Description of cover: The background image is acquired on August 24, 2011 with the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Terra satellite by NASA, (NASA Earth Observatory Images). Heavy rainfall led to severe flooding in Cambodia, the Mekong River is extended south of the Tonle Sap Lake. The front image is taken on August 6, 2003 with kite aerial photography. The Luang Say is traveling the Mekong River in the Golden Triangle, courtesy to Pierre Lesage. The ovals are manually added to illustrate ICESat/GLAS footprints crossing the river and representing height profiles.

MSc Geomatics
Thesis

February, 2012
Delft, the Netherlands

USING ICESAT/GLAS LASER ALTIMETRY FOR WATER LEVEL ESTIMATIONS IN THE MEKONG RIVER

Marjolein Koudijs |1264664
koudijs.marjolein@gmail.com

Delft University of Technology
Faculty of Civil Engineering
Department of Earth and Atmospheric Sciences
Section of Remote Sensing

SUPERVISING PROFESSOR
Prof.dr. Massimo Menenti

SUPERVISOR
Dr. Roderik Lindenbergh

TU Delft Delft University of Technology

Challenge the future

ABSTRACT

Keywords: Mekong River, water levels, GLAS, river cross-section, classification

This study evaluates the potential of the Ice, Cloud and Elevation Satellite (ICESat) with the onboard Geoscience Laser Altimeter System (GLAS) instrument for the monitoring of water levels in river networks. The monitoring of water levels in river networks provides information about the continental water resources. Furthermore, information about changes in water levels can support flood forecasting systems and hydrological models. The number of hydrological stations declined over the years. Remote sensing from space is used to observe water levels in rivers with altimeters and river width with high resolution images. Satellite radar altimetry has limited applicability for the monitoring of rivers because of coarse spatial resolution and precision related to the size of the target. The ICESat/GLAS instrument is based on the principles of satellite laser altimetry and has an appropriate footprint size of 70 m which implies the potential in monitoring smaller targets such as rivers.

The determination of water level estimations with GLAS altimetry data in the Mekong River (Southeast Asia) is evaluated. The Mekong River is affected by recurrent flooding events with extreme seasonal variation of water levels. The research questions in this study aimed to develop a method to derive water level estimations and eventually water level trends in the Mekong River. Furthermore, the corresponding strengths and weaknesses of GLAS altimetry data for hydrological monitoring of the Mekong River are evaluated.

The method to obtain GLAS derived river footprints in the Mekong River consists of two classification steps. First the candidate river footprints are identified with use of a customized water mask created by the combination of the USGS HydroSHEDS River Network and the MODIS 250m land water mask. The geographical boundaries of these data sets were set with the IWMI drainage basin of the Mekong River. The GLAS footprints that intersect the customized watermask were selected as candidate river footprints. The candidate river footprints result in possible river cross-sections over the Mekong River. Based on the GLA14 data product, height profiles of these river cross-sections were obtained. Besides the height, waveform variables that are derived from the GLA14 data product do contribute to the land cover classification. To evaluate the geometric and waveform variables as classification criteria, a sampling set of 12 river profiles is created representing the river cross-sections over the Mekong River. With Landsat images, the footprints located in the Mekong River are identified. The Landsat-derived river footprints are used to determine the efficiency of candidate classification criteria and test the classification method.

Second, the river footprint identification evaluated the geometric and waveform variables of the candidate river footprints resulting in a weighted geometric and waveform score. The classification method is applied to select river footprints in the Mekong River. GLAS derived water level estimations are obtained by combining the mean height levels of the river footprints. This resulted in 785 GLAS derived water level estimations with an overall standard deviation of 0.087 m. Only 6.6 percent of the GLAS derived water level estimations are located in the Upper Mekong Basin with a standard deviation of 0.41 m. This showed the difficulty of river footprint identification in narrow rivers. Furthermore, only 31 percent of the GLAS derived water level estimations are obtained in the wet season caused by a mismatch of the monsoon climate in the Mekong River Basin and the ICESat operational periods.

Water level trends are calculated with use of nearby water level estimations created by repeated ground tracks of ICESat between 2003 and 2009. Two case study areas are considered in the Lower Mekong Basin: The determination of water level trends in the Upper Mekong Basin is very hard due to poor availability of nearby water level estimations. The two case study areas show inconsistencies in the seasonal amplitude caused by missing values of GLAS derived water level estimations especially in the wet season. Furthermore, the scheduling of the operational periods of ICESat result in only one

measurement per season which does not result in added value for monitoring systems in river networks.
The validation of the GLAS derived water levels is supported with downstream river profiles derived from the USGS HydroSHEDS digital elevation model (DEM) and in situ measurements provided by the Mekong River Commission (MRC). Two downstream river profiles are constructed corresponding with the mainstream of the Mekong River in the Upper Mekong Basin and the Lower Mekong Basin. The GLAS derived water level estimations are plotted over the downstream river profiles to check for extreme outliers. In general, the GLAS derived water level estimations matched the downstream river profiles. The MRC provides in situ data since 2008 in the wet season at 22 hydrological stations located in the Lower Mekong Basin. The GLAS derived water level estimations close by in situ measurements are compared. Only 11 GLAS derived water level estimations were suitable for validation. However due to the complex delta system it was difficult to compare the GLAS derived water level estimations with the in situ measurements. However, the seasonal amplitude derived from the location with the most nearest water level estimations was comparable to the seasonal amplitude derived from the in situ measurements.
It was concluded that the classification method in this study has potential to observe water level estimations derived from GLAS altimetry data in the Mekong River. However, the scheduling of the ICESat operational periods showed limited applicability of the monitoring of water levels during the wet season and over time. Furthermore the geographical coverage is related to the width of the river which resulted in less GLAS derived water level estimations in the Upper Mekong Basin. The water level trends were generally based on one measurement per season which cannot contribute to a river monitoring system based on hydrological stations providing daily (or even hourly) in situ measurements. Nonetheless, the seasonal amplitudes derived from nearby GLAS derived water level estimations at unobserved locations in the river monitoring system can supplement the in situ measurements. The GLAS instrument is able to observe water levels in narrow rivers but with less precision: The behavior of the GLAS laser does not always show consistent waveform variables.
Recommendations for further research include the use of higher resolution classification data. The customized water mask consists of multiple buffers due to limited information in the USGS HydroSHEDS river network about the width of rivers and geolocation errors after converting the MODIS 250m land water mask to a feature. A vertical buffer based on a digital elevation model could be used to reduce extreme outliers in the water level estimations. The weather conditions, saturation effects and waveform errors in the GLAS altimetry data are suggested as subject in further studies related to the potential in hydrological applications. The classification method could be improved by exploring multiple channels in river cross-sections and the consequences of nearly located water level estimations parallel to steep rivers. For future studies, study rivers located in a monsoon climate that match the operational periods of ICESat are preferable as well as the availability of sufficient in situ data to validate the results.

PREFACE

This research will evaluate the potential of the Ice, Cloud and Elevation Satellite (ICESat) with the onboard Geoscience Laser Altimeter System (GLAS) instrument for the monitoring of water levels in river networks. Monitoring water levels of surface-water is relevant for better understanding of the fresh water supply and resources on large scales. Besides, water level estimations based on satellite remote sensing can supplement in situ measurements to optimize flood forecasting systems or other hydrological models. To contribute to the hydrological applications of ICESat/GLAS laser altimetry, this explorative study will assess GLAS altimetry data (GLA14 data product) for water level estimations in the Mekong River.

With this study, the contribution of satellite laser altimetry (ICESat) to the monitoring of water levels in river networks is assessed. It is great that nowadays aerospace scientists are able to study the earth from space in amazing detail. However, scientists in other fields are not always aware of the possibilities with these remote sensing methods to their research interests. This study is therefore written from a broader perspective and incorporates a lot of background readings to support application-oriented researchers that may have benefits from the use of ICESat.

The MSc Geomatics (TU Delft) is concluded with the thesis research. The student will use the knowledge and skills gathered and apply this during an individual study. This research was launched at June 20, 2011 and achieved at the Section of Optical and Laser Remote Sensing (Faculty of Aerospace Engineering, Department of Remote Sensing). The Section of Optical and Laser Remote Sensing moved physically to the Faculty of Civil Engineering & Geosciences in January 2012 at which this thesis was finished.
The process of writing a thesis is a great learning experience although it sometimes takes blood, sweat and tears. Discipline and enthusiasm were the basic ingredients I put in this work. The last two months of the research all pieces of the puzzle finally felt in place. I finished my results, wrote my thesis and learned that I was most efficient when I was enjoying it.

I would like to express my gratitude to some people that supported me through the delivery of a thesis. First of all, I would like to thank my parents for always encouraging me to become the best version of me. Going to university and becoming an engineer felt thereby as a natural course in my life.
Furthermore I would like to thank Dr. Roderik Lindenbergh for the weekly meetings and his positive contribution to my work. Sometimes you hear from fellow students that their supervisors are driving them nuts with their support. With Roderik it was a great pleasure to work on this topic. He never forgot the objectives and focus while I was lost in programming and unintelligibly for the rest of the world. Furthermore, he understood my objectives in life and supported me when I was focusing on things that are not related to the thesis but for my future career.
Vu Phan Hien shared MATLAB codes that helped me to defeat the trouble of pre-processing GLAS binary data files. Furthermore he helped me getting started with GLAS altimetry data, which initially was rocket science to me. My roommate at the Faculty of Aerospace Engineering, Hamid Rea Ghafarian Malamiri, provided a great working atmosphere. Our room was calm with positive energy that helped us both to work hard. I enjoyed the chats about our lives and everything that comes with it. Last but not least, Ali Mousivand, for challenging me to wake up early every morning to be the first at work. He definitely won.

CONTENTS

Abstract iii

Preface v

List of Figures ix

List of Tables xv

List of Abbreviations xvii

1 Introduction 1
1.1 Research context 1
1.2 Research objective and research questions 2
1.3 Research methodology and report outline 3

2 The hydrology of the Mekong River Basin 5
2.1 Introduction to hydrology 5
2.1.1 Definition and terminology 5
2.1.2 Methods for monitoring water levels 7
2.2 Mekong River Basin 8
2.2.1 Study area 8
2.2.2 Meteorology 10

3 The determination of the Mekong River location 12
3.1 Mekong River classification data 12
3.1.1 USGS HydroSHEDS 12
3.1.2 MODIS 250m land water mask 16
3.2 Customized water mask of the Mekong River 18
3.3 Validation of customized water mask 19
3.3.1 Landsat images 19
3.3.2 Customized water mask overlapping Landsat images 21

4 Principles of satellite laser altimetry 23
4.1 ICESat mission 23
4.2 GLAS instrument 23
4.3 Satellite laser altimetry 25
4.3.1 Laser scanning 25
4.3.2 Full waveform laser scanning 26
4.3.3 Interaction of GLAS instrument with water surfaces 28
4.4 Parameters derived from the GLA14 data product 31
4.4.1 Workflow of parameterization 31
4.4.2 Original GLA14 variables 32
4.4.3 Ellipsoidal and temporal conversions 33
4.4.4 Customized GLA14 variables 34

5 Methodology for river footprint identification 38

5.1 Candidate river footprint identification 38
5.1.1 Workflow of candidate river footprint identification 38
5.1.2 Descriptive statistics of candidate river footprints 40
5.1.3 Sampling set of river profiles 43
5.2 Classification criteria 48
5.2.1 Candidate classification variables 48
5.2.2 Efficiency of candidate classification criteria 51
5.3 River footprint identification 52
5.3.1 Workflow of river footprint identification 53
5.3.2 Testing of classification scores 57
5.3.3 Weighting of classification scores 59
5.3.4 Implementation procedure for the Mekong River 61

6 Results and discussion 62
6.1 Water level estimations 62
6.1.1 Water levels 62
6.1.2 Precision of water levels 65
6.2 Water level trends derived from water level estimations 68
6.2.1 Study area 1: Mekong River Mainstream 68
6.2.2 Study area 2: Nam Phong Tributary (Ubolratana Dam) 70
6.3 Validation of results in Mekong River Basin 72
6.3.1 Downstream river profiles 72
6.3.2 In situ measurements 77

7 Conclusions and recommendations 84
7.1 Conclusions 84
7.2 Recommendations 86

References 89

Appendix 93
Appendix 1: Waveform variables of sampling set of river profiles 93
Appendix 2: Geometric and waveform scores for the sampling set of river profiles 100
Appendix 3: Alternative weights of geometric and waveform scores on sampling set .. 104

LIST OF FIGURES

Figure 1.1: Schematic overview of the report outline. The number between the brackets refers to the corresponding section in the report. 4

Figure 2.1: Terminology related to the physical geography of a stream. 6

Figure 2.2: Terminology related to the river cross-section. 6

Figure 2.3: Six countries intersecting the Mekong River Basin. 8

Figure 2.4: Six river components represented with different colors. Components are defined on river parts between cities. 8

Figure 2.5: Longterm (1960-2004) monthly average rainfall, based on Mekong River Commission [26]. Average rainfall is based on six measurements (Chiang Rai, Pakse, Khon Kaen, Pleiku, Phnom Penh, and Chau Doc). Green indicates the wet season, grey indicates the dry season. 11

Figure 2.6: Longterm (1960-2004) monthly average discharge, based on Mekong River Commission [26]. Average discharge is based on six measurements (Chiang Saen, Vientiane, Nakhon, Mukdaha, Pakse, and Kratie) along the Mekong River. Green indicates the wet season, grey indicates the dry season. 11

Figure 3.1: USGS HydroSHEDS, selecting tiles covering the area of interest (yellow box). . 12

Figure 3.2: USGS HydroSHEDS drainage basin and river network in Mekong River Basin, streams with standard threshold of 100 upstream cells. 14

Figure 3.3: Example of disconnected streams with reservoirs, Nam Ngum Dam is the upper turquoise area. 14

Figure 3.4: USGS HydroSHEDS derived Mekong River with thresholds of upstream cells, coverage of complete Mekong River Basin. 14

Figure 3.5: USGS drainage basin compared with IWMI drainage basin, distributaries in the Mekong River Delta (green lines) are excluded by USGS. 15

Figure 3.6: Absence of width information in USGS river network, Tonle Sap Lake (Cambodia), therefore complementing data is necessary. 15

Figure 3.7: MODIS 250m land water mask data in the Mekong River Basin. 16

Figure 3.8: MOD44W raster (yellow), converted to feature format (blue) and enhanced feature with 250m buffer (turquoise) to cover original data. 17

Figure 3.9: MOD44W raster (yellow), and enhanced feature with 250m buffer (turquoise) closes feature gaps but also islands and narrow mid-channel banks. 17

Figure 3.10: Buffer sizes for MOD44W and USGS River Network. 18

Figure 3.11: Landsat ETM images referenced to WGS84, intersecting the Mekong River Basin. The Path/Row indices are used to select specific Landsat images. 20

Figure 3.12: Landsat image (PRS 135053) original color composite (RGB = 321), located in Mekong River Delta. 20

Figure 3.13: Landsat image (PRS 135053) false color composite (RGB = 541), located in Mekong River Delta. 20

Figure 3.14: Customized water mask (yellow boundary) over Landsat 5 ETM image covers the Mekong River (dark blue) and MOD44W raster with gaps (turquoise). 21

Figure 3.15: Customized water mask (yellow boundary) over Landsat 5 ETM image cover the Mekong River in and preserves islands in the Mekong River Delta. 21

Figure 3.16: Errors in USGS HydroSHEDS River Network (RIV) causes absence of the complete left branch in the Mekong River Delta. 22

Figure 3.17: Final vector representation of Mekong River based on a combination of USGS HydroSHEDS, IWMI and MOD44W data. 22

Figure 4.1: Along-track and across-track spacing applicable for the Mekong River Basin. . 25

Figure 4.2: Multiple return echos. (a) Laser signal intercepting at different height levels. (b) Return signal as a function of time. Breakdown of the return signal into different echos: (c) Only first and last echo and (d) multiple return echos. 26

Figure 4.3: Example of Gaussian curve described by 27

Figure 4.4: Example of multiple Gaussian curves described by y = N(x,3.75,0.75)+N(x,6,0.5). 27

Figure 4.5: Illustration of the GLAS laser altimeter measurement, geolocation and surface elevation determination, taken from Duong [12]. 27

Figure 4.6: Waveform as referenced in GLA14 data with the centroid as GLA14 elevation data for the average height of the features in the footprint. The height of the ground surface is represented by the centroid of the last return pulse (i_gpCntOff1). 28

Figure 4.7: Characteristics of return laser pulse as a function of surface type. Presence of surface slope and roughness both broaden the pulse, taken from Brenner [6]. 29

Figure 4.8: Land cover classification based on waveforms of bare land (left) and water (right), taken from Duong [12]. 30

Figure 4.9: Example of a non-saturated waveform (left plot) and a saturated waveform (right plot). Notice the clipped waveform when saturated. Taken from Molijn [32]. 30

Figure 4.10: Overview of parameterization of GLA14 data. 32

Figure 4.11: Model of the Earth, taken and enhanced from Esri [14]. 35

Figure 4.12: Orthometric height (H), ellipsoidal height (h) and geoid height (N). 35

Figure 4.13: Waveform variables as described by the return pulse. 36

Figure 4.14: Example probability density functions with varying kurtosis and skewness values. In both subfigures, the red line represents the normal distribution having zero kurtosis and zero skewness. All distributions in each subfigure have the same standard deviation and are normalized with respect to the area under the distribution. The figure is only meant for visualization purposes, taken from Molijn [32]. 37

Figure 5.1: Overview of candidate river footprint identification. 39

Figure 5.2: Example of missing GLAS footprint in a river cross-section. The large gap in the river cross-section is a result of the spatial selection of the GLAS footprints with the customized water mask. 40

Figure 5.3: Example of false candidate river footprints (red dots) as a result of spatial selection of the GLAS footprints with the customized water mask. 40

Figure 5.4: Temporal variation per year (left plot) and per month (right plot) of candidate river footprints. 41

Figure 5.5: Geographical coverage of candidate river footprints in the Upper Mekong Basin (left plot) and the Lower Mekong Basin (right plot). The number of footprints per river cross-section is illustrated with the colored dots. 42

Figure 5.6: Histogram of the number of GLAS footprints per river cross-section (left plot) and with less than 50 footprints (right plot). 43

Figure 5.7: Example of a river cross-section described by the latitude [deg] and longitude [deg] (top-view). The arrow indicates the descending latitude direction. 44

Figure 5.8: Example of a corresponding river profile (see Figure 5.7) described by the height [m] and along-track profile [m] (side-view). 44

Figure 5.9: Sampling set of river profiles in the Mekong River. The label (ID) refers to the identifier of the river profiles. 44

Figure 5.10: Sampling set of river profiles part 1. The graph represents mean and corresponding standard deviations of the height values (red). The dots below represent the classification of river footprints in which green = land, grey = land-water transition, blue = water. 46

Figure 5.11: Sampling set of river profiles part 2. The graph represents mean and corresponding standard deviations of the height values (red). The dots below represent the classification of river footprints in which green = land, grey = land-water transition, blue = water. 47

Figure 5.12: Waveform variables of river profiles with ID = 650 (left plot) and ID = 1542 (right plot) with corresponding Google Earth image. 51

Figure 5.13: Example of a river cross-section from the sampling set with 2 groups of adjacent inliers (blue boxes). The adjacent inliers are defined by the adjacent height difference between consecutive footprints. 55

Figure 5.14: Workflow of the calculation of the geometric score. The yellow box represents a decision that has consequences for the next steps. The grey boxes represent the input and output of the workflow. 55

Figure 5.15: Workflow of the calculation of the waveform score. The yellow box represents a decision that has consequences for the next steps. The grey boxes represent the input and output of the workflow. 57

Figure 5.16: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 650. The turquoise dots indicate Landsat-derived river footprints. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference. 59

Figure 5.17: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 1542. The turquoise dots indicate Landsat-derived river footprints. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference. 59

Figure 6.1: River cross-section over Landsat ETM image with multiple land-water transitions. The turquoise dot represents the simplified geolocation determination by median latitude/longitude of river footprints. 63

Figure 6.2: Effect of simplified geolocation determination by median latitude/longitude of river footprints (turquoise dots) and the determination of the water level (turquoise striped line). 63

Figure 6.3: Temporal variation per year (left plot) and per month (right plot) of water level estimations. 64

Figure 6.4: Locations of water level estimations in the Upper Mekong Basin (left plot) and the Lower Mekong Basin (right plot). The number of river footprints per water level estimation is indicated with colors. 64

Figure 6.5: Number of nearby water level estimations within a radius of 2 km in the Upper Mekong Basin (left plot) and the Lower Mekong Basin (right plot). 65

Figure 6.6: Histograms of the standard deviation of water level estimations (left plot) and with selected standard deviation smaller than 0.2 m (right plot). 66

Figure 6.7: River cross-section with ID = 2640 corresponding to the maximum standard deviation over Landsat ETM image. The blue dot refers to the geolocation of the water level estimation. 67

Figure 6.8: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 2640. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference. 67

Figure 6.9: Nearby water level estimations (green dots) corresponding to repeated ground track (red line) in the Mekong River Delta over Landsat ETM image. The labels indicate the campaigns. 69

Figure 6.10: Water level trend between 2004-2008 in the Mekong River Delta. The turquoise bars indicate measurements in the wet season; the grey bars indicate measurements in the dry season. The labels indicate the campaigns 70

Figure 6.11: Geometric (left plot) and waveform (right plot) scores of the river profile with ID =936. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference. 70

Figure 6.12: Water level estimations (green dots) corresponding to repeated ground track (red line) in the Nam Phong Tributary over Landsat image. The Nam Phong Tributary connects the Ubolratana Dam Reservoir with the Mekong River. 71

Figure 6.13: Close-up image of the water level estimations (green dots) in the Nam Phong Tributary. The labels refer to the campaigns. 71

Figure 6.14: Water level trend between 2003-2009 in the Nam Phong Tributary. The turquoise bars indicate measurements in the wet season; the grey bars indicate measurements in the dry season. The labels indicate the campaigns 72

Figure 6.15: Channel constructed manually (red line) with GLAS derived water level estimations within a distance of 5 km (green dots) in the Upper Mekong Basin. 73

Figure 6.16: Channel constructed based on flow accumulation (red line) with GLAS derived water level estimations within a distance of 5 km (green dots) in the Lower Mekong Basin. 73

Figure 6.17: Downstream river profile created with a 9th order polynomial (grey line) based on the values of the USGS HydroSHEDS DEM at the location of the water level estimations within a distance of 5 km of the channel in the Upper Mekong Basin. 74

Figure 6.18: Downstream river profile created with a 9th order polynomial (grey line) based on the values of the USGS HydroSHEDS DEM at the location of the water level estimations within a distance of 5 km of the channel in the Lower Mekong Basin. 74

Figure 6.19: Downstream river profile created with a 9th order polynomial (grey line) and the 43 GLAS derived water level estimations within a distance of 5 km of the channel in the Upper Mekong Basin. 75

Figure 6.20: Downstream river profile created with a 9th order polynomial (grey line) and the 348 GLAS derived water level estimations within a distance of 5 km of the channel in the Lower Mekong Basin. 75

Figure 6.21: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 1693. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference. 76

Figure 6.22: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 547. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference. 76

Figure 6.23: Outlying GLAS derived water level estimation of river cross-section with ID = 1693 (pink dots) and nearby river cross-sections (green dots) over a Landsat ETM image. The labels refer to the height levels [m]. 77

Figure 6.24: Outlying GLAS derived water level estimation of river cross-section with ID = 0547 (pink dots) and nearby river cross-sections (green dots) over a Landsat ETM image. The labels refer to the height levels [m]. 77

Figure 6.25: Location of the 18 hydrological stations in the Mekong River with expected geolocations. 78

Figure 6.26: Downstream river profile created with a 9th order polynomial (grey line) and the hydrological stations within a distance of 5 km of the channel in the Lower Mekong Basin. 79

Figure 6.27: Location of the water level estimations suitable for validation (green dots) located and the nearest hydrological stations (pink triangles). The labels refer to the ID of the water level estimations. 80

Figure 6.28: The left plot shows the group of 19 GLAS derived water level estimations (green dots) located near the hydrological station Stung Treng. The right plot is a close-up of the water level estimations (labeled with ID) and the corresponding ground tracks (red lines) over Landsat ETM images. 82

Figure 6.29: Gauge height at Stung Treng in the wet season, taken from the MRC [28]. The gauge height should be added to the zero gauge level of 36.79 m. 82

Figure 6.30: Seasonal amplitude for a group of 19 GLAS derived water level estimations near Stung Treng. Only 17 water level estimations are plotted due to overlapping observation times. 83

LIST OF TABLES

Table 2.1: Intersection of the six Mekong River Basin countries with the catchment, taken from the MRC [26]. 9

Table 3.1: USGS HydroSHEDS data layers. 13

Table 3.2: Values for the MOD44W raster. 16

Table 3.3: Landsat 5 TM sensor characteristics. 19

Table 4.1: ICESat campaigns and operational periods, taken from NSIDC [42]. 24

Table 4.2: Original variables selected from the GLA14 data product, based on NDISC GLAS Altimetry Data Dictionary [38]. 33

Table 4.3: Summary of GLAS ellipsoid and WGS84 ellipsoid, taken from NSIDC [40]. 33

Table 4.4: Customized GLA14 variables, grouped in standard, geometric and waveform variables. 34

Table 5.1: Summary of candidate river footprints per campaign. 41

Table 5.2: Overview of the 12 river profiles in the sampling set. The ID refers to the index of the river profile. N refers to the number of footprints per river cross-section. The Landsat WRS refers to the index of the Landsat ETM image. 45

Table 5.3: Categories and number of samples (N) for each class. 51

Table 5.4: Class statistics for values per land cover type (water/land). 51

Table 5.5: Results of the Independent Samples T-Test for variables. 52

Table 5.6: Summary of geometric and waveform scores of the sampling set of river profiles 58

Table 5.7: Different results of hypothesis testing. 60

Table 5.8: Alternative scenarios for weighting geometric and waveform scores. 60

Table 5.9: Comparing mean and standard deviation of heights for 4 alternatives. 61

Table 6.1: Standard deviation of water level estimations. N represents the number of water level estimations, the median, minimum and maximum refer to the descriptive statistics of the standard deviations. 66

Table 6.2: Standard deviation of water level estimations per season. N represents the number of water level estimations, the Median [m], Minimum [m] and Maximum [m] refer to the descriptive statistics of the standard deviations. 68

Table 6.3: Standard deviation of water level estimations per basin. N represents the number of water level estimations, the Median [m], Minimum [m] and Maximum [m] refer to the descriptive statistics of the standard deviations. 68

Table 6.4: Characteristics of water level estimations in the Mekong River Delta. The Mean [m] and the Std [m] refer to the water level estimation. The Length [m] is the perpendicular distance between the water level estimation and the ground track. 68

Table 6.5: Characteristics of water level estimations in the Nam Phong Tributary. The Mean [m] and the Std [m] refer to the water level estimation. The Length [m] is the perpendicular distance between the water level estimation and the ground track. 71

Table 6.6: Expected geolocations of hydrological stations. The ID refers to the index of the station. The Lat [deg] and Lon [deg] are used as geolocation. The zero gauge level [m] is the low flow level at the location. The grey-shaded stations are influenced by the tidal system in the Southeast China Sea. 78

Table 6.7: Summary of water level estimations for validation. N represents the number of river footprints. The Mean [m] and Std [m] refer to the estimated water level. The Length [m] refers to the perpendicular distance between the water level estimation and the hydrological station. 80

Table 6.8: Comparison of the water level estimations for validation and in situ measurements at corresponding dates. WLE [m] refers to the water level estimations. The height difference (ΔH) with and without the height correction are included. 81

Table 6.9: Seasonal amplitudes (Amp) at Stung Treng obtained and enhanced from the MRC [28]. The Mean [m], Std [m], Min [m] and Max [m] refer to the GLAS derived water level estimations. 82

LIST OF ABBREVIATIONS

ATBD	Algorithm Theoretical Basic Document
DEM	Digital Elevation Model
EOS	Earth Observation System
ETM	Enhanced Thematic Mapper
GCLF	Global Land Cover Facility
GLA14	GLAS Level 2 Global Land Surface Altimetry Data
GLAS	Geoscience Laser Altimeter System
GLOVIS	Global Visualization Viewer
ICESat	Ice, Cloud and Elevation Satellite
IR	infrared
IWMI	International Water Management Institute
LEGOS	Laboratoire d'Etudes en Géophysique et Océanographie Spatiales
MOD44W	MODIS 250m land water mask
MRC	Mekong River Commission
NASA	National Aeronautics and Space Administration
NSIDC	National Snow and Ice Data Center
PAD	Precision Altitude Document
POD	Precision Orbit Determination
rms	root mean square
SRTM	Shuttle Radar Topography Mission
UNDP	United Nations Development Programme
USGS	U.S. Geological Survey
WGS84	World Geodetic System 1984
WRS	Worldwide Reference System

1 INTRODUCTION

This chapter will introduce the topic. In section 1.1, the research is described. Hydrological processes and limitations of traditional monitoring of water levels in river networks will create motivation for this study. The research objective and accompanying research questions proceed naturally from this motivation and will be outlined in section 1.2. A brief overview of the research methodology as described by the outline of this report will be provided in section 1.3.

1.1 RESEARCH CONTEXT

The monitoring of water levels in river networks is relevant for several reasons. First, continental water resources contribute to a balanced fresh water supply. Measurements of the temporal variations in water levels of rivers, lakes and reservoirs are necessary because many developing countries have insufficient networks of hydrological stations [2]. Second, flood forecasting systems are developed to predict flooding events based on the real time observations of water levels.

The Mekong River Basin (Southeast Asia) is familiar with recurrent flooding events causing severe damage. The flood in 2000 resulted in US$ 400 million costs of damage and 800 people's lives [25]. This triggered the development of a flood forecasting and river monitoring system based on water level data and rainfall data provided by a number of hydrological stations.

The monitoring of rivers with hydrological stations is challenging in remote areas. Furthermore, rivers intersect countries which results in an involuntarily dependence of the downstream countries for water supply from inland water resources. To monitor the water resources, water level data sharing between countries is not always coordinated which requests an independent data source. This highlights the limitations of hydrological stations: although they do provide water levels with daily temporal intervals, they cannot provide complete coverage of large rivers.

Satellite remote sensing methods can supplement hydrological stations. The strength of river network monitoring from space is naturally the global coverage. Satellite radar altimetry is used for hydrological observations and provides all-weather operability, global coverage, and temporal repetivity (weeks) [2]. However, the coarse spatial resolution (few hundred meters to a few kilometers) implies that satellite radar altimetry can only monitor large water surfaces. The accuracies of the water level estimations are directly related to the size of the water body and for the monitoring of river networks the river's width is therefore an unambiguous limitation.

Satellite laser altimetry such as the Ice, Cloud and Elevation Satellite (ICESat) with the onboard Geoscience Laser Altimeter System (GLAS) has potential to monitor smaller water bodies because of a smaller footprint size (70 meters). It is stated that the ICESat/GLAs instrument provides the longest, densest, and most repeatable records of inland river elevations from space [51].

This study focuses on the potential of ICESat/GLAS laser altimetry for the monitoring of water levels in river networks. The study is limited in three dimensions: first of all, the altimetry sensor used is the GLAS instrument. Second, the data will be used for hydrological altimetry applications, more specific the derivation of water level estimations. Third, the area of interest is the Mekong River. Take altogether, the topic of GLAS derived water level estimations in the Mekong River Basin is established.

The study has an evaluative nature since ICESat is decommissioned since August 14, 2010. However, the research has also an explorative element since river networks are not yet thoroughly explored in studies with use of ICESat/GLAS altimetry. Furthermore, a second-generation satellite (ICESat-2) is scheduled for launch in 2016 for which this study might be considered as motivation for future use of satellite laser altimetry in hydrological applications.

1.2 RESEARCH OBJECTIVE AND RESEARCH QUESTIONS

The research objective originates from the relevance of monitoring of water levels and the potential of GLAS altimetry data to monitor small water bodies such as rivers. This study aims to contribute to the monitoring of water levels by exploring another method to acquire water levels in river networks based on satellite remote sensing. The scope of this study is explorative in a sense that the challenges and opportunities of ICESat for the specific hydrological application of monitoring water levels in river networks will be enquired. ICESat carried the GLAS instrument with the ability to measure height based on laser altimetry. The research objective is formulated as follows:

The research objective is to assess GLAS altimetry data for the monitoring of water levels in river networks.

The study is limited by the geographical boundaries of the Mekong River in Southeast Asia. The Mekong River belongs to the largest rivers in the world. Furthermore, flooding in the Mekong River Basin is a recurrent event affecting the entire Basin [25]. The flood forecasting and river monitoring system in the Mekong River is still developing and tools that contribute are very welcome. To support the research objective, this study evaluates the determination of water level estimations with GLAS altimetry data in the Mekong River. From this, the main research question that will be answered is:

How can GLAS altimetry data be optimally used for water level estimations in the Mekong River Basin?

To answer this research question, three sub questions are formulated to contribute to a complete assessment of the GLAS altimetry data in the Mekong River Basin. Each sub question will be briefly described:

- *How to select footprints corresponding with water?*
  The first sub question is very explorative formulated. The GLAS altimetry data that overlaps the Mekong River needs to be identified. Furthermore the variables obtained from the GLAS altimetry data will be evaluated for the classification of land cover types. The characteristics of the mainstream will be explored to develop a method to estimate the location and the corresponding water levels in the Mekong River. This method will be tested on specific study areas to assess the strengths and weaknesses of the use of GLAS altimetry data in river networks.

- *How can nearby water level estimations be used to assess water level trends?*
  The contribution to river monitoring is only possible when nearby GLAS derived water level estimations are available. Nearby water level estimations in the Mekong River are identified. The water level trends for two study areas are evaluated on their seasonal amplitude.

- *What are the strengths and weaknesses of GLAS derived water levels estimations?*
  The performance of GLAS in the monitoring of river networks depends on correct classification of GLAS altimetry data corresponding with the Mekong River and the validity of the GLAS derived water level estimations. Both aspects will be discussed in this study.

The first sub question is most relevant in this study as it covers the actual assessment of GLAS altimetry data in river networks. The second sub question naturally flows from the results in this study and evaluates water level trends which will show the contribution to river monitoring systems. The assessment is concluded with the third sub question by including the validity of the results.

1.3 RESEARCH METHODOLOGY AND REPORT OUTLINE

The report is outlined consistent with the research methodology. A schematic overview of the report is illustrated in Figure 1.1. The number between the brackets refers to the corresponding section of the report. For best understanding of the research it is recommended to read the complete report from the beginning until the end. Each chapter is briefly described.

Chapter 2 describes the hydrology of the Mekong River Basin. The relevant terminology and theory used in hydrologic studies will be explained (section 2.1) and used to describe relevant hydrological characteristics of the Mekong River Basin for the monitoring of river networks (section 2.2).

Chapter 3 elaborates the determination of the location of the Mekong River. Only GLAS altimetry data located near the mainstream is relevant for this study. With use of a customized water mask of the Mekong River, the GLAS altimetry data is selected. Additional data is necessary. The additional data sets are introduced (section 3.1) and combined to create a customized water mask of the Mekong River (section 3.2). The customized water mask is validated with Landsat satellite images (section 3.3).

Chapter 4 describes the sensor of interest in this study: ICESat with the onboard GLAS instrument. After briefly discussing the ICESat mission (section 4.1) and the onboard GLAS instrument (section 4.2), the principles of satellite laser altimetry are explained (section 4.3). The GLAS altimetry data (GLA14 data product) includes variables useful for the river footprint identification in the Mekong River. The parameters derived from GLA14 data and the applicable conversions and computations are discussed (section 4.4).

Chapter 5 describes the methodology of river footprint identification. The methodology consists of two parts. First, candidate river footprints are obtained by selecting the GLAS footprints that intersect the customized water mask (section 5.1). Second, the candidate river footprints are further classified into actual river footprints based on a geometric score and a waveform score and the results are evaluated for a sampling set of 12 river profiles (section 5.2). The implementation procedure of the river footprint identification in the Mekong River is discussed (section 5.3).

Chapter 6 presents the results of the river footprint identification and discusses the opportunities and weaknesses of the findings in the Mekong River Basin. The GLAS derived water level estimations and the quality of the results is analyzed (section 6.1). Water level trends based on closely located water level estimations are determined to evaluate possibilities for river monitoring (section 6.2). The water level estimations are validated with a comparison to downstream river profiles based on a digital elevation model and in situ water levels (section 6.3).

Chapter 7 discusses the conclusions of this study and recommendations with respect to the research question and further research.

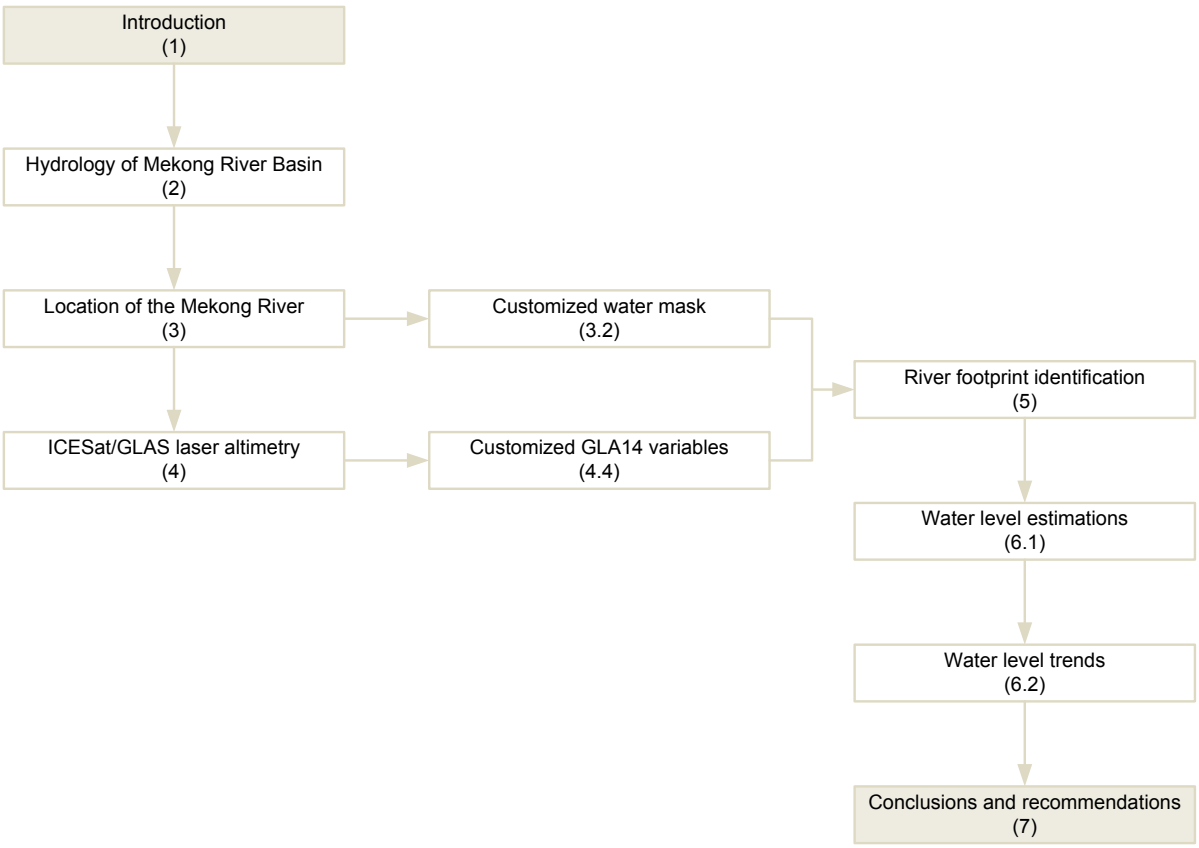


Figure 1.1: Schematic overview of the report outline. The number between the brackets refers to the corresponding section in the report.

2 THE HYDROLOGY OF THE MEKONG RIVER BASIN

This chapter will give an introduction to the hydrology of the Mekong River Basin. The topic is related to hydrology (section 2.1). In section 2.1.1 the relevant terminology and theory used in hydrologic studies will be explained. Furthermore, a short introduction to the various methods for the monitoring of water levels will be described in section 2.1.2. This includes satellite radar altimetry which grounds the motivation of this study. The study area is the Mekong River Basin (section 2.2). Section 2.2.1 will give a description of the study area and the relevance of monitoring water levels. Flooding and the development of dams are shortly discussed because they are responsible for changes in the hydrological regime of the Mekong River. Section 2.2.2 will describe the meteorology of the study area. The monsoon climate results in heavy rainfall during the wet season which influences the water levels in the Mekong River. The vulnerability of the Mekong River to climate change will be briefly discussed.

2.1 INTRODUCTION TO HYDROLOGY

The terms described in this section are taken from the U.S. Geological Survey (USGS) [21]. In the first place the definition of surface-water hydrology will be provided. Terms related to the physical geography of a stream, related to the river cross-section, and related to surface-water monitoring are introduced. The methods used for surface-water monitoring will only briefly discuss in situ measurements and focus on remote sensing measurements.

2.1.1 DEFINITION AND TERMINOLOGY

This study focuses on surface-water hydrology. Surface-water hydrology is the study of the origin and processes of water in streams and lakes, in nature, and as modified by man [21]. Water-surface hydrology is related to the field of meteorology and geomorphology. Meteorology comprehends among other things the study of precipitation. The amount of water that is carried by a river is correlated with the precipitation in its basin. Geomorphology can be defined as the study of the shape, size and number of river channels. This is related to hydrology because river channels are formed as a consequence of the rates and quantities of water they must carry.
The surface-water feature observed in this study is the Mekong River. The Mekong River can be described as a stream. The terminology related to the physical geography of a stream is visualized in Figure 2.1. The catchment can be described as the area which consists of the surface stream together with all tributaries. When traveling from the source to the mouth of the mainstream, the mainstream can be split into diverging branches which can eventually reenter the mainstream (anabranches). A tributary is a stream that merges the mainstream at the confluence. A distributary is the opposite and is a stream that flows away from the mainstream.
The source of the stream originates from overland flow which is the flow of rainwater or snow and glacier meltdown over the land surface toward stream channels. After it enters the river it becomes runoff, together with groundwater supply. The stream flow is the discharge that occurs in a natural channel. Besides water, a stream transports sediments. Sediments consist of fragmental material that originates from weathering of rocks. The mouth of the river results in distributary networks called river deltas. This is caused by the decrease in stream flow and the corresponding deposition of sediments, sand and gravel.

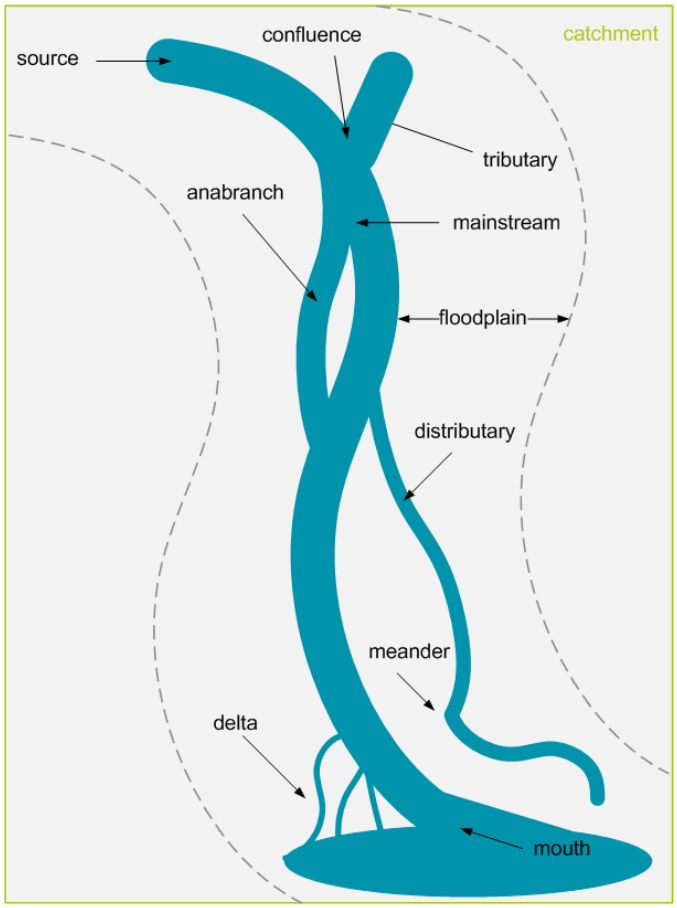


Figure 2.1: Terminology related to the physical geography of a stream.

An illustration of an example of a river cross-section with the corresponding terms is given in Figure 2.2. The water level or stage is the height of the water surface above an established datum plane where the stage is zero. The zero level is arbitrary. The bankfull width is determined in accordance to the bankfull stage or the moment when the water level meets the maximum capacity before overflow occurs. To determine a river cross-section, one needs to know at least the water level, stream width, and the bankfull width. As soon as the water flows over the banks, as happens after heavy rainfall, the Riparian stage is reached. When even floodplains or the lowland that borders the river get subject to flooding, the flooding stage is reached.

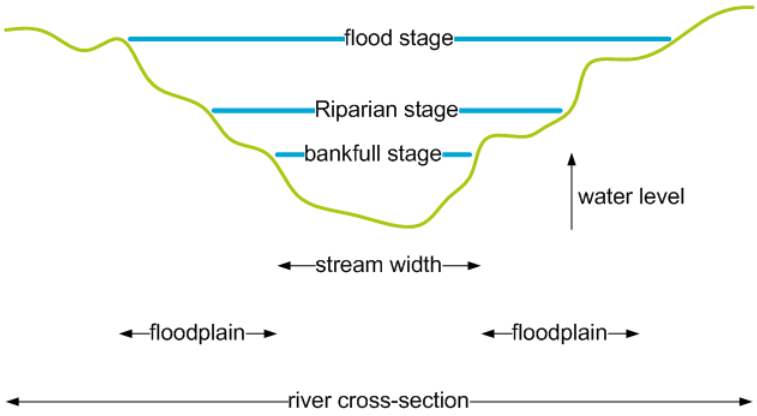


Figure 2.2: Terminology related to the river cross-section.

Based on the length of a stream with respect to the stream flow, water level, area and slope, the river can be characterized in components. For every component, the hydrological regime, or the habits of the stream with respect to velocity, volume, form and changes in the channel, capacity to transport sediment and the amount of material supplied for transportation will vary. The hydrological regime can be influenced by changes in the water cycle, wind, earth movements, climate change and human factors like dredging and diversions [44].

2.1.2 METHODS FOR MONITORING WATER LEVELS

Surface-water monitoring is relevant for the study of hydrological regimes. Traditional hydrological measurements are conducted with in situ gauging stations. The gauging stations can provide water levels and flows in water bodies. However, gauging networks are often organized on a local basis to monitor water resources. The gauging stations provide temporal variations in the water levels, yet the spatial distribution is related to freshwater needs and does therefore not offer complete coverage of water bodies [2]. To establish a comprehensive global river monitoring network with a gauging network, technological, economic and institutional obstacles need to be overcome. This has resulted in a worldwide decline in the number of gauging stations since 1980 [4] [13].
With satellite remote sensing, a global surface-water monitoring system can be established. During the last decennia, several satellites carried sensors that are able to map water bodies. Water level observations from space can be determined with radar altimetry and laser altimetry. The radar altimetry systems that are used for observations in large water bodies are TOPEX/Poseidon, ERS-2, ENVISAT and Jason-1/2. While the satellite orbits the earth, the instrument emits microwave pulses toward the surface. The distance between the satellite and the earth's surface can be derived from the time-delay between the emitted pulse and the echo pulse [3]. Radar altimetry is very suitable for hydrological observations since they operate under all weather conditions, providing global data coverage and have temporal repetivity of up until 10 days (Topex/Poseidon and Jason-1/2) [2]. Other advantages are the potential to detect water beneath canopy or vegetation cover, the ability to provide stage data at ungauged locations, its use of a single fixed reference datum, and its potential ability to measure the depths of water within floodplains and wetlands [3].
Unfortunately, the footprint size of radar altimetry systems is quite large, in the order of a few hundred meters to a few kilometers, which requests for water bodies of similar dimensions to detect water. Consequently, this method is more suitable for the monitoring of large water bodies. In general, it appears that with larger water bodies the probability of a correct detection of water by radar increases. This results in more echo pulses and therefore better accuracy. Topex/Poseidon data in time-series between 1992-1999 is examined in the Amazon Basin [3]. This study resulted in water levels for river with a theoretical minimum width of a 580 to 1160 m with an overall accuracy of 1.1 m root mean square (rms), where the rms value is deduced from the mean difference between the observed height and the gauged height values over the observation period. The variety in accuracy depends on both target size and season. Best accuracy values (0.4-0.6 m rms) were obtained in the larger rivers which again highlights the relevance of the size of the target. Other satellite remote sensing studies executed in the Amazon Basin show reasonable results. The ERS-2 and ENVISAT satellites are evaluated and the overall performance of these systems results in accuracies of 0.7 and 0.3 m respectively [8].
Another satellite remote sensing technique is based on laser altimetry. The advantage of laser altimetry is the use of a laser beam instead of microwaves which results in smaller footprint resolutions. This will result in more correct water detections and therefor higher accuracies. The footprint resolution of 70 m and an along-track spacing of 175 m make ICESat a potential sensor for surface-water monitoring [18]. ICESat is evaluated over three coastal scenarios: continental coast, open ocean island, and an inland river [51]. The accuracy of water levels (rms) over the inland river (Tapajos River, Brazil) is 0.03 m under clear conditions, 0.08-0.15 m under partly clouded skies, and 0.25 m under heavy clouds. This is a great improvement compared to satellite radar altimetry although the different accuracies make ICESat not operable under all weather conditions.

2.2 MEKONG RIVER BASIN

The study area is limited to the Mekong River Basin. This section will first describe the physical geography of the study area. The Mekong River is split in six components for better understanding of the different hydrological regimes. A short remark about the political, social, and economical situation of the people living in the Mekong River Basin will be provided. Furthermore the meteorology and the vulnerability of the Mekong River Basin to climate change will be briefly discussed.

2.2.1 STUDY AREA

The Mekong River Basin (9°61' – 33°82' N, 93°86' – 108°78' E) is located in Asia and encompasses the Mekong River. With a length of approximately 4800 km, the Mekong River runs from the Tibetan Plateau through six countries: China, Myanmar, Lao PDR, Thailand, Cambodia and Vietnam (Figure 2.3). With a catchment area of 795,000 km^2 and a mean annual water discharge of 470 km^3, the Mekong River is one of the largest rivers of the world [24].

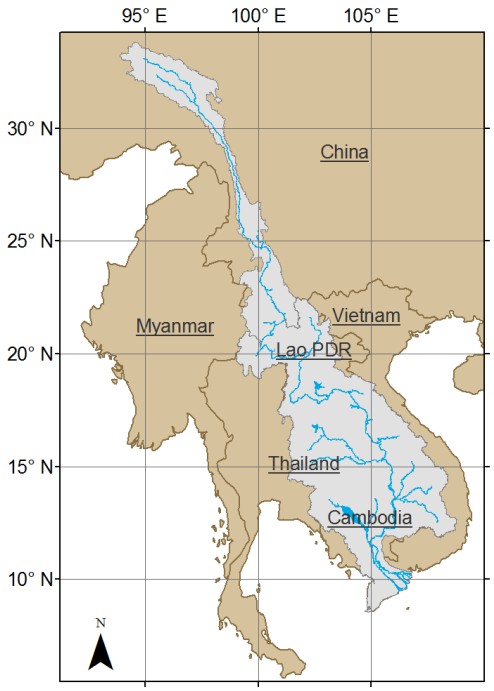


Figure 2.3: Six countries intersecting the Mekong River Basin.

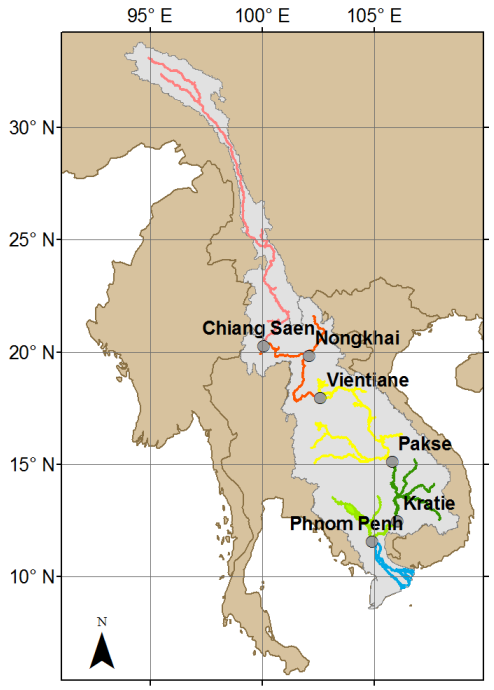


Figure 2.4: Six river components represented with different colors. Components are defined on river parts between cities.

The Mekong River Basin (total catchment area of 795,000 km^2) is split in the Upper Mekong Basin (24 percent of the catchment area) and the Lower Mekong Basin (76 percent of the catchment area). The mainstream in the Upper Mekong Basin (almost 2200 km) is characterized as steep and narrow. The mainstream decreases by nearly 4500 m in altitude before it enters the Lower Mekong Basin at the intersection of the borders from Thailand, Lao PDR, China and Myanmar [26]. The mainstream continues for the remaining 2600 km with only a decrease in altitude of 750 m. In the Lower Mekong Basin, the mainstream becomes wider and contains several tributaries resulting in a complex delta system before entering the South China Sea in Vietnam.

The Mekong River is divided in six components based on the hydrological regimes, physiography, land use, and resource developments [26]. The components are defined as the part of the river between cities. An overview of the components is presented in Figure 2.4. The Upper Mekong River is referred to as the Yunnan Component and originates from melting snow on the Tibetan Plateau. This component is responsible for 30 percent of the average dry season flow. The planning and development of dams in this component is of concern since this will influence the downstream hydrology. The second component is running from Chiang Sean to Vientiane and Nongkhai, the mainstream splits into two distributaries. The component is dominated by the Yunnan Component. However, the third component running from Vientiane to Pakse is increasingly influenced from the tributaries in Lao PDR and Thailand that enter the mainstream. The tributaries from Thailand have low runoff potential and significant reservoir storage for dry season irrigation. The tributaries from Lao PDR have potential for hydropower. The fourth component is located from Pakse to Kratie. This component is the largest hydrological component of the Lower Mekong Basin. Over 95 percent of the flow has entered the Mekong River in the fifth component from Kratie to Phnom Penh. This component is characterized by hydrodynamic complexity that is revealed by the Tonle Sap Lake system and extreme seasonal flooding. The Tonle Sap Lake system is the phenomenon that the direction of water movement is determined by the water levels of the Mekong River. The Mekong River is connected to the Tonle Sap Lake (Great Lake) and during the dry season the Tonle Sap Lake drains into the Mekong River. During the wet season, the water levels of the Mekong River rise above the lake level and results in an inflow in the Tonle Sap Lake [24]. After Phnom Penh, the mainstream alters into a network of branches and canals after which the river runs into the South China Sea. This distributary network is referred to as the Mekong Delta and also has yearly problems with flooding due to heavy rainfall.
A summary of the intersection of the six countries with the catchment in terms of catchment and water flow is provided in Table 2.1.

Table 2.1: Intersection of the six Mekong River Basin countries with the catchment, taken from the MRC [26].

	China	Myanmar	Lao PDR	Thailand	Cambodia	Vietnam
Area [km^2]	165,000	24,000	202,000	184,000	155,000	65,000
Catchment [%]	21	3	25	23	20	8
Water flow [%]	16	2	35	18	18	11

The Mekong River functions as major water resource for the population in Southeast Asia. In contrast to the Upper Mekong Basin, the Lower Mekong Basin is highly populated. Approximately 55 million people are living from the resources delivered by the Lower Mekong River. The population is assumed to increase to 90 million people in 2025 [24]. A low standard of living is maintained by a large part of population. The countries in the Lower Mekong Basin are characterized as 'medium human development' countries according the Human Development Index developed by the United Nations Development Programme (UNDP) [27]. The UNDP relates water resources to poverty by using access to safe water as indicator for the standard of living in developing countries. The percentage of the population without access to safe water is respectively 39 percent for Cambodia and 43 percent for Lao PDR [48]. More than 25 million people live in a 15 km corridor either side of the mainstream in the Lower Mekong Basin [27]. In Cambodia, this equals 70 percent of the population. Furthermore, 79 percent of the 15 km corridor population lives within 5 km of the mainstream. The 5 km corridor is predominantly present in Vietnam (51 percent) and Cambodia (34 percent).
The hydrological dependence in the Lower Mekong Basin has important consequences for the monitoring of water resources; the countries that rely mostly on the Mekong (Cambodia, Lao PDR and Vietnam) are located downstream. However, the basin is shared with other countries that do not have similar basic needs and will use water resources for other purposes than fresh water supply and sanitation.

For water governance in the Mekong River Basin, the Mekong River Commission (MRC) was established in 1995 with four member countries: Cambodia, Lao PDR, Thailand and Vietnam. The mission of the MRC is stated as: *To promote and coordinate sustainable management and development of water and related resources for the countries' mutual benefit and the people's well-being* [31]. Myanmar and China are the dialogue partners of the MRC. The dialogue with the two upstream countries is established and they participate in little basin initiatives. This is very important because during the dry season, 50 percent of the inflow of water in Lao PDR originates from China [31]. The dependence on China for water supply creates concern in the Lower Mekong Basin.
Furthermore, hydropower is a booming development in the Mekong River Basin. Although the construction of dams appears in each country, all eyes are on the cascade of dams in China of which four dams are already operational and belong to the largest dams in the world. The influence of damming in China on the downstream hydrological regime can be both positive and negative. The storage capacity of dams can be used to regulate extreme water levels and support irrigation during the dry season and prevent flooding during the wet season. Research has concluded that the impact of dams on the water level variations is only significant on small timescales [23]. China has shown some sympathy towards the other countries in 2002 to agree to share water level data in the wet season from two stations that contribute to the flood forecasting system that is managed by the MRC [29].

2.2.2 METEOROLOGY

The Lower Mekong Basin is characterized by a temperate to tropical monsoon climate which results in a wet season from June to October and a dry season during the rest of the year [24]. The wet season caused by the Southwest Monsoon is characterized by heavy rainfall and tropical storms. The rest of the year is dominated by lower temperatures and less rainfall caused by the Northeast Monsoon.
The average rainfall for the full Mekong River Basin is presented in Figure 2.5. The wet season results in an average rainfall of 1625 mm per year. The geographical distribution of rainfall in the Mekong River basin shows a distinct east-to-west gradient [30]. In the Lower Mekong Basin, most rain falls in the Eastern countries and less rain in the Western countries. Mean annual precipitation varies from 1000-1600 mm in Thailand, 600-1700 mm in the Northern Region to over 3000mm in Lao PDR and Cambodia. The left bank tributaries from Lao PDR generate more flow into the Mekong River because of more precipitation. When comparing the left and right bank tributaries, the flow contributions to the mainstream are respectively 60 and 24 percent. The remaining 16 percent comes from China. This implies that during the wet season, possible flow regulations caused by damming in China will have little influence on the hydrological regime.
The trend of the average water flow is related to the precipitation illustrated in Figure 2.6. The discharge of the Mekong River shows peak flows in August and September due to heavy rainfall. However, the magnitudes of the seasonal flows are fluctuating from year to year. Based on data derived at 10 mainstream sites, the average annual discharge is 14,500 m^3/s. Average peak flows result in 45000 m^3/s while the lowest water levels result in average minimum flows of 1500 m^3/s [24].
Heavy rainfall during the wet season as a consequence of tropical storms and cyclones has great impact on the rainfall climate. During the wet season, the number of tropical storms increases and is the cause of flooding, especially in the Lower Mekong River Basin where the slopes are generally flat.
The Mekong River Basin is vulnerable to climate change. While impacts of climate change are difficult to forecast, the possible consequences cannot be ignored. An increase in temperature will cause an increase in melting snow from glaciers at the Tibetan Plateau running into the Mekong River during the dry season. The southern catchments will face more droughts during the dry season. Furthermore, the total annual runoff is expected to increase with 21 percent. This will make especially the Lower Mekong River Basin vulnerable to meteorological flooding.

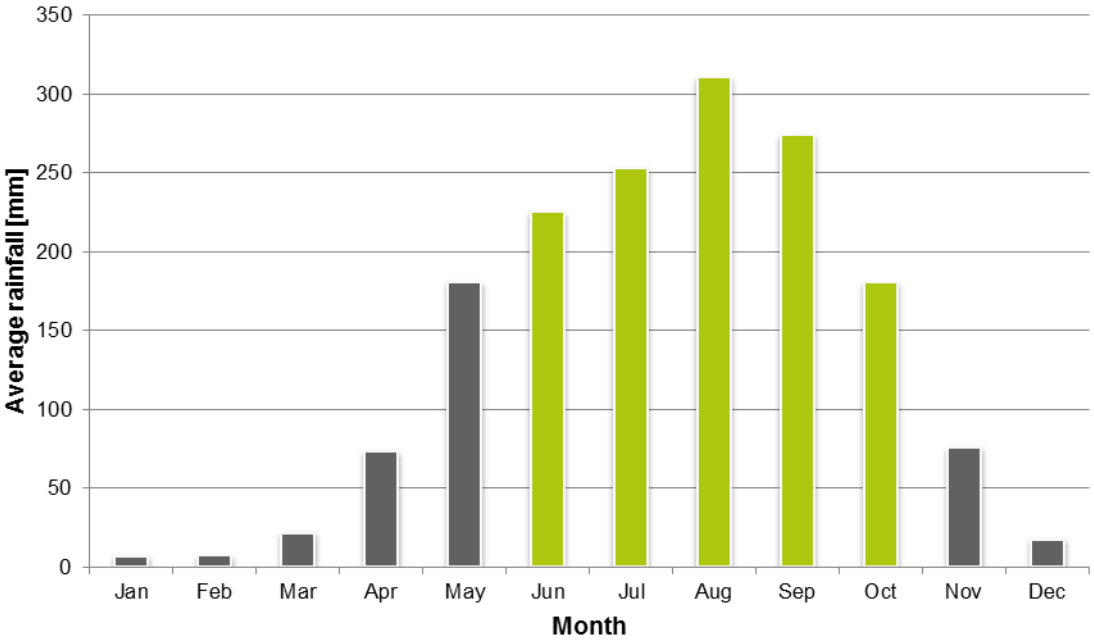


Figure 2.5: Longterm (1960-2004) monthly average rainfall, based on Mekong River Commission [26]. Average rainfall is based on six measurements (Chiang Rai, Pakse, Khon Kaen, Pleiku, Phnom Penh, and Chau Doc). Green indicates the wet season, grey indicates the dry season.

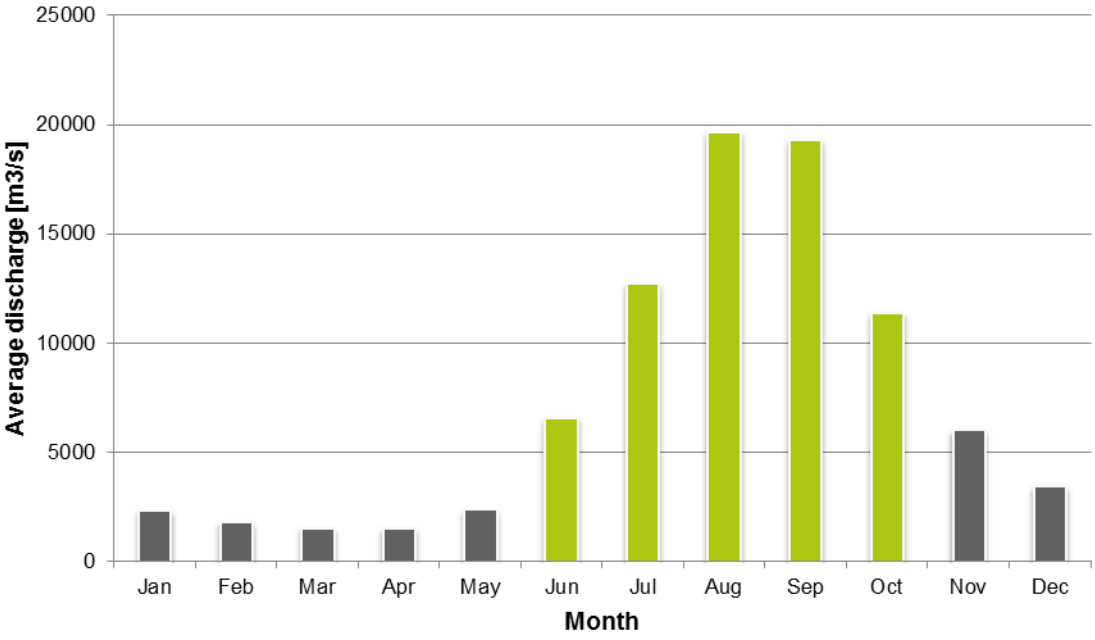


Figure 2.6: Longterm (1960-2004) monthly average discharge, based on Mekong River Commission [26]. Average discharge is based on six measurements (Chiang Saen, Vientiane, Nakhon, Mukdaha, Pakse, and Kratie) along the Mekong River. Green indicates the wet season, grey indicates the dry season.

3 THE DETERMINATION OF THE MEKONG RIVER LOCATION

This chapter will present the approach to determine the location of the Mekong River. In section 3.1, the data sets used to create a customized water mask of the Mekong River are described. The boundary of the basin and the Mekong mainstream are determined using the USGS HydroSHEDS data collection (section 3.1.1). The MODIS 250m land water mask (section 3.1.2) provides a global map of surface water. This data is necessary to identify water pixels within the Mekong River Basin that may overlap GLAS footprints. The classification is achieved by combining the two data sets to create a smooth river topology which will be discussed in section 3.2. The validation of the customized water mask is discussed in section 3.3. The quality of the geolocation of the water mask is evaluated with use of Landsat images (section 3.3.1). The findings based on visual inspection of the water mask over Landsat images will be presented in section 3.3.2.

3.1 MEKONG RIVER CLASSIFICATION DATA

Two data sets are used to determine the location of the Mekong River: the USGS HydroSHEDS data and the MODIS 250m land water mask. The sources of the data sets, the method to process the data sets to usable information and considerations for the study of river networks will be discussed.

3.1.1 USGS HYDROSHEDS

Data
The USGS HydroSHEDS collection is based on Shuttle Radar Topography Mission (SRTM) data and results in a collection of both vector and raster data sets providing hydrographic information at various scales with global coverage [49]. Examples of the data that is provided by the USGS are Digital Elevation Models (DEM), drainage basins, flow directions and flow accumulations, and river networks with resolutions varying between 3 arc-second (approximately 90 m at the equator) and 5 minutes (approximately 10 km at the equator. The data is accessed with a viewer to select the area of interest at the USGS HydroSHEDS Download Site (Figure 3.1).

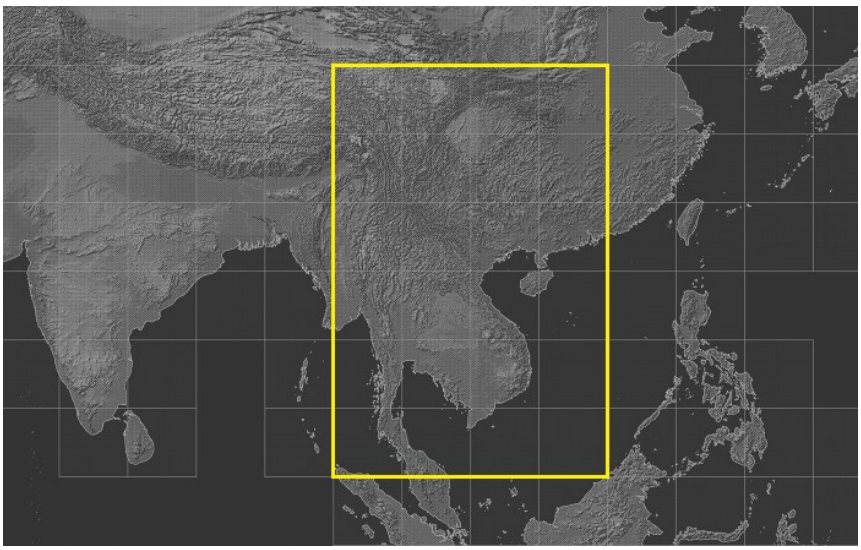


Figure 3.1: USGS HydroSHEDS, selecting tiles covering the area of interest (yellow box).

Three USGS HydroSHEDS data layers are used in this study: drainage basin (catchment as described in section 2.1.1) and river network and the digital elevation model (DEM). A summary of the data layers is presented in Table 3.1. The USGS HydroSHEDS data layers are distributed based on continental extent which implies some simple processing in ArcGIS to select the area of interest.

Table 3.1: USGS HydroSHEDS data layers.

Data layer	Drainage basin	River network	Digital elevation model
Data label	BAS	RIV	DEM
Data format	Vector (polygons)	Vector (lines)	Raster
Values	Surface area [km^2]	Flow accumulation [upstream cells]	Elevation [m], WGS84/EGM96
Projection	WGS84 datum	WGS84 datum	WGS84 datum
Resolution	15 arc-second	15 arc-second	15 arc-second

Only the drainage basin and river network are used for the customization of a water mask. The DEM is used in section 6.3 for creating downstream river profiles.
The drainage basin layer provides a vector format consisting of polygons corresponding to the watersheds in Asia. The layer is referenced to the World Geodetic System 1984 (WGS84) horizontal datum in latitude/longitude with 15 arc-second (approximately 500 m at the equator) resolution. For the purpose of this study, only the Mekong River Basin is needed for further analysis. The Mekong River Basin is selected and exported as a feature. The layer contains the surface area as attribute, which corresponds to 774272 km^2 for the Mekong River Basin.
The river network layer is derived from the drainage directions and flow accumulations as defined in the USGS HydroSHEDS Technical Documentation [22]. The drainage directions define the direction of flow from each cell in the conditioned Digital Elevation Model (DEM) to its steepest down-slope neighbor. The unconditioned DEM differs from the original DEM particularly by the filling of no-data cells by interpolation. The flow accumulation is defined as the amount of upstream area (in number of cells) draining into each cell. This implies that an increase in flow accumulation corresponds with a cell located more downstream in the catchment. Rivers are calculated as pixels with more than 100 upstream cells. The river network layer is in vector format with lines representing river components with the maximum flow accumulation as attribute. The river network is also referenced to the WGS84 datum with 15 arc-second resolution. As mentioned before, the data layers are distributed based on continental extent which results in a dense network of all rivers in Asia.

Method
The Mekong River is obtained in two steps: First, the rivers located within the Mekong River Basin are selected by restricting the USGS HydroSHEDS river network layer to the USGS HydroSHEDS drainage basin of the Mekong (Figure 3.2). Second, a threshold for the number of upstream cells is used to select river components belong to the mainstream of the Mekong River. The stream determination based on a threshold is ambivalent: A small threshold will return streams that do not even exist or may not belong to the mainstream. A large threshold might disconnect the stream and remove streams that are of relevance to the mainstream. As a rule of thumb, 1 percent of the maximum flow accumulation can be used as threshold [52]. The maximum flow accumulation is obtained from the USGS HydroSHEDS river network and equals 14,353,619 upstream cells. A threshold of 143,536 upstream cells is however not suitable for this study due to the disconnection to large lakes and reservoirs. This is illustrated in Figure 3.3. The river network with a threshold of 143,536 upstream cells disconnects the mainstream from the Nam Ngum Dam Reservoir in Lao PDR. The Nam Ngum Dam Reservoir is a 250 km^2 reservoir which is connected to the Mekong River by the Nam Ngum River. The information about the Nam Ngum River is lost with the larger threshold. Consequently, a smaller threshold is chosen of 50,000 upstream cells. This threshold is somewhat arbitrary and is specifically chosen for the Mekong River.

A complete overview of the differences between the two tested thresholds for the complete basin is illustrated in figure 3.4. Especially in the north of the Upper Mekong Basin, the use of a lower threshold results in better coverage of the mainstream.

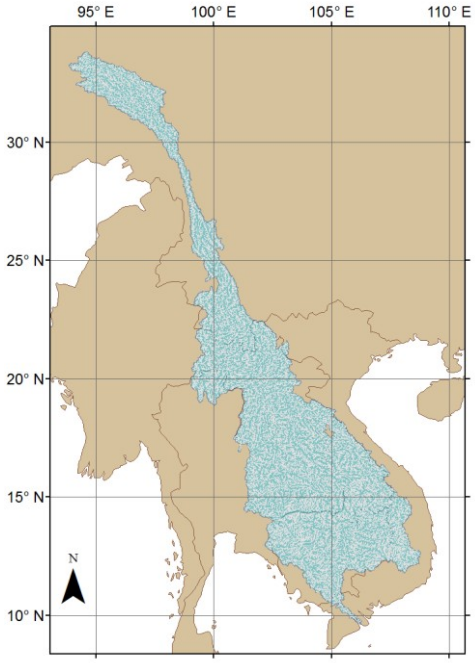


Figure 3.2: USGS HydroSHEDS drainage basin and river network in Mekong River Basin, streams with standard threshold of 100 upstream cells.

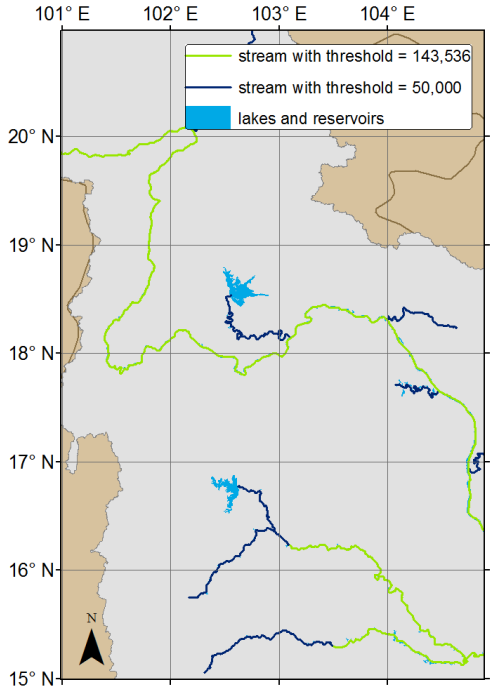


Figure 3.3: Example of disconnected streams with reservoirs, Nam Ngum Dam is the upper turquoise area.

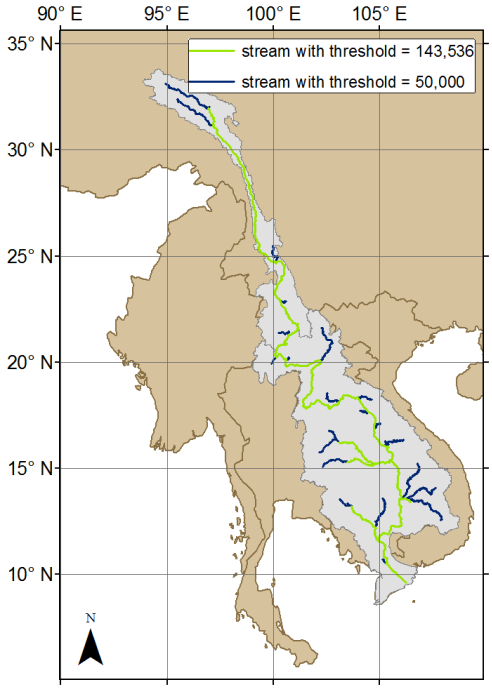


Figure 3.4: USGS HydroSHEDS derived Mekong River with thresholds of upstream cells, coverage of complete Mekong River Basin.

Considerations

The use of USGS HydroSHEDS data results in considerations for further research. The most important limitation of the drainage basin is the insufficient representation of the Mekong Delta (Phnom Penh to the South China Sea). The drainage basin does only include the mainstream and does not cover all branches and canals. In the technical documentation it is stated that the drainage basins are not yet finalized [22]. Indeed, a closer look at the river network showed a disconnection between the mainstream and the upper branch in the Mekong River Delta, probably due to a lower value than the standard threshold of 100 upstream cells. The drainage basin is therefore not correctly estimated and will result in loss of relevant GLAS footprints when using this basin for spatial selection. The search for a better drainage basin was thus incorporated in this study. A drainage basin created by the International Water Management Institute (IWMI) provided a more sufficient outlet of the river into the South China Sea [20]. This drainage basin is used instead to select the river components from the USGS HydroSHEDS river network that intersect the Mekong River Basin. In Figure 3.5 the dissimilarity between the USGS (grey area) and IWMI (turquoise area) basins is clearly visualized. The branches (green lines) that are excluded by the USGS basin are a large part of the Mekong River Delta.

A limitation of the use of river networks in general is the partial information about the extent of the river. It should be noted that based on flow accumulations only information about the downstream profile of the Mekong River is provided. The width of the river is still unknown. This becomes clearly apparent when looking at the Tonle Sap Lake in Cambodia (Figure 3.6). This lake is represented with a line element located on the downside of the lake. Additional data for the identification of complete water bodies is therefore necessary.

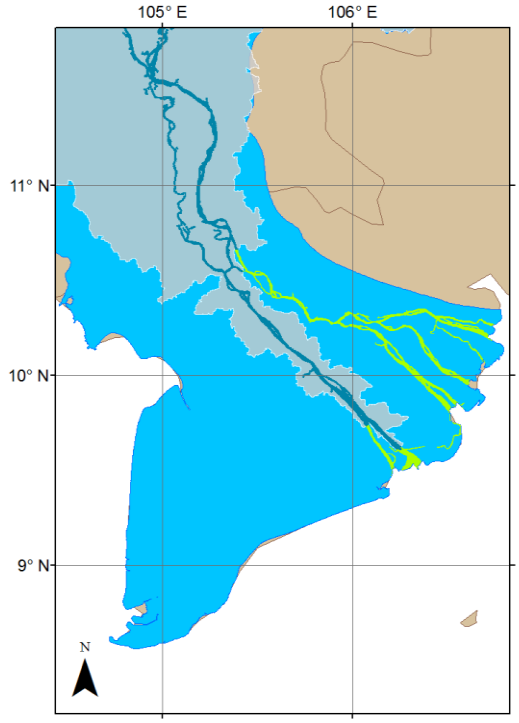


Figure 3.5: USGS drainage basin compared with IWMI drainage basin, distributaries in the Mekong River Delta (green lines) are excluded by USGS.

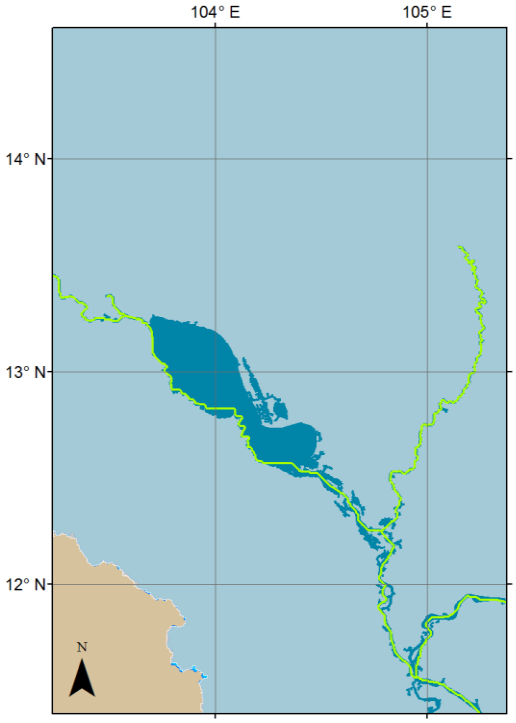


Figure 3.6: Absence of width information in USGS river network, Tonle Sap Lake (Cambodia), therefore complementing data is necessary.

3.1.2 MODIS 250M LAND WATER MASK

Data
The MODIS 250m land water mask (MOD44W) is created from the Shuttle Radar Topography Mission (SRTM) water body data supplemented with MODIS (MOD44C) 250m resolution data [7]. The data is provided by the Global Land Cover Facility (GLCF) which delivers the data in GeoTIFF format which is a raster, referenced to the WGS84 datum, and a tiling system suitable in combination with Landsat imagery [16]. The MOD44W raster has values related to the land cover type and an additional category used for values outside the projection (Table 3.2).

Table 3.2: Values for the MOD44W raster.

Value	Label
0	Land
1	Water
255	Fill

The MOD44W raster has a resolution of 250 m, which is not very appropriate for the purpose of identifying rivers due to their variation in width. A narrow river will not be apparent in the MOD44W raster. The coarse resolution results in gaps in the land water mask (due to smaller river width), while a river should in theory appear as a connected feature. When using this data to select GLAS footprints corresponding with water, many footprints might be falsely excluded. A recommendation for research towards rivers is therefore to use classification data with appropriate spatial resolution.
The MOD44W raster covering the complete Mekong River Basin is presented in Figure 3.7. Note that the water bodies are included in contrast with the river network described in the previous section. However, the Upper Mekong Basin contains barely any water pixels because of the narrow river width.

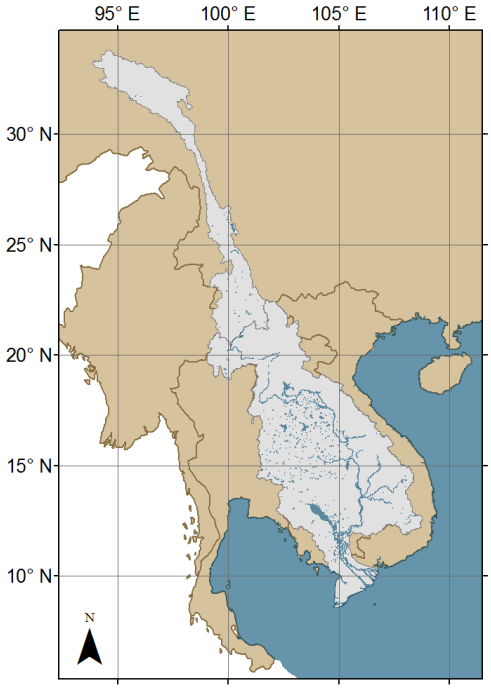


Figure 3.7: MODIS 250m land water mask data in the Mekong River Basin.

Method
The MOD44W raster is processed for the purpose of spatial selection in two steps: first the raster data is converted to vector format. This will enable the intersection with the GLAS footprints. Based on visual inspection, a geolocation error occurs of a maximum of 250m. Comparing this error with the spatial resolution makes the MOD44W raster not very appropriate to use for selecting streams. The objective of the land-water mask is yet the identification of GLAS candidate river footprints and the priority is therefore to include as many footprints as possible. The second step is therefore to buffer the MOD44W vector with 250m. As illustrated in Figure 3.8, the geolocation error is now not resulting in loss of data.

Considerations
The buffering of the MOD44W vector has consequences for further processing. Although no footprints will be falsely excluded because of the buffer (decrease of type II error), the probability that more footprints that are further away from the river zones are included is larger (increase of type I error). Type I and II errors are discussed in more detail in section 5.2.4. In this study, the purpose of the spatial selection is to identify candidate river footprints and then this buffering can be considered as an adequate method to counterbalance the geolocation error.
Another consequence is the loss of information about small islands and neighboring tributaries with a mid-channel bank of less than 500 m. The buffer of 250 m might enclose the islands and mid-channel banks and presume the overlapping footprints as candidate river footprints. This is inefficient because in a later stage these footprints will be removed again. In Figure 3.9, the two squares in the circle highlight the case of removed features due to the buffering. An advantage is the closing of feature gaps in the MOD44W vector that contributes to a connected river feature.
To conclude, a comment on the use of spatial selection for the identification of river footprints is necessary. For river networks, only spatial selection based on a land water mask will mostly not be appropriate because of the variety in a river's width. Only for wide rivers it might be sufficient. Then the land water mask can be narrowed with a negative buffer instead and the loss of data occurs only near the river banks.

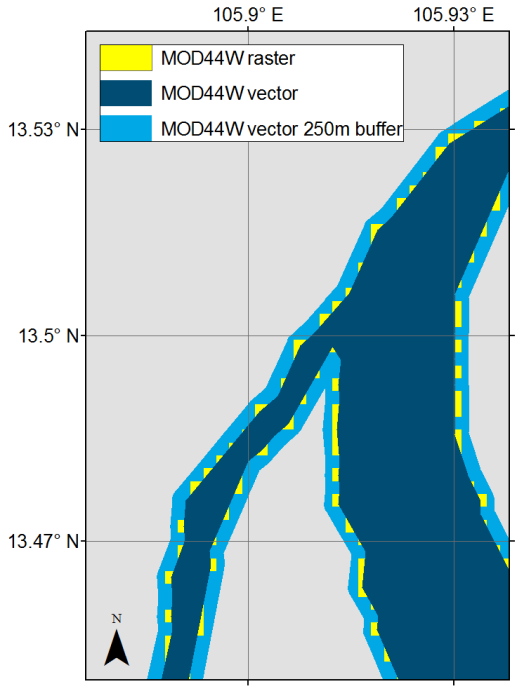


Figure 3.8: MOD44W raster (yellow), converted to feature format (blue) and enhanced feature with 250m buffer (turquoise) to cover original data.

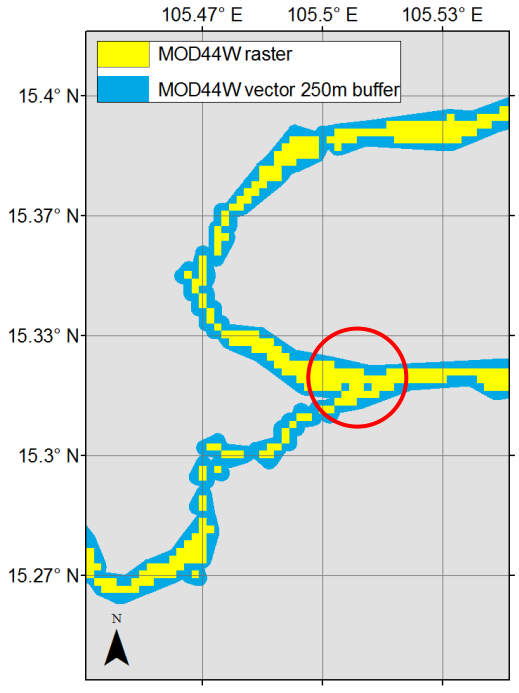


Figure 3.9: MOD44W raster (yellow), and enhanced feature with 250m buffer (turquoise) closes feature gaps but also islands and narrow mid-channel banks.

3.2 CUSTOMIZED WATER MASK OF THE MEKONG RIVER

The selection of footprints corresponding to rivers based on solely the MODIS 250m land water mask (see section 3.1.2) is insufficient due to the gaps in the geometry caused by the coarse spatial resolution. On the other hand, for information about the river's width, the USGS HydroSHEDS river network (see section 3.1.1) cannot be used as standalone dataset. To create a connected feature, the MOD44W needs to be enhanced to ensure that footprints are not falsely ignored. With use of a combination of the MOD44W with the USGS HydroSHEDS River Network, all footprints within a distance of 500 m from the mainstream of the Mekong River can be selected. This requires some processing in ArcGIS.
First, the MOD44W is converted to a vector to work with consistent representations. However, due to geolocation errors that occur with the conversion, a buffer of 250m is used to compensate data losses. The MOD44W has a spatial resolution of 250m and an absolute geolocation accuracy of approximately 9m, which makes this enhancement very inefficient but necessary to obtain all candidate river footprints in the Mekong River [7]. The USGS HydroSHEDS River Network is already represented as a vector (polylines). Since only the mainstream is important for the analysis, all river components with more than 50,000 upstream cells are included. Note that the river network provides only vectorized river components, which gives no information about the river's width. To connect the USGS HydroSHEDS River Network to the MOD44W data to fill the gaps, a buffer of 500 m is used. This will be sufficient since the gaps only occur at places where the river's width is less than the spatial resolution of the MOD44W (250 m). Even if the line is positioned near the boundary of the MODD4W pixels, the river pixels can still be identified (Figure 3.10).

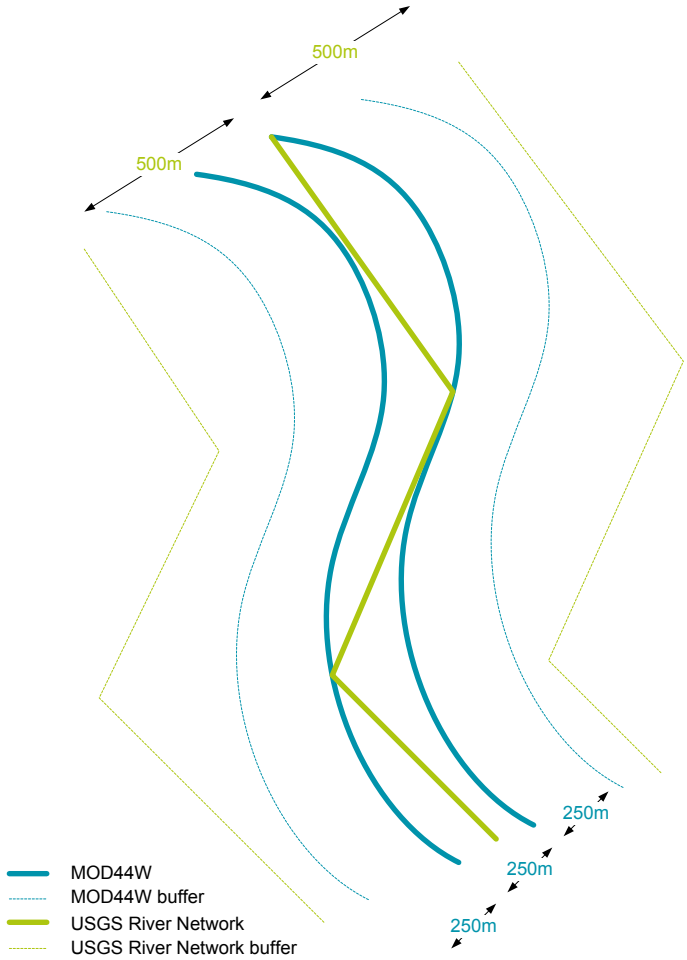


Figure 3.10: Buffer sizes for MOD44W and USGS River Network.

The next step is to aggregate polygons of MOD44W (including the buffer) that are within a distance of 500 m from each other. In other words, polygons that are supposed to be connected will be joined. This will reduce the number of polygons of MOD44W. An iterative process is executed to include as many possible footprints corresponding with water. The idea is to enlarge the river polygon with MOD44W in every step to connect river components belonging to the mainstream and filling gaps. First, all aggregated MOD44W polygons that intersect the buffered USGS River Network are selected. The selection is merged with the USGS HydroSHEDS River Network. The gaps are filled by dissolving the vector. Now this process of intersect>merge>dissolve can be executed with only MOD44W polygons that intersect with the product of the first iteration are selected. After 3 iterations, most of the MOD44W polygons that belong to the mainstream are connected and the gaps are filled.

3.3 VALIDATION OF CUSTOMIZED WATER MASK

Landsat images are used to evaluate the customized water mask of the Mekong River. First, satellite images from the Landsat program are introduced. Second, the quality of the customized water mask is evaluated by visual inspection.

3.3.1 LANDSAT IMAGES

The Landsat program is launched in 1972 and collects continuous and consistent spectral information from the Earth's surface until today [35]. Nowadays, the archive of Landsat images is immense and freely distributed with help of the USGS Global Visualization Viewer (GLOVIS) [43]. Seven Landsat programs have been launched from which only the Landsat 6 program failed to reach orbit. The GLAS data is recorded between 2003-2009; the Landsat 7 satellite that was launched in 1999 provides the most accurate images. Unfortunately, in 2003 the scan line corrector (SLC) failed which results in stripes in the data. Therefore for this study the Landsat 4-5 Enhanced Thematic Mapper (ETM) collection was used instead.
The TM sensor supplies six spectral bands with a resolution of 30 m and a thermal band of 120 m (Table 3.3). Band 1, 2 and 3 detect respectively visible blue, green, and red wavelengths. Band 4, 5, and 7 cover the infrared (IR) wavelengths and are distinguished in near IR, mid IR and far IR. Band 6 shows the thermal reflection.

Table 3.3: Landsat 5 TM sensor characteristics.

Band	Spectral range (nm)	Ground resolution (m)
1	0.45-0.52	30
2	0.52-0.60	30
3	0.63-0.69	30
4	0.76-0.90	30
5	1.55-1.75	30
6	10.40-12.50	120
7	2.08-2.35	30

The Landsat ETM images are referenced with the Worldwide Reference System (WRS) that is basically a Path/Row indicator. The Landsat TM sensor has a global coverage of 199 rows and 233 paths. The WRS is referenced to the WGS84 to obtain the images that intersect the Mekong River Basin (Figure 3.11). The Landsat images that intersect the Mekong River Basin are included. This results in 65 Landsat images.
With GLOVIS, Landsat ETM images can be searched based on the Path/Row index and the temporal information of the scene. However, the ETM images are gathered between 1999 and 2003 and do not match the operational periods of the GLAS instrument. Because the extent of rivers is depending on the monsoon season, this is a weakness in the use of Landsat images. Nonetheless the Landsat images will be sufficient for estimating the

location of the Mekong River within the customized water mask as well as the classification of river footprints in section 5.1.3.
Combining three spectral bands and assigning the colors red, green and blue (RGB) to the selected band will result in a color composite. The true-color composite is RGB = 321, which uses the three bands that detect the visible light (Figure 3.12). This can be interpreted as assigning band 3 to the color red (R), band 2 to the color green (G) and band 1 to the color blue (B). The purpose of this study is to detect water and therefore a false-color composite can support better interpretation of the Landsat images. With a false-color composite features of interest can be better distinguished in the image. For this study, the false color-composite that is used is RGB = 541 (Figure 3.13) [50]. Water bodies absorb the light of near IR (band 4) and mid IR (band 5) and therefore result in darker colors.

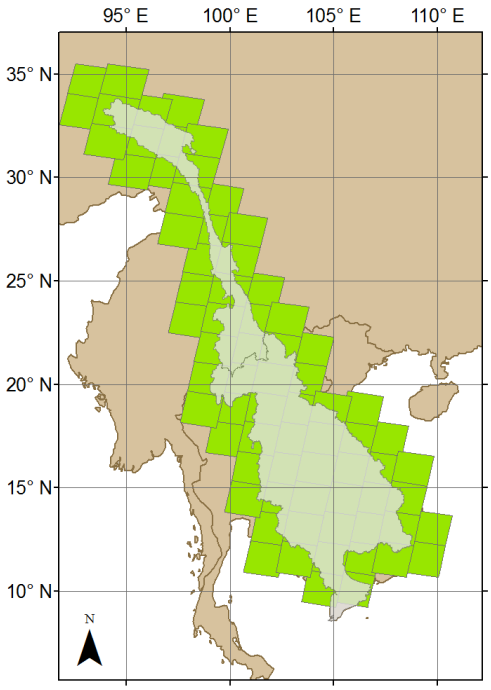


Figure 3.11: Landsat ETM images referenced to WGS84, intersecting the Mekong River Basin. The Path/Row indices are used to select specific Landsat images.

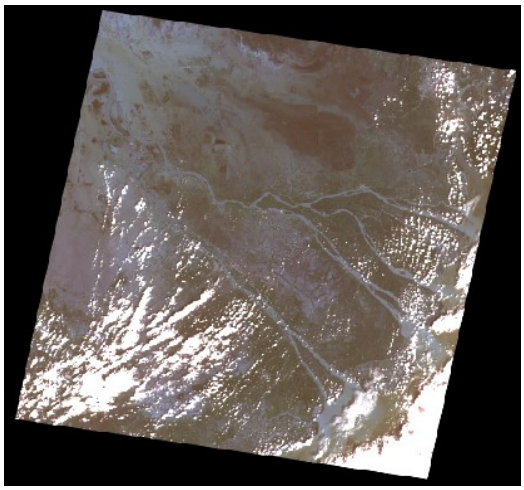


Figure 3.12: Landsat image (PRS 135053) original color composite (RGB = 321), located in Mekong River Delta.

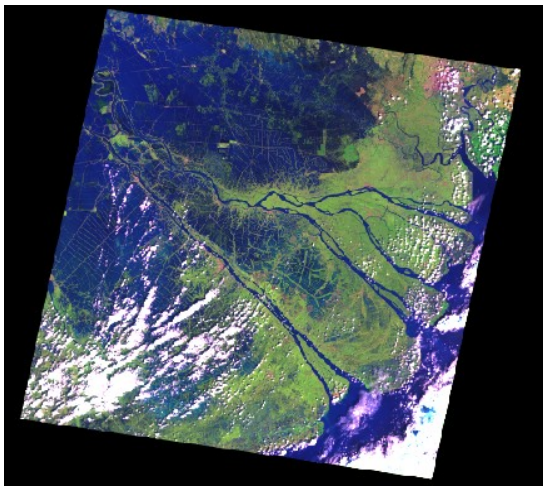


Figure 3.13: Landsat image (PRS 135053) false color composite (RGB = 541), located in Mekong River Delta.

3.3.2 CUSTOMIZED WATER MASK OVERLAPPING LANDSAT IMAGES

The customized water mask of the Mekong River is drawn over the Landsat ETM images. The water mask is evaluated by visual inspection for coverage of the river. Because of the buffer of 500m, the Mekong River is indeed completely covered by the custom water mask. The river is especially validated in areas were the MODIS 250 m land water mask contains gaps. It is necessary to check whether this problem is solved by the addition of the USGS HydroSHEDS river network. This is illustrated in Figure 3.14 and 3.15 for two areas. In Figure 3.14, the gaps in the MOD44W raster (turquoise) are solved by the water mask. Furthermore, the customized water mask (yellow boundary) covers the Mekong River as represented by the Landsat 5 ETM image (dark blue). In Figure 3.15, islands in the Mekong River Delta are preserved by the customized water mask. In the southwest stream, some small islands are ignored.

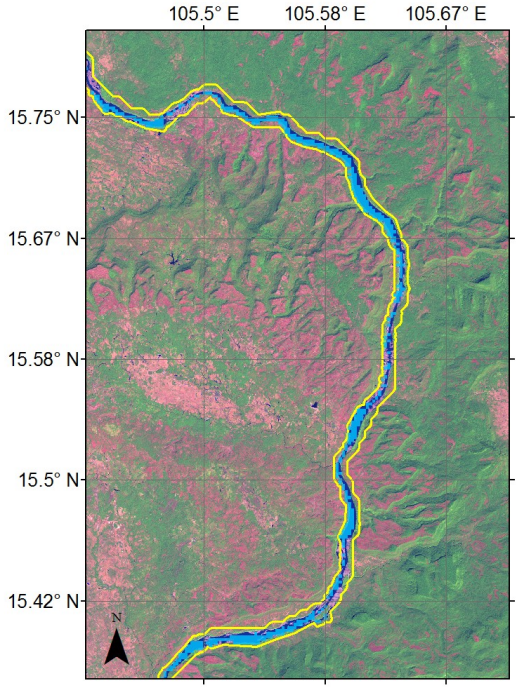


Figure 3.14: Customized water mask (yellow boundary) over Landsat 5 ETM image covers the Mekong River (dark blue) and MOD44W raster with gaps (turquoise).

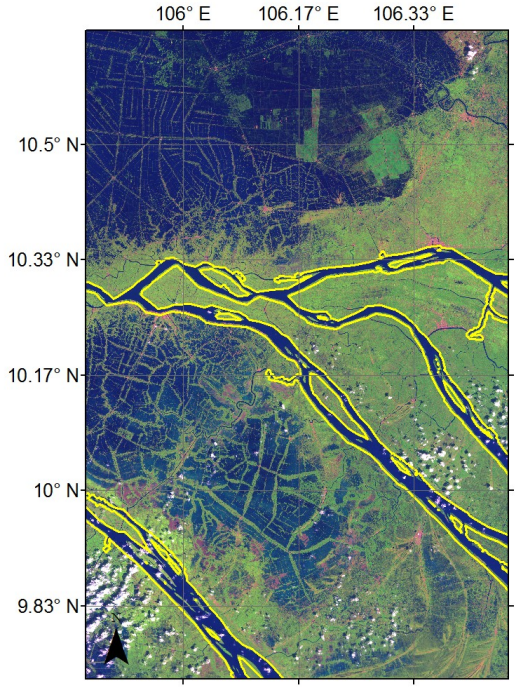


Figure 3.15: Customized water mask (yellow boundary) over Landsat 5 ETM image cover the Mekong River in and preserves islands in the Mekong River Delta.

Unfortunately, visual inspection identified an error in the USGS HydroSHEDS River Network. The complete left branch of the Mekong River Delta is not included because of a gap in the data set (Figure 3.16). The cause of this gap is unknown but is probably related to the threshold of 100 upstream cells to include a component to the USGS River Network. The left branch needs to be corrected manually. This is done by merging the MODIS 250m land water mask left branch with the feature created after the iterations.
The result of the ArcGIS processing is showed is Figure 3.17. The representation consists of a single feature covering the Mekong River mainstream. This includes all footprints that possibly correspond with water. The USGS HydroSHEDS River Network results in added value in the Upper Mekong Basin where the river is narrow and not apparent in the MOD44W. The MOD44W comprehends lakes and reservoirs that cannot be derived based on the vectorized river components as provided with the USGS HydroSHEDS data. These applications make both data sets mandatory and mutual complementary for river identification.

The drawback of using these data sets is the introduction of uncertainty about the river's geolocation. In the Upper Mekong Basin, the MOD44W is almost non-existent due to the narrow width of the Mekong River. A profile can be extracted with the USGS HydroSHEDS River Network under the assumption that the river is within 500 m of the line representation. This will include false positive footprints, in terms of footprints that do not correspond with water. The same happens with the MOD44W. Although the MOD44W is very accurate by itself, after including a buffer and aggregating nearby polygons, also footprints that not correspond with water will be included. Furthermore, islands located in channels smaller than 500m are ignored due to the maximum buffer size of 500m. However, for this step the focus is to rather include footprints corresponding with land than ignoring footprints corresponding with water. For future research in which river identification is necessary, it is recommended to consider the use of a land-water mask with smaller resolution.

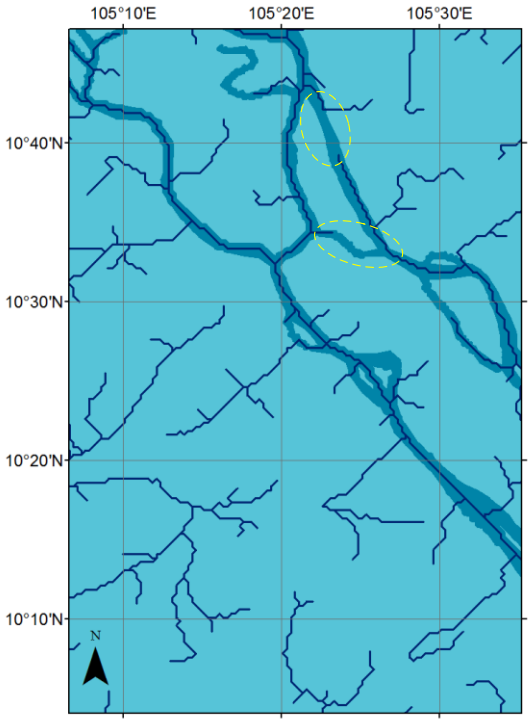


Figure 3.16: Errors in USGS HydroSHEDS River Network (RIV) causes absence of the complete left branch in the Mekong River Delta.

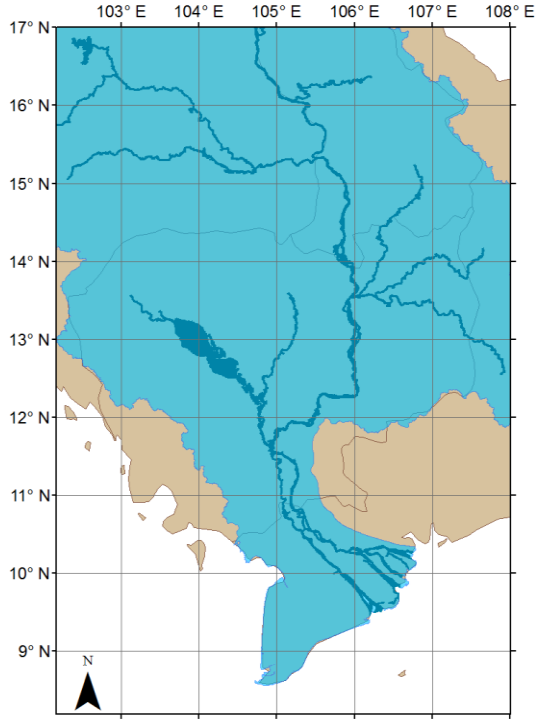


Figure 3.17: Final vector representation of Mekong River based on a combination of USGS HydroSHEDS, IWMI and MOD44W data.

4 PRINCIPLES OF SATELLITE LASER ALTIMETRY

This chapter will give information about satellite laser altimetry, the platform (ICESat) and sensor (GLAS instrument). With the ICESat mission, a space-based laser scanning system was established. An overview of the ICESat mission and its objectives will be given in section 4.1. The GLAS instrument carried aboard on ICESat and the operational history will be described in section 4.2. In section 4.3, the concept of satellite laser altimetry will be elaborated. This will start with a short introduction to laser scanning continue to full waveform laser scanning by GLAS and the interaction of the laser pulse with the surface type (more specific water bodies). In section 4.4, the use of the GLA14 data product to obtain variables will be described. This will result in the input data set for GLAS derived water level estimations of the Mekong River.

4.1 ICESAT MISSION

The ICESat satellite mission was launched on January 12, 2003 by the National Aeronautics and Space Administration (NASA). The ICESat mission is part of the Earth Observation System (EOS) which can be referred to as a coordinated series of polar-orbiting and low-inclination satellites for long-term global observations of the land surface, biosphere, solid Earth, atmosphere and oceans [33]. The ICESat mission aims to contribute to understanding of the Earth's climate and ultimately predict the impact of future climate change on ice sheets and sea level. The objectives of the ICESat mission are explicitly described in the GLAS Science Requirements [47]. The science goals of the ICESat mission are referring to the polar ice sheets, land surface, and the atmosphere of the earth's system. This study focuses on surface-water and therefore only the description of the science goals with respect to land processes is relevant and is described as follows:

Land processes: *"to conduct topographic measurements of the Earth's land surface on a global basis in order to contribute to a global grid of ground control points for georeferencing of topographic maps and digital elevation models. The secondary land processes science goal is to detect topographic change at the meter per year level or better in selected regions of limited spatial extent."*

The ICESat mission is primarily designed based on the cryosphere requirements. This will have some consequences for the specific application of monitoring water levels in river networks that will become clear in this study. Nonetheless, the topographic measurements and topographic change as objectives described for land processes correspond with the monitoring of water level estimations in the Mekong River.
Although ICESat is decommissioned since August 14, 2010, a second-generation satellite (ICESat-2) is scheduled for launch in 2016. The instrument will have different specifications such as a much higher pulse rate (10 kHz instead of 40 Hz) resulting in a denser along-track spacing (70 cm instead of 175 m) [34]. Lessons learned from this study may already provide some inside knowledge for further research about satellite laser altimetry for hydrological applications.

4.2 GLAS INSTRUMENT

The ICESat mission is known as the first Earth-orbiting satellite based on laser scanning measurements. The GLAS instrument was a laser altimeter having three identical lasers aboard, operated one at a time. The laser was designed to generate along-track surface profiles with repeated ground tracks in 183-day cycles. Each laser produced 1064nm pulses at an initial energy level of 75mJ for the altimeter measurements and additional

532nm pulses at 35mJ to provide atmospheric backscatter measurements which are used for the study of the vertical distribution of clouds and aerosols [12].
The first campaign was started at February 20, 2003 but was prematurely discontinued due to the failure of laser 1 on March 29, 2003. This has led to a modified operation plan with 33-day to 56-day operation periods, three times per year, to extend the mission life. An overview of the campaigns and the operational periods is given in Table 4.1. During the campaigns, ground tracks were repeated in 8-day and 91-day cycles. After the failure of the remaining lasers (laser 3 failed on October 19, 2008 and laser 2 on October 11, 2009), the satellite was decommissioned in 2010.

Table 4.1: ICESat campaigns and operational periods, taken from NSIDC [42].

Campaign	Start date	End date	Days in operation
L1A	2003-02-20	2003-03-29	38
L2A	2003-09-24	2003-11-19	55
L2B	2004-02-17	2004-03-21	34
L2C	2004-05-18	2004-06-21	35
L3A	2004-10-03	2004-11-08	37
L3B	2005-02-17	2005-03-24	36
L3C	2005-05-20	2005-06-23	35
L3D	2005-10-21	2005-11-24	35
L3E	2006-02-22	2006-03-27	34
L3F	2006-05-24	2006-06-26	33
L3G	2006-10-25	2006-11-27	34
L3H	2007-03-12	2007-04-14	34
L3I	2007-10-02	2007-11-05	37
L3J	2008-02-17	2008-03-21	34
L3K	2008-10-04	2008-10-19	16
L2D	2008-11-25	2008-12-17	23
L2E	2009-03-09	2009-04-11	34
L2F	2009-09-30	2009-10-11	12

The GLAS instrument operated at an altitude of 600 km with an inclination of 94 degrees. The laser produces 40 pulses per second, resulting in an along-track spacing of approximately 175 m between two adjacent footprints on the Earth's surface. The across-track spacing between two ascending or descending tracks is a function of latitude as the ICESat mission was primarily designed for studies in the cryosphere. The spacing is therefore smaller near the poles (approximately 5 km) than near the equator (approximately 30 km). The resolution of the footprints is illustrated in Figure 4.1.
The GLAS instrument provides 15 data products, which are distributed by the National Snow and Ice Data Center (NSIDC). During this study, the GLAS L2 Global Land Surface Altimetry data (GLA14) will be used, which comprehends surface elevations for land including the footprint geolocation and reflectance, as well as geodetic, instrument, and atmospheric corrections for range measurements [41]. The GLA14 data product includes waveform parameters that are already an initial product from pre-processing by the NSIDC [19]. This parameterization based on a waveform fitting procedure is described in an Algorithm Theoretical Basic Document (ATBD) provided by the NSIDC [6].

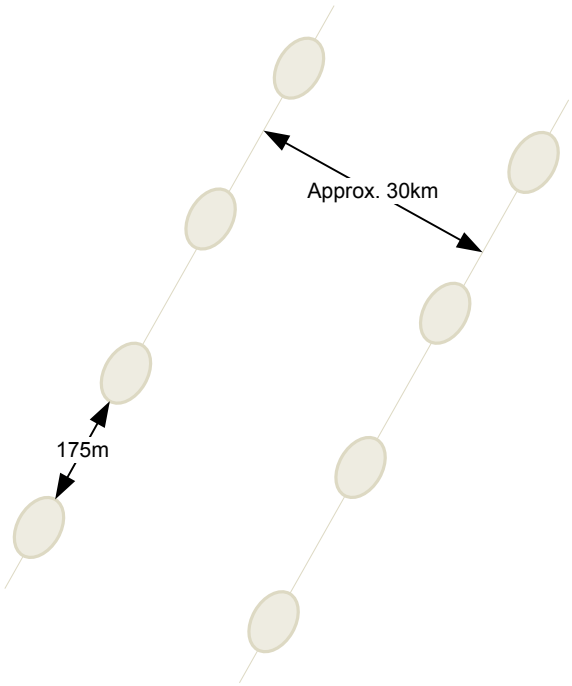


Figure 4.1: Along-track and across-track spacing applicable for the Mekong River Basin.

4.3 SATELLITE LASER ALTIMETRY

In this section, laser altimetry is briefly introduced. The waveform fitting procedure based on Gaussian curve fitting as described by the NSIDC is discussed for the purpose of understanding the waveform parameters included in the GLA14 data product (see section 4.4). The referencing of the waveforms to obtain the GLA14 height point is described in more detail as well as the interaction of the waveform with water surfaces.

4.3.1 LASER SCANNING

Laser altimetry is a method of measuring the height or elevation of the terrain from a platform such as a satellite [15]. The instrument used is a laser altimeter. The laser altimeter will emit pulses of laser light to the earth's surface where the pulse will be partly reflected. The instrument measures the time of pulse return. The time of a round trip can be multiplied by the speed of light to obtain the two-way distance [55]. The range measurement, or the distance between the instrument and the surface, can be obtained as half the round trip distance.
The range measurements provided by the laser altimeter is a scalar. However, in order to obtain range measurements with respect to a reference frame, the position of the sensor and the laser pointing direction should be known [15]. This will be explained in detail for referencing of full waveforms in section 4.4.
The laser pulse has a certain width. For example, the GLAS instrument produced illuminated spots (footprints) of approximately 70m–diameter on the ground. It is imaginable that these footprints will include scatterers of different height and material. When the return signal is described as a function of time, this will lead to multiple return echos. Full waveform systems are capable of recording the full return signal and thereby provide information about the vertical structure of the footprint. An example is given in Figure 4.2. When the laser signal intercepts a tree, the first return echo corresponds with the canopy height and the last return echo represents the ground surface.

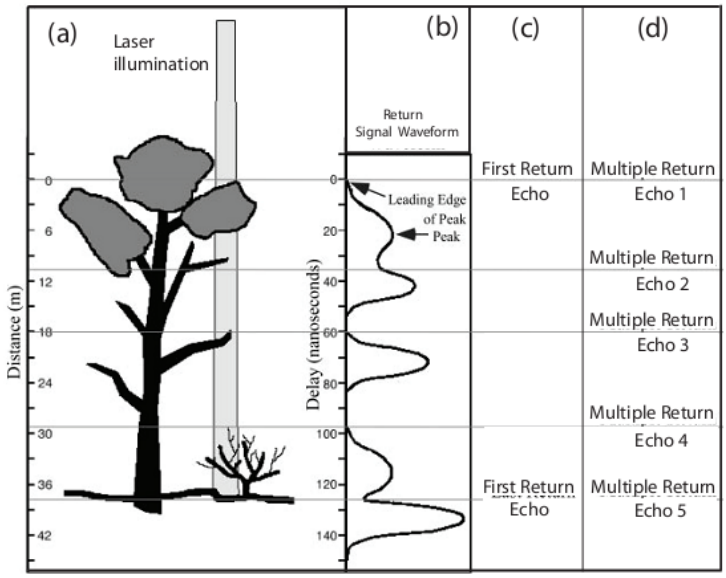


Figure 4.2: Multiple return echos. (a) Laser signal intercepting at different height levels. (b) Return signal as a function of time. Breakdown of the return signal into different echos: (c) Only first and last echo and (d) multiple return echos. Figure taken from Duong (2010) [12].

4.3.2 FULL WAVEFORM LASER SCANNING

Waveform fitting

The full waveform that is recorded by the GLAS instrument represents the return energy as a function of time. In the GLA14 data product, the return signal of the full waveform is described by a mathematical model based on Gaussian decomposition: each return echo is considered to correspond to a Gaussian curve parameterized by amplitude (A_m), position of the peak (t_m), and standard deviation (σ_m) as follows:

$$w_m(t) = A_m \, e^{\frac{(t-t_m)^2}{2\sigma_m^2}} \tag{4.1}$$

This results in an individual Gaussian curve as illustrated in Figure 4.3. However, for the GLA14 data product, a maximum of six Gaussian curves are included representing the multiple return echos as illustrated in Figure 4.4, where the case of two return echos is visualized [6]. The bias of the waveform (ε) fitting procedure is added to the signal. This results in a waveform decomposed as a sum of Gaussian curves together with a bias term as follows:

$$w(t) = \varepsilon + \sum_{m=1}^{N_p} A_m \, e^{\frac{(t-t_m)^2}{2\sigma_m^2}} \tag{4.2}$$

Note that the axes are for better understanding set as y and x. The GLAS instruments records the amount of return energy (fJ) as a function of time (t). The width of the transmitted pulse is 4 nanoseconds which is equivalent to 0.60 m in surface elevation [6]. The width of the return pulse is translated to the surface elevation. The waveform fitting procedure itself is less relevant for this research, since the data product includes only the GLA14 altimetry data. .

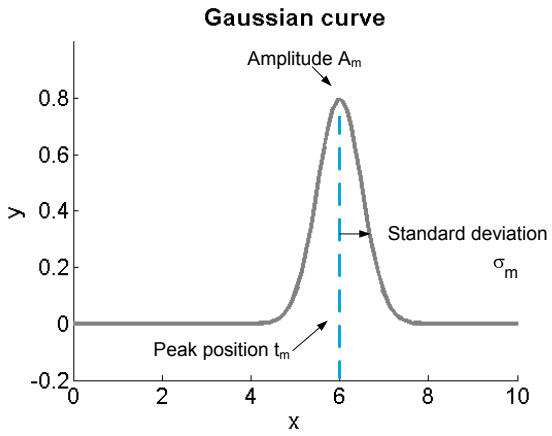


Figure 4.3: Example of Gaussian curve described by y = N(x,6,0.5).

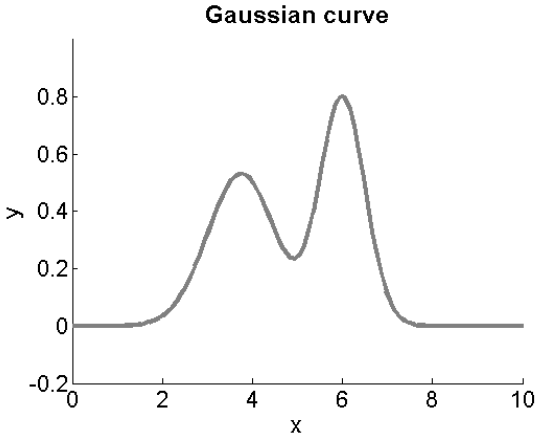


Figure 4.4: Example of multiple Gaussian curves described by y = N(x,3.75,0.75)+N(x,6,0.5).

Waveform referencing

It is important to understand how the waveform is referenced to the Earth's surface. This is described by Duong [12]. The geolocation and the surface elevation are presented in Figure 4.5. The laser altimeter vector (L) is basically the range measurement or the distance between the sensor and the ground surface. It also provides information about the laser pointing angle. The determination of the range measurement and the laser pointing angle is described in the Precision Altitude Document (PAD) Algorithm Theoretical Basis Document (ATBD) [1]. The GLAS geocentric position vector (G) represents the position of the sensor with respect to the center of the Earth and is described in the Precision Orbit Determination (POD) ATBD [46]. The spot position vector (I) or footprint is inferred by the sum of the vector L and G. Note that the vector I is referenced to the centroid of the waveform which represents the average height of the features in the footprint. This corresponds with the range measurement from the sensor to the centroid (t_c).

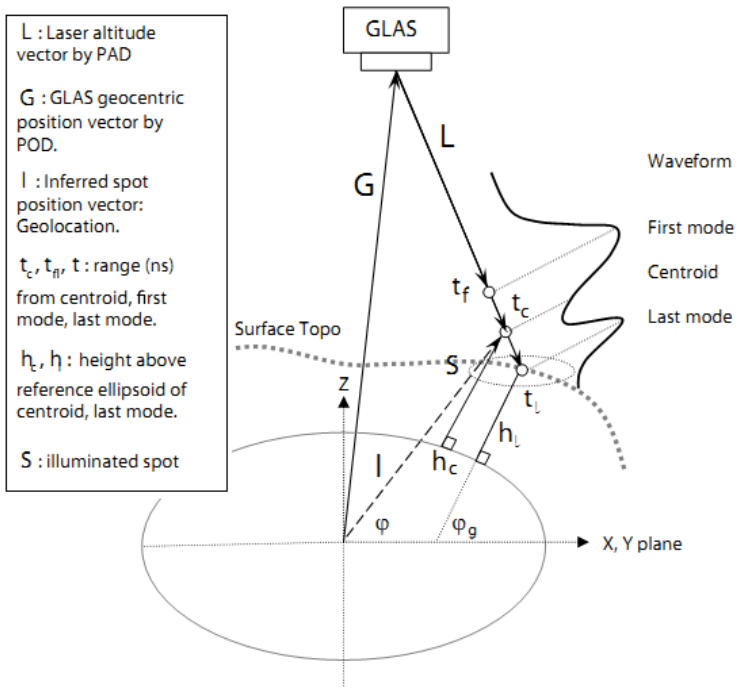


Figure 4.5: Illustration of the GLAS laser altimeter measurement, geolocation and surface elevation determination, taken from Duong [12].

The GLA14 altimetry data is referenced to the centroid of the waveform. However, for better inside in the vertical structure within the footprint, several variables are included to derive the first and last mode of the waveform and obtain the corresponding height measurements. The variables and the interpretation of the waveform are illustrated in Figure 4.6. The reference range (i_refRng) is defined as the distance between the sensor and the land return pulse (i_SigEndOff). Since, the GLA14 data is referenced to the centroid of the waveform, the distance between the sensor and the centroid is referred to as the land range offset (i_ldRngOff). The position of every peak in the waveform is described by the corresponding centroid, up to six centroids due to the assumption of a maximum of six peaks in the waveform fitting procedure. In the GLA14 data, the first centroid, (i_gpCntOff1) is corresponding with the ground surface [17]. This will be further discussed in section 4.5.4 and 5.3 for the purpose of water level estimations.

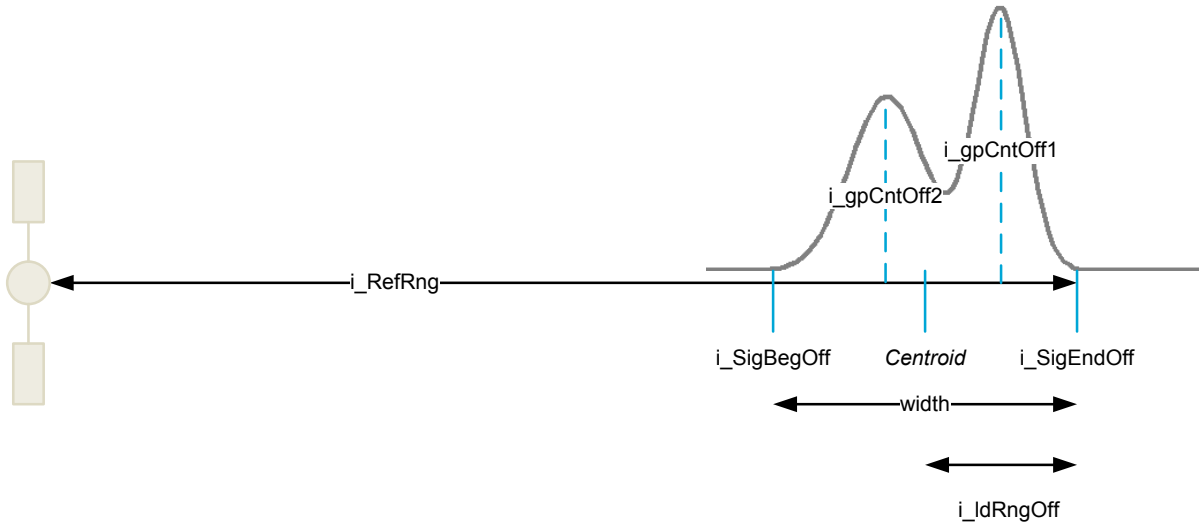


Figure 4.6: Waveform as referenced in GLA14 data with the centroid as GLA14 elevation data for the average height of the features in the footprint. The height of the ground surface is represented by the centroid of the last return pulse (i_gpCntOff1).

4.3.3 INTERACTION OF GLAS INSTRUMENT WITH WATER SURFACES

General behavior of the pulse

The emitted laser pulse by the GLAS instrument is backscattered from the target surface. Depending on the surface characteristics, varying levels of return energy are backscattered. Water surfaces almost completely absorb the transmitted energy while bare land results in high return energy [12]. The shape of the return pulse reveals therefore a lot of information about the surface characteristics. For the determination of footprints corresponding with water, the behavior of the laser on water surfaces should be known to distinguish the Gaussian curve of water from other land cover types. In general, the surface height distribution as represented by the Gaussian curve is defined by a mean surface slope and a surface roughness within the footprint of 70 m [6]. The effect of slope and roughness on the shape of the return pulse is illustrated in Figure 4.7. It appears that both slope and roughness result in a wider return pulse. For water bodies it is expected that the surface is generally flat. The absence of slope and roughness should result in a small width of the return pulse corresponding to the width of the emitted pulse.

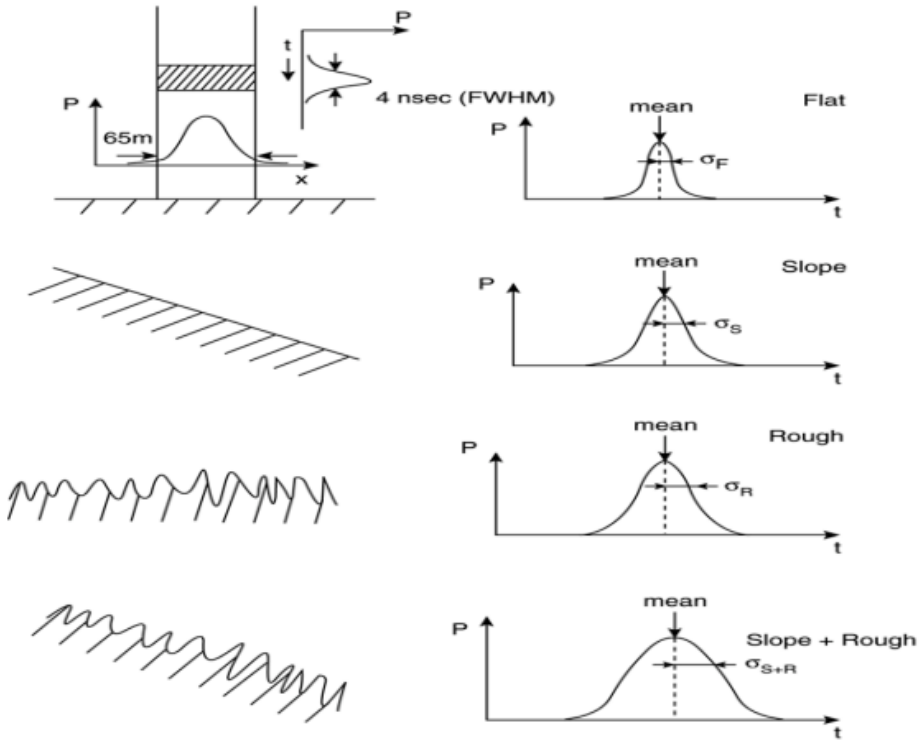


Figure 4.7: Characteristics of return laser pulse as a function of surface type. Presence of surface slope and roughness both broaden the pulse, taken from Brenner [6].

Behavior of the pulse on water surfaces

The interaction of the pulse with water surfaces can be further described by waveform parameters. A complete description of the waveform parameters is provided in section 4.4.4. The use of full waveform analysis for land cover classification is investigated [12]. Based on information enclosed by the Gaussian components four land cover types can be distinguished: high vegetation, urban areas, bare land/low vegetation and water. It was found that water absorbs more energy of the emitted pulse than other land cover types. Therefore, water was distinguished from other land cover types only based on a threshold of the amount of the emitted return energy (emitted return energy ≤ 45 fJ). The emitted return energy was calculated by taking the area under the waveform. When applying this threshold, it was difficult to distinguish water from bare land/low vegetation. Both land/cover types are flat surfaces resulting in a single narrow peak. Additional peaks may occur due to low vegetation, but they will have low amplitude and will be few. The waveform corresponding with bare land and water are visualized in Figure 4.8. The green peaks in the left figure belong to low vegetation.

A remark about classification based on the return energy should be made. The return energy is not an adequate input variable because it is sensitive for the performance of the laser aboard. The initial energy level of the laser pulse was 75mJ but declined with its number of operation days. The return energy is correlated with the transmitted energy. Because the transmitted energy of the laser fluctuates, it would be more appropriate to use the reflectivity (return energy / transmitted energy) instead. The reflectivity is usable when saturation effects do not occur. However, saturation caused by land-water transitions is very common and therefore it will result in false values of the reflectivity. It appears that the reflectivity can exceed 100 percent when the saturation energy correction is very large. Of course it is possible to delete footprints with a reflectivity larger than 100 percent and assign these pixels as biased but this will result is many deletions. This is further discussed in section 5.1.3.

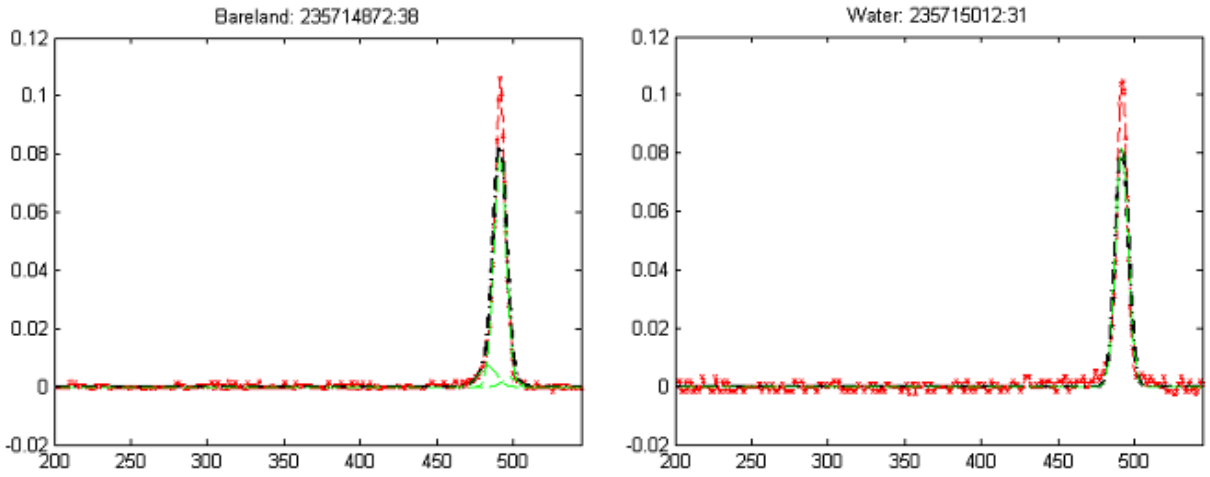


Figure 4.8: Land cover classification based on waveforms of bare land (left) and water (right), taken from Duong [12].

Other studies have evaluated the performance of GLAS over water surfaces; however the number of studies that are explicitly focusing on river networks is yet limited. The use of GLAS for dense river networks was investigated in France [2]. In this study, only 46 footprints were selected in rivers with a width greater than 50 m which already shows the drawback of using laser altimetry for rivers with smaller width. The study resulted in an accuracy of 1.14 m (rms). Furthermore, 61 percent of the shots were saturated.

Saturation of the waveform

Saturation occurs when the return energy overshoots the threshold value of the gain. The gain will enhance the amplitude of the return signal when the return energy is low. In the case of water surfaces, the return energy is generally low because water absorbs more energy of the emitted pulse than other land cover types [12]. This phenomenon occurs when the ICESat track changes land-cover, for example a land-water transition. The response time of the gain value is too slow and not yet adjusted to the measured land-cover transition. The amount of return energy will not be properly recorded because the waveform is clipped [32]. An example of a non-saturated waveform and a saturated waveform are illustrated in Figure 4.9. The consequence for the range measurement is an underestimation of the elevations.

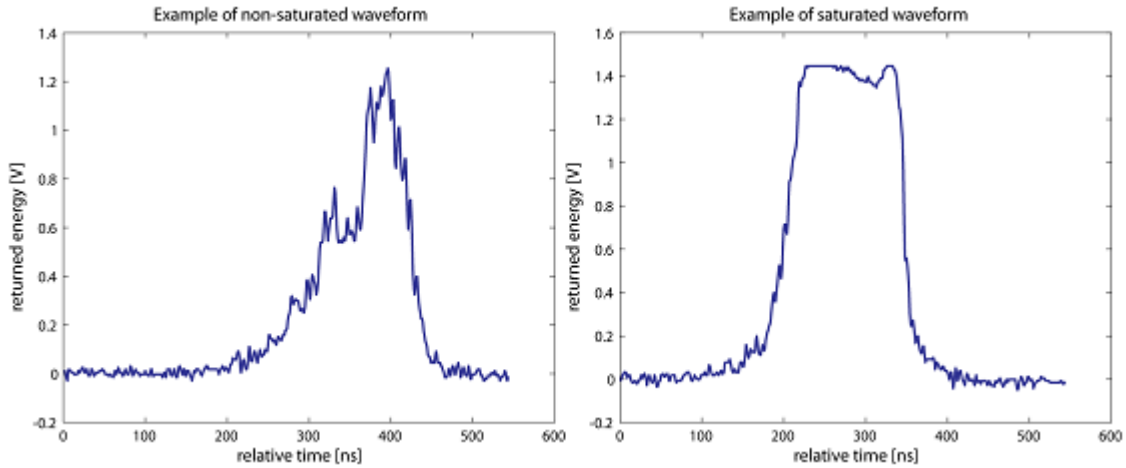


Figure 4.9: Example of a non-saturated waveform (left plot) and a saturated waveform (right plot). Notice the clipped waveform when saturated. Taken from Molijn [32].

The NSIDC provides saturation correction variables to enhance the elevation quality. However, it seems that the use of the saturation elevation correction (i_satElevCorr) results in overestimation of the elevations [2]. In the study of monitoring water levels in France it is therefore proposed to ignore the biased elevation measurements in the first ten footprints that correspond with water following from land, the so-called transition shots [2]. This points out the difficulty of ICESat for small rivers because this statement requires a minimum river width of several hundreds of meters.
A comment about the use of the gain value should be made. Note that the gain value can also be used as a pseudo-cloud filter. When the gain value is smaller than 100, the return energy is very large because of the presence of clouds. In this study it is chosen to ignore the impact of clouds and saturation.

4.4 PARAMETERS DERIVED FROM THE GLA14 DATA PRODUCT

The original variables are directly obtained from the GLA14 binary files ordered by the NSIDC. Some of the original variables are converted or recomputed in customized variables. The derivations and descriptions of these customized variables are elaborated in this section.

4.4.1 WORKFLOW OF PARAMETERIZATION

The translation from the GLA14 data product to parameters is essential for the derivation of variables useful for water level estimations. A complete workflow of the parameterization of the GLA14 data product is presented in Figure 4.10.
The GLA14 data can be ordered with the ICESat/GLAS Data Subsetter available at the NSIDC website [36]. As input the bounding coordinates of the subset need to be defined. In this study, the Mekong River Basin is chosen as the primary study area and therefore the spatial references are set by the latitude/longitude ranges in degrees of respectively [92.5 110] and [8.0 35]. The data set can be retrieved from a FTP server and consists of binary files that can be further processed with IDL and MATLAB.
The binary files are converted to the ASCII format for better use with MATLAB. The GLA14 data product contains many variables that are computed by the NSIDC and described in a dictionary [38]. Only 22 variables are included in this study. For the extraction of the selected variables from the ASCII files, the ASCII file needs to be explored to understand the data storage. Some records occur to have a failure in the determination of the geolocation (latitude/longitude position of the footprint). These records are removed from the data set. A standard geolocation filter is used to select latitude [-90 90] and correct longitude values [0 360].
The GLA14 data product is referenced to the Topex/Poseidon ellipsoid. For consistency with other data, an ellipsoidal conversion to the WGS84 datum is performed. Because this conversion has no significant consequences for the geolocation, the ellipsoidal conversion is only executed in the vertical direction. The theory behind the ellipsoidal conversion is described in section 4.4.3. The time stamp of the footprints is appointed with the J2000 epoch. This epoch started on noon on January 1, 2000 and counts in seconds. The time is manually converted to UTC time expressed in the common known DD/MM/YYYY-format. This is also elaborated in section 4.4.3. After some recalculations, the customized GLA14 variables are converted in Shapefile-format for further processing in ArcGIS.

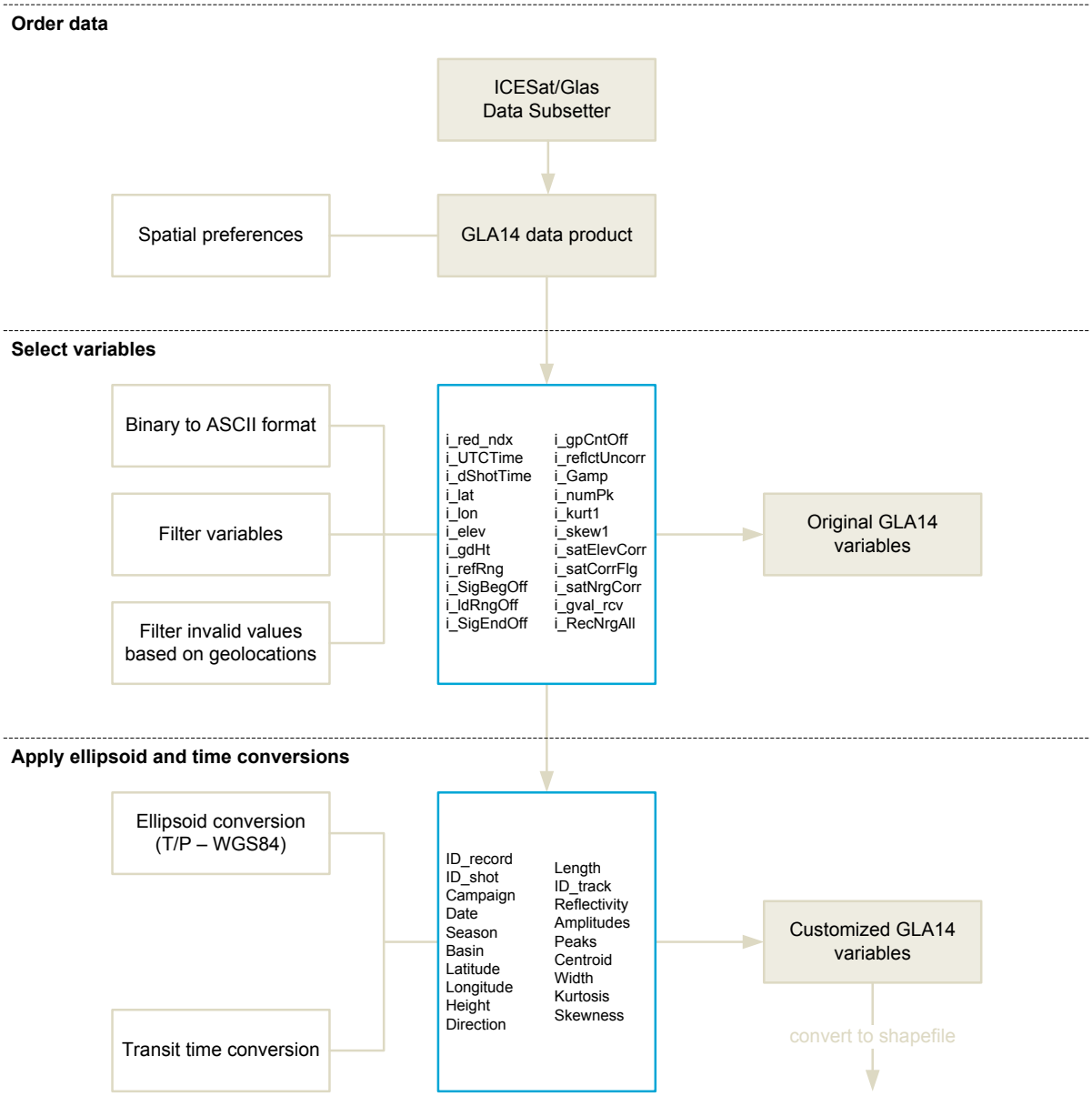


Figure 4.10: Overview of parameterization of GLA14 data.

4.4.2 ORIGINAL GLA14 VARIABLES

For this study, GLAS Release-31 Global Land Surface Altimetry Data (GLA14) was used. At the time of this study, NSIDC was re-processing GLAS data and the upcoming release-33 was not yet completely available. Release-33 includes several important changes such as improvements of the saturation correction for Laser 3, additional flags for more consistency in atmospheric effects, and improved calibration of the atmospheric product parameters [39]. Moreover, for the GLA14 data product a WGS84 ellipsoidal surface height is included. The GLA14 products contain campaign, repeat cycle and track numbers which will make comparison in temporal variations easier. For further research it is recommended to use GLAS Release-33 data products.
The GLA14 data product is provided in binary format and consists of variables per record and per shot (40 Hz). To evaluate temporal variations in the water levels, the data set contains all campaigns between 2003-2009. The 22 variables that are selected from the GLA14 data product are listed in Table 4.2.

Table 4.2: Original variables selected from the GLA14 data product, based on NDISC GLAS Altimetry Data Dictionary [38].

№	Variables	Units	Description
01	i_rec_ndx	N/A	Unique index that relates this record to the corresponding record(s)
02	i_UTCTime	[s]	Transmit time in UTC of the 1st shot in the 1 second frame
03	i_dShotTime	[s]	Time deltas of pulses 2 through 40 to i_UTCTime.
04	i_lat	[microdeg]	North latitude of the forty laser spots in the 1 second time frame.
05	i_lon	[microdeg]	East longitude of the forty laser spots in the 1 second time frame.
06	i_elev	[mm]	Surface elevation with respect to the ellipsoid (TOPEX/Poseidon)
07	i_gdHt	[cm]	Height of the geoid above the ellipsoid
08	i_refRng	[mm]	Reference range
09	i_SigBegOff	[mm]	Signal begin range increment
10	i_ldRngOff	[mm]	Land range offset
11	i_SigEndOff	[mm]	Signal end range increment
12	i_gpCntRngOff	[mm]	Centroid range increment for all six peaks
13	i_reflctUncorr	Unitless * 1E6	Reflectivity (calculated as received energy / transmitted energy)
14	i_Gamp	[0.01 volts]	Amplitude of each Gaussian for all six peaks
15	i_numPk	N/A	Number of peaks in the waveform produced by Gaussian filtering
16	i_kurt1	Unitless * 100	Kurtosis of the received echo from signal begin to signal end
17	i_skew1	Unitless * 100	Skewness of the received echo from signal begin to signal end
18	i_satElevCorr	[mm]	Correction to elevation for saturated waveforms
19	i_satCorrFlg	N/A	Saturation Correction Flag
20	i_satNrgCorr	[0.01 fJ]	Correction to energy for saturated waveforms
21	i_gval_rcv	N/A	Gain value used for received pulse
22	i_RecNrgAll	[0.01 fJ]	Received energy signal begin to signal end

4.4.3 ELLIPSOIDAL AND TEMPORAL CONVERSIONS

The footprints are spatially and temporally referenced. To avoid inconsistencies with other data sets, the GLA14 data product is converted to an often used reference system. The spatial referencing is discussed first: The ICESat/GLAS data products are referenced to the Topex/Poseidon ellipsoid. Because the additional data sets are referenced to the WGS84 ellipsoid, inconsistencies should be avoided and an ellipsoidal conversion is performed. The difference between the two ellipsoids is summarized in Table 4.3.

Table 4.3: Summary of GLAS ellipsoid and WGS84 ellipsoid, taken from NSIDC [40].

	ICESat/GLAS	WGS84
Equatorial radius (a)	6378136.300000	6378137.000000
Polar radius (b)	6356751.600563	6356752.314245
Reciprocal flattening (1/f)	298.25700000	298.25722356
Eccentricity (e)	0.081819221456	0.081819190843

The two ellipsoids are very similar: the radius of the Topex/Poseidon ellipsoid is only 70cm smaller than the radius of the WGS84 ellipsoid [40]. The horizontal shift (latitude/longitude) will therefore fall below the horizontal accuracy and is ignored in the ellipsoidal conversion. The vertical shift (height) is adjusted based on a standard MATLAB code provided by the NSIDC.
The GLA14 data contains two variables to describe the time at which the laser pulse was emitted. This is defined as the transit time and it relates to the variables i_UTCTime and i_dShotTime (Table 5.1). As the laser produces 40 pulses per second, i_UTCTime provides the time at the first pulse (t = 1) in the second frame and i_dShotTime provides the time deltas for the remaining pulses (t = 2 until t = 40). To calculate the time for each pulse, the time deltas should be added to the time of the first pulse. However, both variables are expressed in seconds (and even microseconds for i_UTCTime) and are referenced to the epoch J2000 that is Noon January 1, 2000. To convert the transit time to the common DD/MM/YYYY format, first the individual time for each pulse is calculated by:

$$UTC_{pulse} = i_UTCTime + i_dShotTime \quad (4.3)$$

The conversion involves the number of seconds apparent in each day (60*60*24) from the start of the J2000 epoch (UTC_{J2000}) and is expressed with the following formula:

$$UTC_{DD/MM/YYYY} = \frac{UTC_{pulse}}{(60 \cdot 60 \cdot 24)} + UTC_{J2000} \quad (4.4)$$

4.4.4 CUSTOMIZED GLA14 VARIABLES

Some variables are recomputed to provide better information for the derivation of water level estimations. The variables are grouped in standard, geometric and waveform variables. Furthermore, some variables are exactly taken from the GLA14 data product and do not involve any actions. The standard variables are identifiers and provide general characteristics of the footprints. The geometric variables provide the geolocation of the footprint as well as information about the topology of the river cross-section. The waveform variables are the characteristics of the return pulse of the footprint. An overview of the variables is provided in Table 4.4 and will be discussed in the same order. Each variable is in bold style for quicker reading.

Table 4.4: Customized GLA14 variables, grouped in standard, geometric and waveform variables.

№	Variables	Units	Description	Custom
Standard variables				
1.1	ID_record	N/A	Index that relates each record (40 shots per record)	No
1.2	ID_shot	N/A	Index per second frame in each record (range [1-40]	Yes
1.3	Campaign	N/A	Name of the campaign (see Table 3.1)	No
1.4	Date	DD-MM-YYYY	Transmit time in UTC	No
1.5	Season	0 = dry, 1 = wet	Indicator for seasonal variation	Yes
1.6	Basin	0 = lower, 1 = upper	Indicator for geographical variation	Yes
Geometric variables				
2.1	Latitude	[Deg]	North latitude	No
2.2	Longitude	[Deg]	East longitude	No
2.3	Height	[m]	Height (MSL) referenced to WGS84 datum	Yes
2.4	Direction	-1 = down, 1 = up	Direction of the ground track	Yes
2.5	Length	N/A	Number of footprints in river cross-section	Yes
2.6	ID_track	N/A	Index of river cross-section	Yes
Waveform variables				
3.1	Reflectivity	Unitless	Reflectivity corrected for saturation	No
3.2	Amplitude_ratio	N/A	Maximum amplitude ratio of last return peak	Yes
3.3	Peaks	N/A	Number of peaks in waveform (max of 6 peaks)	No
3.4	Centroid	[m]	Position of peaks (max of 6 peaks)	No
3.5	Width	[m]	Width of the received echo	Yes
3.6	Kurtosis	Unitless	Kurtosis of received echo	No
3.7	Skewness	Unitless	Skewness of received echo	No

Standard variables

The variable **ID_record** is exactly copied from the original variable i_rec_ndx and is just an identifier to recognize the record. However, the GLAS instrument has a frequency of 40 pulses per second and therefore another identifier is included namely **ID_shot** which numbers the footprints from 1 to 40 corresponding to the number of shots in each record. The time stamp of the footprints is described by **Campaign** as given in Table 3.1. A more specific time stamp is derived in the previous section and resulted in the variable **Date**. The variable Date is reclassified in the binary categorical variable **Season** which consists of two

categories to verify the seasonal variation in the data. All data between May and November are reclassified as Season = 1, corresponding with the wet season. The dry season is reclassified as Season = 0. Another binary categorical variable is **Basin**. With use of the IWMI drainage basin, a spatial selection is performed to select all footprints in the Upper Mekong Basin. All footprints located in the Upper Mekong Basin have values Basin = 1, all footprints in the Lower Mekong Basin have values Basin = 0.

Geometric variables

The geolocation of the footprints are described with the variables **Latitude** and **Longitude** with unit Degrees. These variables are directly taken from the GLA14 data products i_lat and i_lon although a small unit conversion has taken place (from micro degrees to degrees). The ellipsoidal conversion as described in the previous section in ignored in the horizontal direction.

The **Height** is transformed to the WGS84 datum but is not yet suitable for water level estimations. The GLA14 elevation (i_elev) is defined as the surface elevation with respect to the Topex/Poseidon ellipsoid (Table 4.3). Obviously the earth is not perfectly described by an ellipsoidal shape and the irregularities in the shape are defined by the geoid [14]. This is illustrated in Figure 4.11. The perfect shape of the ellipsoid does not completely match the geoid. The surface of the earth contains mountainous areas when above the geoid. Large water surfaces correspond to surfaces where the earth falls below the geoid. To determine the water levels, the height obtained by the GLAS instrument is corrected for the geoid height. This is illustrated in Figure 4.12. This can be obtained by simply subtracting the geoid height from the ellipsoidal height as described in:

$$height_{MSL} = height_{ellipsoid} - height_{geoid} \tag{4.5}$$

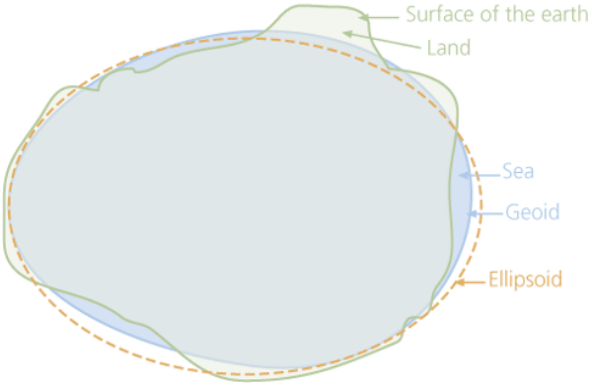


Figure 4.11: Model of the Earth, taken and enhanced from Esri [14].

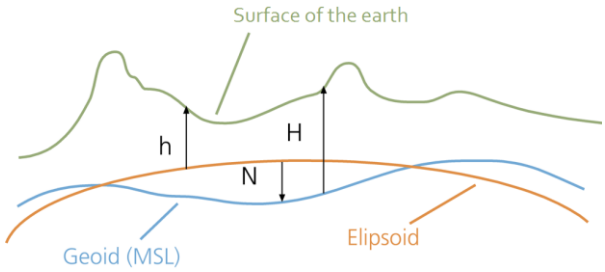


Figure 4.12: Orthometric height (H), ellipsoidal height (h) and geoid height (N).

As stated in section 4.3.2, the GLA14 data product is providing an average height of all features in the footprint. However, for water level estimations the last return pulse is relevant because this represents the ground surface. With the GLA14 data referenced to the land range offset (i_ldRngOff), the range increment provided by the centroid of the last return pulse ($i_gpCntRngOff_{(1)}$) is used to define the ground surface height by:

$$height_{ground} = height_{MSL} + (i_ldRngOff - i_gpCntRngOff_{(1)}) \tag{4.6}$$

After the identification of river cross-sections based on the customized land-water mask as created in the section 3.2, three additional variables are calculated. The **Direction** of the movement of the GLAS instrument can be ascending or descending. This is calculated by

comparing the latitude of the first footprint and the last footprint in the track. If the latitude of the first footprints is larger, the direction is downward (Direction = -1), otherwise the direction is upward (Direction = 1). Furthermore, the **Length** of the river cross-section is calculated by the number of footprints that are apparent in an intersection with the Mekong River as represented by the customized water mask. Each river cross-section is given a numerical identifier **ID_track** to evaluate each river cross-section individually.

Waveform variables

The return pulse is described in the GLA14 data product by a waveform decomposed into a maximum of six Gaussian components. An example of the waveform with its descriptive variables is provided in Figure 4.13. The directly obtained variable representing the **Reflectivity** (i_reflctUncorr) is defined by the NSIDC as the return energy calculated as the area under the waveform divided by the transmitted energy, not corrected for saturation effects [38]. As discussed in section 4.3.3, the clipping of waveforms caused by saturation results in underestimated values of the return energy. Fortunately, the saturation energy correction (i_satNrgCorr) is applicable for all shots with saturation flag value (i_satCorrFlg) above 2. The Reflectivity for saturated waveforms is then calculated by:

$$reflectivity_{sat\ corrected} = \frac{i_reflctUncorr\ * (i_RecNrgAll + i_satNrgCorr)}{i_RecNrgAll} \tag{4.7}$$

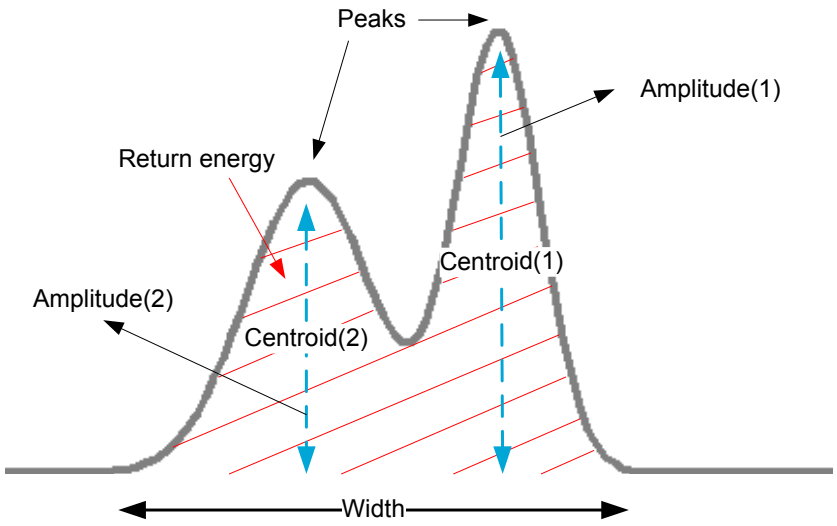


Figure 4.13: Waveform variables as described by the return pulse.

Again, note that after the Reflectivity is corrected, the return energy is often overestimated and still results in biased measurements (see section 4.3.3). The **Amplitude_ratio** is derived from the information of the amplitude the Gaussian curves. The last return peak is of interest for the estimation of ground levels. When the fitting procedure resulted in more peaks, these peaks could be relatively small in comparison with the last return pulse. This is tested with the amplitude-ratio that divides the amplitude of the last return pulse by the maximum amplitude of the other peaks in the signal. The number of **Peaks** or the modes of the waveform are directly obtained for the GLA14 data products. Since a fitting procedure is applied, the number of peaks has a maximum of six modes. The position of each peak is described by the **Centroid** and is organized in a matrix of six columns. If less peaks occur, the missing values are recognized by the value = 2.147483647e+07. This is the standard procedure derived from the NISDC. The **Width** represents the difference between the

beginning of the waveform (i_SigBegOff) and the end of the waveform (i_SigEndOff). This is simply calculated by:

$$width = i_SigEndOff - i_SigBegOff \tag{4.8}$$

Two descriptive statistics are included to define the shape of the waveform. These waveform variables are directly obtained from the GLA14 data product. The **Kurtosis** is described as the "peakedness" of the waveform. When observing Figure 4.13, the peak of the first centroid, Centroid(1) is higher than the second. The kurtosis is related to the roughness of the surface. When the laser interacts with a rough surface, the return signal will be less peaked and result in in a lower kurtosis. The interaction of the laser with water surfaces is expected to result in high kurtosis values due to the flatness of the surface. The **Skewness** is related to the asymmetry of the waveform. If the distribution of the return signal is normal, the skewness = 0. However, the footprint has a diameter of 70m so the features will most likely result in non-uniform height distributions. This variety in the footprints is described with the skewness. An example for the kurtosis and skewness with different values is illustrated in Figure 4.14.

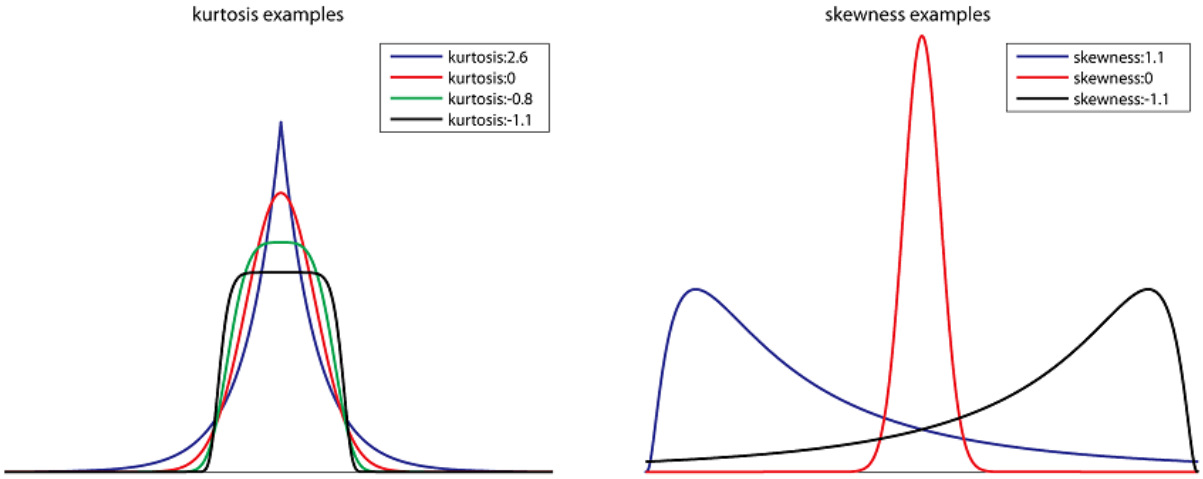


Figure 4.14: Example probability density functions with varying kurtosis and skewness values. In both subfigures, the red line represents the normal distribution having zero kurtosis and zero skewness. All distributions in each subfigure have the same standard deviation and are normalized with respect to the area under the distribution. The figure is only meant for visualization purposes, taken from Molijn [32].

5 METHODOLOGY FOR RIVER FOOTPRINT IDENTIFICATION

This chapter will present a methodology to identify river footprints in the Mekong River. The first step is to determine candidate river footprints using the customized water mask of the Mekong River. The candidate river footprints identification is described in section 5.1. The classification criteria based on the customized GLA14 variables. This is described in section 5.2. The second step is to determine the river footprints based on classification criteria. This is elaborated in section 5.3. Two workflows for river footprint identification are described and tested on a sampling set of river profiles. The implementation procedure in the Mekong River will be discussed.

5.1 CANDIDATE RIVER FOOTPRINT IDENTIFICATION

The intersection of the GLAS footprints with the customized water mask (see section 3.2) results in candidate river footprints. To assess the characteristics of the GLAS footprints that intersect the customized water mask, this section will elaborate the definitions of river cross-sections and river profiles as used in this study. A sampling set of 12 river profiles will be presented to derive descriptive statistics of river footprints and to evaluate height profiles and waveform variables as classification criteria.

5.1.1 WORKFLOW OF CANDIDATE RIVER FOOTPRINT IDENTIFICATION

In Figure 5.1, an overview of the steps to obtain candidate river footprints is provided. The 1,372,762 footprints in the Mekong River Basin are gathered by the spatial preferences (see section 4.4.1) with the ICESat/GLAS data Subsetter. Only a very small part of this selection will be considered relevant for the determination of water levels in the Mekong River. With candidate river footprints identification, only the footprints located within a distance of 500 m from the mainstream of the Mekong River as derived from the customized water mask are selected (see section 3). This resulted in 24992 candidate river footprints.
The next step is to group the candidate river footprints in river cross-sections. Each river cross-section represents a track of GLAS footprints intersecting the customized water mask. To only keep likely river cross-sections in the Mekong River, candidate river footprints are included in a river cross-section if the along-track spacing between consecutive footprints is smaller than 500 m. The along-track spacing is under normal circumstances approximately 175 m (see section 4.2). However, sometimes a footprint is missing because the GLAS laser altimeter failed to obtain a range measurement. This is illustrated in Figure 5.2. On the other hand, a larger distance between two adjacent footprints can include footprints that do not correspond to the Mekong River. With this in mind, the maximum along-track spacing is therefore set 500 m. Because it is not possible to derive relevant geometric information about height profiles from less than three footprints, river cross-sections consisting of 1 or 2 footprints are removed.
A remark about the spatial selection of the GLAS footprints should be taken into account. In this study, it is assumed that each river cross-section derived from the intersection with the customized water mask overlaps the Mekong River. Spatial selection simply includes all tracks of GLAS footprints within the customized water mask. Therefore it is possible that river cross-sections selected near the boundary of the customized water mask actually do not overlap the Mekong River at all. The intersection results in false river cross-sections. An example is provided in Figure 5.3. This can be resolved by improving the spatial selection operation or enhancement of the water mask. However, the effect of false river cross-sections on the water level estimations is expected to be insignificant because the candidate

river footprints will probably not meet the classification criteria to be discussed in section 5.1.3. After having explained this shortcoming, it will be assumed that an intersection of the GLAS footprints with the customized water mask will always intersect the Mekong River.

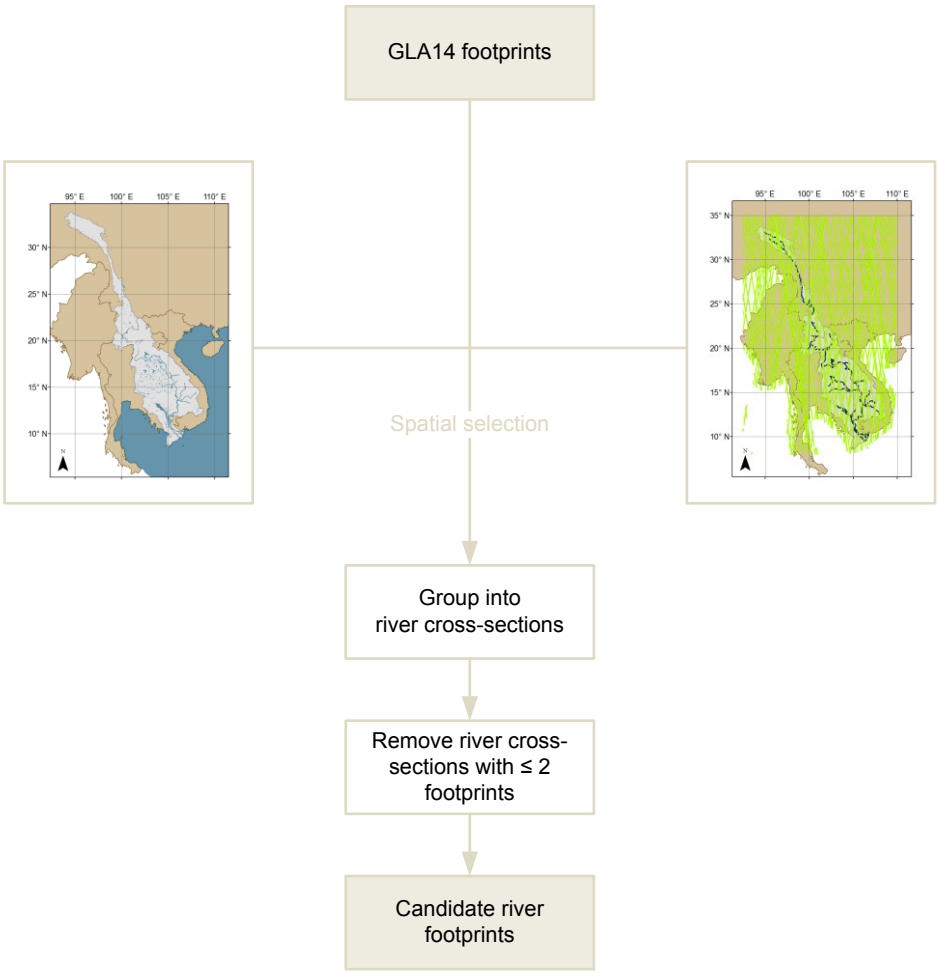


Figure 5.1: Overview of candidate river footprint identification.

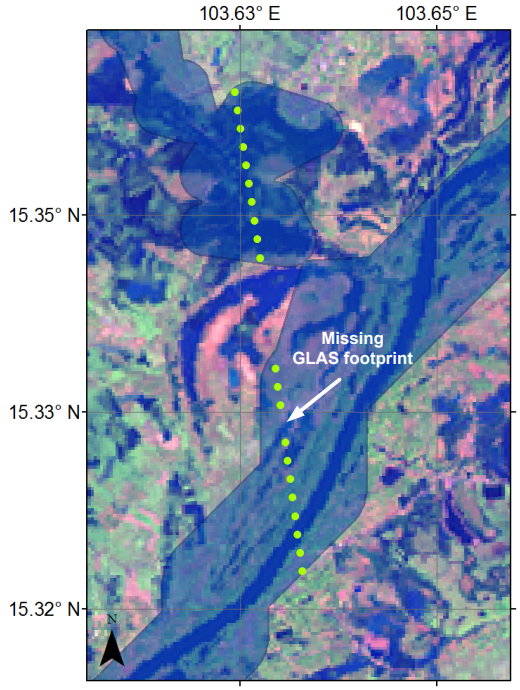


Figure 5.2: Example of missing GLAS footprint in a river cross-section. The large gap in the river cross-section is a result of the spatial selection of the GLAS footprints with the customized water mask.

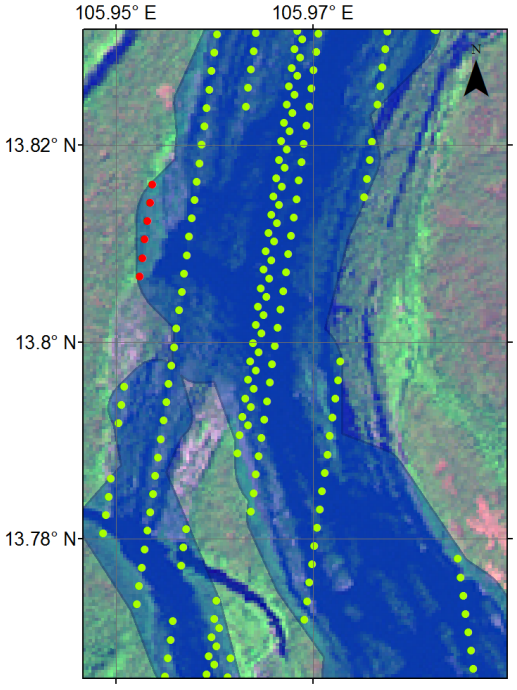


Figure 5.3: Example of false candidate river footprints (red dots) as a result of spatial selection of the GLAS footprints with the customized water mask.

5.1.2 DESCRIPTIVE STATISTICS OF CANDIDATE RIVER FOOTPRINTS

Histograms

Histograms are plotted to analyze data distributions to get a first impression of the availability of GLAS footprints in the Mekong River. A histogram is a useful graphical representation in which the range of data is divided into bins of equal width [10]. The choice of the bin width is relevant for a correct representation of the data. When the bin width is set too small, isolated peaks can occur which makes it hard to interpret the histogram. On the other hand, a bin width set too large will result in less detail and loss of information. For this study, the bin width (k) is set as the square root of the number of data points in the sample (n) [58]:

$$k = \sqrt{n} \tag{5.3}$$

However, for discrete variables such as the years (2003-2009) and the months (Jan-Dec) it is chosen to manually set the bin width accordingly.

Temporal coverage

The temporal coverage of the Mekong River by GLAS footprints is summarized in Table 5.1. In total 23966 footprints are identified as candidate river footprints. The candidate river footprints are grouped in 2242 river cross-sections. The frequency distributions of the footprints over time are shown in Figure 5.4. On average, 3424 footprints were obtained each year from 2003 until 2009. In 2009, only 1904 footprints were recorded by the GLAS instrument due to failure of the third laser aboard (left plot). A more interesting result is the temporal variation per month (right plot). No footprints were obtained in the months January and August. Furthermore, in July and September, the number of footprints is far below average, respectively 19 and 1 footprints. The monthly distribution of footprints can

be converted to the seasonal variation by selecting the footprints between May and November. Only 27.1 percent of the footprints that are selected as candidate river footprints are obtained in the wet season. Monitoring of water levels with respect to flooding is especially relevant during the wet season.
The main reason for this lack of measurements in the wet season is the scheduling of the operational periods as described in Table 5.1. These campaigns were predefined based on the laser capability and do therefore not match the monsoon climate in the Mekong River Basin. This might be taken into account for the planning of the ICESat-2 mission.

Table 5.1: Summary of candidate river footprints per campaign.

Campaign	Start date	End date	№ of cross-sections	№ of footprints
L1A	2003-02-20	2003-03-29	119	1633
L2A	2003-09-24	2003-11-19	196	1978
L2B	2004-02-17	2004-03-21	145	1571
L2C	2004-05-18	2004-06-21	78	743
L3A	2004-10-03	2004-11-08	89	997
L3B	2005-02-17	2005-03-24	95	811
L3C	2005-05-20	2005-06-23	38	238
L3D	2005-10-21	2005-11-24	182	2369
L3E	2006-02-22	2006-03-27	166	2133
L3F	2006-05-24	2006-06-26	95	819
L3G	2006-10-25	2006-11-27	127	1369
L3H	2007-03-12	2007-04-14	146	1584
L3I	2007-10-02	2007-11-05	134	1285
L3J	2008-02-17	2008-03-21	176	1929
L3K	2008-10-04	2008-10-19	168	1705
L2D	2008-11-25	2008-12-17	98	898
L2E	2009-03-09	2009-04-11	122	1290
L2F	2009-09-30	2009-10-11	68	614

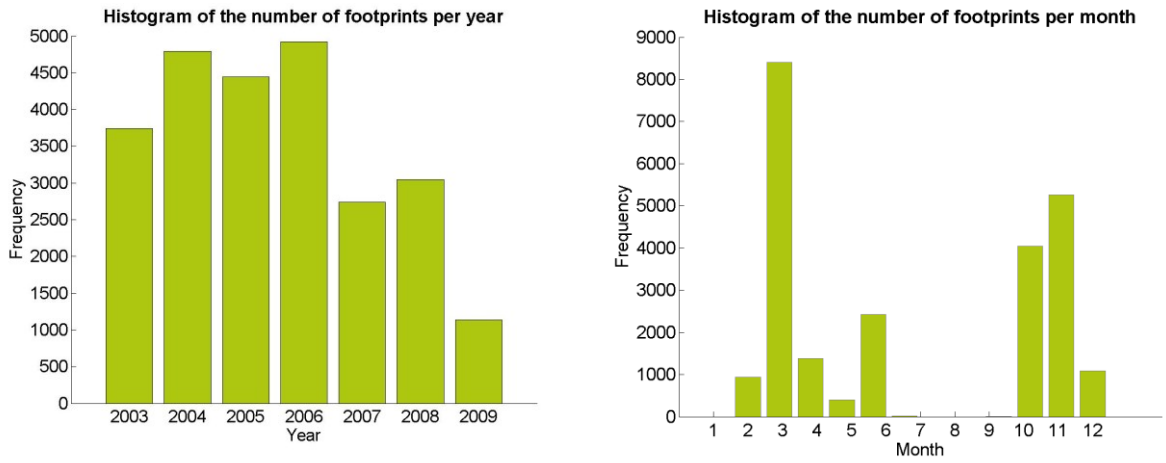


Figure 5.4: Temporal variation per year (left plot) and per month (right plot) of candidate river footprints.

Geographical coverage

The geographical coverage of the candidate river footprints is visualized in Figure 5.5. The number of footprints that were obtained in the Upper Mekong Basin is only 3683, which equals 15.4 percent (Figure 5.5, left plot). The narrow width of the Mekong River in this area makes it difficult to select candidate river footprints. The Lower Mekong River is densely covered by candidate river footprints (Figure 5.5 right plot). The first impression is that especially the Lower Mekong River will have sufficient geographical coverage to identify river footprints and evaluate seasonal fluctuations.

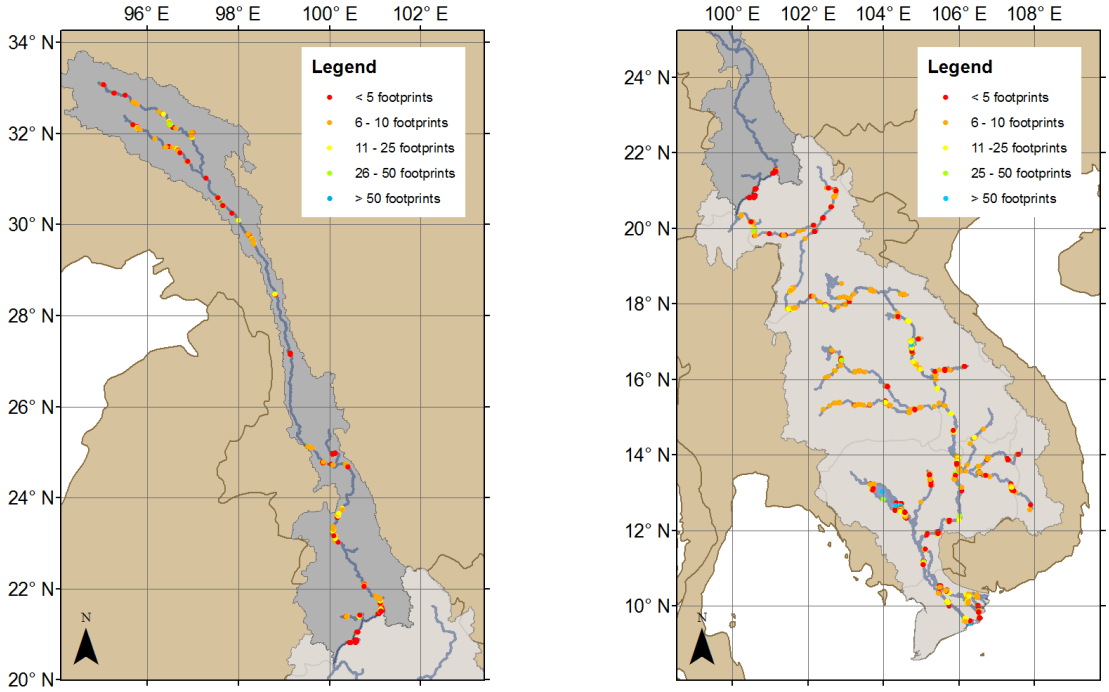


Figure 5.5: Geographical coverage of candidate river footprints in the Upper Mekong Basin (left plot) and the Lower Mekong Basin (right plot). The number of footprints per river cross-section is illustrated with the colored dots.

Number of footprints per river cross-section

Histograms of the number of GLAS footprints per river cross-section are plotted in Figure 5.6. The left histogram shows that 98.4 percent of the river cross-sections contain less than 50 footprints. The few river cross-sections with over 50 footprints are located in the Tonle Sap Lake or are located parallel to the Mekong River. However, in general the distribution of the number of footprints per river cross-section is varying between 3 and 20 footprints (Figure 5.6, right plot). Note that the river cross-sections containing only 1 or 2 footprints were removed because it is not possible to derive accurate water level estimations.

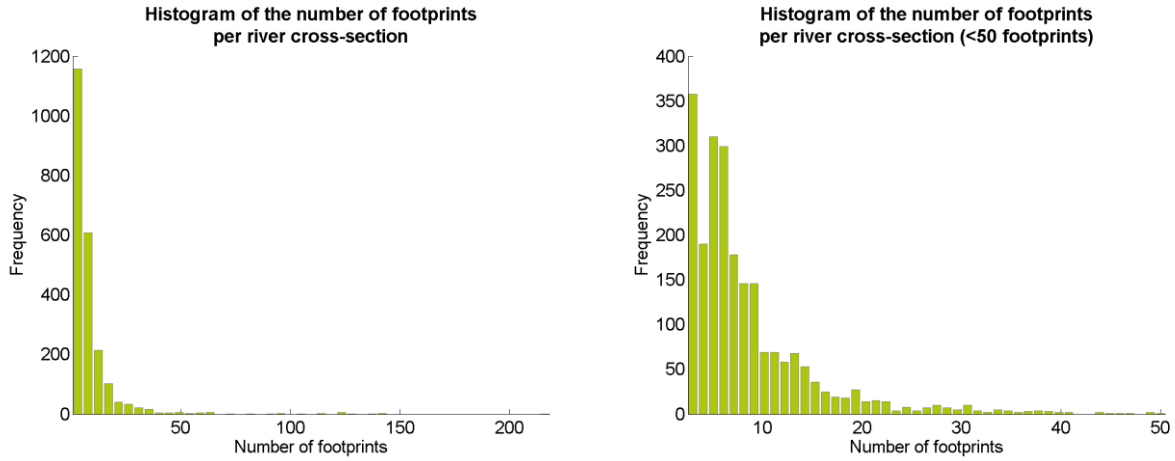


Figure 5.6: Histogram of the number of GLAS footprints per river cross-section (left plot) and with less than 50 footprints (right plot).

5.1.3 SAMPLING SET OF RIVER PROFILES

The intersections of the GLAS footprints with the customized water mask are visualized from two viewpoints: top-view and side-view. This is pointed out in Figure 5.7 and 5.8 with an example of a perpendicular aligned intersection of GLAS footprints with the customized water mask. The top-view as illustrated in Figure 5.7 shows a possible river cross-section. In this example, the river cross-section is aligned perpendicular to the river. The arrow indicates a descending direction of the ground track. The ICESat ground tracks have a descending or ascending direction. All river cross-sections are described in the descending direction for consistency reasons.

A remark about the alignment of the river cross-sections and the corresponding river profiles should be made. The examples of the two viewpoints as described in Figure 5.7 and 5.8 correspond to a perpendicular alignment of the GLAS footprints with the customized water mask. Since the candidate river footprints are obtained by spatial selection, it is possible that the river cross-sections are parallel aligned to the Mekong River. The river profiles will differ accordingly.

It is also possible to describe the intersection of GLAS footprints with the customized water mask as channel geometries including the banks. The channel margins are defined by the height difference between the stream bed and the bank. This side-view results in river profiles as illustrated in Figure 5.8. The height (expressed in meters) of the river profiles is situated on the y-axis. The distance between the consecutive footprints is named the along-track profile (expressed in meters) and is derived from the latitude/longitude position. For consistency, the along-track profile is described by descending latitude values. Furthermore, the along-track profile should not be mistaken for the along-track spacing which indicates the distance between two consecutive footprints of approximately 175 m. When reading the river profile, this implies that the left always correspond with the highest latitude value in the river cross-section.

The diversity of the river cross-sections available in the Mekong River makes it difficult to develop a uniform method to determine which footprints correspond to the actual river. However, the objective is to explore the behavior of the GLAS instrument in river networks and the diversity can provide information about optimal river geometries. For the purpose of understanding the variety of river profiles, a sampling set of 12 river profiles (including 321 candidate river footprints) is selected. The river profiles are geographically spread in the Mekong River Basin (Figure 5.9). For better interpretation, the river profiles are ordered based on their mean height corresponding with traveling downstream on the Mekong River. The characteristics of the river profiles are summarized in Table 5.2.

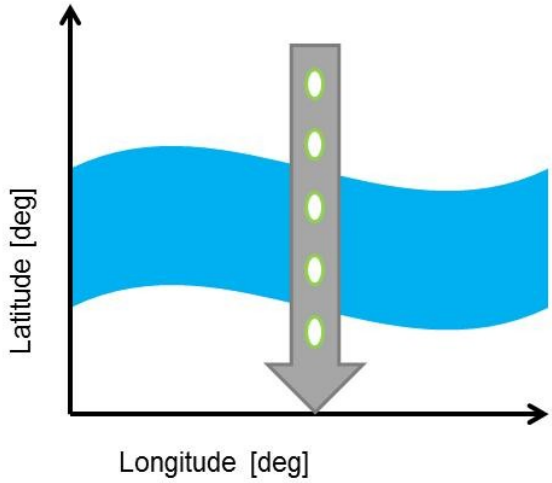


Figure 5.7: Example of a river cross-section described by the latitude [deg] and longitude [deg] (top-view). The arrow indicates the descending latitude direction.

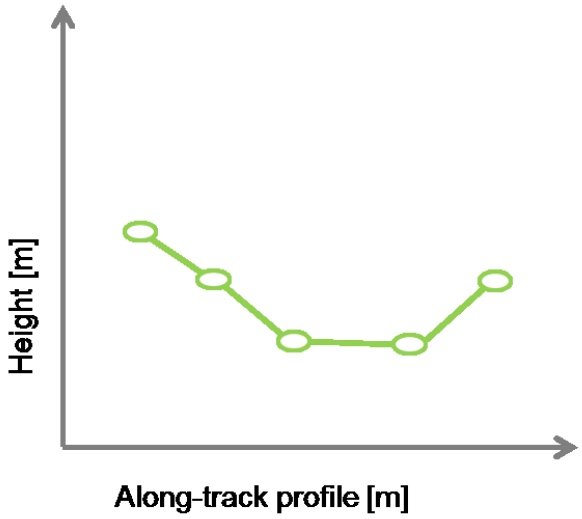


Figure 5.8: Example of a corresponding river profile (see Figure 5.7) described by the height [m] and along-track profile [m] (side-view).

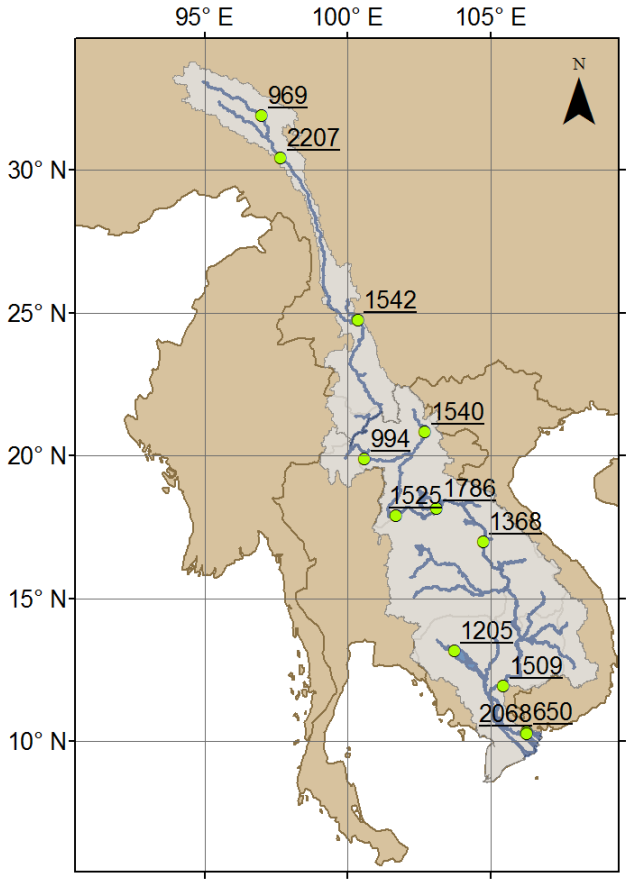


Figure 5.9: Sampling set of river profiles in the Mekong River. The label (ID) refers to the identifier of the river profiles.

Table 5.2: Overview of the 12 river profiles in the sampling set. The ID refers to the index of the river profile. N refers to the number of footprints per river cross-section. The Landsat WRS refers to the index of the Landsat ETM image.

ID	Campaign	Date	Season	Basin	N	Landsat WRS
0969	L3B	2005	Dry	Upper	14	134038
2207	L3H	2007	Dry	Upper	9	134039
1540	L3E	2006	Dry	Lower	7	129046
1542	L3E	2006	Dry	Upper	10	131043
1786	L3F	2006	Wet	Lower	9	128047
0994	L3B	2005	Dry	Lower	52	130046
1525	L3E	2006	Dry	Lower	9	129048
1368	L3D	2005	Dry	Lower	64	127048
1509	L3E	2006	Dry	Lower	20	126052
1205	L3C	2005	Wet	Lower	107	127051
2068	L3H	2007	Dry	Lower	11	125052
0650	L2C	2004	Wet	Lower	10	125052

The 12 river profiles in the sampling set are visualized in Figures 5.10 and 5.11. The river profiles are evaluated with Landsat ETM images (see section 3.3.1). Each river cross-section is overlapping the corresponding Landsat image in Table 5.2. The Landsat WRS index refers to the World Reference System Path/Row index. Landsat provides information about the land cover type. Furthermore, the images have a spatial resolution of 30 m with a horizontal accuracy of 15 m which can be used to identify the river footprints of each river profile beforehand by visual inspection. The Landsat-derived river footprints are therefore defined as the footprints that overlap the Mekong River in Landsat images. The river footprints located completely in the Mekong River according to Landsat are represented by blue dots. For some footprints it was difficult to classify them as water or land because they were located at a land-water transition. These footprints will be referred to as boundary river footprints and are represented by grey dots. The land footprints are represented by green dots.

The diversity of the Mekong River is reflected by the river profiles in the sampling set. The number of footprints per river cross-section varies from a minimum of 7 (Figure 5.10c) to a maximum of 107 (Figure 5.11d) footprints. The mean height per river profile varies between 3620 m (Figure 5.10a) to 0.4 m below Sea Level (Figure 5.11f). The standard variation of the height per river profile is also varying from 0.17 m (Figure 5.11e) to 854 m caused by the appearance of a measurement error of the GLAS instrument (Figure 5.10e). This is checked with a comparison between the height profile and the USGS HydroSHEDS DEM (see section 3.1.1).

The Landsat-derived river footprints correspond sometimes with the minima of the river profiles (Figure 5.10a), while in other cases, the minima correspond to outliers (Figure 5.11d). It is even possible that the river footprints correspond to the maximum height value in a river profile (Figure 5.10c). There are two possible explanations: First, the spatial selection with the customized water mask can touch the boundary of the actual Mekong River. The extent of the channel geometry is not fully captured by the river cross-section. Second, tracks with an along-track spacing between consecutive candidate river footprints larger than 500 m are split in multiple river cross-sections resulting in a discontinued track. The maximum height values are actually located lower but this is not visible due to the discontinuity. For river cross-sections with small width, only one river footprint is estimated (Figure 5.10a) while a wider part of the river or even a lake might result in a group of adjacent river footprints (Figure 5.11d).

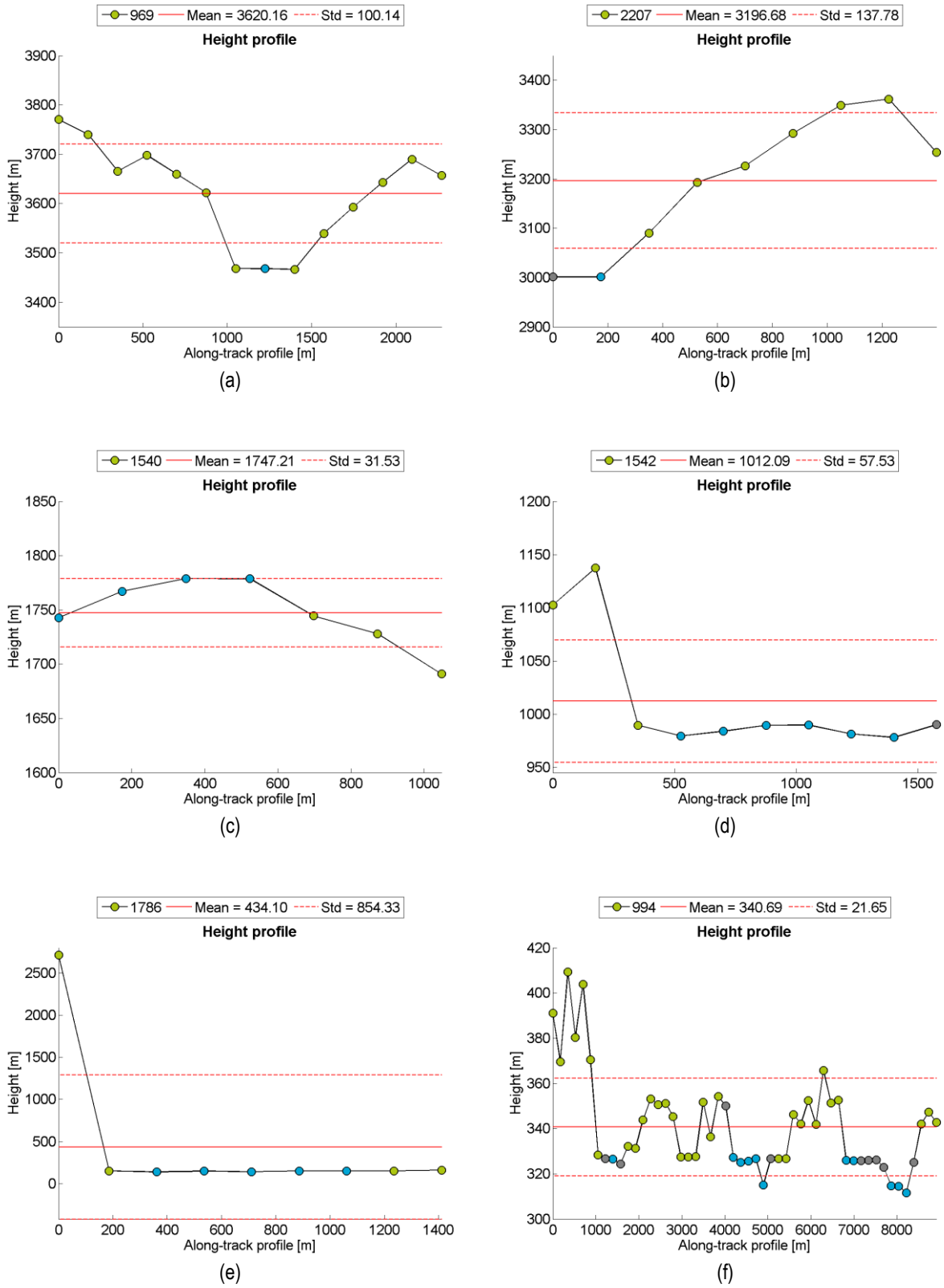


Figure 5.10: Sampling set of river profiles part 1. The graph represents mean and corresponding standard deviations of the height values (red). The dots below represent the classification of river footprints in which green = land, grey = land-water transition, blue = water.

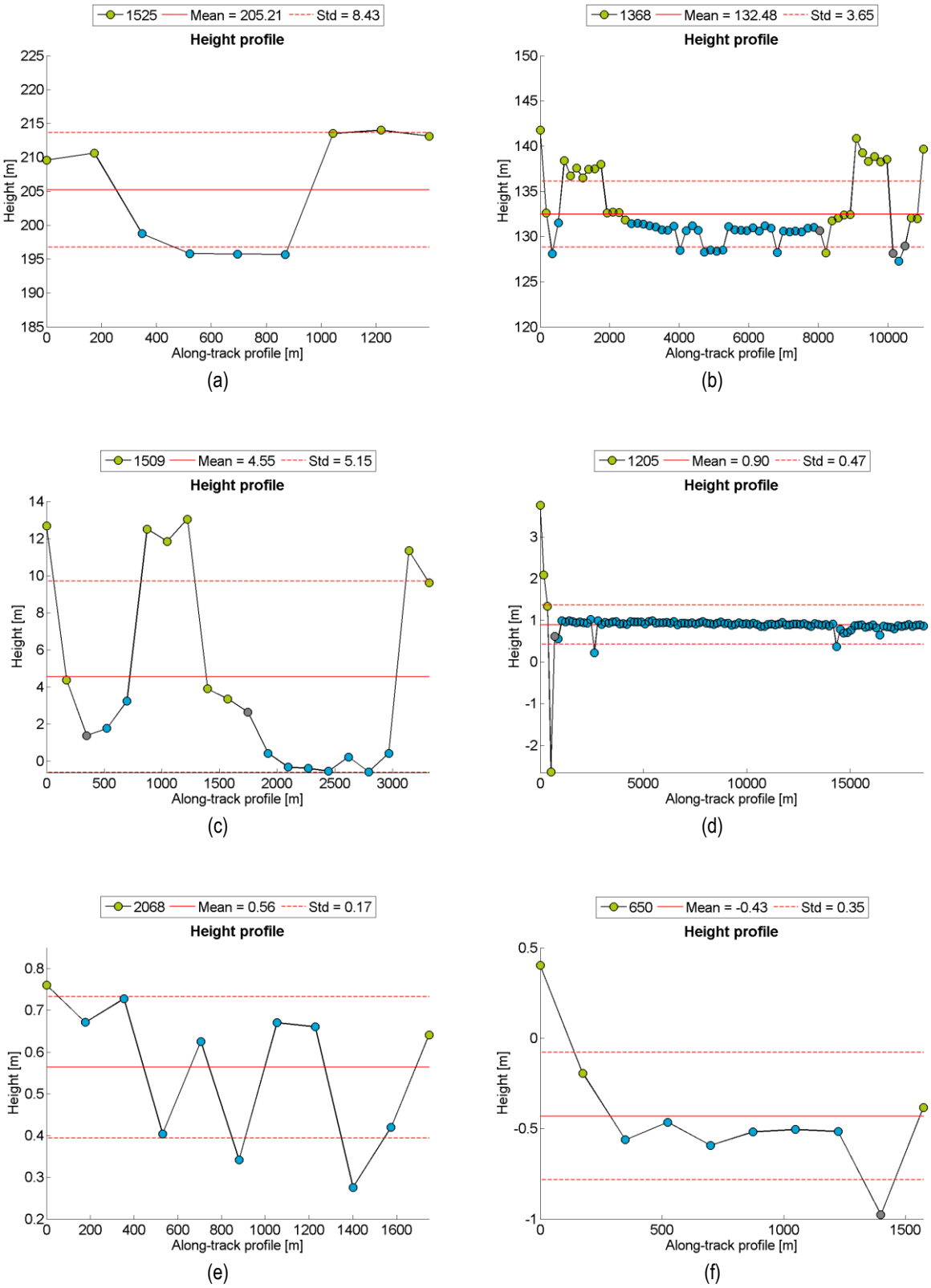


Figure 5.11: Sampling set of river profiles part 2. The graph represents mean and corresponding standard deviations of the height values (red). The dots below represent the classification of river footprints in which green = land, grey = land-water transition, blue = water.

5.2 CLASSIFICATION CRITERIA

In this section, a classification methodology for river footprint identification is developed. Candidate classification variables are tested at the sampling set of river profiles. This resulted in the development of two workflows to support the river footprint identification in the Mekong River.

5.2.1 CANDIDATE CLASSIFICATION VARIABLES

The candidate classification criteria are derived from the theory of the behavior of the GLAS laser pulse on water surfaces (see section 4.3.3.) and the customized GLA14 variables (see section 4.4.4). This resulted in a set of candidate classification variables listed as follows:

- height
- reflectivity
- number of peaks
- amplitude-ratio
- width
- kurtosis
- skewness

The sampling set of river profiles is described by the height profiles and waveform variables that are derived for each footprint. For two river profiles (ID = 650 and ID = 1542) the characteristics of the most relevant waveform variables with potential for river footprint identification are provided in Figure 5.12. Because of the large amount of images, the complete sampling set of river profiles is presented in Appendix 1. The waveform variables, height profiles and an image of each river cross-section in Google Earth are included. Although the horizontal accuracy of Google Earth in Southeast and East Asia is approximately 45.9 meters (rms), the imagery supports better interpretation of the environment and visualization of the river cross-sections [45]. Note that the waveform variables and the height profiles are described in descending latitude direction. Reading the footprints from left to river corresponds with descending footprints in the Google Earth image.
Analyzing the information enclosed in the waveform variables and height profiles result in opportunities for classification of river footprints. The height profile can provide geometric information of the river while the waveform variables can tweak the results with additional information about the laser interaction with the ground surface. This will be shortly discussed below.

Geometric variables
Based on the height profiles in Appendix 1, river cross-sections in the Mekong River have small height differences between consecutive river footprints. This is obviously related with the assumption that water surfaces are generally flat. The along-track spacing of 175 m is very small and therefore the water levels will be considerably equal. However, this is depending on the alignment of the river cross-section with the topology and decline of the river. The height between consecutive river footprints will be equal when the river cross-section is nearly perpendicular to the river. The decline does not have any influence. However, a parallel aligned river cross-section is influenced by the decline of the river. Especially in the upper course of the Mekong River this is important to consider.
Furthermore, the number of footprints that have small height differences contributes to the certainty of the presence of a water surface. More footprints with similar height levels indicate a wider river width. For example in the height profile of the left plot of Figure 5.12, the river cross-section consists of six footprints with similar height levels. The standard deviation in the river cross-section is equal to 0.35 m. It is very likely that the six footprints with similar height levels correspond to the river. The other footprints have elevations that

deviate from this height levels. This is further illustrated with an example in section 5.2.2. The certainty of river footprints can therefore be defined by the number of footprints that have small height differences between consecutive footprints.

Waveform variables

The behavior of the pulse on water surfaces is used in full waveform analysis for land cover classification (see section 4.3.3). This leads to presumptions about the contribution of waveform variables to the river footprint identification. A set of waveform variables is derived from the GLA14 data product and considered as relevant for classification purposes (see section 4.4.4).

The most relevant waveform variable to distinguish water from other land cover types is the reflectivity [12] [32]. The reflectivity is very usable if saturation effects are non-existent (see section 4.3.3). In the right plot of Figure 5.12, the reflectivity values for some footprints are not apparent due to saturation effects. A suggestion would be to exclude all footprints with a reflectivity larger than 100 percent. However, 41.4 percent of the candidate river footprints is saturated (saturation flag > 0). It was also suggested to remove biased reflectivity values > 1. This will result in a data loss of 31.9 percent. The reflectivity is therefore not further considered in the Mekong River Basin.

In Figure 5.12, the left plot shows a river cross-section in the Mekong River Delta. In this area, the land surface is generally flat and therefore the return signal is generally a single peak. The intersections with the Mekong River in the Upper Mekong basin are expected to have more footprints included as river footprints that do not correspond to the narrow river. Because of the narrow width, more mixed footprints caused by different height scatterers occur. Although it is preferred to only consider single peaks, it cannot be taken as strict selection criterion for river footprints.

The amplitude-ratio is defined by the ratio between the amplitude of the last return pulse and the maximum amplitude of the other peaks in the signal. In Figure 5.12 in the right plot, the amplitude-ratio shows some correlation with the number of peaks. The third footprint measurement from the right has 5 peaks and an amplitude-ratio of almost 1:7. This footprint is assumed to correspond with the Mekong River. An explanation for this behavior can be the presence of an object in the water that leads to a mixed footprint or the influence of saturation or clouds. It should be briefly noted that the saturation can result in a ringing peak that does not represent the surface [37]. This is not further considered in this study but is discussed as recommendation.

Both plots show an increasing width for footprints that do not correspond with water. However, in the left plot the water footprints have a minimum width of 2.7 m while the right plot's minimum width is equal to 5.0 m. A possible explanation for this difference can be the sediment load in the Mekong River. The amount of suspended sediment concentration is correlated with the stream flows in the Mekong River [53]. The changes in sediment load have diverse causes: Confluences in the river, the gravitational forces of river decline, as well as periods of heavy rainfall influence the sediment transport. When flooding occurs, erosion of the channels causes a remobilization of the sediments. Furthermore, the snowmelt from the Tibetan Plateau is also suggested as a dominant cause. The amount of sediment transport is more apparent in the Upper Mekong Basin. The difference in width is also even more extreme for the river cross-sections in the Upper Mekong Basin (see Appendix 2). Again, this waveform variable cannot be strictly used to exclude footprints.

Although the kurtosis and the skewness are both used in previous studies for the classification of water, in these plots they do not make any sense [32][54]. This will be further assessed in section 5.2.1.

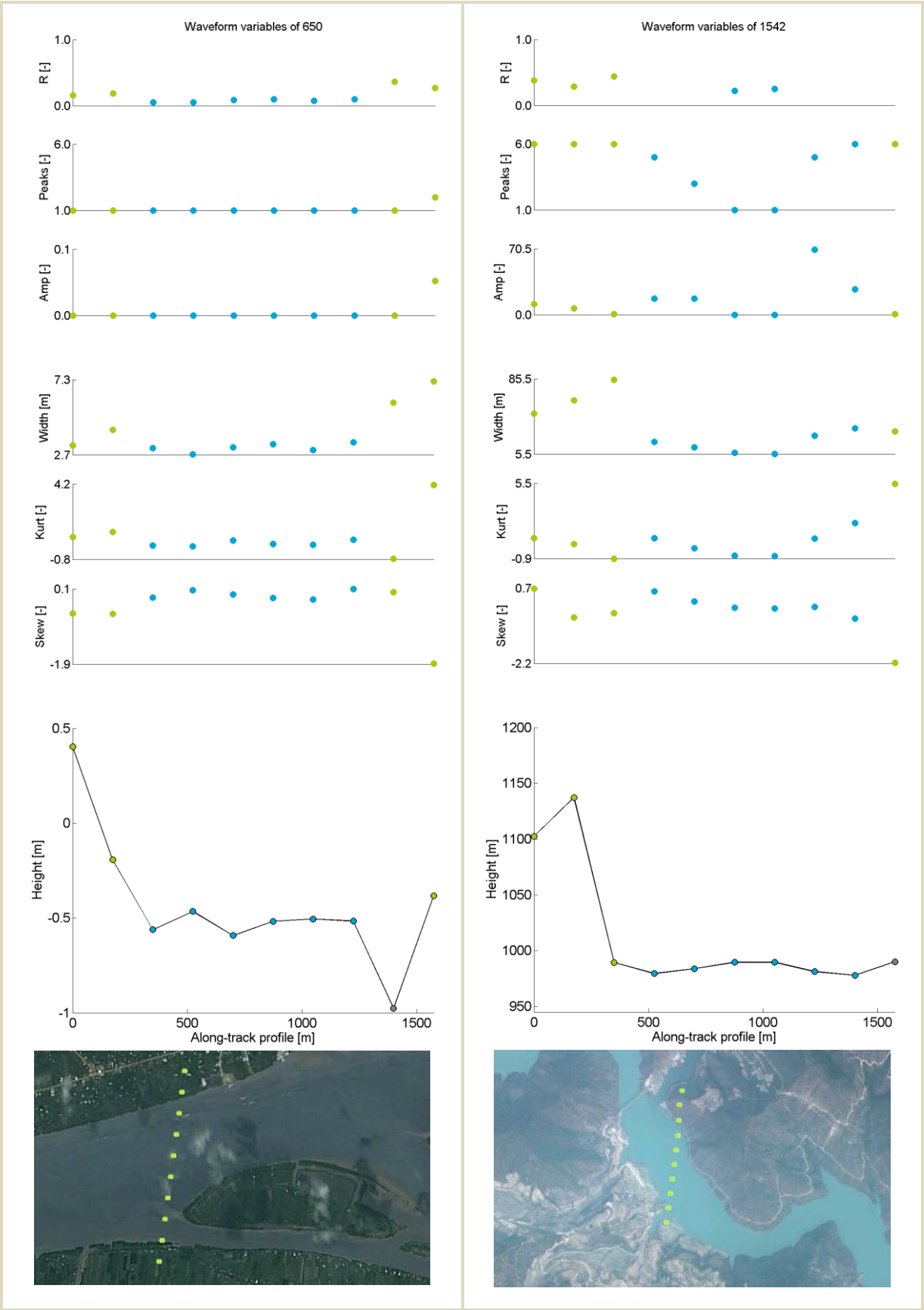

Waveform variables of 650
R [-]
1.0
0.0
Peaks [-]
6.0
1.0
Amp [-]
0.1
0.0
Width [m]
7.3
2.7
Kurt [-]
4.2
-0.8
Skew [-]
0.1
-1.9
Height [m]
0.5
0
-0.5
-1
0
500
1000
1500
Along-track profile [m]
Waveform variables of 1542
R [-]
1.0
0.0
Peaks [-]
6.0
1.0
Amp [-]
70.5
0.0
Width [m]
85.5
5.5
Kurt [-]
5.5
-0.9
Skew [-]
0.7
-2.2
Height [m]
1200
1150
1100
1050
1000
950
0
500
1000
1500
Along-track profile [m]

Figure 5.12: Waveform variables of river profiles with ID = 650 (left plot) and ID = 1542 (right plot) with corresponding Google Earth image.

5.2.2 EFFICIENCY OF CANDIDATE CLASSIFICATION CRITERIA

Mean and standard deviations
The sampling set with the Landsat-derived river footprints based on the visual inspection of the river cross-sections over the Mekong River with Landsat images is used to determine the mean and standard deviation of footprints into three categories: the land cover type (water/land), the season (wet/dry) and the basin (Upper Mekong Basin/Lower Mekong basin). The categories and the number of samples in each class are presented in Table 5.3.

Table 5.3: Categories and number of samples (N) for each class.

Categories	Group = 1	N	Group = 0	N
Land cover type	Water	194	Land	127
Season	Wet	125	Dry	196
Basin	Upper	33	Lower	288

A comparison of the mean and standard deviations per class indicates whether variables can be used for river footprint identification. This is illustrated in Table 5.4 for the category of land cover type. Note that the reflectivity is excluded based on the considerations in section 5.2.1. For the sampling set of river profiles two groups are defined: water is defined by the Landsat-derived river footprints and land is defined by all other footprints. This resulted in 194 water footprints and 127 land footprints.

Table 5.4: Class statistics for values per land cover type (water/land).

Variables	Water (N=194)		Land (N=127)	
	Mean	Std. dev.	Mean	Std. dev.
Adjacent height difference	1.46	8.42	8.10	14.90
Number of peaks	1.72	1.33	2.87	2.09
Maximum amplitude-ratio	1.68	9.32	5.97	14.80
Width	9.29	12.38	19.75	18.70
Kurtosis	1.20	2.87	1.34	6.10
Skewness	-0.13	0.65	-0.27	1.00

At first glance, the adjacent height difference, number of peaks, maximum amplitude-ratio and the width of the return have different means in the sampling set of river profiles. This corresponds with the theoretical expected behavior of the laser pulse with water surfaces. The water surface is in general flat and therefore will show smaller height differences between consecutive footprints in comparison with a more relieved land surface. The number of peaks over water is lower as well as the maximum amplitude-ratio and the width compare Figure 4.7. The kurtosis and the skewness show very similar means. This was not expected from previous studies resulting in ICESat classification schemes including both waveform parameters [32] [54]. However the sampling set of river profiles did already indicate that these parameters provided unstable information.

Student's t-test
To analyze whether the different means in the sampling set of the river profiles are significant, the equality of the means is tested with the statistical Student's t-test [10]. Although it is not necessary to have equal sample sizes in each class for this test, 127 random water footprints were selected. The Student's t-test as implemented in SPSS for two

independent samples has only one prerequisite condition: The two classes should have equal variances. It should be noted that the standard deviation is not a robust statistic and outliers can influence the results which can be repaired with equal variances in the classes. This is analyzed with Levene's test of equal variances [9]. The null hypothesis assumes (H_0) equal variances in the two classes while the alternative hypothesis (H_a) assumes unequal variances (σ^2) as formulated in the following two formulas:

$$H_0 : {\sigma_{water}}^2 = {\sigma_{land}}^2 \tag{5.1}$$

$$H_a : {\sigma_{water}}^2 \neq {\sigma_{land}}^2 \tag{5.2}$$

A significance level of $\alpha = 0.05$ is used. This implies that the null hypothesis is rejected for all significance values below $\alpha = 0.05$. If the classes have equal variances, the Student's t-test is based on the pooled variance instead of the individual variances of the classes. The null hypothesis (H_0) of the Student's t-test assumes equal means (μ) in the two classes, while the alternative hypothesis (H_a) assumes unequal means as formulated in the following two formulas:

$$H_0 : \mu_{water} = \mu_{land} \tag{5.3}$$

$$H_a : \mu_{water} \neq \mu_{land} \tag{5.4}$$

The results of the Student's t-test for independent samples are summarized in Table 5.5. The kurtosis is the only variable for which the null hypothesis of equal variances is not accepted (significance value = 0.036). To compare the means of the classes, the kurtosis (significance value = 0.829) and the skewness (significance value = 0.183) results in equal means. For all other variables, the null hypothesis of equal means is rejected.
This has consequences for the design of the classification criteria: Only variables with significant differences in mean values will contribute to the river footprint identification. Therefore, the four variables that are taken into account are the adjacent height difference, and the waveform information described by the number of peaks, the maximum amplitude-ratio and the width of the return pulse.

Table 5.5: Results of the Independent Samples T-Test for variables.

Variables	Equality of variances		Equality of means	
	F	Significance	t	Significance
Adjacent height difference	38.122	0.000	4.052	0.000
Number of peaks	57.352	0.000	5.079	0.000
Maximum amplitude-ratio	15.970	0.000	2.719	0.007
Width	19.100	0.000	5.106	0.000
Kurtosis	3.206	0.075	0.216	0.829
Skewness	4.451	0.036	-1.336	0.183

5.3 RIVER FOOTPRINT IDENTIFICATION

The candidate river footprints are rated based on their geometric and waveform characteristics. This is tested on the sampling set of river profiles. The geometric and

waveform scores are eventually weighted and applied to determine the river footprints in the Mekong River.

5.3.1 WORKFLOW OF RIVER FOOTPRINT IDENTIFICATION

The candidate footprints in the river cross-sections are evaluated using the variables defined as classification criteria. The relation of a candidate river footprint with its consecutive footprints is administrated with a geometric score. The individual footprints can also be evaluated based on the interaction of the laser with the surface. This will be described with a waveform score. In this section, the workflow for the calculation of the geometric and waveform scores is provided. The illustrations of the workflows in Figure 5.14 and 5.15 consist of several components. The components are in bold style for quicker reading.
The geometric and waveform scores are expressed in normalized scores. With normalization, the score (x) is rated with respect to a set minimum (min) and maximum (max). The normalization formula is described as:

$$Normalized\,(x) = \frac{x - min}{(\max - min)} \tag{5.5}$$

The advantage of using a normalized score is that footprints with waveform variables that do not perform as expected by river footprints, are not excluded from further evaluation but only rated as less likely to be a river footprint. With normalization of scores, the inconsistency of the laser interaction with the river due to sediment load, roughness of the water surface, different river width, et cetera will have less impact. This is as well a disadvantage: the value of the score is depending on the other candidate river footprints in the river cross-section. In this study, the normalization is implemented that the score is between 0 and 1: 0 is the lowest score and 1 is the best score.

Geometric score
Based on the assumptions, candidate river footprints are rated on the geometric information enclosed in the river cross-section. The workflow for calculating geometric scores is illustrated in Figure 5.14. The components will be discussed.

Decision 1: threshold on adjacent height difference
First, the candidate river footprints that are located in the Upper Mekong Basin are recognized. This is necessary for the threshold setting of the adjacent height difference. The adjacent height difference was formerly introduced in section 5.1.3 and is defined by the height difference between consecutive footprints in a river cross-section. It was assumed that low adjacent height difference flat ground surface such as water bodies.
The threshold for the adjacent height difference is defined by the theoretical vertical precision of ICESat and the maximum height difference as defined in the Upper Mekong Basin. The vertical precision of ICESat is evaluated and related to the land cover type [12]. The terrain height based on ICESat is compared with airborne laser scanning data over the Netherlands. Over bare land and urban areas, the mean terrain difference is below 0.25 m with a standard deviation of 0.23 m. For the purpose of water levels, it is assumed that in the Lower Mekong Basin the height difference between consecutive footprints should be very low because of the flat surface and the absence of large decline in the river profile. Therefore, two times the standard deviation (0.5 m) as derived from the assessment over bare land and urban areas will used as first threshold.
In the Upper Mekong Basin, the river decline is larger because the river declines by 4500 m in the initial 2200 km, resulting in an average decline of 2.0 m/km [26]. A maximum along-track spacing of 500 m between consecutive footprints is used for plausible river cross-

sections (see section 5.1.1) and therefore the theoretical maximum height difference between two consecutive footprints would be 1.0 m. It should be noted that the assumption of the maximum height difference in the Upper Mekong Basin depends on the alignment of the river cross-sections (parallel or perpendicular to the course of the Mekong River), and the height profile of the river. Therefore the height threshold for the Upper Mekong Basin is set at 2.0 m.

Decision 2: height difference ≤ height threshold
If candidate river footprints do not meet the height threshold (2.0 m in the Upper Mekong Basin, 0.5 m in the Lower Mekong Basin), the footprints are excluded from further processing. The excluded footprints have a flagged geometric score of -1. This is the only step which excludes footprints.

Geometric score 1: normalized height difference
Based on the assumption that river footprints have small adjacent height differences, each candidate river footprint is equipped with a score for the height difference with consecutive footprints. The absolute height difference is calculated. The adjacent height difference is compared with the height threshold and a minimum height value of 0.005. The minimum height value is included for the purpose of normalized scores.

Decision 3: > 1 group of adjacent inliers
Based on the previous step, the candidate river footprints that are selected based on the adjacent height difference are grouped as adjacent inliers. If more than 1 group is apparent, the grouped inliers can be rated on the number of footprints per group. An example of a river cross-section with multiple groups of inliers is presented in Figure 5.13.

Geometric score 2 (optional): normalized number of footprints per group
Groups of adjacent inliers with many footprints are more likely to be part of a flat surface such as a water body. In Figure 5.13, the left group consists of 2 footprints while the right group consists of 7 footprints. Although the left group can possibly correspond to a small channel, the probability of correct classification is more likely in the right group. It is assumed that a group of adjacent inliers consist of minimal 2 footprints (see section 5.1.1). An example of the normalization of this variable is as follows: The left group consisting of 2 footprints will receive the score = 0 (the normalized score as defined in formula 5.5. leads to 2-2/7-2 = 0). The right group with 7 footprints will receive the score = 1 (7-2/7-2 = 0).

The average geometric score is defined by either the normalized score of the adjacent height difference (if only one group of adjacent inliers is available) or the average of the normalized adjacent height difference and the normalized number of footprints per group of adjacent inliers (multiple groups of adjacent inliers are available).

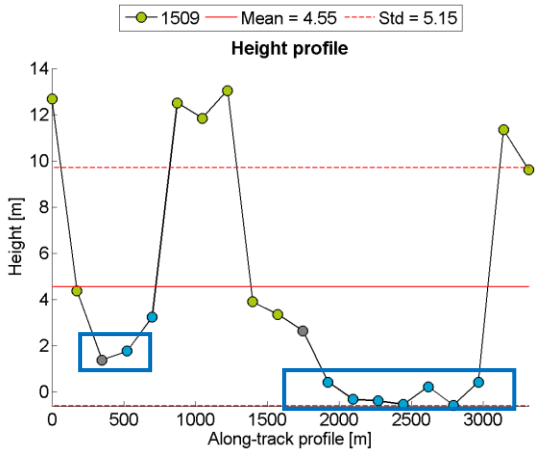


Figure 5.13: Example of a river cross-section from the sampling set with 2 groups of adjacent inliers (blue boxes). The adjacent inliers are defined by the adjacent height difference between consecutive footprints.

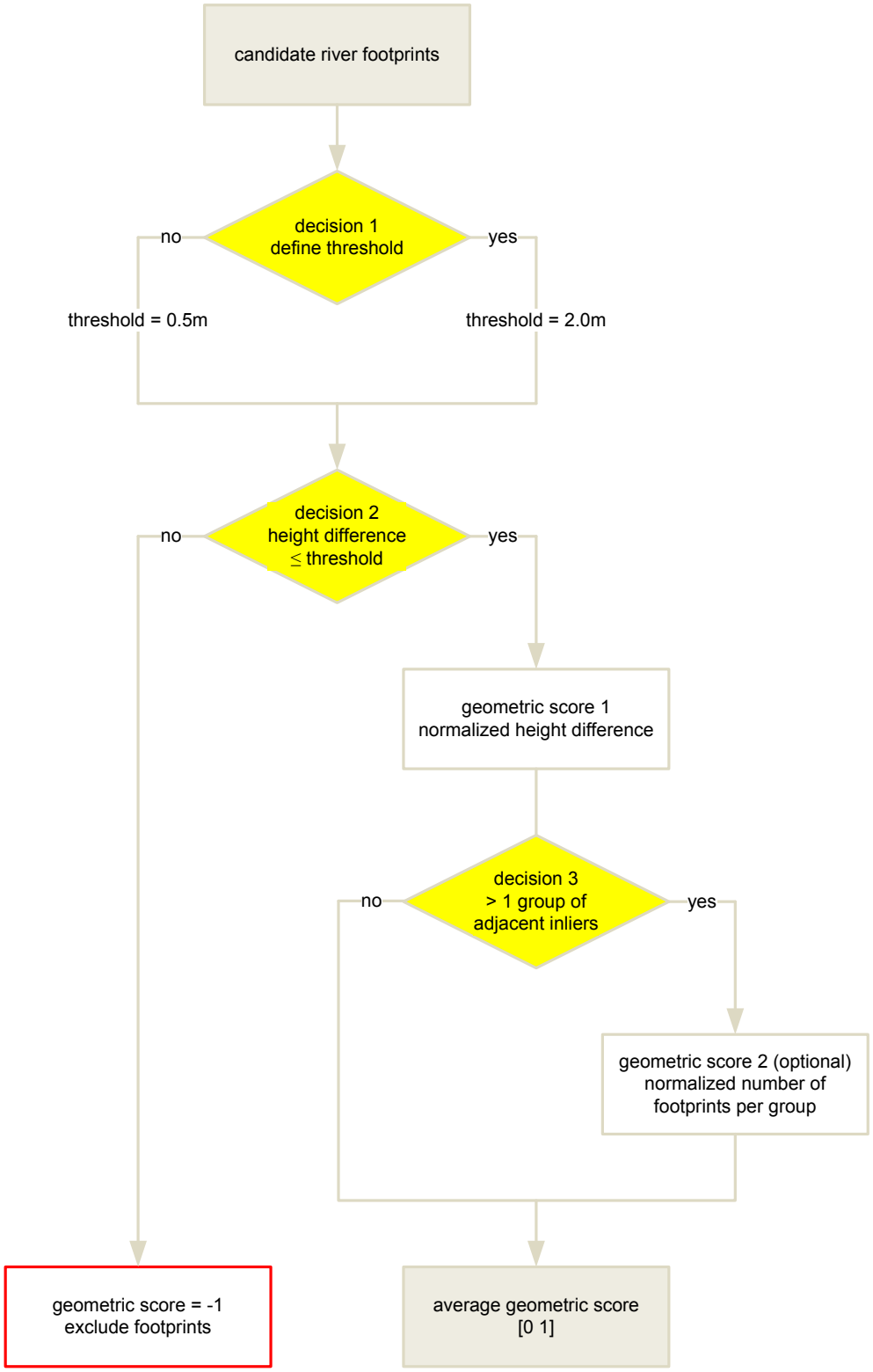


Figure 5.14: Workflow of the calculation of the geometric score. The yellow box represents a decision that has consequences for the next steps. The grey boxes represent the input and output of the workflow.

Waveform score
Candidate river footprints are also rated individually based on the waveform variables included in the customized GLA14 variables (see section 4.4.4). Recall that the GLA14 data product already included information about the waveform. However, some of the variables for example the amplitude-ratio are customized. The workflow for calculating waveform scores is illustrated in Figure 5.15. The components will be discussed.

Waveform score 1: normalized number of peaks
First the number of peaks in the return pulse is rated. Water surfaces are likely to have a single peak; however the waveform variables in the sampling set of river profiles show large variation. The number of peaks is therefore normalized in which a single peak is rated with the maximum score (score = 1) and the pre-defined maximum of six peaks is rated with the minimum score (score = 0).

Decision 1: return signal consists of single peak
It should be decided whether the return signal consists of a single peak or multiple peaks. If multiple peaks are included, the maximum amplitude-ratio is taken into account as a score. For water surfaces it is assumed that the pulse corresponding to water is more peaked than the other pulses. Therefore a single peak is expected but not consistently apparent in the sampling set of waveform variables (Appendix 2).

Waveform score 2 (optional): maximum amplitude-ratio
The maximum amplitude-ratio evaluates the amplitude of the last return pulse with regards to the other pulses. The amplitude of the last return pulse should at least be five times larger than the amplitudes of the remaining pulses. The threshold of 1:5 is somewhat arbitrary. Furthermore, ringing peaks caused by saturation effects are ignored. The maximum amplitude-ratio is not normalized which implies that footprints with a maximum amplitude-ratio larger than 1:5 obtain value 1 and otherwise value 0.

Waveform score 3: normalized width
The width of the signal is calculated as the difference between the end and begin of a return signal. Smaller widths correspond with flat surfaces such as water. The width is normalized and does not use any thresholds. For every footprint in the river cross-section, the width is compared with the minimum and maximum width values of the corresponding river cross-section. The smallest width gets the highest score.

The waveform score is calculated by the average of the scores. The maximum amplitude-ratio is optional and depends on the number of peaks in the waveform. The waveform score is therefore based on the number of peaks and the width if the return pulse is only a single peak. The waveform score has a minimum value of 0 and a maximum value of 1.
The geometric and the waveform score provide an index per footprint considering the probability that the footprints corresponds to the river. The candidate river footprints are updated with the corresponding geometric and waveform scores. Alternative weights of the scores can be used to identify river footprints. This is tested in section 5.2.4.
Including optional scores in the classification strategy results in a bias in the scores of the footprints. While some candidate river footprints have a score based on 5 classification variables, others can be based on the minimum of 3 classification variables. This is a disadvantage of the method. However, not including the optional scores will result in a loss of information that is contributing to the river footprint identification. This is a trade-off that is not further discussed in this study.

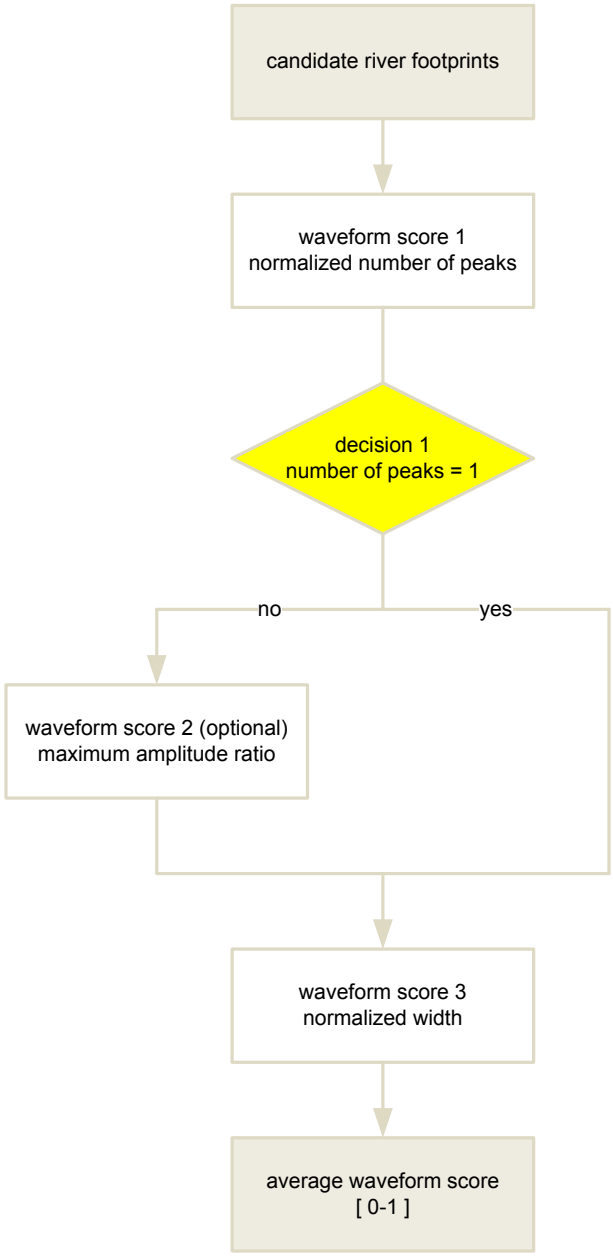


Figure 5.15: Workflow of the calculation of the waveform score. The yellow box represents a decision that has consequences for the next steps. The grey boxes represent the input and output of the workflow.

5.3.2 TESTING OF CLASSIFICATION SCORES

The workflows for the geometric score and waveform score are applied on the sampling set of river profiles. The geometric scores have values between -1 and 1, in which -1 is assumed to be an outlier based on the adjacent height difference. When removing the footprints with value = -1, the score is between 0 and 1 and the minimum, maximum and mean scores are calculated. For the waveform scores, no special handle for outliers is included. This will immediately result in a score between 0 and 1. The scores are multiplied with 100 to get a score between 0-100 percent which is more natural to interpret. The summary of the scores for each river profile in the sampling set are presented in Table 5.6.

Table 5.6: Summary of geometric and waveform scores of the sampling set of river profiles

ID	Geometric score			Waveform score		
	min	max	mean	min	max	mean
0969	38%	92%	74%	33%	100%	67%
2207	98%	98%	98%	3%	100%	55%
1540	68%	68%	68%	7%	33%	24%
1542	93%	93%	93%	0%	100%	39%
1786	80%	81%	81%	0%	100%	60%
0994	33%	91%	63%	0%	100%	48%
1525	2%	90%	54%	50%	100%	86%
1368	23%	97%	65%	1%	100%	58%
1509	11%	48%	32%	7%	100%	88%
1205	47%	100%	93%	32%	100%	93%
2068	44%	99%	74%	50%	100%	85%
0650	8%	99%	72%	60%	100%	87%

It appears that the waveform scores are often heading towards the maximum score of 100%. This is caused by the normalization of the waveform variables with respect to the minimum and maximum values within a footprint. This implies that two footprints from different river cross-sections, with different width values but with a single peak and a maximum amplitude-ratio larger than 1:5 can have both a waveform score of 100 percent. This is a drawback of the use of normalized scores.
Two river profiles of the sampling set with corresponding geometric and waveform scores are visualized in Figures 5.16 and 5.17. The results of the complete sampling set of river profiles are included in Appendix 2.
In Figure 5.16 , the left plot shows the geometric score of the river profile with ID = 650. Since the river profile is located in the Lower Mekong Basin, the threshold of the adjacent height difference between footprints is 0.5 m. The outer footprints do not meet this threshold and are therefore excluded from the geometric scores (red circles dots). The Landsat-derived river footprints do have a larger geometric score than the remaining included footprints. This is caused by the normalization of the adjacent height difference. With this criterion, the lower height value that is an outlier is therefore rated with a very low geometric score. The waveform score is calculated for all footprints independent of the results of the geometric scores. In Figure 5.16, the right plot shows the waveform scores. The footprints that do not correspond with the Mekong River have slightly lower waveform scores because of the larger width in comparison with the Landsat-derived river footprints. Nonetheless, the waveform scores are all above 60 percent.
In Figure 5.17, the river profile with ID = 1542 is visualized with the corresponding geometric and waveform scores. Since the river profile is located in the Upper Mekong Basin, the adjacent height difference is evaluated with a threshold of 2 m. Only two footprints meet this condition which implies the exclusion of valid river footprints. These two footprints are located near the centerline of the river. The waveform scores of the footprints show a large difference between the two included river footprints with a waveform score of 100 percent and the other footprints varying between 0 and 50 percent.

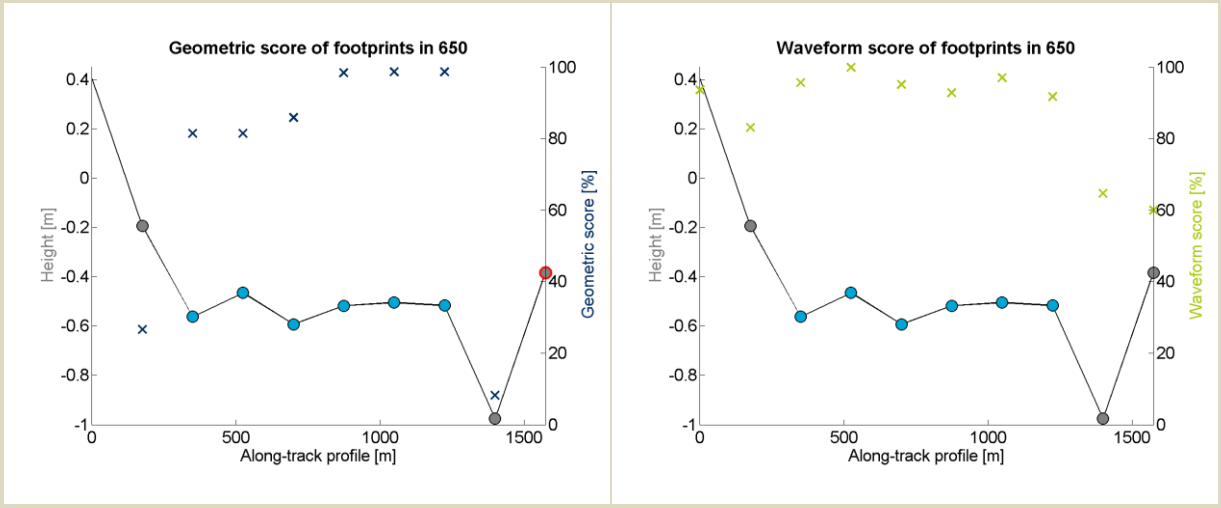


Figure 5.16: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 650. The turquoise dots indicate Landsat-derived river footprints. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference.

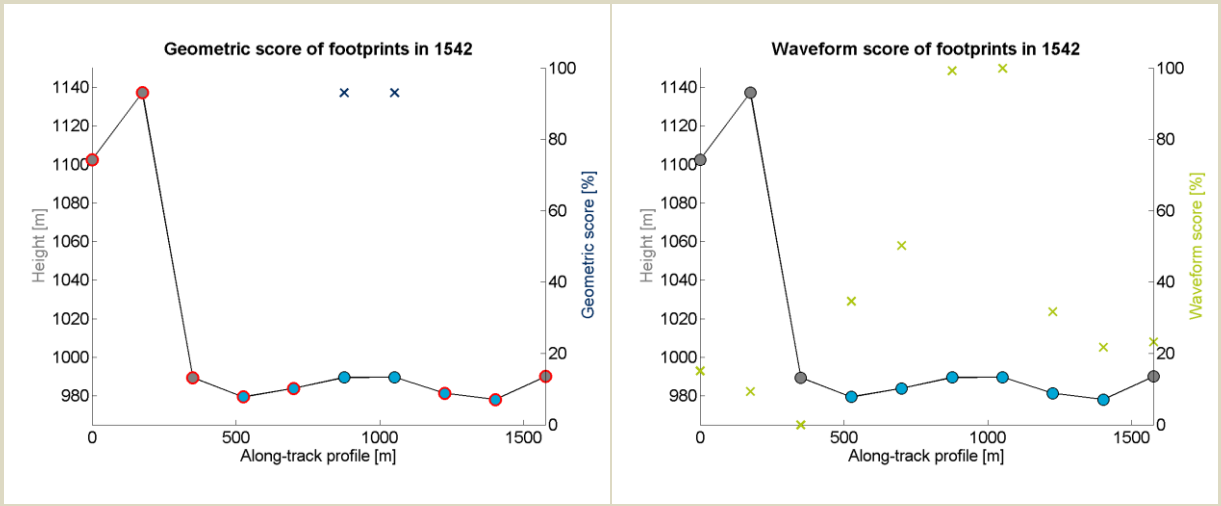


Figure 5.17: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 1542. The turquoise dots indicate Landsat-derived river footprints. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference.

5.3.3 WEIGHTING OF CLASSIFICATION SCORES

To assess the accuracy of different classification scores, statistics to describe the errors are introduced. The null hypothesis in the classification process is formulated as 'a footprint corresponds to water'. This implies that the alternatives hypothesis is formulated by the contrary statement and the footprint does not correspond to water. This can be described by the following formulas:

$$H_0 : footprints = water \quad (5.5)$$

$$H_a : footprints \neq water \quad (5.6)$$

The results of a classification experiment are derived from the theory of hypothesis testing and are illustrated in Table 5.7 [56]. For a footprint classified as water (H_0 is accepted), two options are possible: First, the classified water footprint matches the Landsat-derived river footprints. Then the H_0 is true and correctly accepted. This results in a true positive classification result. If the Landsat-derived river footprint does not correspond to water, the footprint is falsely classified. This will result in a false positive (Type I error). The false positive rate is the percentage of the classified footprints that is falsely rejecting the null hypothesis. The presence of type I errors implies that water level estimations will be partly derived from footprints not corresponding to the river. Depending on the height differences between the adjacent footprints this can have large influence on the water level estimations. It is also possible that a classified footprint is assumed to correspond to land (H_0 is rejected). If H_0 is false based on the Landsat-derived classification, the footprint is correctly classified as a true negative. If the footprint is a Landsat-derived river footprint, the footprint is falsely classified as outlier. This results in a false negative or a type II error. The false negative rate is the percentage of false negatives with respect to the river footprints. The consequence of false negatives is that river footprints are falsely ignored and do not contribute to the water level estimations.

Table 5.7: Different results of hypothesis testing.

	H_0 = true	H_0 = false
Accept H_0	True positive	False positive = Type I error
Reject H_0	False negative = Type II error	True negative

The river footprints are selected based on geometric and waveform scores. Alternatives of strictness in the scores are evaluated to maximize the accuracy of water level estimations. The alternatives are presented in Table 5.8. In the 'easy' alternative, footprints are classified as river footprints only on their geometric score. The 'gentle' alternative requires geometric and waveform scores above 25 percent. The 'mild' alternative selects footprints with a geometric score above 50 percent and a waveform score above 75 percent. The 'strict' alternative selected only footprints with a geometric score above 75 percent and a waveform score above 90 percent.

Table 5.8: Alternative scenarios for weighting geometric and waveform scores.

Alternative	Geometric score	Waveform score
1 (easy)	≥ 0%	Not applicable
2 (gentle)	≥ 25%	≥ 25%
3 (mild)	≥ 50%	≥ 75%
4 (strict)	≥ 75%	≥ 90%

The four alternative weighting methods are tested on the sampling set of river profiles. The results can be found in Appendix 3. Basically a stricter weighting alternative leads to a decrease of Type I errors and an increase of Type II errors. In other words, strict weights of the geometric and waveform score result in more accurate river footprints; however a lot of river footprints are excluded because they do not meet the scores. This is a trade-off between the number of footprints that will be included as water footprints and the accuracy of the water level estimation.
In Table 5.9, the mean and standard deviation of the water level estimations derived from the sampling set of river profiles are presented. With the easy alternative, water level estimations are calculated for all river profiles, while the strict alternative misses 4 river profiles. It is interesting that the use of different alternatives has in general less impact on the mean of the water level estimation. Only the river profile with ID = 1525 (see section

5.1.3) has a mean height of 202.94 m with the easy alternative while the other alternatives results in a mean height of 195.73 m. However, the standard deviations of the water level estimations do decrease with the degree of strictness.

Table 5.9: Comparing mean and standard deviation of heights for 4 alternatives.

ID	Easy		Gentle		Mild		Strict	
	M [m]	Std [m]	M [m]	Std [m]	M [m]	Std [m]	M [m]	Std [m]
0969	3467.77	0.77	3467.77	0.77	n/a	n/a	n/a	n/a
2207	3001.19	0.03	3001.19	0.03	3001.19	0.03	3001.17	n/a
1540	1778.75	0.12	1778.83	n/a	n/a	n/a	n/a	n/a
1542	989.62	0.10	989.62	0.10	989.62	0.10	989.62	0.10
1786	151.41	0.06	151.41	0.06	151.39	0.07	151.39	0.07
0994	325.00	4.27	324.95	3.88	325.93	0.17	325.93	0.17
1525	202.94	9.87	195.73	0.06	195.73	0.06	195.76	0.05
1368	131.64	2.30	131.30	1.79	131.22	0.21	n/a	n/a
1509	0.39	1.10	-0.40	0.12	n/a	n/a	n/a	n/a
1205	0.89	0.09	0.89	0.09	0.90	0.08	0.90	0.05
2068	0.56	0.17	0.56	0.17	0.54	0.17	0.70	0.04
0650	-0.54	0.21	-0.48	0.13	-0.53	0.04	-0.53	0.04

5.3.4 IMPLEMENTATION PROCEDURE FOR THE MEKONG RIVER

The weighting of the geometric and waveform score is a user-specific decision. It depends largely on the accuracy level that is requested. If the study requires very solid water level estimations, the strict alternative should be used. This will result in fewer water level estimations. Because this study is exploring the possibilities of GLAS, it is not required that the highest level of accuracy is reached. On the other hand, it is undesirable to obtain many fuzzy water level estimations. Therefore the mild alternative is selected for the weighting of the scores. This implies that footprints are selected that have a geometric score above 50 percent and a waveform score above 75 percent.

6 RESULTS AND DISCUSSION

This chapter will present and discuss the resulting GLAS derived water level estimations with corresponding precisions. The water level estimations in the Mekong River are presented in section 6.1. Water level trends are estimated based on nearby water level estimations. This is elaborated with two case studies in section 6.2. The validation of the results in the Mekong River Basin is presented in section 6.3. The Upper Mekong Basin and the Lower Mekong Basin have distinct characteristics which requests for a difference in validation methods.

6.1 WATER LEVEL ESTIMATIONS

The water level estimations are determined using the classification criteria (see section 5.2) and the mild alternative of selecting river footprints. This section will present the water level estimations and corresponding quality measures in the Mekong River Basin. The temporal and geographical coverage will be taken into consideration.

6.1.1 WATER LEVELS

The GLAS derived water level estimations are based on selecting all footprints with a geometric score above 50% and a waveform score above 75% according to the mild alternative of classification criteria.
The 2242 river cross-sections that were determined with the customized water mask (see section 3.2) initially resulted in 987 water level estimations. The water level estimations are simply estimated by the mean of the height values of the selected river footprints in each river cross-section. The quality of the water level estimations is estimated by the standard deviation of the selected height values. The standard deviation is thereby a measure for the precision of the water level estimation. The mean and standard deviation are non-robust statistics that are sensitive for outliers. These descriptive statistics are chosen to include information about the variance of the selected footprints. At least two footprints should be included to calculate a standard deviation, therefore 202 water level estimations are excluded because they consist of only one footprint. This results in 785 water level estimations with a quality measure based on the standard deviation. Compared with the number of river cross-sections that were taken as input in this method, 35 percent of the river cross-sections resulted in water level estimations. This can be further specified with respect to the geographical coverage in the basin. In the Upper Mekong Basin, 11.5 percent of the river cross-sections resulted in water level estimations, while in the Lower Mekong Basin 41.0 percent was achieved.
The geolocation of a water level estimation is determined by taking the median location (latitude/longitude) of footprints within each group of selected nearby river footprints. This is a very basic method which is suitable for the assessment of this study but it is recommended to study more precise geolocation determination methods for river cross-sections in future research. For example, when a river cross-section passes an island, two groups of water level estimations might be available. With this simple method, only one water level estimation per river cross-section is estimated. This results in loss of information in the geolocation of the Mekong River. Furthermore, the estimated water levels are based on measurements that do not coincide with the river's geolocation. This is illustrated in Figure 6.1 and 6.2. In Figure 6.1, a river cross-section is selected from the sampling set in which multiple land-water transitions occur. In Figure 6.2, the turquoise dots represent the river footprints that will be included for the water level estimations. The river footprints located on the right will lead to a decrease in the average water level while the geolocation is only set on the red marked footprint. This is a drawback of calculating a single measurement per river cross-section.

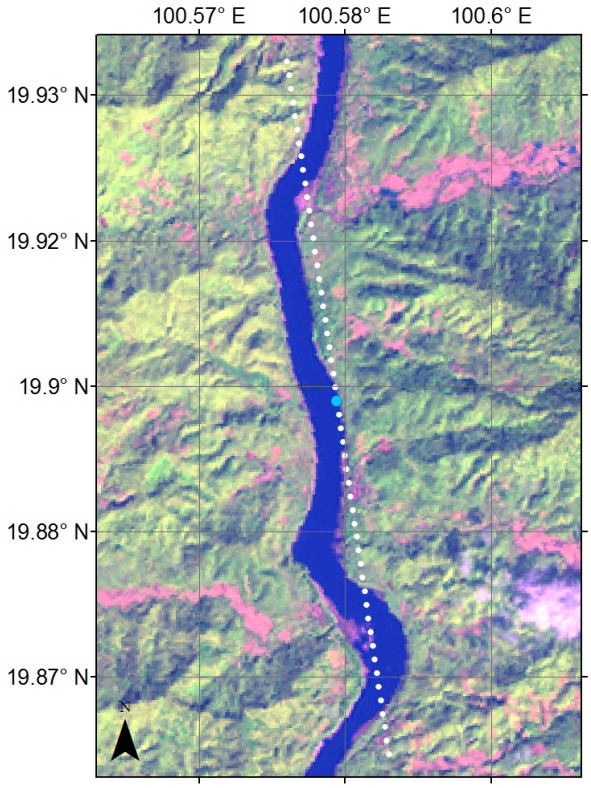


Figure 6.1: River cross-section over Landsat ETM image with multiple land-water transitions. The turquoise dot represents the simplified geolocation determination by median latitude/longitude of river footprints.

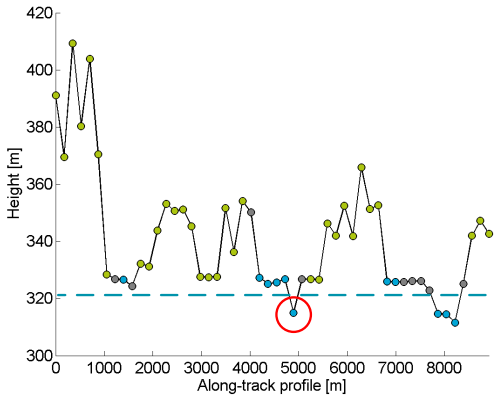


Figure 6.2: Effect of simplified geolocation determination by median latitude/longitude of river footprints (turquoise dots) and the determination of the water level (turquoise striped line).

Temporal coverage

The number of GLAS derived water level estimations per year is distributed equally over the operational periods between 2003 and 2009. This is illustrated in Figure 6.3 (left plot). Again, the number of water level estimations in the wet season is significantly lower caused by the scheduled campaigns (Figure 6.3, right plot). This will make it difficult to derive water level changes over time. Only 32 percent of the water level estimations correspond to wet season measurements. Initially, 27.1 percent of the candidate river footprints were measured during the wet season. This slight increase is likely to be caused by the better waveform characteristics during the wet season. The river stage will increase or even wider river channels will be apparent. The laser will interact accordingly (see section 4.3.3).

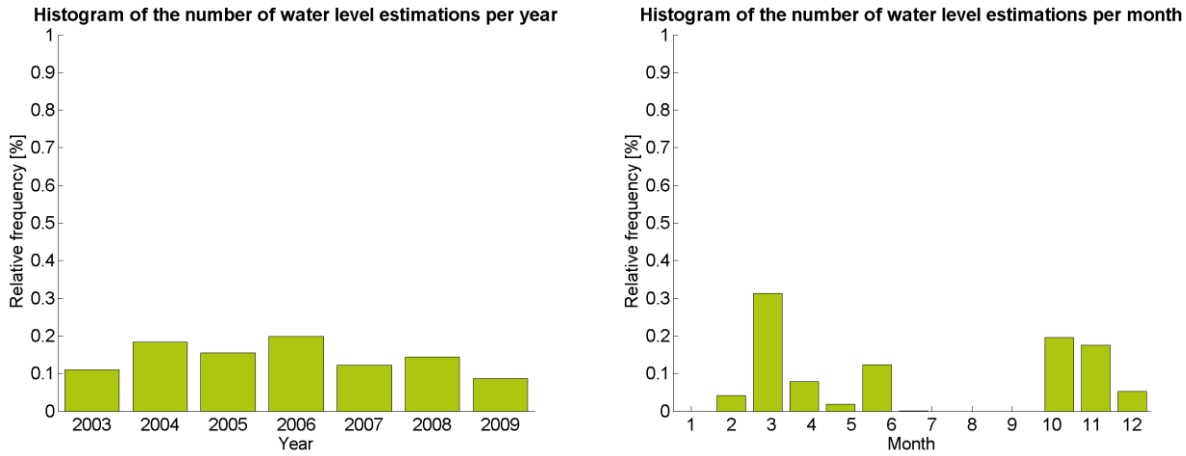


Figure 6.3: Temporal variation per year (left plot) and per month (right plot) of water level estimations.

Geographical coverage

The geographical coverage of the GLAS derived water level estimations in the Mekong Basin is presented in Figure 6.4. Only 52 water level estimations are located in the Upper Mekong Basin, which is 6.6 percent of the complete set of water level estimations. Furthermore, the left plot shows that a large part of the mainstream in the Upper Mekong Basin between 26-30 degrees of latitude is uncovered. This is interpreted as caused by the smaller width of the river. Furthermore, the number of footprints that are considered as river footprints is often based on fewer footprints in the Upper Mekong Basin. In the Upper Mekong Basin, the average number of footprints per water level estimation is 5 in comparison with 9 in the Lower Mekong Basin. Besides, 52% of the water level estimations in the Upper Mekong Basin are based on only 2 river footprints compared to 31% in the Lower Mekong Basin. Again, this is likely to be caused by the narrow width of the river.

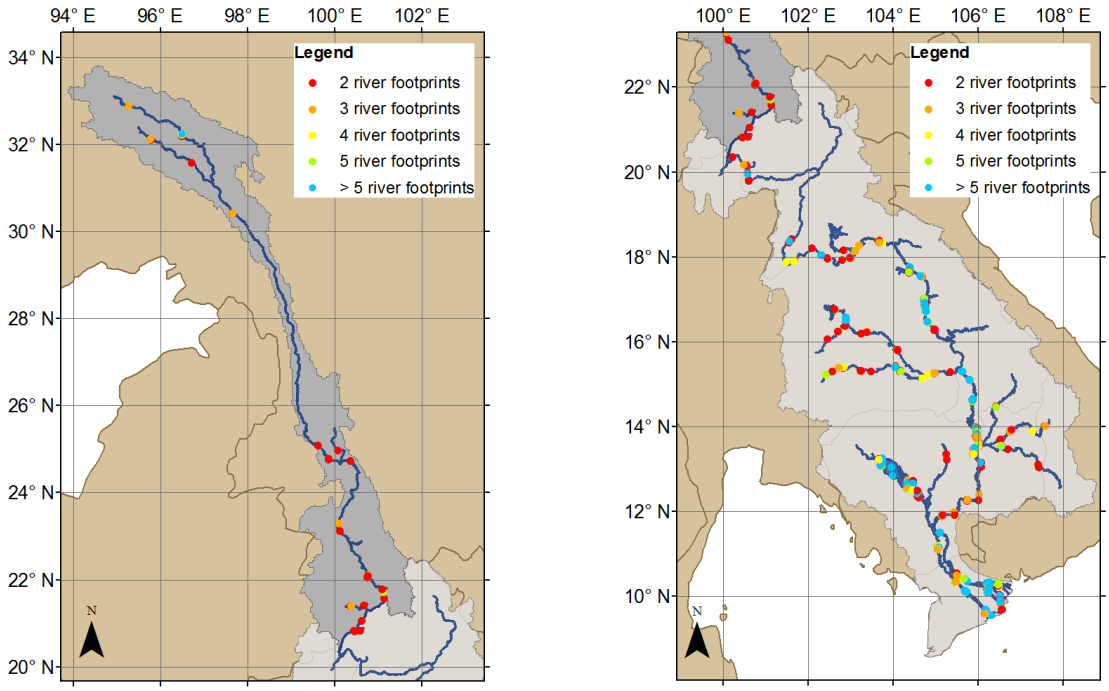


Figure 6.4: Locations of water level estimations in the Upper Mekong Basin (left plot) and the Lower Mekong Basin (right plot). The number of river footprints per water level estimation is indicated with colors.

Nearby water level estimations

For the monitoring of water levels the number of nearby water level estimations is relevant. Nearby water level estimations are found by searching water level estimations within a radius of 2 km. For the water level estimations, all nearby water level estimations are identified to evaluate monitoring opportunities. As noted in section 4.1, the operational periods of the GLAS instrument took place between 2003-2009. Only water level trends in this time period can be derived.

In the Upper Mekong Basin, 29 percent of the water level estimations do not have any nearby water level estimations. This implies that monitoring of water levels is not possible on the corresponding locations in the Mekong River. This is illustrated in Figure 6.5, left plot. The red dots are all single water level estimations. Only a few locations have up to 6 nearby water level estimations. In the Lower Mekong Basin, only 6 percent of the water level estimations have no nearby water level estimations. Furthermore, the maximum is 18 nearby water level estimations, while most are clustered together in sizes of 2 until 12 water level estimations. A closer look at Figure 6.5 (right plot) shows that the Mekong River Delta near the outlet of the river in the South China Sea contains a lot of nearby water level estimations. This is caused by the wider channels. The width of the river is therefore an important characteristic for the suitability of the GLAS instrument for the monitoring of water level estimations.

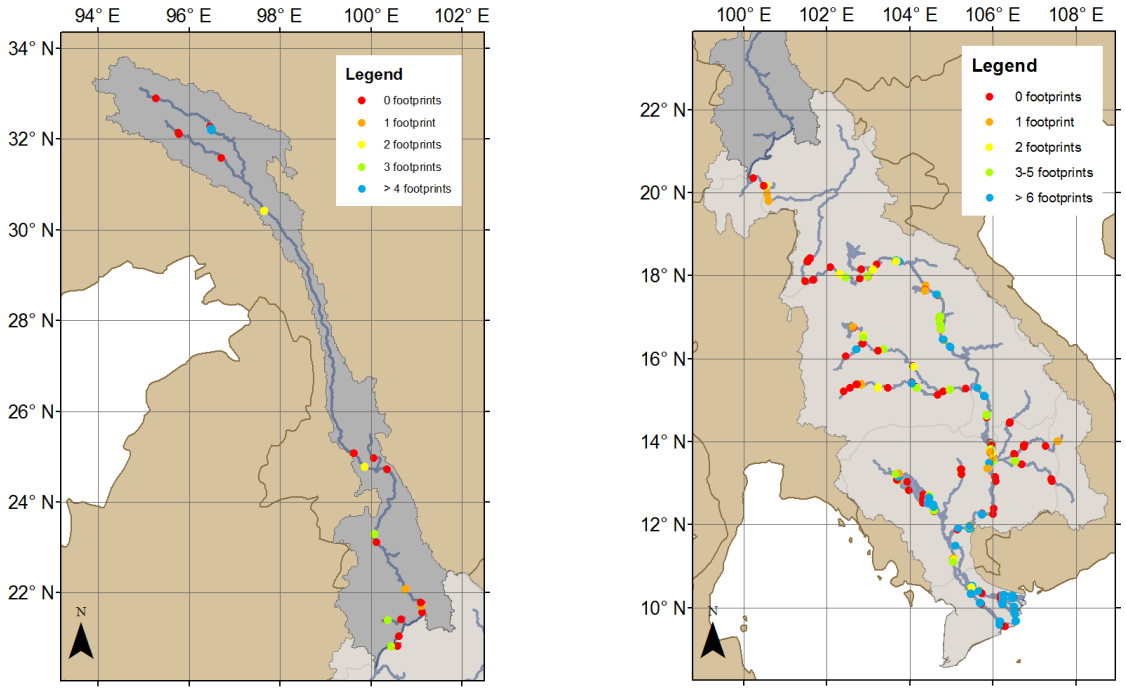


Figure 6.5: Number of nearby water level estimations within a radius of 2 km in the Upper Mekong Basin (left plot) and the Lower Mekong Basin (right plot).

6.1.2 PRECISION OF WATER LEVELS

The standard deviation of the GLAS derived water level estimations is calculated based on the input of all river footprints per river cross-section. The variation in the standard deviation is relevant for the precision of the GLAS derived water level estimations. To analyze the standard deviation in terms of geographical and seasonal variation, outliers of large standard deviations need to be identified. The distribution of the standard deviations is presented in Figure 6.6. Based on the left histogram, it is concluded that most standard deviations have small values; 78.9 percent of the standard deviation have a value smaller

than 0.2 m. The right plot shows that most values of the standard deviation between 0-0.2m are almost equally distributed.

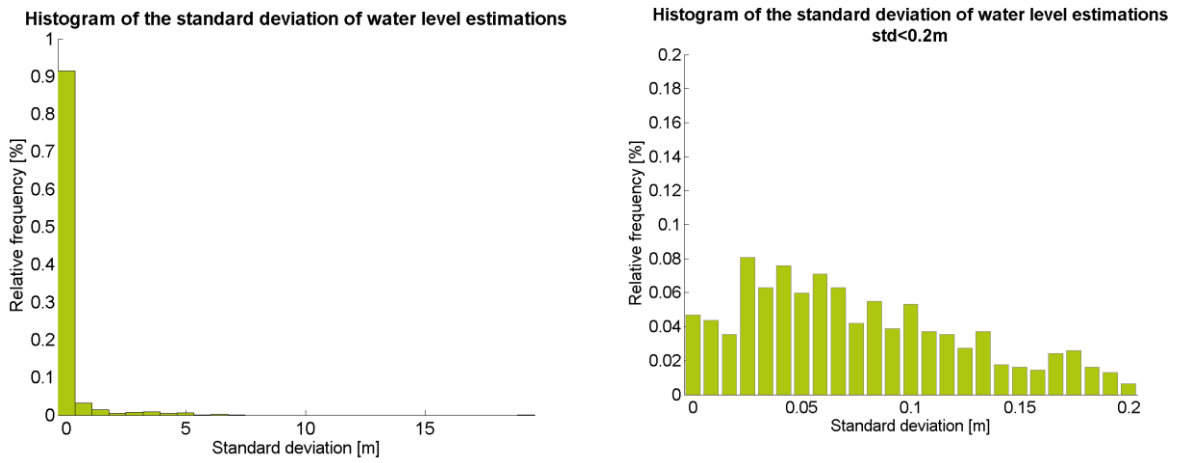


Figure 6.6: Histograms of the standard deviation of water level estimations (left plot) and with selected standard deviation smaller than 0.2 m (right plot).

The overall precision of the GLAS derived water level estimations in the Mekong River is calculated by the median of all standard deviations. The median is a robust estimator of the central tendency which reduces the influence of outliers. The results are presented in Table 6.1. Based on the 785 water level estimations, the median standard deviation is 0.087 m.

Table 6.1: Standard deviation of water level estimations. N represents the number of water level estimations, the median, minimum and maximum refer to the descriptive statistics of the standard deviations.

Overall quality	N	Median [m]	Minimum [m]	Maximum [m]
	785	0.087	0.00007	19.22

In Figure 6.7 and 6.8, the river cross-section (ID = 2640) with the maximum standard deviation of 19.22 m is visualized. This river cross-section is located in the Upper Mekong Basin near the source of the Mekong River. The river cross-section intersects the Mekong River three times. Based on the mild alternative of weighted geometric and waveform scores (Figure 6.8), a footprint corresponding with land is included with a larger height than the river footprints. This results in a rise of the mean height of the water level and the standard deviation. The extreme standard deviation is mainly caused by the extreme height variation in the Upper Mekong Basin.

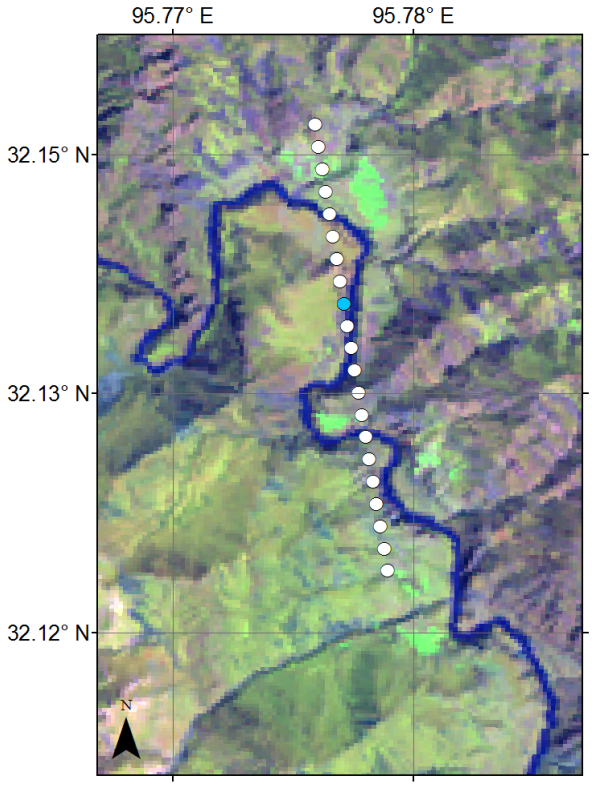


Figure 6.7: River cross-section with ID = 2640 corresponding to the maximum standard deviation over Landsat ETM image. The blue dot refers to the geolocation of the water level estimation.

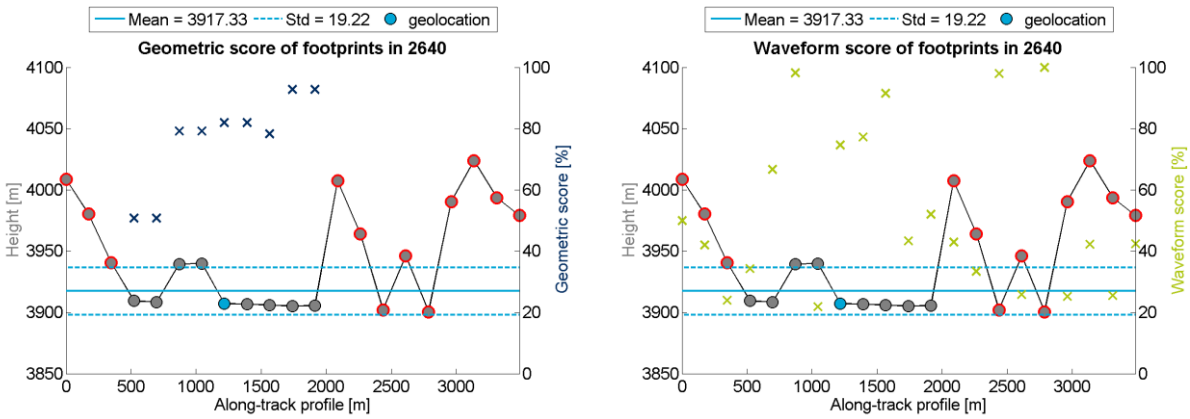


Figure 6.8: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 2640. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference.

It is also checked whether the standard deviation differs by the seasonal and geographical spreading. To do so, the median, minimum, and maximum of the standard deviations are given in Table 6.2 and Table 6.3. In Table 6.2, the median of the standard deviation does not differ in the dry and wet season. In Table 6.3, it is shown that the median of the standard deviation in the Upper Mekong Basin is larger than the Lower Mekong Basin. Again, this highlights the limitations of the GLAS in the Upper Mekong Basin.

Table 6.2: Standard deviation of water level estimations per season. N represents the number of water level estimations, the Median [m], Minimum [m] and Maximum [m] refer to the descriptive statistics of the standard deviations.

Season	N	Median [m]	Minimum [m]	Maximum [m]
Wet season	250	0.087	0.0014	6.40
Dry season	535	0.091	0.00007	19.22

Table 6.3: Standard deviation of water level estimations per basin. N represents the number of water level estimations, the Median [m], Minimum [m] and Maximum [m] refer to the descriptive statistics of the standard deviations.

Basin	N	Median [m]	Minimum [m]	Maximum [m]
Upper Mekong Basin	52	0.41	0.0014	19.22
Lower Mekong Basin	733	0.08	0.00007	6.40

6.2 WATER LEVEL TRENDS DERIVED FROM WATER LEVEL ESTIMATIONS

Nearby water level estimations in the Mekong River provide information about the water level change over time. This supports the monitoring of river networks. Two case studies will show water level trends derived from GLAS altimetry data between 2003 and 2009.

6.2.1 STUDY AREA 1: MEKONG RIVER MAINSTREAM

The first study area is located in the Mekong River Delta, approximately 250 km from the outlet of the Mekong River in the South China Sea. The GLAS derived water level estimations and the ground track are visualized in Figure 6.9. Note that the water level estimation during campaign L3A is located outside the boundaries of the river banks. An overview of the characteristics of these water level estimations is presented in Table 6.4. The perpendicular distance between the water level estimations and the ground track is calculated. Within a maximum perpendicular distance of 1181 m from the ground track, 8 water level estimations between 2004 and 2008 are obtained.

Table 6.4: Characteristics of water level estimations in the Mekong River Delta. The Mean [m] and the Std [m] refer to the water level estimation. The Length [m] is the perpendicular distance between the water level estimation and the ground track.

ID	Campaign	Date	Mean [m]	Std [m]	Length [m]
0599	L2B	17-03-2004	0.39	0.05	300
0901	L3A	07-11-2004	7.18	0.07	635
1107	L3B	24-03-2005	0.16	0.03	956
1245	L3C	26-06-2006	0.20	0.03	468
1588	L3E	23-03-2006	0.42	0.09	481
1984	L3G	30-11-2006	4.15	0.12	1181
2204	L3H	08-04-2007	0.53	0.14	1086
2523	L3J	17-03-2008	0.73	0.03	1019

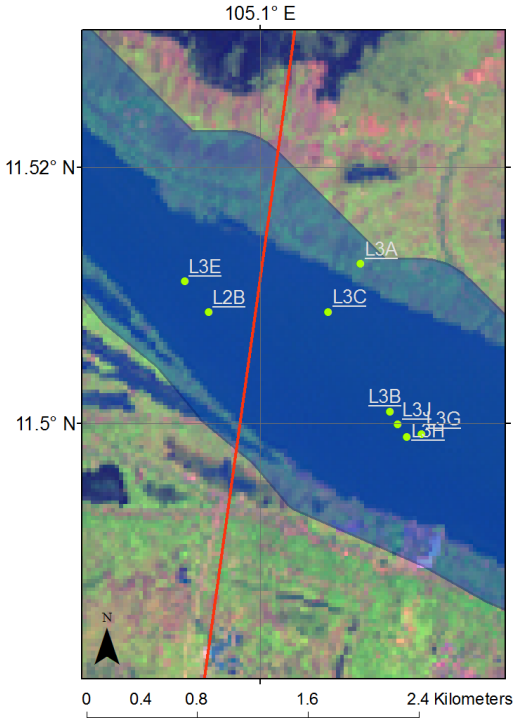


Figure 6.9: Nearby water level estimations (green dots) corresponding to repeated ground track (red line) in the Mekong River Delta over Landsat ETM image. The labels indicate the campaigns.

The GLAS derived water level estimations are plotted as a function of time. This is illustrated with a bar chart in Figure 6.10. Two measurements (L3A and L3G) are obtained during the wet season and show an increase in the water level estimation. In November 2004, the water level was approximately 7 m higher than normal. It could be interpreted that this extreme increase resulted in a flooding since the geolocation of the water level estimation is located outside the river boundaries in Figure 6.9. Whether seasonal amplitudes of 7 m are accurate is evaluated in section 6.3.2.

The correctness of the river footprint identification for the extreme water level estimations is further studied with the geometric and waveform score of the water level estimation obtained during campaign L3A in Figure 6.11. It appears that two outliers (extreme low values) are included which pushes the geolocation more to the side of the channel. Further study is needed to find explanations for the presence of these outliers. The centerline of the river cross-section differs approximately 1 km from the geolocation of the water level estimation. This counterbalances the argument that a flooding was apparent in November 2004. This highlights the limitations of a simple method for the definition of the geolocation.

In Figure 6.10, a second peak corresponds with a water level estimation in the wet season of 2006. Information about the wet seasons of 2005 and 2007 are missing because of the scheduling of the campaigns. This is a disadvantage in the use of GLAS altimetry data for the monitoring of river networks.

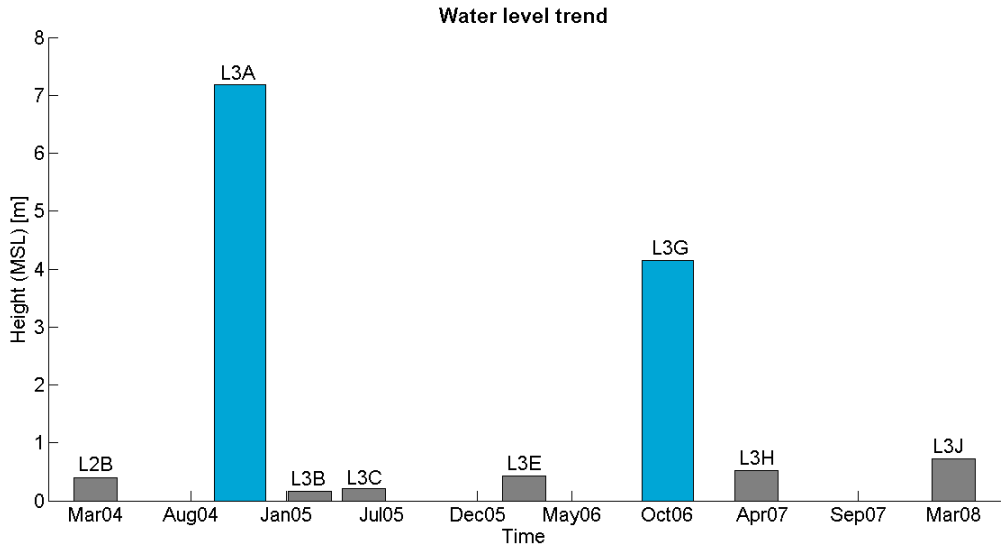


Figure 6.10: Water level trend between 2004-2008 in the Mekong River Delta. The turquoise bars indicate measurements in the wet season; the grey bars indicate measurements in the dry season. The labels indicate the campaigns

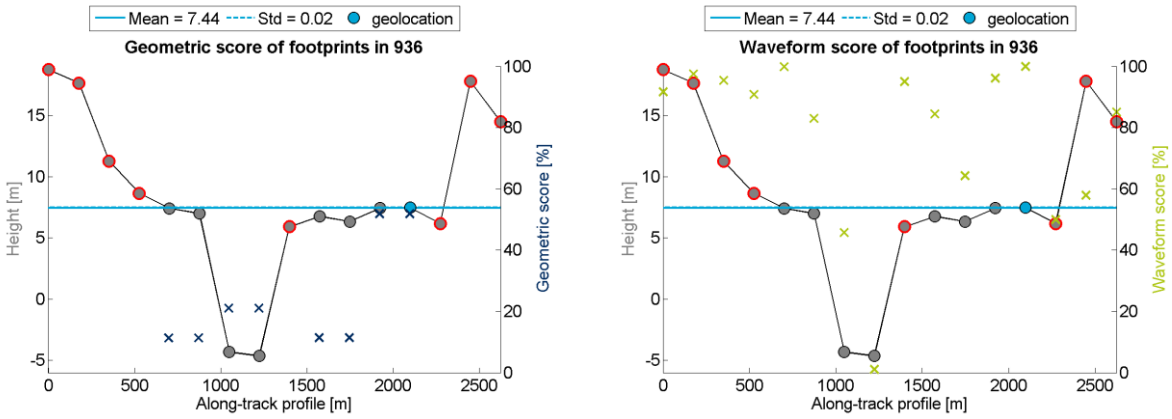


Figure 6.11: Geometric (left plot) and waveform (right plot) scores of the river profile with ID =936. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference.

6.2.2 STUDY AREA 2: NAM PHONG TRIBUTARY (UBOLRATANA DAM)

Another case study is analyzed in the Nam Phong Tributary. This tributary connects the Mekong River with the Ubolratana Dam Reservoir as illustrated in Figure 6.12 and 6.13. The Ubolratana Dam is a multi-purpose dam which besides the generation of hydro-electric power supports flood control. The Ubolratana Dam Reservoir has a storage capacity of 2,599,000m3 [57]. The characteristics of the water level estimations in the Nam Phong Tributary are summarized in Table 6.5. Between November 2003 and April 2009, ten water level estimations are located downstream of the Ubolratana Dam, within a maximum distance of 1405 m of the ground track.

The water level trend derived from the water level estimations is plotted in Figure 6.14. In general, water levels obtained during the wet season lead to an increase in the water level estimations. However, in this study area the water level trends are reversed: The water level estimations obtained in the wet season are lower. This is interpreted as a consequence from the flood control performed at the Ubolratana Dam.

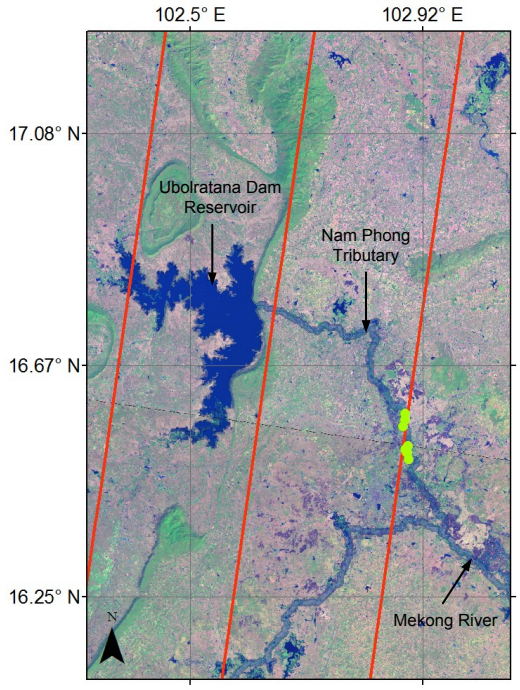


Figure 6.12: Water level estimations (green dots) corresponding to repeated ground track (red line) in the Nam Phong Tributary over Landsat image. The Nam Phong Tributary connects the Ubolratana Dam Reservoir with the Mekong River.

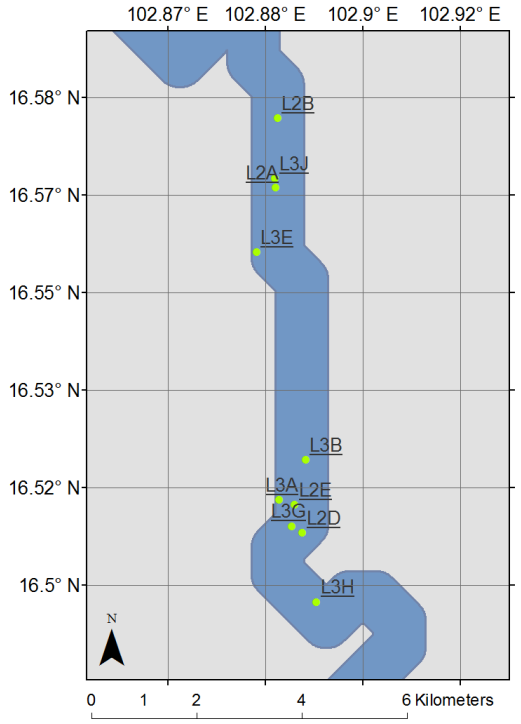


Figure 6.13: Close-up image of the water level estimations (green dots) in the Nam Phong Tributary. The labels refer to the campaigns.

Table 6.5: Characteristics of water level estimations in the Nam Phong Tributary. The Mean [m] and the Std [m] refer to the water level estimation. The Length [m] is the perpendicular distance between the water level estimation and the ground track.

ID	Campaign	Date	Mean [m]	Std [m]	Length [m]
0423	L2A	21-11-2003	154.57	0.04	276
0603	L2B	21-03-2004	155.33	0.20	416
0905	L3A	07-11-2004	150.70	0.45	537
1114	L3B	25-03-2005	151.67	0.48	88
1594	L3E	31-03-2006	154.50	0.10	450
1991	L3G	25-11-2006	151.60	0.57	804
2211	L3H	20-04-2007	152.08	0.13	1405
2532	L3J	20-03-2008	155.15	0.15	318
2760	L2D	16-12-2008	151.61	0.37	1005
2923	L2E	08-04-2009	151.54	0.43	811

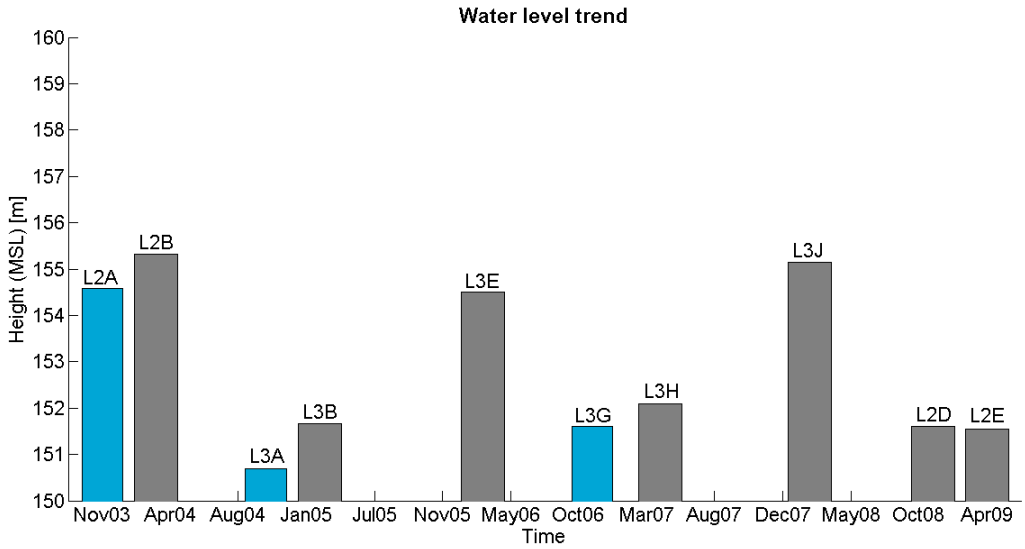


Figure 6.14: Water level trend between 2003-2009 in the Nam Phong Tributary. The turquoise bars indicate measurements in the wet season; the grey bars indicate measurements in the dry season. The labels indicate the campaigns

6.3 VALIDATION OF RESULTS IN MEKONG RIVER BASIN

Ultimately the quality of the GLAS derived water level estimations is evaluated. The water level estimations are internally validated by the construction of downstream river profiles. In the Lower Mekong Basin, in situ data provided by the Mekong River Commission (MRC) is gathered and compared with the water level estimations.

6.3.1 DOWNSTREAM RIVER PROFILES

The GLAS derived water level estimations in the Mekong River are difficult to validate. The absence of in situ measurements in the Upper Mekong Basin makes internal validation necessary. This also highlights the benefits of using GLAS altimetry data for providing at least some water level estimations in remote areas or countries that do not share water level data. To draw conclusions about the internal correctness of the GLAS derived water level estimations, two downstream river profiles are created.

Construction of channels
It is expected that the estimated long profile based on the water level estimations will decrease as a function of the distance from the source. The downstream river profiles in the Mekong River Basin are based on the USGS HydroSHEDS river network and the USGS HydroSHEDS digital elevation model (DEM) (see section 3.1.1). First, the streams represented as line features in the river network are split at the vertices. This results in line segments for which the length is calculated in ArcGIS.
In the Upper Mekong Basin, the channel representing the mainstream of the Mekong River is created in two steps: First, the line features intersecting the Upper Mekong Basin are selected. Second, the branches that correspond with the mainstream are added manually. Only the left floodplain is taken into account, otherwise two channels would be available. To exclude nearby multiple channels in the downstream river profile, the water level estimations within a distance of 5 km from the channel are selected. The channel with the corresponding 43 GLAS derived water level estimations are visualized in Figure 6.15.
To select the mainstream channel of the Mekong River in the Lower Mekong Basin, the streams of the USGS HydroSHEDS river network are selected based on the flow accumulation. Only line features with more than 990,000 upstream cells are selected. The number of upstream cells is chosen based on a plausible fit of the mainstream in the Lower

Mekong Basin. The water level estimations within a distance of 5 km from the channel are selected. The channel with the corresponding 348 GLAS derived water level estimations are visualized in Figure 6.16.
For each of the water level estimations, the nearest stream in the channel is calculated. This enables the construction of a spatial join between the water level estimations and the corresponding nearest stream to include the information about the length of the streams and the number of upstream cells to the water level estimations. Furthermore, the value of the USGS HydroSHEDS DEM at the locations of the selected water level estimations is included.

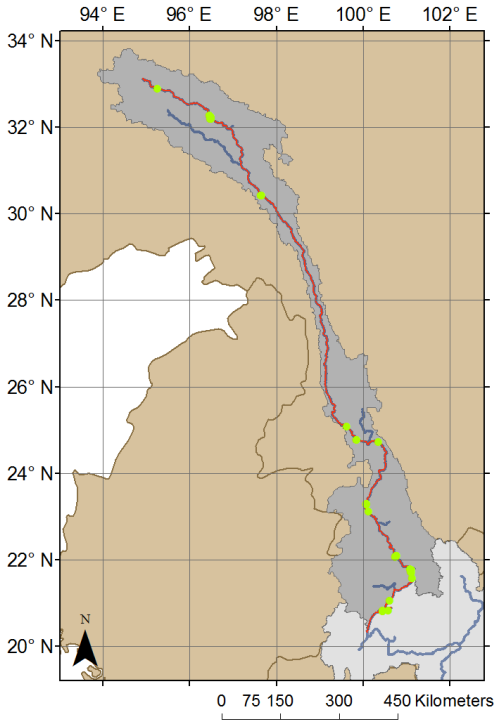


Figure 6.15: Channel constructed manually (red line) with GLAS derived water level estimations within a distance of 5 km (green dots) in the Upper Mekong Basin.

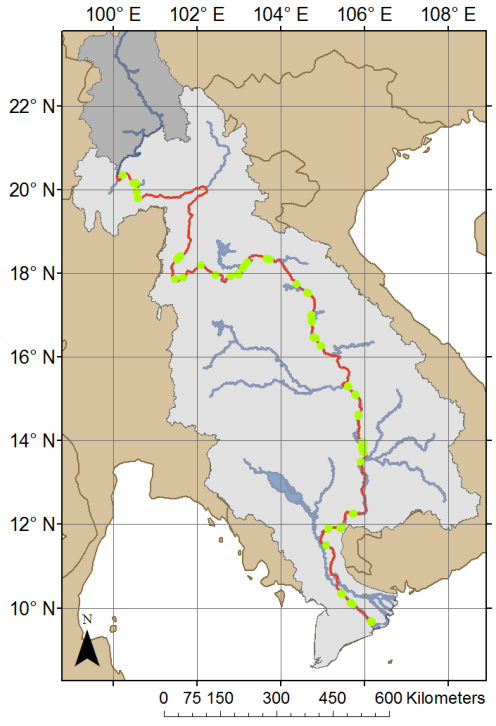


Figure 6.16: Channel constructed based on flow accumulation (red line) with GLAS derived water level estimations within a distance of 5 km (green dots) in the Lower Mekong Basin.

Downstream river profiles based on USGS HydroSHEDS DEM
Since water is flowing downstream, the water level estimations are sorted in descending order based on the number of upstream cells. Less upstream cells indicate a more upstream location and should result in higher water level estimations. Now the cumulative length is calculated which creates the long distance from the start of the channel until the outlet in the Southeast China Sea.
Downstream river profiles for the channels in the Upper Mekong Basin and Lower Mekong Basin are derived with the values of the USGS HydroSHEDS DEM at the locations of the water level estimations. This is visualized in respectively Figure 6.17 and 6.18. Because the DEM has a coarse spatial resolution of 15 arc-second (approximately 500 m at the equator), the values do not completely overlap the Mekong River. Note that in Figure 6.18 an extreme outlier is apparent. For better interpretation, a smoothed downstream profile is created with an interpolated line fitted to the USGS HydroSHEDS DEM value at the long distance. A 9th order interpolating polynomial is used. The influence of the outlier is insignificant on the smoothed downstream profile.

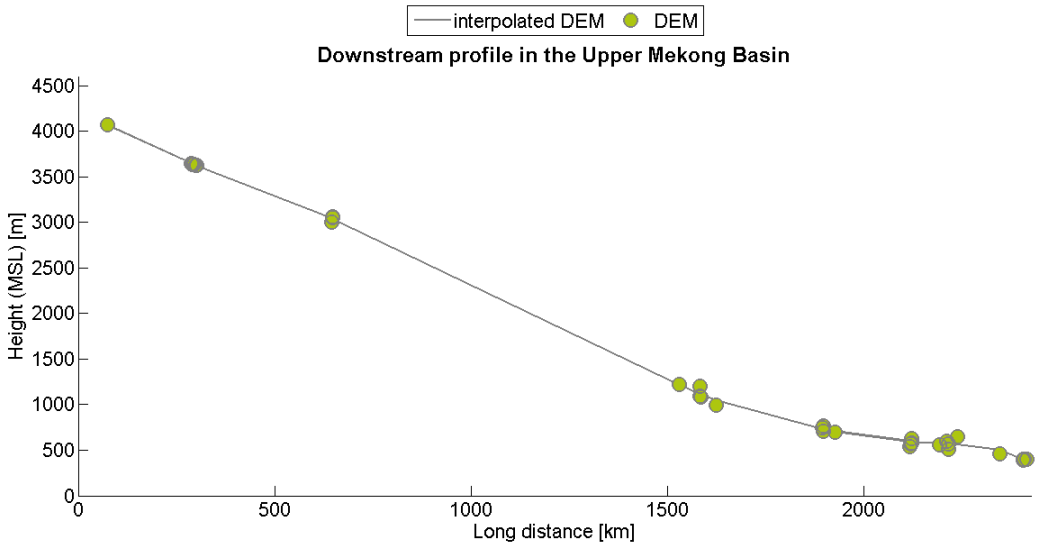


Figure 6.17: Downstream river profile created with a 9th order polynomial (grey line) based on the values of the USGS HydroSHEDS DEM at the location of the water level estimations within a distance of 5 km of the channel in the Upper Mekong Basin.

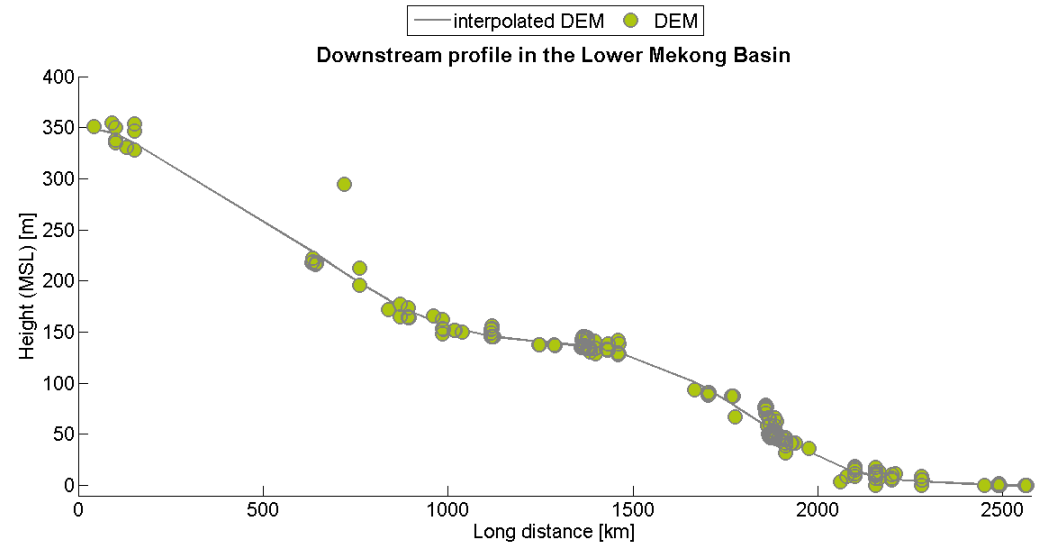


Figure 6.18: Downstream river profile created with a 9th order polynomial (grey line) based on the values of the USGS HydroSHEDS DEM at the location of the water level estimations within a distance of 5 km of the channel in the Lower Mekong Basin.

GLAS derived water level estimations over downstream river profiles

Instead of plotting the USGS HydroSHEDS DEM values, the water level estimations can be visualized over the interpolated DEM to check for consistency in elevation values. This is visualized in Figure 6.19 and 6.20. The long distance of both downstream profiles corresponds with the length of the Mekong River (see section 2.2.1).

Furthermore, it appears that most values do match the interpolated DEM. In Figure 6.19, two extreme outliers labeled with ID = 1693 and ID = 0547 in the mainstream appear: it is interesting to get more details of these water level estimations.

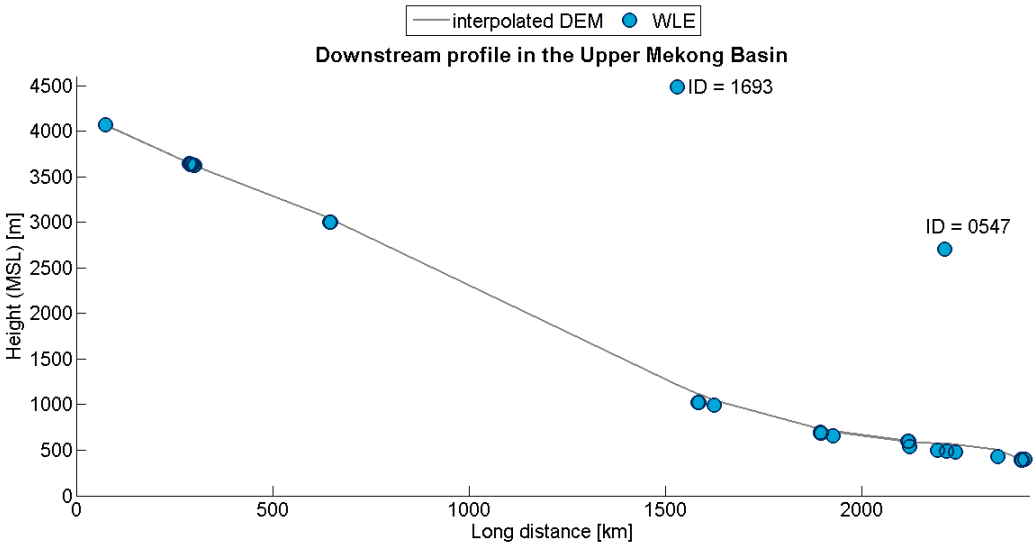


Figure 6.19: Downstream river profile created with a 9th order polynomial (grey line) and the 43 GLAS derived water level estimations within a distance of 5 km of the channel in the Upper Mekong Basin.

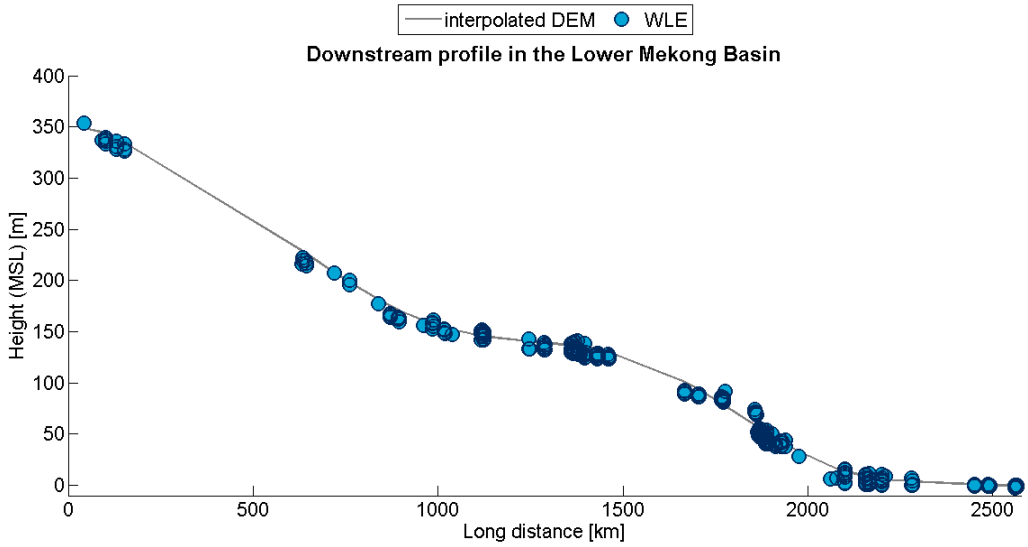


Figure 6.20: Downstream river profile created with a 9th order polynomial (grey line) and the 348 GLAS derived water level estimations within a distance of 5 km of the channel in the Lower Mekong Basin.

To understand the outliers, the corresponding river cross-sections with the geometric and waveform scores are visualized for both outliers (ID = 1693 and ID = 0547) in Figure 6.21 and 6.22. The height profiles do not show clear channel geometries. It is possible that the water levels are estimated coincidently due to small adjacent height differences between the consecutive footprints. Nearby located river cross-sections are plotted in Figure 6.23 and 24 for the both outliers to check consistency in the height levels. In Figure 6.23, the nearby located river cross-section differ around 3000 m from the outlier with ID =1693. For the outlier with ID = 547, the difference is approximately 1500 m (Figure 6.24). Furthermore, both outliers are compared with the USGS HydroSHEDS digital elevation model. The DEM does not show corresponding large height values in the environment. From this it can be concluded that these outliers are not caused by errors in the classification method for river footprint identification but rather by an error in the range measurements obtained with the GLAS altimetry data. The outliers could be solved by introducing a vertical buffer based on

a DEM. This will exclude outliers above a certain threshold compared to the nearest DEM value.
It should be noted that the downstream river profiles do not take repeated ground tracks into account. Some water level estimations are grouped together. Those nearby water level estimations are obtained during different campaigns but correspond to the same ground track. The repeated ground tracks do not completely overlap and show a few hundreds of meters of deviance of the geolocation (see section 6.2). The repeated water level estimations will show some fluctuations related to the season and year of data acquisition.

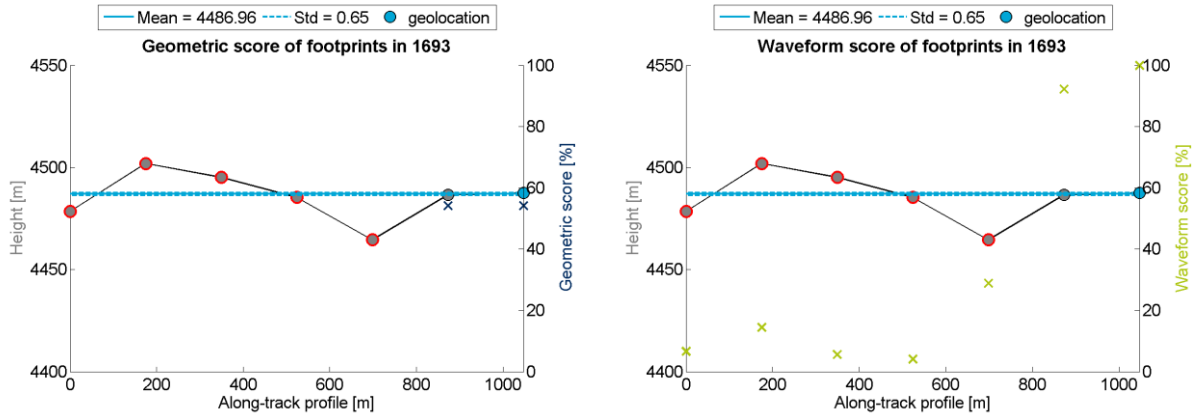


Figure 6.21: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 1693. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference.

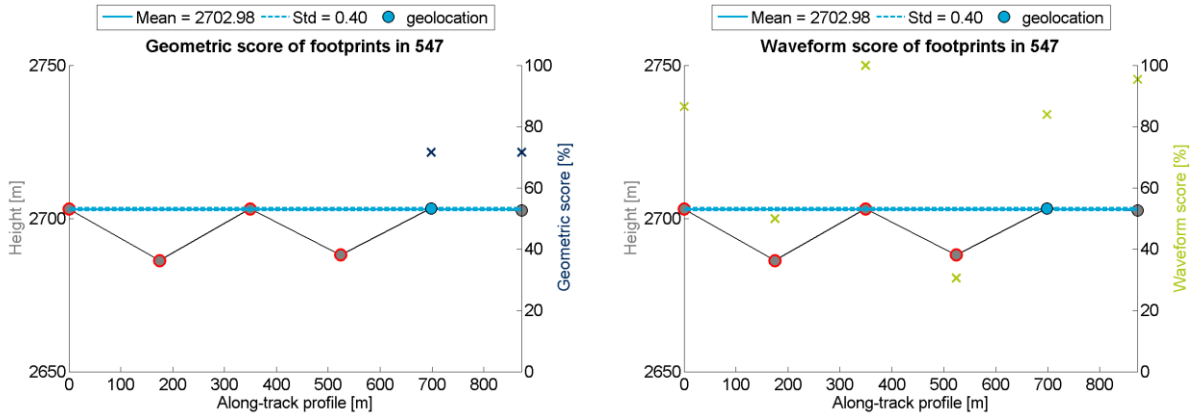


Figure 6.22: Geometric (left plot) and waveform (right plot) scores of the river profile with ID = 547. The blue crosses indicate the corresponding geometric scores while the green crosses indicate the waveform scores. The red circled dots indicate footprints that do not meet the threshold value for adjacent height difference.

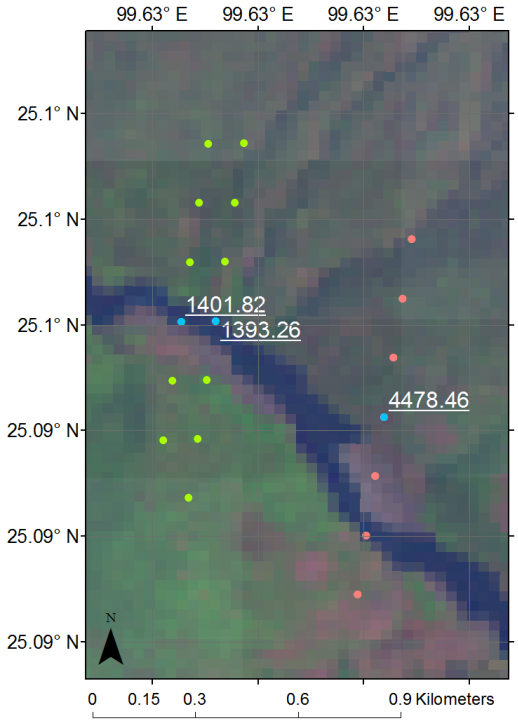


Figure 6.23: Outlying GLAS derived water level estimation of river cross-section with ID = 1693 (pink dots) and nearby river cross-sections (green dots) over a Landsat ETM image. The labels refer to the height levels [m].

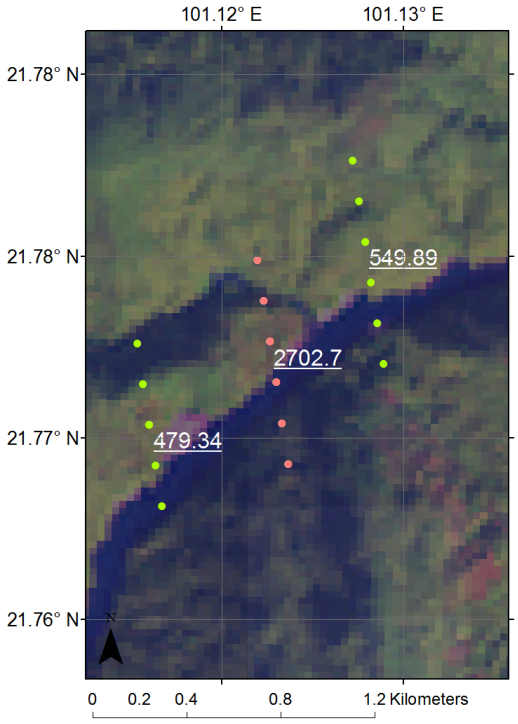


Figure 6.24: Outlying GLAS derived water level estimation of river cross-section with ID = 0547 (pink dots) and nearby river cross-sections (green dots) over a Landsat ETM image. The labels refer to the height levels [m].

6.3.2 IN SITU MEASUREMENTS

The Mekong River Commission provides water level data at 22 telemetry stations during the wet season (June – October) for the purpose of flood forecasting. Historical data (2008-2011) is available on the MRC flood forecasting website [28]. Daily measurements of the observed water levels are provided. The hydrological stations are located in the Lower Mekong Basin: data sharing with China and Myanmar is still under development and only two stations in China (Jing Hong and Manan) provide flood information but less frequent than the other measurements.

The geolocation of 18 hydrological stations is available by a status report of the hydro-meteorological network for river monitoring in the Mekong River [5]. These hydrological stations are listed in Table 6.6 with more detailed information on the geolocation and the zero gauge level. The zero gauge level is used as reference value for the obtained water levels. The geographical distribution of the hydrological stations is visualized in Figure 6.25. Note that more stations are located in the southern Lower Mekong Basin. The Mekong River Delta consists of a complex distributary network (see section 2.2.1) and therefore several monitoring locations are necessary. It should also be noted that the tidal system of the Southeast China Sea will influence the water levels in the Lower Mekong Delta. The stations with a zero gauge level about 0 m will not be suitable for the comparison of water level estimations with in situ measurements since the tidal effects are not considered.

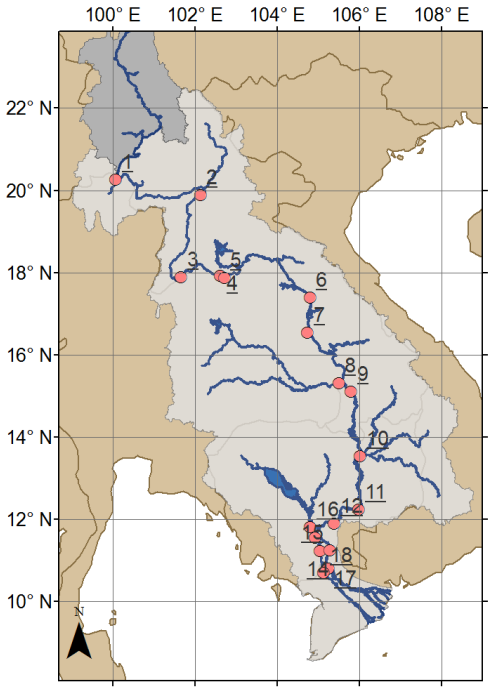


Figure 6.25: Location of the 18 hydrological stations in the Mekong River with expected geolocations.

Table 6.6: Expected geolocations of hydrological stations. The ID refers to the index of the station. The Lat [deg] and Lon [deg] are used as geolocation. The zero gauge level [m] is the low flow level at the location. The grey-shaded stations are influenced by the tidal system in the Southeast China Sea.

ID	Station	Country	Lat [deg]	Lon [m]	Zero gauge level [m]
1	Chiang Saen	Thailand	20.2734	100.0834	357.110
2	Luang Prabang	Laos	19.8917	102.1367	267.195
3	Chiang Khan	Thailand	17.8967	101.6684	194.118
4	Vientiane	Laos	17.9283	102.6200	158.040
5	Nongkhai	Thailand	17.8767	102.7200	153.648
6	Nakhon Phanom	Thailand	17.3984	104.8034	130.961
7	Mukdahan	Thailand	16.5400	104.7367	124.219
8	Khong Chiam	Thailand	15.3184	105.5000	89.030
9	Pakse	Laos	15.1167	105.8000	86.490
10	Stung Treng	Cambodia	13.5451	106.0166	36.790
11	Kratie	Cambodia	12.2398	105.9871	-1.080
12	Kompong Cham	Cambodia	11.9093	105.3877	-0.930
13	Phnom Penh Port	Cambodia	11.5700	104.9300	0.000
14	Koh Khel	Cambodia	11.2396	105.0399	-1.000
15	Neak Luong	Cambodia	11.2609	105.2843	-0.330
16	Prek Kdam	Cambodia	11.8130	104.8040	0.080
17	Tan Chau	Vietnam	10.8030	105.2430	0.001
18	Chau Doc	Vietnam	10.7070	105.1330	0.001

The hydrological stations within a distance of 5 km from the channel derived in the Lower Mekong Basin are selected (see section 6.3.1). The zero gauge levels of these hydrological stations are plotted over the downstream river profile in the Lower Mekong Basin in Figure 6.26. Note that the zero gauge levels fall below the river profile as derived from the USGS HydroSHEDS DEM. The 15 arc-second resolution DEM is based on average value in cells of approximately 500 m near the equator. The zero gauge levels correspond to the lowest ground surface at the location and will therefore show some deviance.

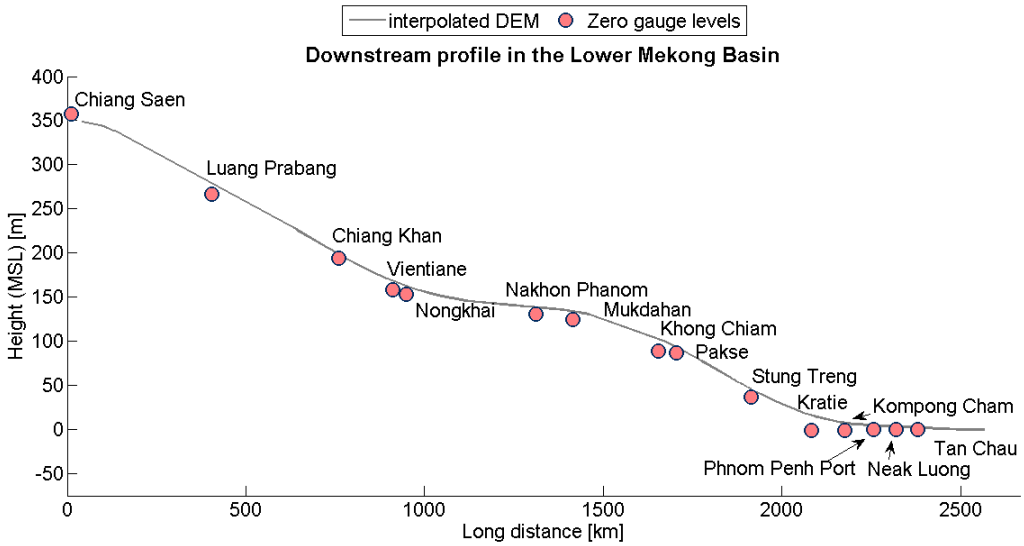


Figure 6.26: Downstream river profile created with a 9th order polynomial (grey line) and the hydrological stations within a distance of 5 km of the channel in the Lower Mekong Basin.

Comparing height levels

The MRC provides only information about the months June – October since the year 2008 on the flood forecasting website. From the 348 GLAS derived water level estimations along the channel in the Lower Mekong River, only the measurements corresponding with this period are selected. Only 15 GLAS derived water level estimations were available. This is caused by the limited number of operational periods in the wet season as well as the failure of the last laser on October 11, 2009 which resulted in a campaign of only 12 days (see section 4.2). As noted before, the water level estimations affected by the tidal system are ignored. This resulted in 11 GLAS derived water level estimations suitable for validation purposes. The locations of the stations and the water level estimations used for validation are visualized in Figure 6.27. In Table 6.8, detailed information about the water level estimations nearby the stations is presented. The nearest channel distance between a water level estimation and a hydrological station is 15 km (ID = 2565) while the largest distance is 169 km (ID = 2962). This makes it very hard to compare the GLAS derived water level estimations with the in situ measurements.

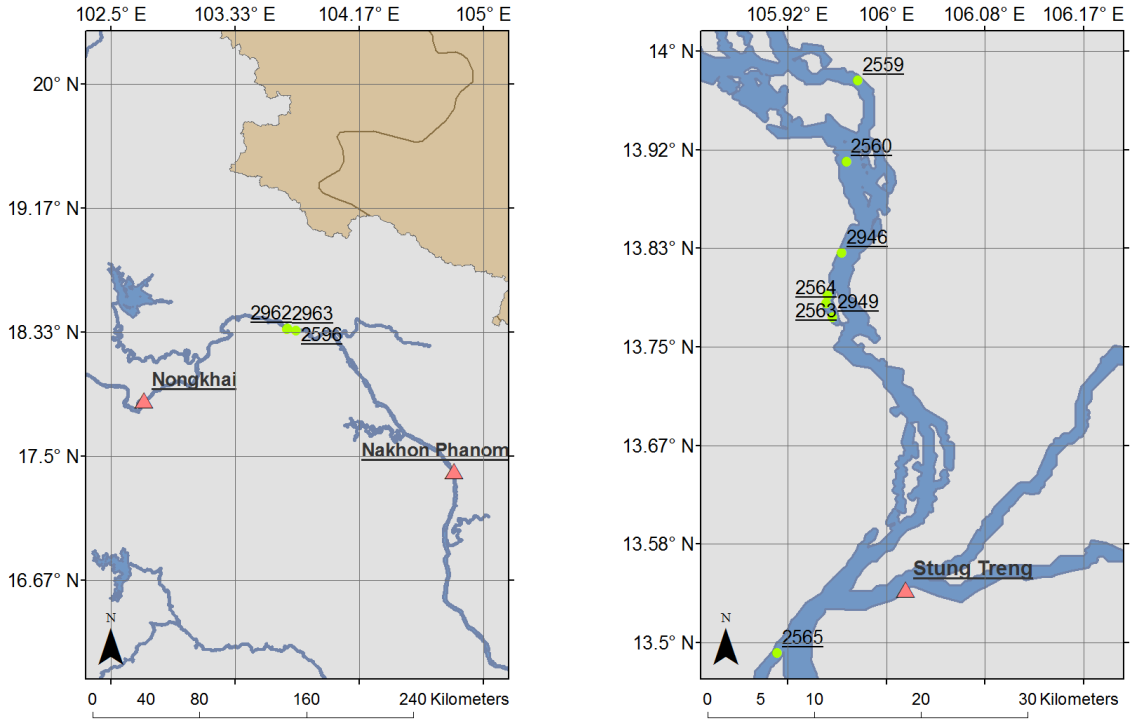


Figure 6.27: Location of the water level estimations suitable for validation (green dots) located and the nearest hydrological stations (pink triangles). The labels refer to the ID of the water level estimations.

Table 6.7: Summary of water level estimations for validation. N represents the number of river footprints. The Mean [m] and Std [m] refer to the estimated water level. The Length [m] refers to the perpendicular distance between the water level estimation and the hydrological station.

ID	Date	N	Mean [m]	Std [m]	Station	Distance [km]
2962	14/10/2009	2	151.67	0.17	Nongkhai	169
2963	07/10/2009	3	148.26	0.11	Nongkhai	169
2596	22/10/2008	4	148.24	0.17	Nongkhai	175
2559	10/10/2008	3	71.42	0.12	Stung Treng	53
2560	11/10/2008	16	54.88	0.1	Stung Treng	44
2946	06/10/2009	8	53.98	0.29	Stung Treng	34
2949	05/10/2009	3	51.71	0.11	Stung Treng	25
2564	11/10/2008	2	50.18	0.05	Stung Treng	29
2563	10/10/2008	2	50.12	0.09	Stung Treng	30
2950	12/10/2009	4	44.34	0.19	Stung Treng	26
2565	08/10/2008	5	43.31	0.16	Stung Treng	15

A comparison of the water level estimations and the in situ measurements is provided in Table 6.8. The in situ measurements at the hydrological stations are collected corresponding to the dates of the water level estimations. Note that the height difference (ΔH) is varying between 0.72 and 27.56 m. An attempt to reduce this deviance with the introduction of a height correction based on the USGS HydroSHEDS DEM values is made. The height correction is calculated by the difference between the nearest DEM values to the water level estimations and the in situ measurements. Using the height difference with the height correction ($\Delta H_{correction}$) should bridge the gap in channel distance. However, the DEM values have a coarse spatial resolution and might not reflect the height levels of the rivers but the nearby land surface. Therefore this correction introduces another bias.
A closer look to the GLAS derived water level estimation nearest to Stung Treng (D = 2565) shows that only a height difference of 0.72 m is apparent. For father water level estimations

located more upstream from Stung Treng, the difference runs up to 27.56 m. Stung Treng is located in a tributary (see Figure 6.27, right plot) which makes it difficult to compare with the upstream located water level estimations.

Table 6.8: Comparison of the water level estimations for validation and in situ measurements at corresponding dates. WLE [m] refers to the water level estimations. The height difference (ΔH) with and without the height correction are included.

ID	WLE [m]	In situ [m]	DEM $_{correction}$[m]	$\Delta H_{correction}$ [m]	ΔH [m]
2962	151.67	159.03	-2	5.36	7.36
2963	148.26	159.74	-12	-0.52	11.48
2596	148.24	159.66	-12	-0.58	11.42
2559	71.42	43.86	22	-5.56	-27.56
2560	54.88	43.62	-2	-13.26	-11.26
2946	53.98	47.12	-3	-9.86	-6.86
2949	51.71	47.35	11	6.64	-4.36
2564	50.18	43.62	6	-0.56	-6.56
2563	50.12	43.86	14	7.74	-6.26
2950	44.34	43.65	-10	-10.69	-0.69
2565	43.31	44.03	-10	-9.28	0.72

Comparing seasonal amplitudes

Another validation method is to evaluate the seasonal amplitude of the in situ measurements based on the minimum and maximum values during the wet season in the Lower Mekong Basin. The group with the most nearby water level estimations is selected to evaluate the seasonal amplitude (see section 6.1.1). In total, 19 GLAS derived water level estimations are located approximately 45 km from the hydrological station Strung Treng. In Figure 6.28, the nearby ground tracks are plotted. This explains the large number of water level estimations: two intersecting ground tracks cross the Mekong River at this location.

Information about the seasonal amplitudes at the hydrological station Stung Treng is obtained from the flood forecasting website of the MRC [28]. A graph with the gauge height during the wet season for several years is included in Figure 6.29. For all periods, the minimum gauge height is around 2.0 m and the maximum gauge height is around 11.0 m resulting in a seasonal amplitude in the wet season of about 9.0 m. Furthermore, based on the historical data reports, the seasonal amplitude in the wet season of 2008 and 2009 is calculated based on the minimum and maximum value of daily water level measurements between June and October. This is summarized in Table 6.10. The seasonal amplitudes in 2008 and 2009 are respectively 6.17 m and 6.40 m. Note that the graph show gauge height levels which need to be added to the zero gauge level of 36.79 m (see Table 6.6 and Figure 6.29).

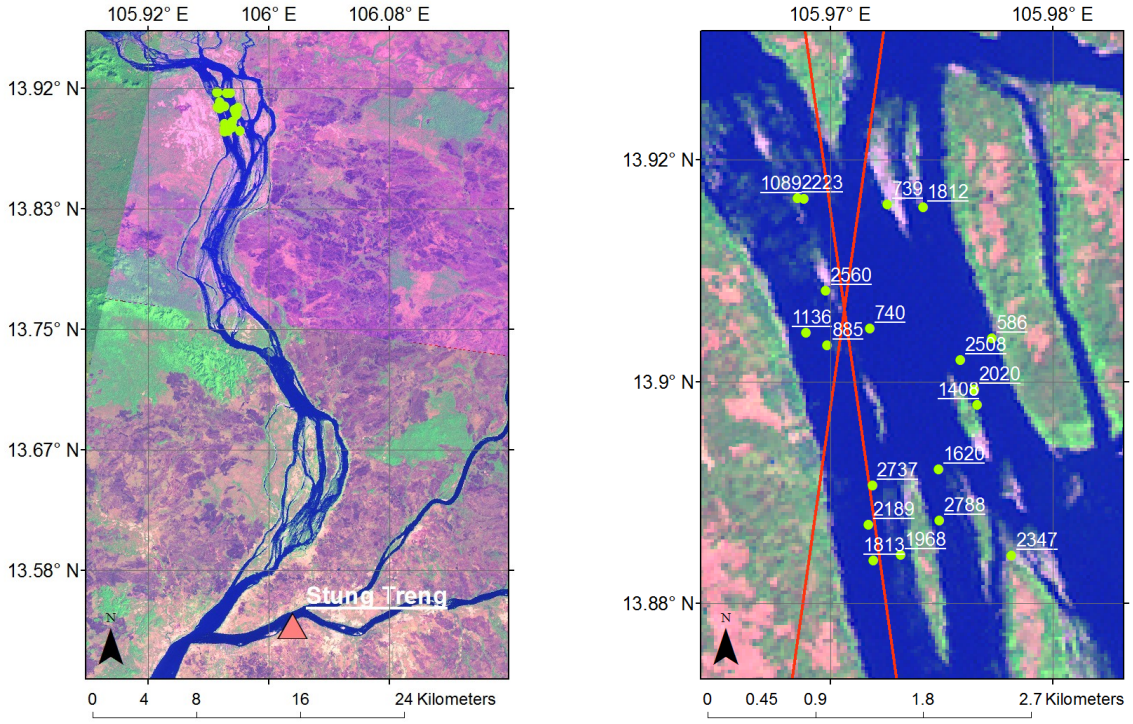


Figure 6.28: The left plot shows the group of 19 GLAS derived water level estimations (green dots) located near the hydrological station Stung Treng. The right plot is a close-up of the water level estimations (labeled with ID) and the corresponding ground tracks (red lines) over Landsat ETM images.

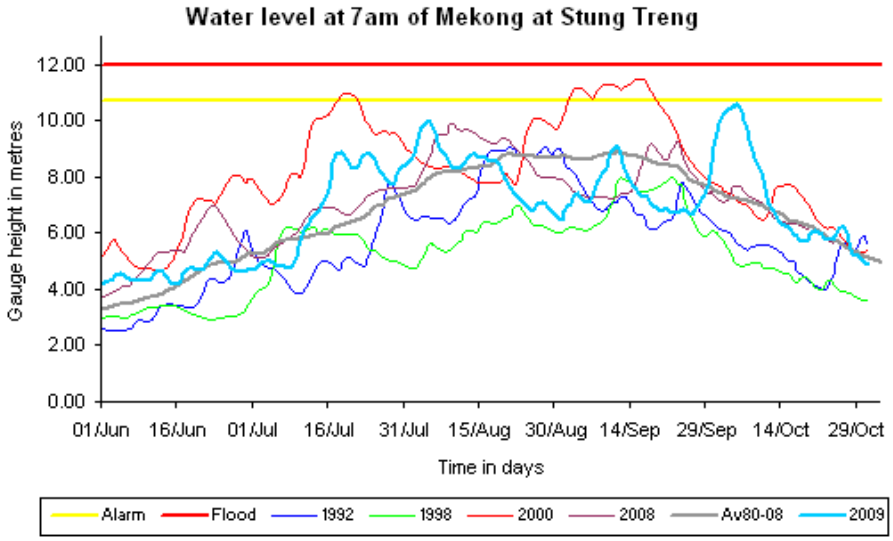


Figure 6.29: Gauge height at Stung Treng in the wet season, taken from the MRC [28]. The gauge height should be added to the zero gauge level of 36.79 m.

Table 6.9: Seasonal amplitudes (Amp) at Stung Treng obtained and enhanced from the MRC [28]. The Mean [m], Std [m], Min [m] and Max [m] refer to the GLAS derived water level estimations.

Year	Mean [m]	Std [m]	Min [m]	Max [m]	Amp [m]
2008	43.83	1.48	40.50	46.67	6.17
2009	43.67	1.72	40.95	47.35	6.40

A water level trend is derived from the 19 GLAS derived water level estimations and visualized in Figure 6.30. Note that only 17 water level estimations are actually visible because two measurements are hidden behind the bars of water level estimations with a similar observation time. The seasonal amplitude in the observed area is calculated by the difference between the minimum and maximum height level. The seasonal amplitude is approximately 9.0 m which corresponds well to the wet season amplitudes at Stung Treng.

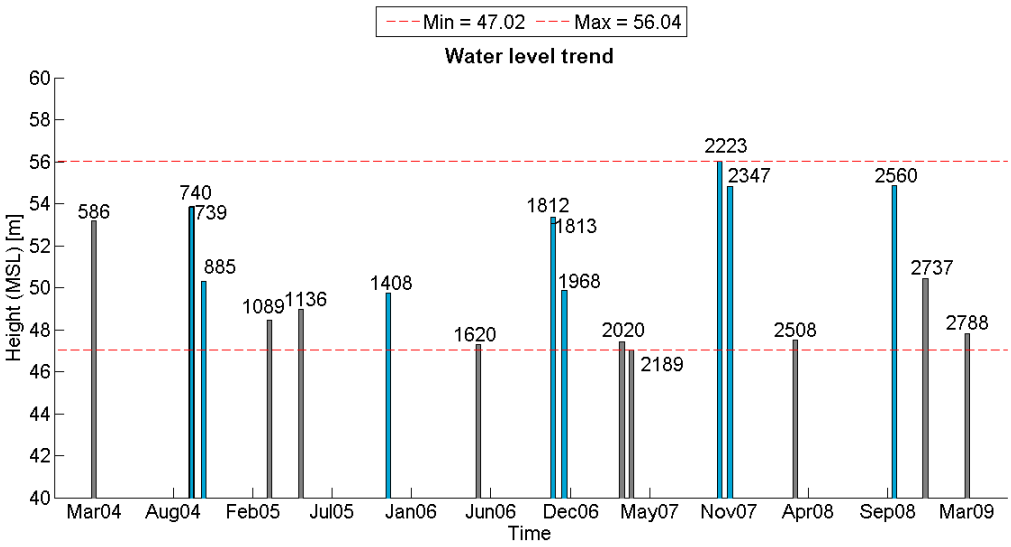


Figure 6.30: Seasonal amplitude for a group of 19 GLAS derived water level estimations near Stung Treng. Only 17 water level estimations are plotted due to overlapping observation times.

The lack of in situ validation data near GLAS derived water level estimations is a difficulty in this study. However it also highlights the contribution of the GLAS altimetry data to fill gaps in the monitoring network. It is recommended to further develop the classification criteria for river footprint identification in rivers in areas with more available validation data.

Other sources of validation data that can be considered in future research is satellite radar altimetry over the Mekong River as studied at the Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (LEGOS) [11]. With the TOPEX/Poseidon mission, 7 virtual stations observe water levels in the Mekong River. With the ENVISAT mission, six virtual stations are obtained.

The quality of the GLAS derived water level estimations is hard to evaluate. This is caused by the large distances from in situ measurements, the presence of multiple channels, and limited number of GLAS derived water level estimations in the wet season. However, the water level estimations do match the downstream river profiles and the magnitudes of the in situ measurements and seasonal amplitudes. The optimal water level trend in Figure 6.30 showed that the behavior of the water level estimations over time is plausible with a clear seasonal pattern without outliers. Nonetheless it is challenging to get consistent validation results.

7 CONCLUSIONS AND RECOMMENDATIONS

This study contributes to the assessment of GLAS altimetry data to observe water levels in rivers. The findings on the Mekong River Basin will be presented in this chapter. In section 7.1, the conclusions about river footprint identification and the water level estimations will be described. Recommendations for further research are elaborated in section 7.2.

7.1 CONCLUSIONS

The research question in this study was: How can GLAS altimetry data be optimally used for water level estimations in the Mekong River Basin? This study explored a method to derive river footprints and corresponding water level estimations. Each sub question as formulated in section 1.2 will be answered.

How to select footprints corresponding with water?

Classification method
The river footprints are identified in two steps: First, candidate river footprints are identified with use of a customized water mask of the Mekong River. Additional data sets from the USGS HydroSHEDS collection, MODIS 250m land water mask and the International Water Management Institute are used. Second, candidate river footprints are evaluated based on suitable geometric and waveform variables. All candidate river footprints with a geometric score above 50 percent and a waveform score above 75 percent are selected as river footprints in the Mekong River. This resulted in 785 water level estimations with an overall standard deviation of 0.087 m.

Minimum river width
The narrow width in the Upper Mekong Basin is leading to fuzzy water level estimations with GLAS altimetry data. Only 6.6 percent of the water level estimations were obtained in the Upper Mekong Basin. Furthermore, higher standard deviations of the heights were obtained in the Upper Mekong Basin (0.41 m).
The suggested minimum requirements of the river width for the use of GLAS altimetry data is 315 m. This is based on the following results: First, river cross-sections with 1 or 2 footprints are removed in the processing of candidate river footprints. With an along-track spacing of 175 m and a footprint size of 70 m, the minimum river width should be 315 m.

Additional data
A water mask is used to select river footprints. For small water bodies such as narrow rivers, the contribution of the USGS HydroSHEDS river network is essential. In the Upper Mekong Basin, only 12 percent of the GLAS derived water level estimations are intersecting the MODIS 250m land water mask. This is in large contrast with the Lower Mekong Basin in which 90 percent of the water level estimations are intersecting the MODIS 250 m land water mask. This implies that most water level estimations in narrow rivers are detected by the intersection of the GLAS footprints with the USGS HydroSHEDS river network. The drawback of using the USGS HydroSHEDS river network is the assumed river width of 500 m. This will result in inaccurate locations of the river and therefore false river footprints and water level estimations.
The MODIS 250m land water mask was buffered by 500 m to compensate the geolocation error that occurred after conversion from raster to feature. This resulted in loss of information about islands and mid-channel banks. Furthermore, the time of observation of the Landsat ETM images do not match the operational period of ICESat (2003-2009). The Landsat images were obtained between 1999-2003. Furthermore, the extent of the river channel is related to the monsoon season which could result in inconsistencies in the Landsat-derived river footprints.

How can nearby water level estimations be used to assess water level trends?

Wet season measurements
GLAS derived water level estimations are easier to obtain during the wet season than during the dry season. Only 27 percent of the candidate river footprints that were obtained with use of the customized water mask correspond to the wet season. After the river footprint identification 31 percent corresponds to wet season measurements. This is ascribed to the better waveform characteristics during the wet season and wider rivers which results in more river footprints.

Availability of nearby water level estimations
The derivation of water level trends based on GLAS altimetry data is difficult. A water level trend can be derived if the repeated ground track of the GLAS altimetry data resulted in nearby water level estimations. In the Upper Mekong Basin, 29 percent of the water level estimations do not have any nearby water level estimations in comparison to 6 percent in the Lower Mekong Basin. Therefore it is almost impossible to derive water level trends in the Upper Mekong Basin.

Seasonal amplitudes
With two case study areas, water level trends are evaluated by selecting all nearby footprints within a distance of 2 km of water level estimations. The seasonal amplitude shows inconsistencies since missing values result in lack of information, especially during the wet season. Furthermore, only one water level estimations per season cannot contribute to a detailed river monitoring system as achieved with in situ measurements. Hydrological stations are more scarcely distributed but provide daily (or even hourly) measurements. On the other hand, the case study in the Nam Phong Tributary seems to indicate a flood control system based on reversed seasonal flows that supply the Ubolratana Dam Reservoir. This can provide basic information about the flow control if access to water level data is not possible due to coordination difficulties with institutions and countries.

What are the strengths and weaknesses of GLAS derived water levels estimations?

Geographical coverage
The Mekong River has diverse characteristics which makes it difficult to achieve a uniform method for river footprint identification and water level estimations based on GLAS altimetry data. First of all, it is hard to derive water level estimations in narrow rivers. This is more obvious for narrow rivers with steep mountains alongside of the channel. A minimum width of 315 m is required. Satellite laser altimetry is still preferred above radar altimetry due to the smaller footprints size and the potential to observe small rivers. Satellite radar altimetry such as Topex/Poseidon uses a theoretical minimum river width of 580 to 1160 m but resulted in precisions of 1.1 m (rms) [3]. ICESat/GLAS altimetry can bridge the gap and provide water level estimations from relatively narrow rivers.

Temporal coverage
A weakness of the GLAS altimetry data is that the operational periods of ICESat do not match the monsoon climate in the Mekong River Basin. A GLAS-based contribution to the river monitoring system of the Mekong River should provide sufficient water level estimations during the wet season (June – November) while most measurements were actually obtained during the campaigns between February and March.

Waveform variables
The GLAS altimetry data (GLA14 data product) does include waveform variables which supports the classification of river footprints and is promising for the future ICESat-2 mission. However, inconsistencies in the waveform variables caused by saturation, laser performance with respect to transmitted energy, and weather conditions such as clouds, should be considered when using GLAS altimetry data for observing water levels in river networks.

7.2 RECOMMENDATIONS

The recommendations provide suggestions of further research related to GLAS altimetry data and the monitoring of river networks as well as considerations that should be noticed about the results of this study.

Water mask
The most important recommendation is that the use of a higher resolution land water mask for the use of GLAS altimetry data in river networks. In this study, a suitable water mask of the Mekong River is customized but also resulted in anomalies; format conversions result in geolocation errors, the MODIS 250m land water mask does not detect small rivers, and the USGS HydroSHEDS river network does not provide information about the width of rivers and larger water bodies such as lakes and reservoirs. A higher resolution (about 15 m) land water mask would be recommendable for future studies with respect to river networks.

Candidate river footprints
The candidate river footprints are derived from GLAS footprints that intersect the customized water mask. Not all identified candidate river footprints cross the centerline of the customized water mask which implies that false candidate river footprints are included. The false candidate river footprints can be avoided by updating the customized water mask with a left and right boundary for improved selection of candidate river footprints. Instead of using spatial selection based on intersection, candidate river footprints can be obtained by those footprints that are located between the footprints that touch the left and right boundary of the feature.

Vertical buffer
An enhancement of the river footprint identification could be the introduction of a vertical buffer based on a DEM. As an example, the USGS HydroSHEDS DEM values that overlap the customized water mask of the Mekong River could be extracted. The GLAS derived water level estimations that overshoot the height derived from the DEM within a certain threshold should be flagged as incorrect. This is not analyzed in this study but might be a useful contribution that should be considered in further research.

Waveform errors
The influence of saturation effects on the waveform parameters is completely ignored in this study. This study assumed that the ground surface corresponds with the last return pulse in the signal. The GLA14 data is normally referenced to the centroid of the signal. For mixed footprints, the signal is biased and the last return pulse will result in more adequate height levels. However, for footprints that do only illuminate water surfaces, the last return pulse might be related with a ringing peak that does not correspond with the ground surface. This could have influenced the results of the maximum amplitude-ratio and the definition of the height above mean sea level in which the ground surface is aligned with the last return pulse.

Classification thresholds
Furthermore, the thresholds used in this study are somewhat arbitrary and need to be evaluated when used in other study areas. The customized water mask is based on the USGS HydroSHEDS river network in which only cells with a flow accumulation above 50,000 upstream cells are selected. This resulted in an adequate representation of the Mekong River but for other river networks this could be incorrect. The USGS HydroSHEDS river network is buffered with 500 m which might not be enough for wider rivers.
The classification criteria are largely based on the height difference between adjacent footprints in river cross-sections. Two height thresholds are used to exclude footprints. In the Upper Mekong Basin footprints are excluded if the height difference with adjacent footprints exceeds 2 m. In the Lower Mekong Basin where the river decline is less evident, a height threshold value of 0.5 m is used. Again, for rivers with different characteristics this threshold needs to be chosen accordingly.

Classification criteria
Another important note is the use of optional scores in the weighting of classification criteria. Two scores can only be obtained under certain circumstances: First, the geometric score can include a score of the normalized number of footprints per group. This can only be calculated if more than 1 group of consecutive inliers is apparent. Second, the maximum amplitude-ratios can only be obtained when the return signal consists of more than 1 peak. These two optional scores do contribute to the classification of river footprints but are not always applicable. Therefore some footprints are stricter evaluated than others. This is a deficiency in the constructed methodology.

Multiple channels
It is recommended to analyze methods to provide information about multiple channels based on the GLAS derived river footprints. Each river cross-section is now only returning single water level estimations while the river cross-section might contain multiple channels. The water level estimations were simply defined by the mean height values of the river footprints and the geolocation was defined by the latitude and longitude of the median river footprint. Furthermore, the width of the channel is unknown to some extent. This information is possible to derive with the GLAS derived river footprints but this has not yet been explored in this study.

Alignment of water level estimations
The alignment of GLAS derived water level estimations in the Mekong River is relevant for the interpretation of water level trends. For steep rivers the distance between nearby water level estimations can result a deviance in the height levels. The water level trend should only consider the temporal variation of nearby water level estimations but a spatial component might disturb the results. This implies that more downstream located nearby water level estimations in combination with a steep river decline will have lower height levels independent from the temporal dimension. For the calculation of water level trends it is recommended to take notice of the influence of the spatial distribution of the nearby water level estimations in steep rivers.

ICESat operational periods
The operational periods of the ICESat mission were rescheduled after laser failure. However, the few repeated ground track measurements do not provide at least a measurement during the dry and wet season per year. This makes it very difficult to derive water level trends without temporal missing values. It would be a great improvement for the monitoring of land processes if the ICESat-2 mission would support better temporal intervals of the campaigns.

Other river
The mismatch between the ICESat operational periods and the monsoon climate in the Mekong River Basin resulted in less water level estimations during the wet season. The South Asia monsoon that better matches the ICESat operational periods is the Northeast monsoon with a wet season from December to March. The Himalayan River system consists of three major rivers: Indus, Ganges, and the Brahmaputra which can be suggested as study area for further research.

Validation data
The GLAS derived water level estimations were difficult to validate with in situ measurements. Although historical data is available since 2008 during the wet season, the number of stations in the Lower Mekong Basin is limited and does not match the locations of the ICESat ground tracks. In the Upper Mekong Basin, no external validation data is available. This shows that the GLAS derived water level estimations contribute to the monitoring of the river network. However, for the purpose of evaluating a classification method, it would be preferred to use study area with more external validation data available.

It is also recommended to use other validation sources such as provided by the Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (LEGOS). These data sources are derived from radar altimetry mission such as TOPEX/Poseidon and ENVISAT.

New release of GLAS altimetry data
The last recommendation is a more practical side note for other studies considering the GLAS altimetry data. At the time of this study, the NSIDC started with the GLAS Data Release-33. For consistency reasons, only data from Release-31 was used. GLAS Release-33 will provide enhancements that can contribute to better results. The most important enhancement in the GLAS altimetry data (GLA14) is that the elevation is referenced to the WGS84 datum and a WGS-84 ellipsoidal surface height is added which makes the ellipsoidal conversions unnecessary.

REFERENCES

[1] S. Bae and B.E. Schutz. Precise altitude determination (PAD). *Geoscience Laser Altimeter System (GLAS). Algorithm Theoretical Basis Document, Version: 2.2*, Center for Space Research, The University of Texas, 2002.

[2] N. Baghdadi, N. Lemarquand, and and J.S. Bailly H. Abdallah. The relevance of GLAS/ICESat elevation data for the monitoring of river networks. *Remote Sensing*, 3: 708–720, 2011. ISSN 2072-4292.

[3] C.M. Birkett, L.A.K. Mertes, T. Dunne, M.H. Costa, and M.J. Jasinski. Surface water dynamics in the amazon basin: Application of satellite radar altimetry. *Journal of Geophysical Research*, 107, 2002.

[4] "D.M. Bjerklie", S.J. Dingman, C.J. Vorosmarty, C.H. Bolster, and R.G. Congalton. Evaluating the potential for measuring river discharge from space. *Journal of Hydrology*, 278: 17–38, 2003.

[5] S. Boonpraesart. Status of MRC hydro-meteorological network for river monitoring. *Mekong River Commission Secretariat (MRCS), Vientiane, Lao PDR*.

[6] A.C. Brenner, H.J. Zwally, C.R. Bentley, B.M.Csatho, D.J. Harding, M.A. Hofton, J. Minster, L. Roberts, J.L. Saba, R.H. Thomas, and D. Yi. Derivation of range and range distributions from laser pulse waveform analysis for surface elevations, roughness, slope, and vegetation heights. *Geoscience Laser Altimeter System (GLAS). Algorithm Theoretical Basis Document, Version: 4.1*, 2003.

[7] M. Carroll, C. DiMiceli, J. Townshend, P. Noojipady, and R. Sohlberg. UMD global 250 meter land water mask user guide. 2009. URL http://landcover.org/data/-guide/technical/techguide_watermask.pdf.

[8] J.S. da Silva, S. Calmant, F. Seyler, O.C.R. Filho, G. Cochonneau, and W.J. Mansur. Water levels in the amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions. *Remote Sensing of Environment*, 114: 2160–2181, 2010.

[9] A. de Vocht. *Basishandboek SPSS 12. Statistiek met SPSS 12*. Bijleveld Press, 2004.

[10] F.M. Dekking, C. Kraaikamp, H.P. Lopuhaa, and L.E. Meester. *A Modern Introduction to Probability and Statistics: Understanding Why and How*. Springer, 2005.

[11] Laboratoire d'Etudes en Géophysique et Océanographie Spatiales. HydroWEB. Lakes, Rivers and wetlands. Water levels from satellite altimetry. Last visited: February 2, 2012. URL http://www.legos.obs-mip.fr/soa/hydrologie/hydroweb/-Cartes/Asie_Sud_Est.html.

[12] V.H. Duong. *Processing and Application of ICESat Large Footprint Full Waveform Laser Range Data*. PhD thesis, Delft University of Technology, 2010.

[13] B.M. Fekete and C.J. Vorosmarty. The current status of global river discharge monitoring and potential new technologies complementing traditional discharge measurements. *International Association of Hydrological Sciences (IAHS). Predictions in Ungauged Basins: PUB Kick-off*, 309, 2007.

[14] W. Fraczek. Mean sea level, GPS, and the geoid. *Esri ArcUser Online, July - September 2003*, Last visited: November 24, 2011. URL http://www.esri.com/-news/arcuser/0703/geoid1of3.html.

[15] GeoLas Consulting. Imaging laser altimetry. Last visited: September 19, 2011. URL http://www.geolas.com/Pages/laser.html.

[16] Global Land Cover Facility (GLCF). MODIS water mask. Last visited: October 11, 2011. URL http://www.glcf.umd.edu/data/watermask/.

[17] P. Gong, and H. Huang Z. Li, G. Sun, and L. Wang. ICESat glas data for urban environment monitoring. *IEEE Transactions on Geoscience and Remote Sensing*, 49: 1158–1172, 2011.

[18] A.C. Hall, G.J. Schumann, J.L. Bamber, and P.D. Bates. Tracking water level changes of the amazon basin with space-borne remote sensing and integration with large scale hydrodynamic modelling: A review. *Physics and Chemistry of the Earth*, 36: 223–231, 2011.

[19] R. Schutz H.J. Zwally, and, and J. Bufton C. Bentley, T. Herring, J. Minster, J. Spinhirne, and R. Thomas. Glas/icesat l2 global land surface altimetry data v031. *Boulder, CO: National Snow and Ice Data Center. Digital media.*, 15 October to 18 November 2003, 2003.

[20] International Water Management Institute (IWMI). Waterdata. Integrated portal for IWMI & CPWF research data. Downloaded on: October 2, 2011. URL http://-waterdata.iwmi.org/DataArchive.php.

[21] W.B. Langbein and K.T. Iseri. General introduction and hydrologic definitions. *Manual of Hydrology, Part 1. General Surface-Water Techniques, Geological Survey Water Supply Paper*, 1541-A: 29, 1960.

[22] B. Lehner, K. Verdin, and A. Jarvis. HydroSHEDS technical documentation version: 1.0. page 27, 2006. URL http://hydrosheds.cr.usgs.gov/datasource.php.

[23] S. Li and D. He. Water level response to hydropower development in the upper mekong river. *A Journal of the Human Environment*, 37: 170–176, 2008.

[24] X.X. Lu and R.Y. Siew. Water discharge and sediment flux changes in the lower mekong river. *Hydrology and Earth System Sciences Discussions*, 2: 2287–2325, 2005.

[25] C. Manusthiparom, C. Apirumanekul, and M. Mahaxy. Flood forecasting and river monitoring in the mekong river basin. *Second Southeast Asia Water Forum, August 29th - September 3rd 2005, Bali Indonesia*, Mekong River Commission Secretariat (MRCS), Vientiane, Lao PDR, 2005.

[26] Mekong River Commission. Overview of the hydrology of the mekong basin. page 73, 2005.

[27] Mekong River Commission. State of the basin report 2010. *Mekong River Commission, Vientiane*, page 123, 2010. URL http://www.mrcmekong.org/assets/-Publications/basin-reports/MRC-SOB-report-2010full-report.pdf.

[28] Mekong River Commission (MRC). River monitoring. Last visited: December 26, 2011. URL http://ffw.mrcmekong.org/.

[29] Mekong River Commission (MRC). Upstream partners. Last visited: November 11, 2011. URL http://www.mrcmekong.org/about-the-mrc/upstream-partners-2/.

[30] Mekong River Commission (MRC). The mekong river basin. climate. Last visited: November 7, 2011. URL http://www.mrcmekong.org/the-mekong-basin/climate/.

[31] Mekong River Commission (MRC). About the MRC. Last visited: October 11, 2011. URL http://www.mrcmekong.org/about-the-mrc/. (Last visited July 2011).

[32] R. Molijn. ICESat full waveform signal analysis for the classification of land cover types over the cryosphere. Master's thesis, Delft University of Technology, Faculty of Aerospace Engineering, Department of Earth Observation and Space Systems, 2009.

[33] National Aeronautics and Space Administration (NASA). About NASA's Earth Observing System. Last visited: November 7, 2011. URL http://-eospso.gsfc.nasa.gov/.

[34] National Aeronautics and Space Administration (NASA). ICESat-2. Last visited: November 7, 2011. URL http://icesat.gsfc.nasa.gov/icesat2/.

[35] National Aeronautics and Space Administration (NASA). Landsat then and now. Last visited: October 11, 2011. URL http://landsat.gsfc.nasa.gov/about/index.html.

[36] National Snow and Ice Data Center (NSIDC). ICESat/GLAS Data Subsetter. URL http://nsidc.org/forms/glas_subset_form.html.

[37] National Snow and Ice Data Center (NSIDC). GLAS Altimetry Product Usage Guidance. 2011. URL http://nsidc.org/data/docs/daac/-glas_icesat_l1_l2_global_altimetry.gd.html. (Last visited July 2011).

[38] National Snow and Ice Data Center (NSIDC). Gla14 records: Release 31. Last visited: December 12, 2011. URL http://nsidc.org/data/docs/daac/glas_altimetry/-gla14_records_v31.html.

[39] National Snow and Ice Data Center (NSIDC). Description of data releases. Last visited: November 7, 2011. URL http://nsidc.org/data/icesat/data_releases.html.

[40] National Snow and Ice Data Center (NSIDC). Frequently asked questions. Last visited: November 7, 2011. URL http://nsidc.org/data/icesat/faq.html.

[41] National Snow and Ice Data Center (NSIDC). GLAS/ICESat L1 and L2 Global Altimetry Data. Last visited: November 7, 2011. URL http://nsidc.org//data/docs/-daac/glas_icesat_l1_l2_global_altimetry.gd.html.

[42] National Snow and Ice Data Center (NSIDC). Laser operational periods. Last visited: November 7, 2011. URL http://nsidc.org/data/icesat/laser_op_periods.html.

[43] Earth Resources Observation and Science Center. Usgs global visualization viewer. Last visited: December 12, 2011. URL http://glovis.usgs.gov/.

[44] Ontario. Ministry of Natural Resources. Water levels and flows. Last visited: October 22, 2011. URL http://www.mnr.gov.on.ca/en/Business/Water/2ColumnSubPage/-STEL02_163598.html.

[45] D. Potere. Horizontal positional accuracy of google earth's high-resolution imagery archive. *Sensors*, 8: 7973–7981, 2008.

[46] H.J. Rim and B.E. Schutz. Precision orbit determination (POD). *Geoscience Laser Altimeter System (GLAS). Algorithm Theoretical Basis Document, Version: 2.2*, Center for Space Research, The University of Texas, 2002.

[47] B.E. Schutz, H.J. Zwally, J. Bufton, C. Bentley, T. Herring, J. Minster, J. Spinhirne, and R. Thomas. Geoscience laser altimeter system GLAS science requirements. *GLAS Mission Requirements, Version: 2.01*, 1997.

[48] United Nations Development Programme (UNDP). Human development report 2010. the real wealth of nations: Pathways to human development. page 238, 2010. URL http://hdr.undp.org/en/media/HDR_2010_EN_Complete_reprint.pdf.

[49] United States Geological Survey (USGS). Hydrosheds. Last visited: October 11, 2011. URL http://hydrosheds.cr.usgs.gov/. (Last visited July 2011).

[50] Author Unknown. Band combinations. Last visited: January 7, 2012. URL http://-web.pdx.edu/~emch/ip1/bandcombinations.html.

[51] T.J. Urban, B.E. Schutz, and A.L. Neuenschwander. A survey of ICESat coastal altimetry applications: Continental coast, open ocean island, and inland river. *Terrestrial Atmospheric And Oceanic Sciences*, 19: 1–19, 2008.

[52] V. Merwade. Watershed and stream network delineation. Last visited: November 13, 2011. URL https://web.ics.purdue.edu/~vmerwade/education/-terrain_processing.pdf.

[53] E. Walling. The changing sediment load of the mekong river. *A Journal of the Human Environment*, 37: 150–157, 2008.

[54] X. Wang, X. Cheng, Z. Li, and X. Li H. Huang, and Z. Niu, and P. Gong. Lake water footprint identification from time-series ICESat/GLAS Data. *IEEE Geoscience and Remote Sensing Letters*, 99: 1–5, 2011.

[55] A. Wehr and U. Lohr. Airborne laser scanning - an introduction and overview. *ISPRS Journal of Photogrammetry & Remote Sensing*, 54: 68–82, 1999.

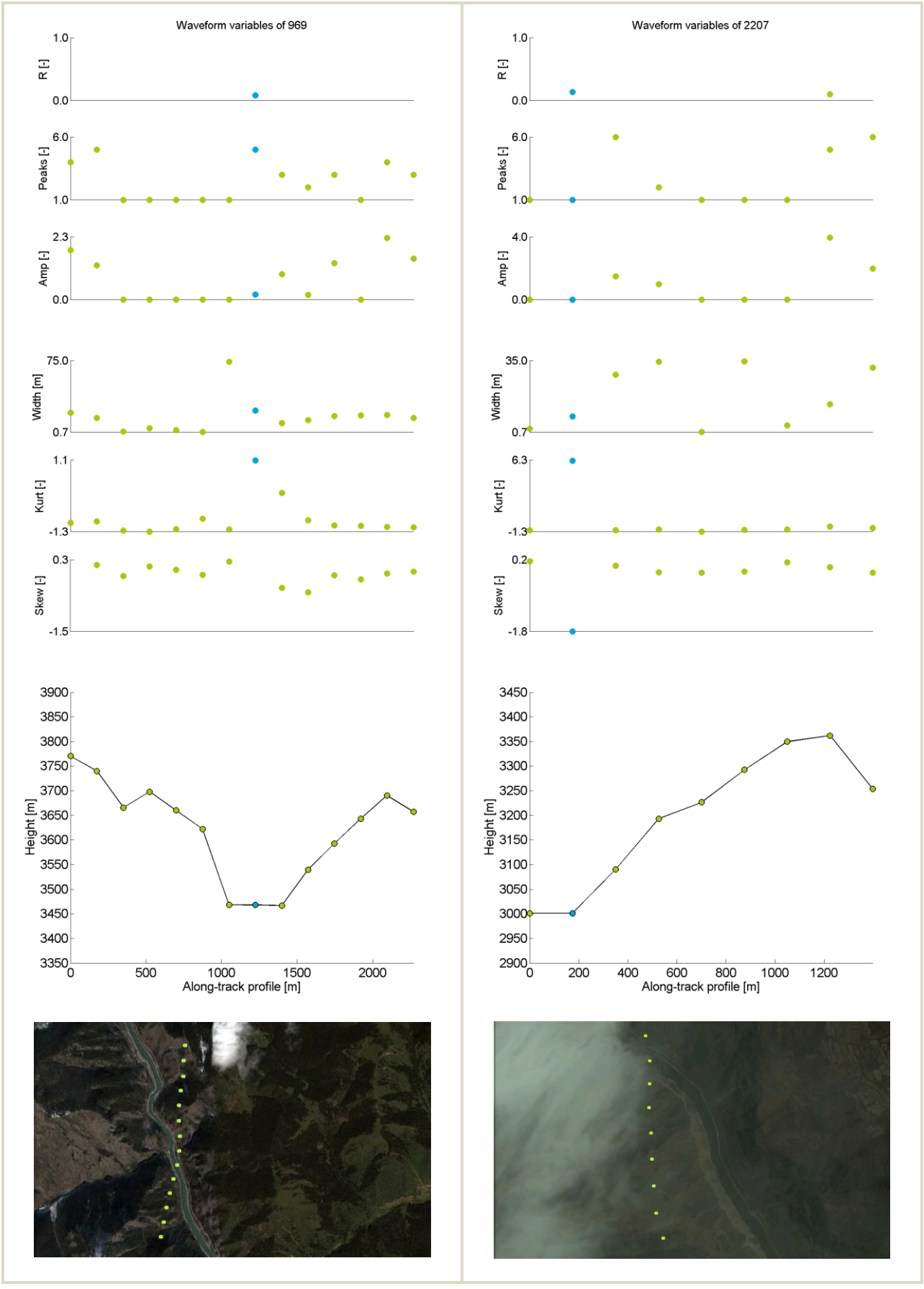
[56] E.W. Weisstein. Hypothesis testing. *From MathWorld–A Wolfram Web Resource*, Last visited: January 13, 2011. URL http://mathworld.wolfram.com/-HypothesisTesting.html.

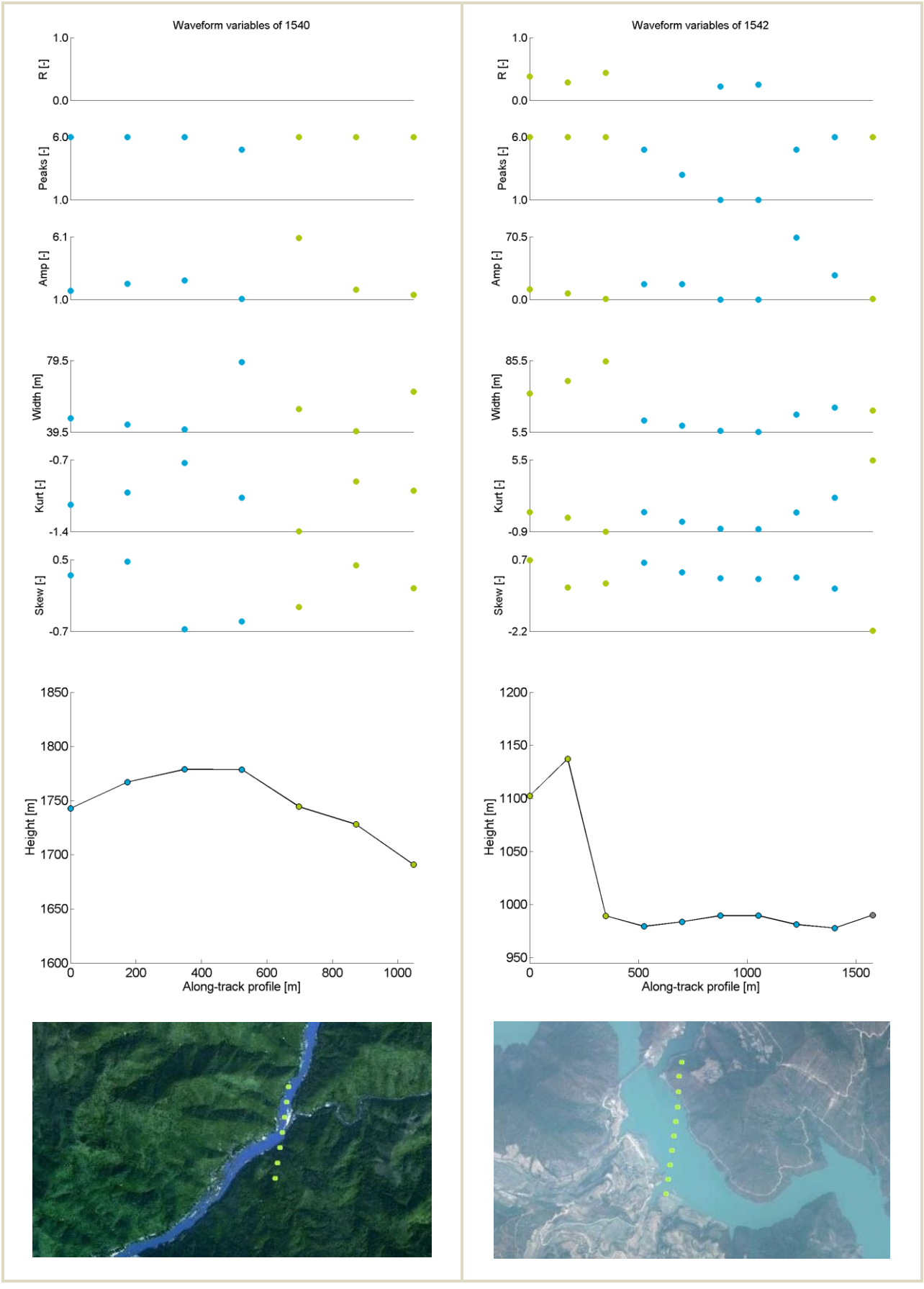
[57] Wikipedia. Ubol ratana dam. Last visited: January 6, 2012. URL http://-en.wikipedia.org/wiki/Ubol_Ratana_Dam.

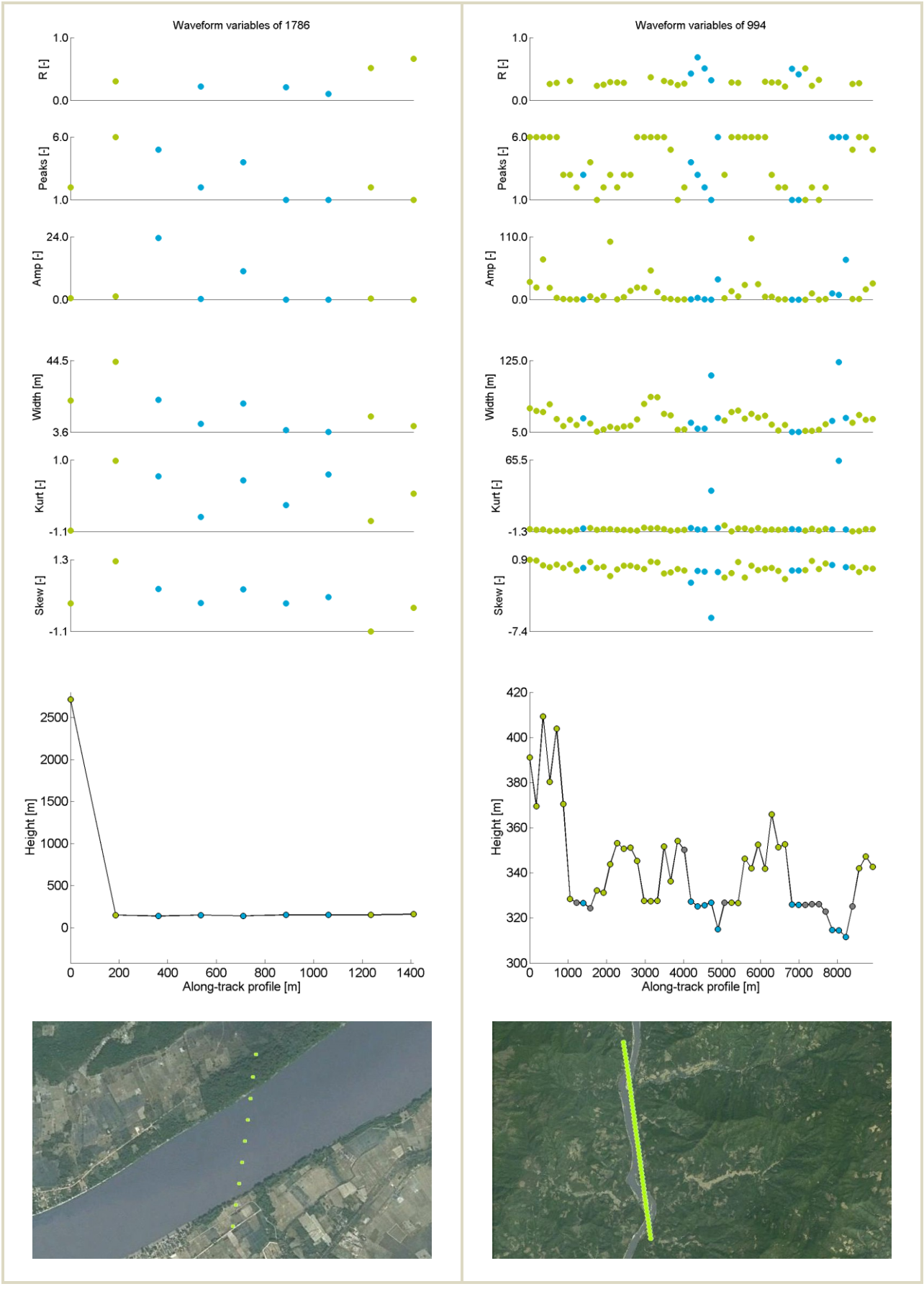
[58] Wikipedia. Histogram. http://en.wikipedia.org/wiki/Histogram, Last visited: January 7, 2012.

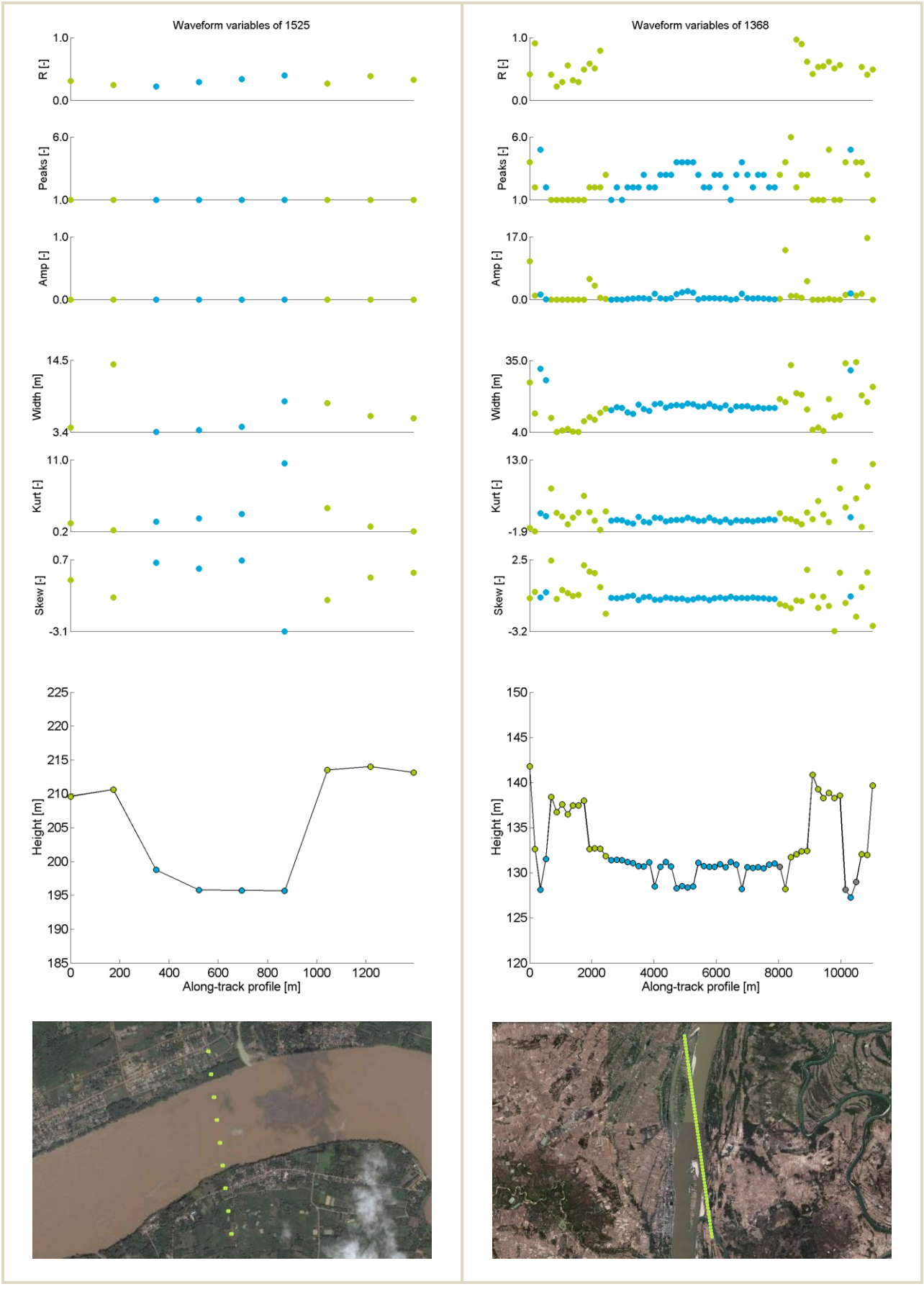
APPENDIX

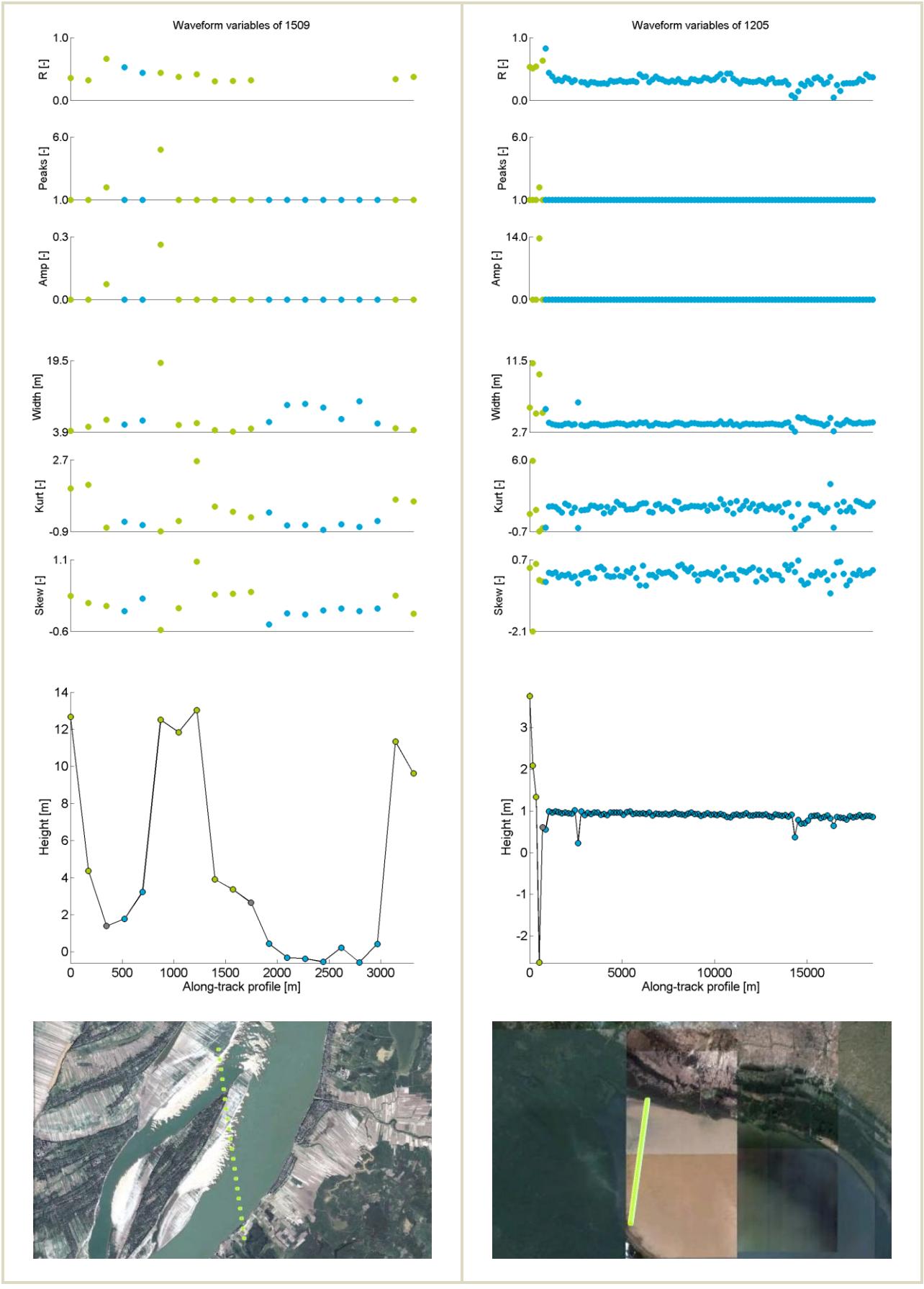
APPENDIX 1: WAVEFORM VARIABLES OF SAMPLING SET OF RIVER PROFILES

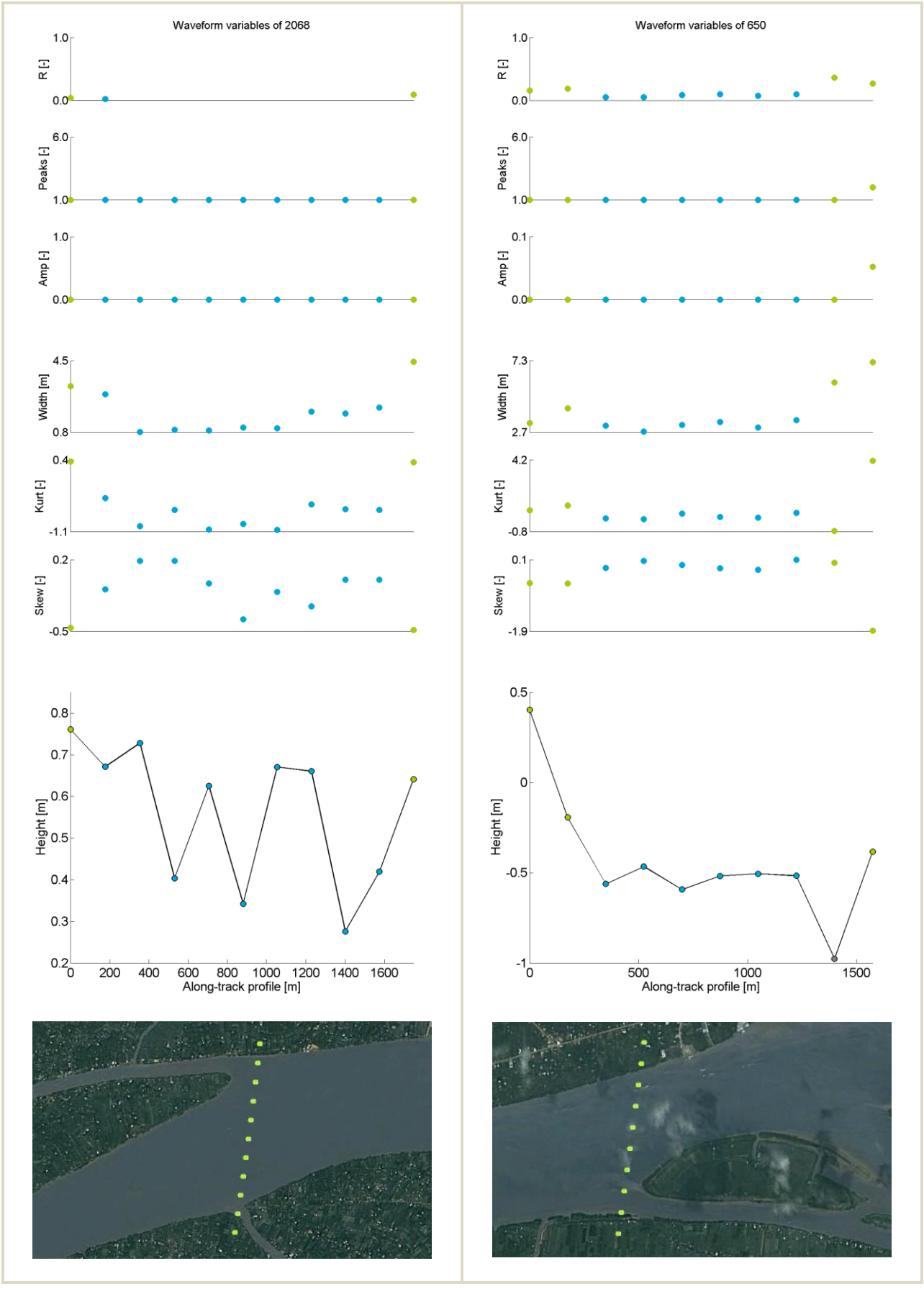

Waveform variables of 969
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]
Waveform variables of 2207
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]


Waveform variables of 1540
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]
Waveform variables of 1542
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]

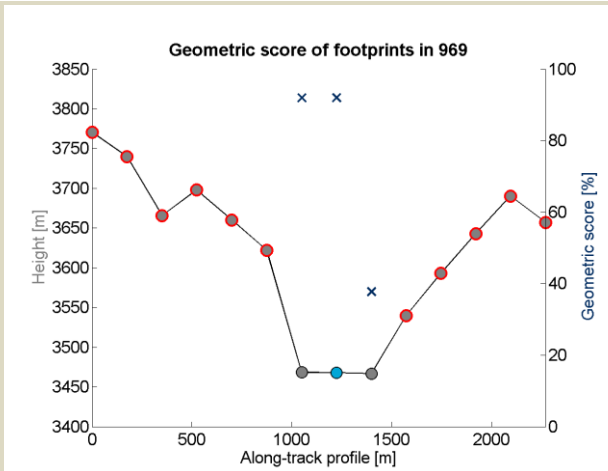

Waveform variables of 1786
R [-]
1.0
0.0
Peaks [-]
6.0
1.0
Amp [-]
24.0
0.0
Width [m]
44.5
3.6
Kurt [-]
1.0
-1.1
Skew [-]
1.3
-1.1
Height [m]
2500
2000
1500
1000
500
0
0
200
400
600
800
1000
1200
1400
Along-track profile [m]
Waveform variables of 994
R [-]
1.0
0.0
Peaks [-]
6.0
1.0
Amp [-]
110.0
0.0
Width [m]
125.0
5.0
Kurt [-]
65.5
-1.3
Skew [-]
0.9
-7.4
Height [m]
420
400
380
360
340
320
300
0
1000
2000
3000
4000
5000
6000
7000
8000
Along-track profile [m]

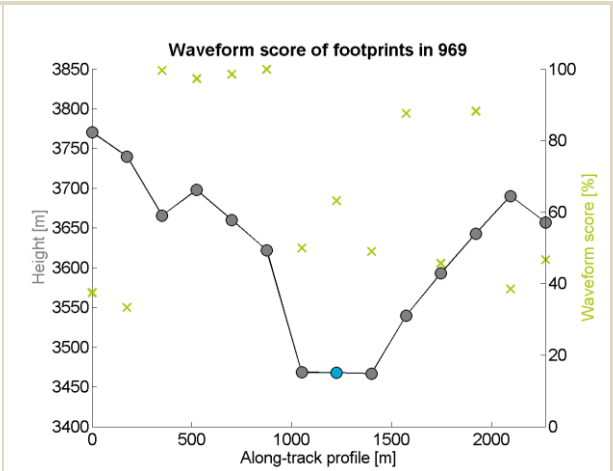

Waveform variables of 1525
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]
Waveform variables of 1368
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]

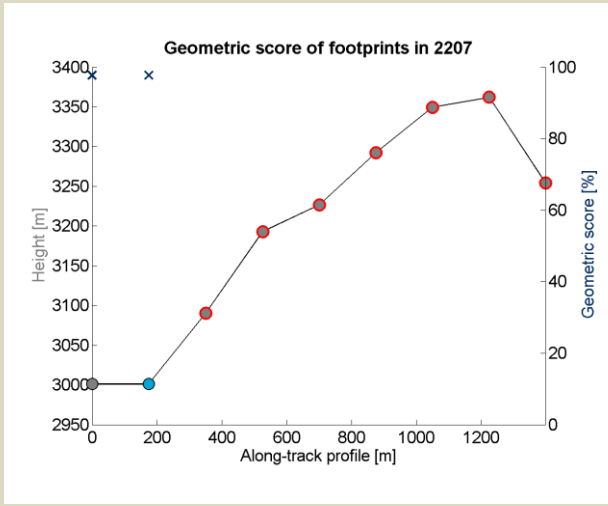

Waveform variables of 1509
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]
Waveform variables of 1205
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]

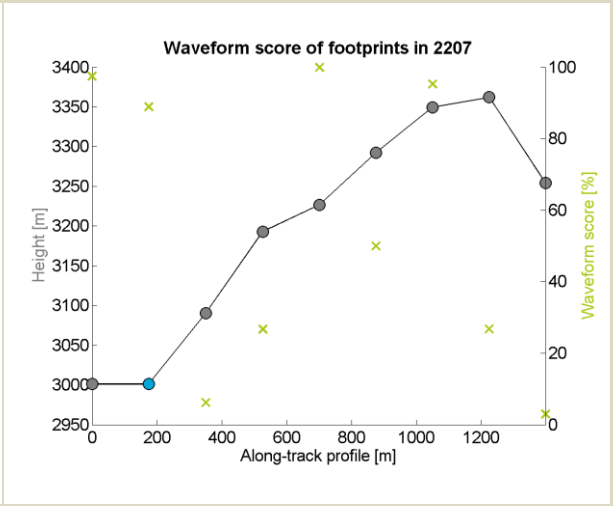

Waveform variables of 2068
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]
Waveform variables of 650
R [-]
Peaks [-]
Amp [-]
Width [m]
Kurt [-]
Skew [-]
Height [m]
Along-track profile [m]

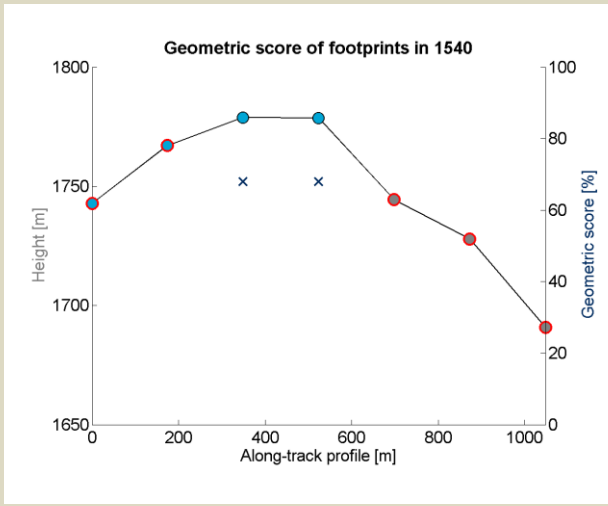
APPENDIX 2: GEOMETRIC AND WAVEFORM SCORES FOR THE SAMPLING SET OF RIVER PROFILES

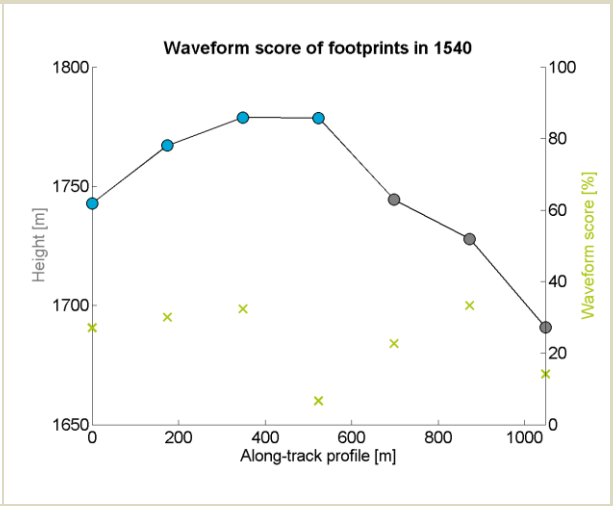


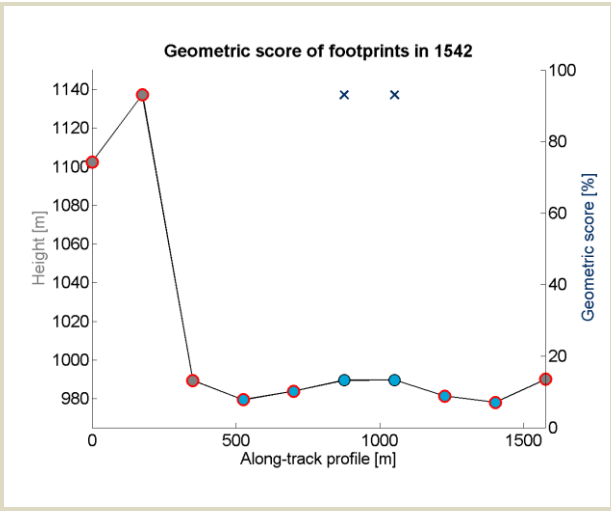


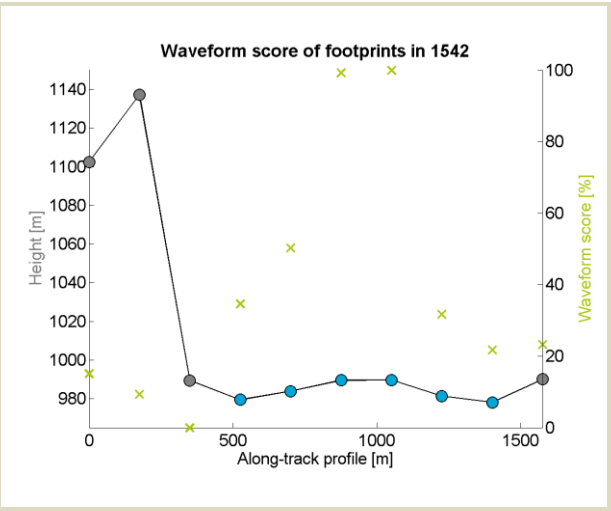


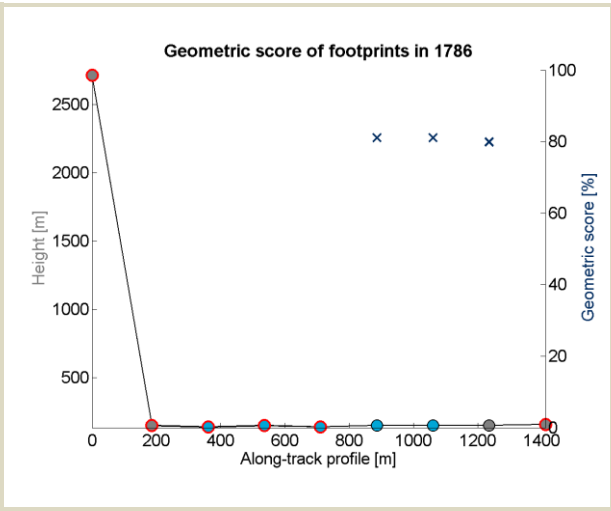


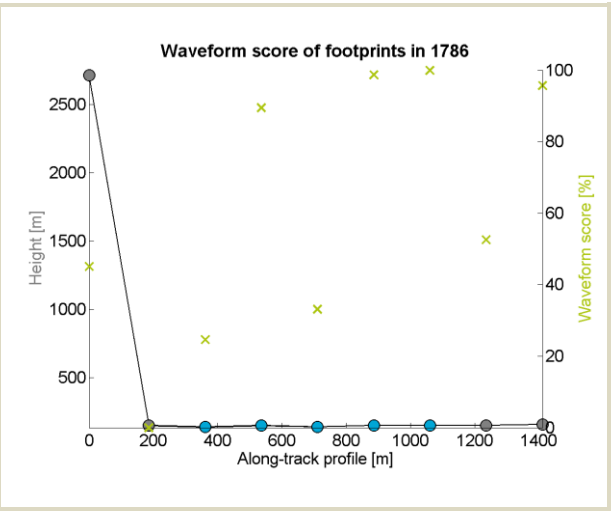


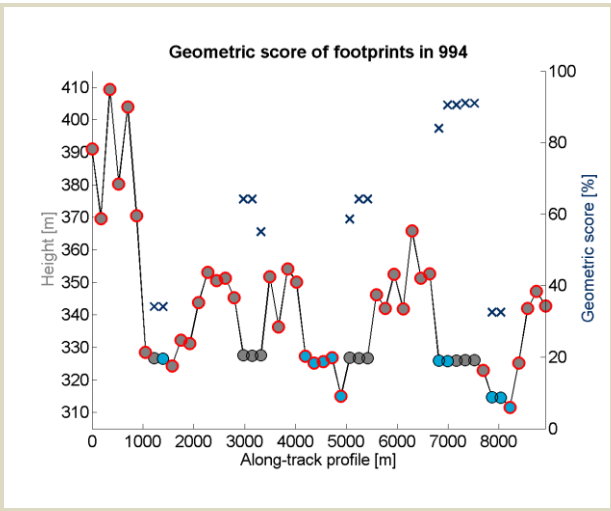


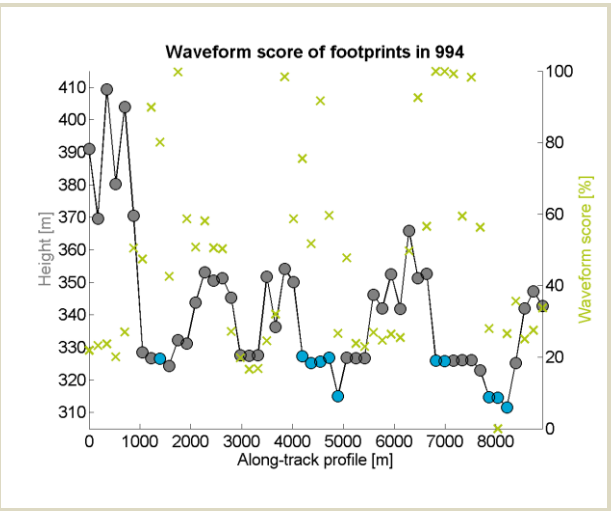

Geometric score of footprints in 1542
Height [m]
Geometric score [%]
Along-track profile [m]

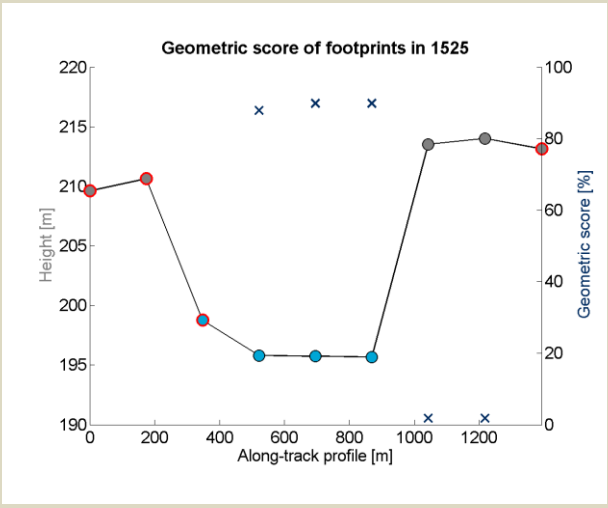

Waveform score of footprints in 1542
Height [m]
Waveform score [%]
Along-track profile [m]

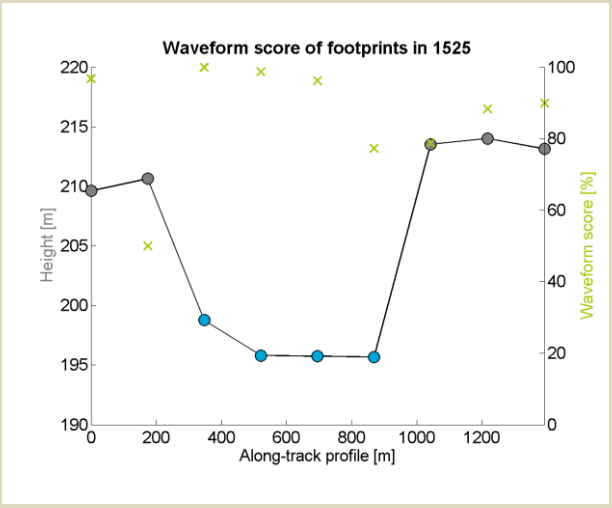

Geometric score of footprints in 1786
Height [m]
Geometric score [%]
Along-track profile [m]

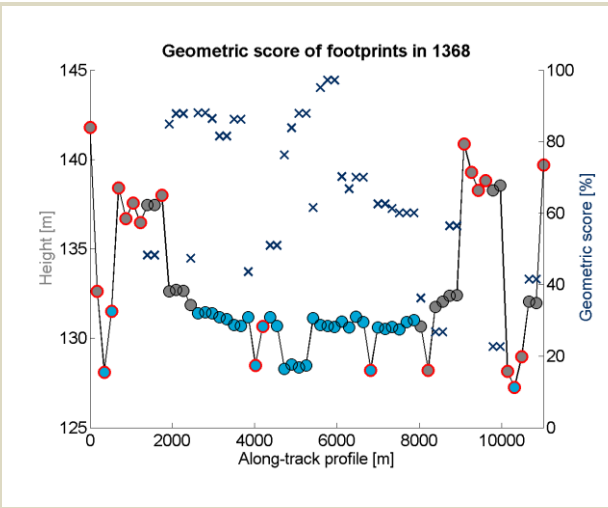

Waveform score of footprints in 1786
Height [m]
Waveform score [%]
Along-track profile [m]

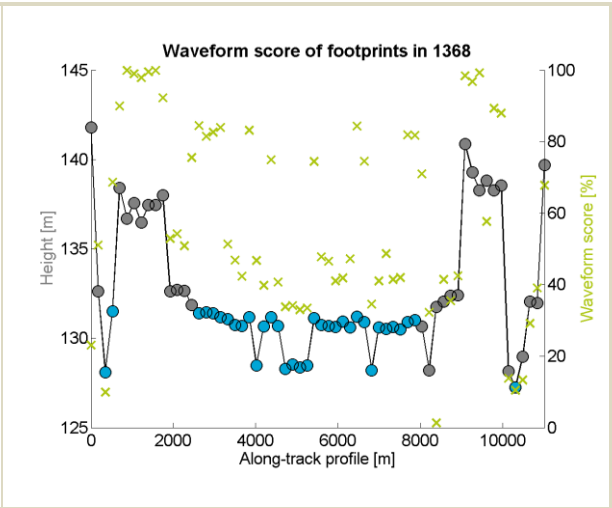

Geometric score of footprints in 994
Height [m]
Geometric score [%]
Along-track profile [m]

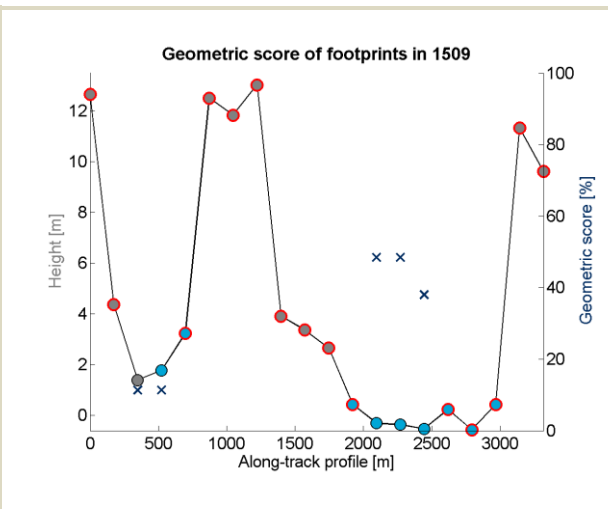

Waveform score of footprints in 994
Height [m]
Waveform score [%]
Along-track profile [m]

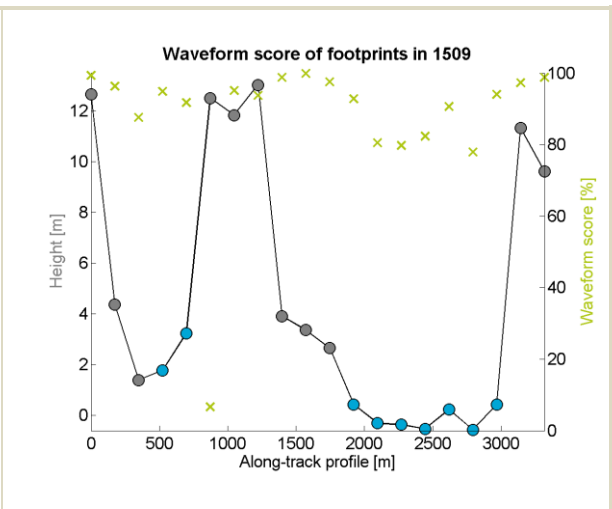

Geometric score of footprints in 1525
Height [m]
Geometric score [%]
Along-track profile [m]

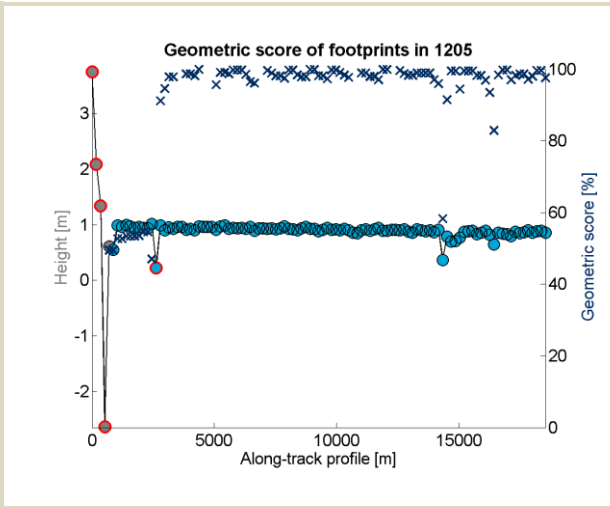

Waveform score of footprints in 1525
Height [m]
Waveform score [%]
Along-track profile [m]

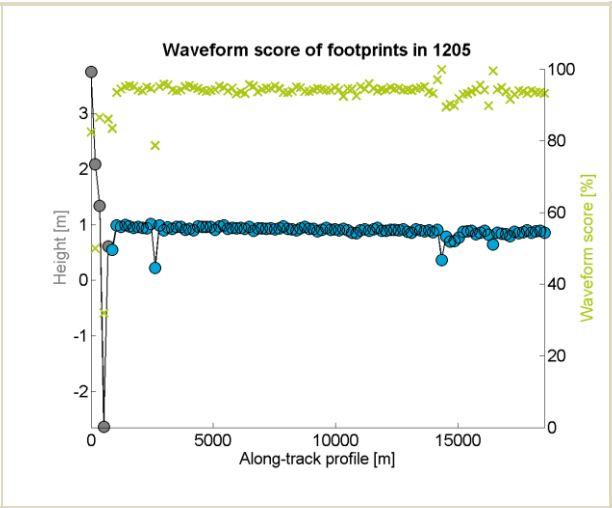

Geometric score of footprints in 1368
Height [m]
Geometric score [%]
Along-track profile [m]

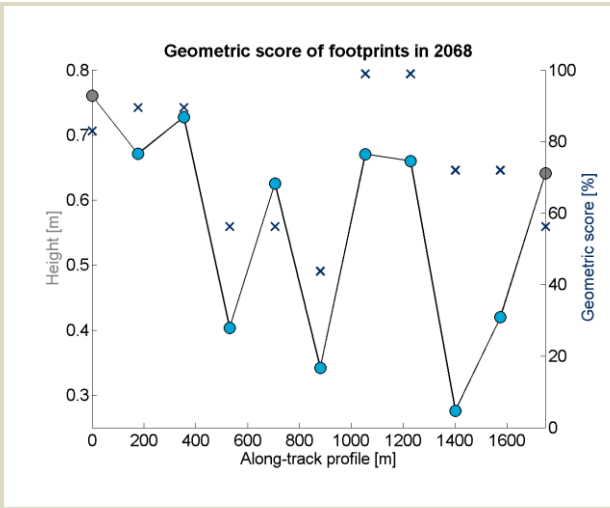

Waveform score of footprints in 1368
Height [m]
Waveform score [%]
Along-track profile [m]

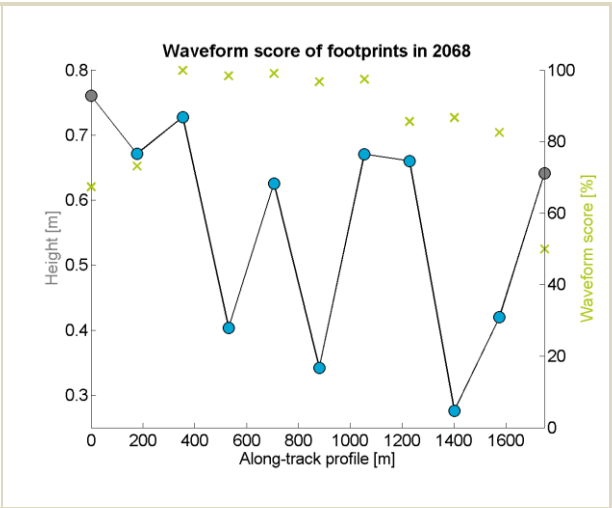

Geometric score of footprints in 1509
Height [m]
Geometric score [%]
Along-track profile [m]

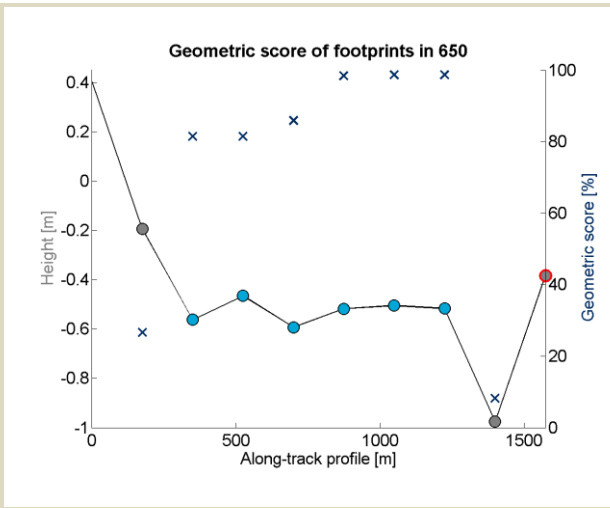

Waveform score of footprints in 1509
Height [m]
Waveform score [%]
Along-track profile [m]

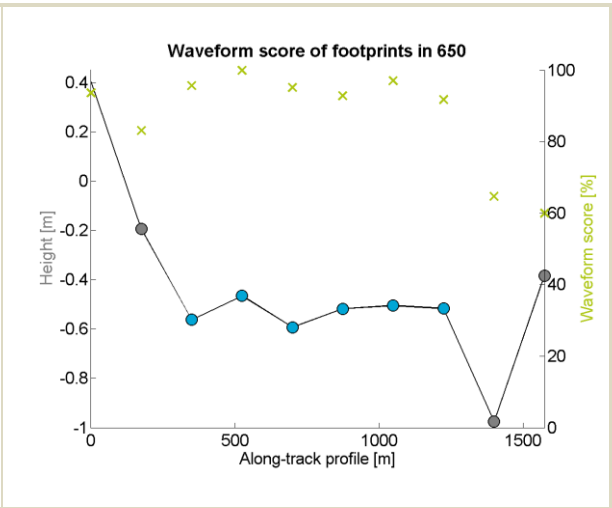

Geometric score of footprints in 1205
Height [m]
Geometric score [%]
Along-track profile [m]


Waveform score of footprints in 1205
Height [m]
Waveform score [%]
Along-track profile [m]


Geometric score of footprints in 2068
Height [m]
Geometric score [%]
Along-track profile [m]


Waveform score of footprints in 2068
Height [m]
Waveform score [%]
Along-track profile [m]


Geometric score of footprints in 650
Height [m]
Geometric score [%]
Along-track profile [m]


Waveform score of footprints in 650
Height [m]
Waveform score [%]
Along-track profile [m]

APPENDIX 3: ALTERNATIVE WEIGHTS OF GEOMETRIC AND WAVEFORM SCORES ON SAMPLING SET

Table A3.1: Results of Alternative 1: Easy (geometric score ≥ 0% and waveform score is not applicable)

ID	N (in)	Mean height [m]	Std height [m]	Type I	Type II
0969	3	3467.77	0.77	0%	0%
2207	2	3001.19	0.03	50%	0%
1540	2	1778.75	0.12	0%	50%
1542	2	989.62	0.10	0%	67%
1786	3	151.41	0.06	33%	60%
0994	15	325.00	4.27	67%	55%
1525	5	202.94	9.87	40%	25%
1368	43	131.64	2.30	35%	18%
1509	5	0.39	1.10	20%	56%
1205	102	0.89	0.09	1%	1%
2068	11	0.56	0.17	18%	0%
0650	8	-0.54	0.21	25%	0%

Table A3.2: Results of Alternative 2: Gentle (geometric score ≥ 25% and waveform score ≥ 25%)

ID	N (in)	Mean height [m]	Std height [m]	Type I	Type II
0969	3	3467.77	0.77	0%	0%
2207	2	3001.19	0.03	50%	0%
1540	1	1778.83	n/a	0%	75%
1542	2	989.62	0.10	0%	67%
1786	3	151.41	0.06	33%	60%
0994	9	324.95	3.88	56%	64%
1525	3	195.73	0.06	0%	25%
1368	40	131.30	1.79	30%	18%
1509	3	-0.40	0.12	0%	67%
1205	102	0.89	0.09	1%	1%
2068	11	0.56	0.17	18%	0%
0650	7	-0.48	0.13	14%	0%

Table A3.3: Results of Alternative 3: Mild (geometric score ≥ 50% and waveform score ≥ 75%)

ID	N (in)	Mean height [m]	Std height [m]	Type I	Type II
0969	0	n/a	n/a	n/a	100%
2207	2	3001.19	0.03	50%	0%
1540	0	n/a	n/a	n/a	100%
1542	2	989.62	0.10	0%	67%
1786	2	151.39	0.07	0%	60%
0994	4	325.93	0.17	50%	82%
1525	3	195.73	0.06	0%	25%
1368	7	131.22	0.21	0%	79%
1509	0	n/a	n/a	n/a	100%
1205	99	0.90	0.08	0%	3%
2068	7	0.54	0.17	0%	22%
0650	6	-0.53	0.04	0%	0%

Table A3.4: Results of Alternative 4: Strict (geometric score ≥ 75% and waveform score ≥ 90%)

ID	N (in)	Mean height [m]	Std height [m]	Type I	Type II
0969	0	n/a	n/a	n/a	100%
2207	1	3001.17	n/a	100%	100%
1540	0	n/a	n/a	n/a	100%
1542	2	989.62	0.10	0%	67%
1786	2	151.39	0.07	0%	60%
0994	4	325.93	0.17	50%	82%
1525	2	195.76	0.05	0%	50%
1368	0	n/a	n/a	n/a	100%
1509	0	n/a	n/a	n/a	100%
1205	87	0.90	0.05	0%	15%
2068	2	0.70	0.04	0%	78%
0650	6	-0.53	0.04	0%	0%