) , root mean squared error (RMSE) , bias, coverage of the nominal 95% confidence interval (a measure of reliability), sharpness and continuous ranked probability score (CRPS). Coverage, sharpness and the Continuous Ranked probability Score were computed from the ensemble. For definition and details of these different indicators, please see Bauer-Gottwein et al. (2015).

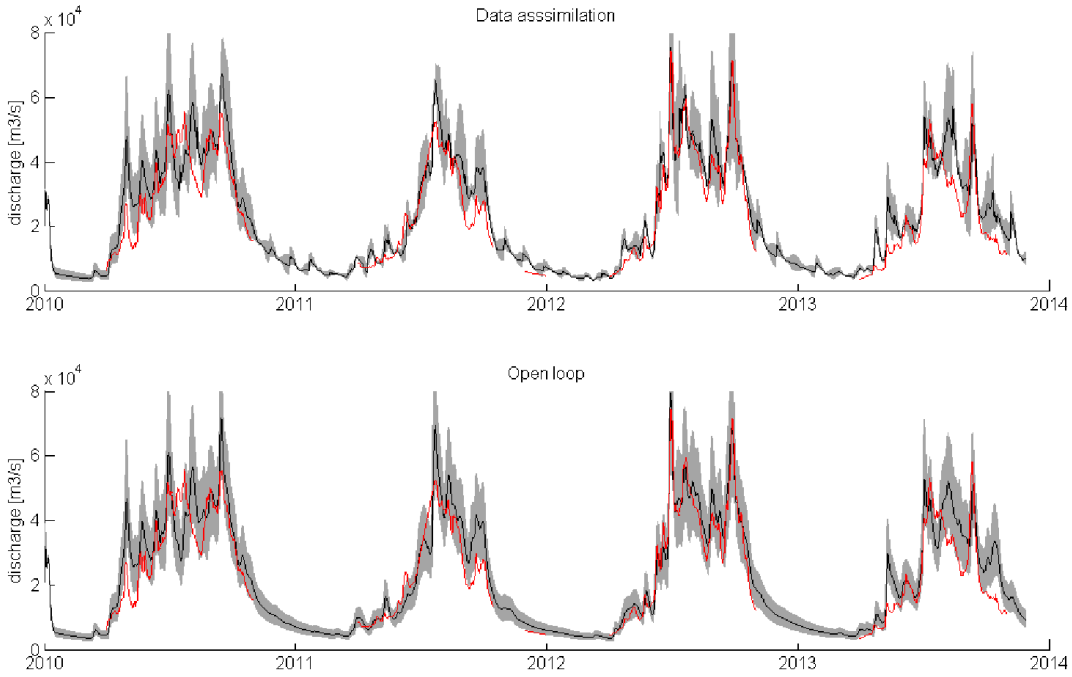



**Figure 47: Comparison of probabilistic discharge simulations produced with and without assimilation of CryoSat-2 data. The central model prediction is shown as a solid line, grey shades indicate 95% confidence intervals, the in-situ observations are in red.**

**Table 6: Comparative performance of the deterministic model run, the probabilistic model run without data assimilation ("open loop") and the run with assimilation of CryoSat-2 data.**

| Run | MAE [m3/s] | RMSE [m3/s] | Bias [-] | Coverage [-] | Sharpness [m3/s] | CRPS [m3/s] |
|---|---|---|---|---|---|---|
| Deterministic | 4548 | 6265 | 0.0418 | | | 4548 |
| Open Loop | 5564 | 7269 | 0.1298 | 0.7878 | 1821 | 4001 |
| Data Assimilation | 5622 | 7230 | 0.1516 | 0.5299 | 1452 | 4429 |

As can be seen from Table 6, in these initial runs, assimilation of CryoSat-2 data does not improve the discharge predictions yet. More detailed analysis is required to identify the reasons, which

could have to do with the parameterization of the model and/or observation errors, with the selection of CryoSat-2 data points or with the rapid wash-out of CryoSat-2 induced updates from the model space. We are confident that performance can be significantly improved by fine-tuning the assimilation scheme and we will use the remainder of the project resources to achieve this.

## 6.7 Conclusion

A 1D hydrologic-hydrodynamic model of the Brahmaputra river basin in South Asia has been set up using mainly remote sensing data. The discharge calibration at the outlet of the basin resulted in a NSE of 0.91 for the calibration period. Besides discharge also water levels in the model's downstream part, the Assam valley, have been calibrated. This calibration 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 is assimilated to the hydrodynamic part of the Brahmaputra model by the help of the DHI Data Assimilation Framework, using an ETKF. This is the first time that CryoSat-2 data have been assimilated into a river model. We have shown that our DA scheme is operational and working as expected. Clearly, the parameters of the assimilation have to be fine-tuned in order to obtain the best result. This has not been done yet, but will be done in the final phase of the project. The main issues are to find optimal parameterizations for model and observation errors.

# 7. References

APHRODITE's Water Resources, 2014. APHRODITE's Water Resources - Home [WWW Document].

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

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.

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_4

Google, 2015. Google Earth Engine - Landsat 7 32-Day NDVI Composite [WWW Document].

ESA-ESRIN. (2013) Cryosat Product Handbook

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

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.

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

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

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

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

Villadsen, H., Andersen, O.B., Stenseng, L., Nielsen, K., Knudsen, P., 2015. CryoSat-2 altimetry for river level monitoring - Evaluation in the Ganges-Brahmaputra basin. Remote Sens. Environ. 168, 80–89.

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