

# TanDEM-X- and TerraSAR-X-based Spring Flood Simulation on the Lena Delta, Siberia

Master's Thesis  
in the Department of Civil, Geo, and Environmental Engineering  
Technical University of Munich.

**Supervised by** Achim Roth  
Department of Land Surface Dynamics, DLR  
Punit Kumar Bhola  
Chair of Hydrology and River Basin Management, TUM  
Timo Schaffhauser  
Chair of Hydrology and River Basin Management, TUM

**Submitted by** Avi Putri Pertiwi  
03681883

**Submitted in** Munich, 20.11.2019

# Appendix I

## Declaration of Authorship

I hereby declare that this submitted thesis is my own work. All used direct or indirect sources are acknowledged as references.

---

Munich, 20.11.2019, Avi Putri Pertiwi

## Abstract

Arctic watersheds heavily influence the oceans and the mean global temperature. They are the largest source of freshwater to the Arctic Ocean and affecting the sea ice cover and the ocean conveyor belt. An example of these impacts is the 7% rise of the average annual freshwater discharge from the six largest Eurasian rivers to the Arctic Ocean over the last century. The interaction between the Arctic water budget and climate change makes the study of arctic surface water important for climate- and environment-related research.

The river Lena is one of the four main contributors to the Arctic Ocean freshwater, the others being Yenisei, Ob, and Mackenzie. It is the 10<sup>th</sup> longest natural river in the world, located in eastern Siberia with a basin the size of  $2.4 \times 10^6$  km<sup>2</sup> and an average annual water runoff of > 500 km<sup>3</sup>. Extreme spring flood events take place annually in Lena, due to the melting of the accumulated snow from the previous winter. During this flood event, around 40% of its annual discharge is released to the ocean.

The study area consisted of the central part of the Lena river delta where the main channel diverges into its major distributaries, namely the Trofimovskaya, Bykovskaya and Olenekskaya channels. This area consists of a braided river system with a vast floodplain area, surrounded by cliffs on its east and west side. Most of the floodplain area was flooded during the spring flood. This remote area cannot be easily accessed, making data acquisition challenging and costly.

The aim of this thesis is to perform a feasibility study in implementing hydraulic modeling methods using the TerraSAR-X and TanDEM-X datasets during the annual spring flood events between 2013 and 2019. This approach integrated two different methods: remote sensing and river hydraulic methods. This integration is widely known as remote sensing of rivers and has been an emerging sub-discipline on river hydraulic studies. The remotely-sensed datasets provided the parameters for hydraulic measures on poorly gauged regions, whereas hydraulic modeling helped derive parameters which could not be acquired by remote sensing.

HEC-RAS by the US Army Corps of Engineers was selected as the hydraulic modeling tool. The selected module for the simulation was the 1-dimensional unsteady flow (dynamic wave) analysis. The resulting simulated inundation boundaries were validated using multi-temporal TerraSAR-X images. The other hydraulic variables, namely flow velocity and flood depth, were also extracted from the simulation. Additionally, the outcomes from the feasibility study, the strengths and weaknesses of the approach, were also expected. By learning from the

weaknesses of the approach, solutions to improve and extend this study in the future were devised.

With very limited field data availability (no field data of the river bathymetry and surface roughness), the assessed accuracy of the model ranged between 65-95%. Considering these results were achieved with many approximations of parameters with high uncertainties, the approach was regarded applicable. This means, with better field data acquisition, the accuracy of the model can be pushed even higher, with better calibration and validation processes. The modeled flow velocity before the flood event was lower than 1.5 m/s. Then the velocity built up to 6 m/s on some parts of the channel during the peak of the flood. During the peak flood of 2014 (the most severe flood during of the simulated timeframe), the flood depth on the floodplain ranged between 0.001 – 2.5 m.

In the future, the proposed method of this feasibility study can be extended for an improved hydraulic study of Lena delta, such as sediment transport, water quality, freshwater-seawater interaction, or ecological modeling. The accuracy can be improved with better field acquisition and methods with more precision, such as 2D and 3D modeling.

# Table of Contents

<b>1.</b>	<b>Introduction .....</b>	<b>1</b>
<b>2.</b>	<b>Project Description .....</b>	<b>3</b>
2.1.	Objectives and Expected Outcomes .....	3
2.2.	Study Area .....	5
2.3.	Simulation Period .....	8
2.4.	Study Limitations and Assumptions.....	9
2.5.	Required Tools and Licenses .....	10
<b>3.</b>	<b>Materials.....</b>	<b>12</b>
3.1.	Remotely-sensed Datasets.....	12
3.1.1.	TerraSAR-X and TanDEM-X Mission Datasets .....	14
3.1.2.	RapidEye Ortho – Level 3A.....	19
3.1.3.	Landsat 8 Level 2 – Surface Reflectance.....	20
3.2.	Ground-Truth Materials.....	22
3.2.1.	Upstream Discharge at Kyusyur .....	22
3.2.2.	Surface Water Level at Khabarova.....	22
3.2.3.	Land Cover.....	23
<b>4.</b>	<b>Remote Sensing Methodology.....</b>	<b>25</b>
4.1.	TerraSAR-X, RapidEye, and Landsat 8 Dataset Pre-processing.....	25
4.2.	TanDEM-X DEM Correction.....	28
4.3.	Bathymetry Approximation .....	29
4.4.	Flood Mapping.....	33
4.5.	Manning’s Surface Roughness Coefficient .....	35
4.6.	Converting the GIS Data into HEC-RAS Input.....	40
<b>5.</b>	<b>Hydraulic Simulation .....</b>	<b>47</b>
5.1.	Discharge Data Preparation .....	50
5.2.	Simulation Setup .....	52
5.3.	Stabilizing the Unsteady Simulation.....	55

<b>6.</b>	<b>Results and Discussions.....</b>	<b>60</b>
6.1.	Bathymetry .....	60
6.2.	Land Cover .....	62
6.2.1.	TerraSAR-X and RapidEye Stack.....	62
6.2.2.	Landsat 8 .....	66
6.3.	Manning’s Surface Roughness Coefficient.....	69
6.4.	HEC-RAS Sensitivity Analysis .....	70
6.4.1.	The model sensitivity to the cross-sectional line spacing ( $\Delta x$ ).....	70
6.4.2.	The model sensitivity to the time step ( $\Delta t$ ) .....	71
6.4.3.	The model sensitivity to the downstream slope .....	71
6.5.	Inundation Boundary .....	71
6.6.	Flow Velocity.....	78
6.7.	Flood Depth .....	79
<b>7.</b>	<b>Conclusions and Future Works.....</b>	<b>81</b>
7.1.	Conclusions.....	81
7.2.	Future Works .....	84
	<b>List of Figures.....</b>	<b>86</b>
	<b>List of Tables .....</b>	<b>89</b>
	<b>List of Abbreviations .....</b>	<b>90</b>
	<b>List of Symbols.....</b>	<b>91</b>
	<b>References.....</b>	<b>92</b>

# 1. Introduction

Arctic watersheds have a big influence on the oceans and the mean global temperature. They are the largest source of freshwater to the Arctic Ocean, heavily affecting the sea ice cover and the ocean conveyor belt (Papa, Prigent, and Rossow 2008). Changes in the ocean conveyor belt have significant impacts on the Earth's radiation budget. An example of these impacts is the 7% rise of the average annual freshwater discharge from the six largest Eurasian rivers to the Arctic Ocean over the last century (Peterson et al. 2002). The relation between the Arctic water budget and climate change makes arctic surface water important for climate- and environment-related researches.

Lena is one of the four main Arctic Ocean freshwater contributors, among other rivers e.g. Yenisei, Ob, and Mackenzie (Vuglinsky 2002). It is located in eastern Siberia with the basin the size of  $2.4 \times 10^6$  km<sup>2</sup> and the annual average water runoff of  $> 500$  km<sup>3</sup> (Ma et al. 2005). It is the 10<sup>th</sup> longest river in the world without human-made construction (Sakai et al. 2015). Lena experiences annual extreme flooding events due to the melting of the accumulation of the winter snow.

Lena typically remains completely frozen from early December to late April. On the upstream reaches, the ice thawing begins in late April and extends to the downstream reaches around late June, normally coinciding with the annual peak flow (Ma et al. 2005). The maximum transport of sediment, organic carbon, and trace metals occurs during this period affecting the floodplain ecology (Pavelsky and Smith 2004). To better understand the impacts this has, monitoring of the river is critical.

Despite of the importance, the number of hydrological monitoring stations in the Arctic regions has decreased dramatically over the past few decades (Pavelsky and Smith 2004) (Sakai et al. 2015). While some open source datasets are made available openly by public and private institutions, the number of stations is limited compared to the vast extent of the prospective study area. Remote sensing techniques can be utilized to compensate for the lack of in-situ data.

Remote sensing of rivers is an emerging sub-discipline in the river sciences (Marcus and Fonstad 2010). Remote sensing techniques are the alternative to gauged data in remote regions and have the advantage of large area coverage. In this study, remotely-sensed datasets in the form of satellite imagery provided the parameters for hydraulic measures on poorly gauged regions, whereas hydraulic modeling helped derive parameters which couldn't be acquired by

remote sensing. Therefore, these two methods complement each other. The utilized satellite images provided water extent and Digital Elevation Model (DEM). On the other hand, there was limited knowledge regarding the water flow due to the inexistence of river bathymetry and flow velocity. DEM acquired by radar sensors doesn't give the surface elevation below the water because the pulse doesn't penetrate water body. Additionally, up until now, water flow velocity cannot be directly measured by remote sensing methods. These unknowns were to be acquired through hydraulic modeling.

The aim of this study is to setup a method to approximate both the unknown river bathymetry and flow velocity during the Spring flood on Lena Delta by performing a hydraulic modeling using the input parameters from the TanDEM-X, TerraSAR-x, RapidEye and Landsat 8 datasets. TanDEM-x datasets were used for the topography mapping. TerraSAR-X, RapidEye, and Landsat 8 datasets were used to acquire the surface roughness, which is an important input parameter for hydraulic modeling, and to validate the resulting inundation boundary. Each of these satellite imageries has different characteristics, therefore having their own advantages and disadvantages to be generated as the input for the hydraulic modeling.

The selected hydraulic modeling tool was HEC-RAS, which is a Windows-based open-source software developed by the US Army Corps of Engineers, specifically the Hydrologic Engineering Center (CEIWR-HEC) (CEIWR-HEC 2019). Specifically, the selected modeling method for this study was the 1D unsteady flow analysis. In the future, the hydraulic modeling methods can be extended for further study of the Lena river, such as sediment transport, water quality, freshwater-seawater interaction, ecological modeling, among others.

## 2. Project Description

This thesis emerged from a cooperation between the German Aerospace Center (DLR), specifically the Land Surface Dynamics department of the German Remote Sensing Data Center Institute (DFD), and the Chair of Hydrology and River Basin Management of the Technical University of Munich (TUM). Both parties contributed to the key components of this study. Most datasets and processing tool licenses were provided by the DLR. Additional open datasets and open source tools were also utilized. This chapter describes the study area, simulation period, materials, study limitations, and methodology.

### 2.1. Objectives and Expected Outcomes

This thesis is part of the permafrost river delta study in the DFD, DLR. The DFD aims to utilize TerraSAR-X (TSX) and TanDEM-X (TDX) products for this study to see the potentials of those products, especially in multidisciplinary studies. Lena delta is one of the area of interest in this study. The final goal of the river delta study is to understand better the hydraulic processes on Lena delta, especially during the annual Spring flood.

TSX and TDX products were already previously used to derive hydraulic parameters such as the flood map, the surface water level (so far, on a relative degree), and water volume. However, up until now, the flow velocity, a hydraulic parameter that is of importance in the hydraulic processes, cannot yet be quantified directly from these products. Therefore, the DFD was interested in incorporating hydraulic simulation into their study. This thesis was devised as a feasibility study to assess the practicality of this desired approach.

As a feasibility study, this thesis aims to:

- devise an approach to integrate remote sensing methods (by means of TSX and TDX products) and hydraulic modeling methods for the Lena delta study
- declare and find the input parameters that are necessary for the simulation
- approximate the necessary input parameters that were not available prior to this study
- demonstrate a method to convert the GIS-based datasets (from the remote sensing methods) into a required input format for the hydraulic modeling
- assess the strengths and weaknesses of the proposed approach
- define the factors that are contributing to the strengths and the weaknesses of the proposed approach
- recommend the solutions to improve this approach for future works

The expected outcomes from this thesis are:

- a proposed method to use TSX and TDX products to perform hydraulic simulation with HEC-RAS
- the accuracy assessment of the proposed method, by comparing the modeled inundation boundary to the remote-sensing-derived flood map
- the flow velocity of the river delta
- the depth of the flood on the floodplain
- the strengths and weaknesses of the proposed approach
- the solution to improve and extend this study in the future

The approach of this thesis is by combining the remote sensing method (satellite image pre-processing, land cover classification, and flood map generation) with hydraulic simulation (using HEC-RAS). This thesis depended heavily on remotely-sensed datasets, given the very limited availability of field data of this remote and deserted area. From these remotely-sensed datasets, some hydraulic parameters were derived. Hydraulic simulation then was performed to compute the hydraulic parameters that could not be derived from remote sensing.

Table 1 shows the hydraulic parameters that could be derived from TSX-TDX products and produced by hydraulic modeling, that were related to this thesis. Flood maps were derived from the TSX-TDX products and hydraulic modeling. The surface water level was relatively approximated from TSX-TDX products. The TDX DEM provided the topography. Land cover map was derived from TSX products (with the addition from RapidEye and Landsat 8 imageries as well). Aside from the flood map, hydraulic modeling was also used in this thesis to quantify the surface water level, flow velocity, and flood depth. The accuracy of the hydraulic simulation results then would be assessed by comparing the flood map that was generated from the TSX and TDX products with the resulting flood map from the hydraulic simulation.

However, there are three unknown parameters that were required to perform the hydraulic simulation; the river bathymetry, the surface roughness (both of channel and floodplain), and the friction slope. These three unknown parameters were to be approximated throughout the simulation.

**Table 1 The hydraulic parameters derived from the TerraSAR-X and TanDEM-X products, the hydraulic simulation, and the unknown parameters in this thesis**

<b>Remotely-sensed information</b>	<b>Hydraulic modeling result</b>	<b>Unknown parameters</b>
<ul style="list-style-type: none"> <li>- Flood map</li> <li>- Relative surface water level</li> <li>- Topography (DEM)</li> <li>- Land cover</li> </ul>	<ul style="list-style-type: none"> <li>- Inundation boundary</li> <li>- Flow velocity</li> <li>- Flood depth</li> </ul>	<ul style="list-style-type: none"> <li>- River bathymetry</li> <li>- Surface roughness</li> <li>- Friction slope</li> </ul>

## 2.2. Study Area

As previously mentioned, Lena is the 10<sup>th</sup> longest natural river in the world. In this study, only a subset of Lena was analyzed, which can be seen in Figure 1, precisely on Lena delta. The study area included a portion of the Lena reach, Bulkurskaya reach, and Bykovskaya reach. This area is remote, deserted, and of challenging weather conditions, making field data acquisition difficult and costly. Lena delta consists of an anastomosing river system with, enclaved by cliffs on its east and west side. Anastomosing river system is indicated by a river system having two or more channels with high stability that are subject to avulsions (Sutter 2008). This kind of river system is common on wetlands. The upstream of the river located on the south part of the reach. The river flows northwards, from a single main channel, then distributed along the anastomosing river system, supplying water runoff to the delta and then to the Arctic ocean.

A river discharge gauging station is located upstream in Kyusyur (green dot in Figure 1). The simulated area (red polygon in Figure 1) started 124.65 km downstream of the Kyusyur gauging station. The total simulated study area is 1,561.15 km<sup>2</sup>, with the length of 93.14 km. On the Bykovskaya reach in the simulated area, the Khabarova surface water level gauging station is located. Downstream between the Sardakh and Trofimovskaya island (orange polygon in Figure 1), a study of sediment transports (Fedorova et al. 2015) was carried out. This area will be of relevance in further process (in subchapter 4.3). Table 2 provides a summary of the common information of Lena delta and the simulated area.

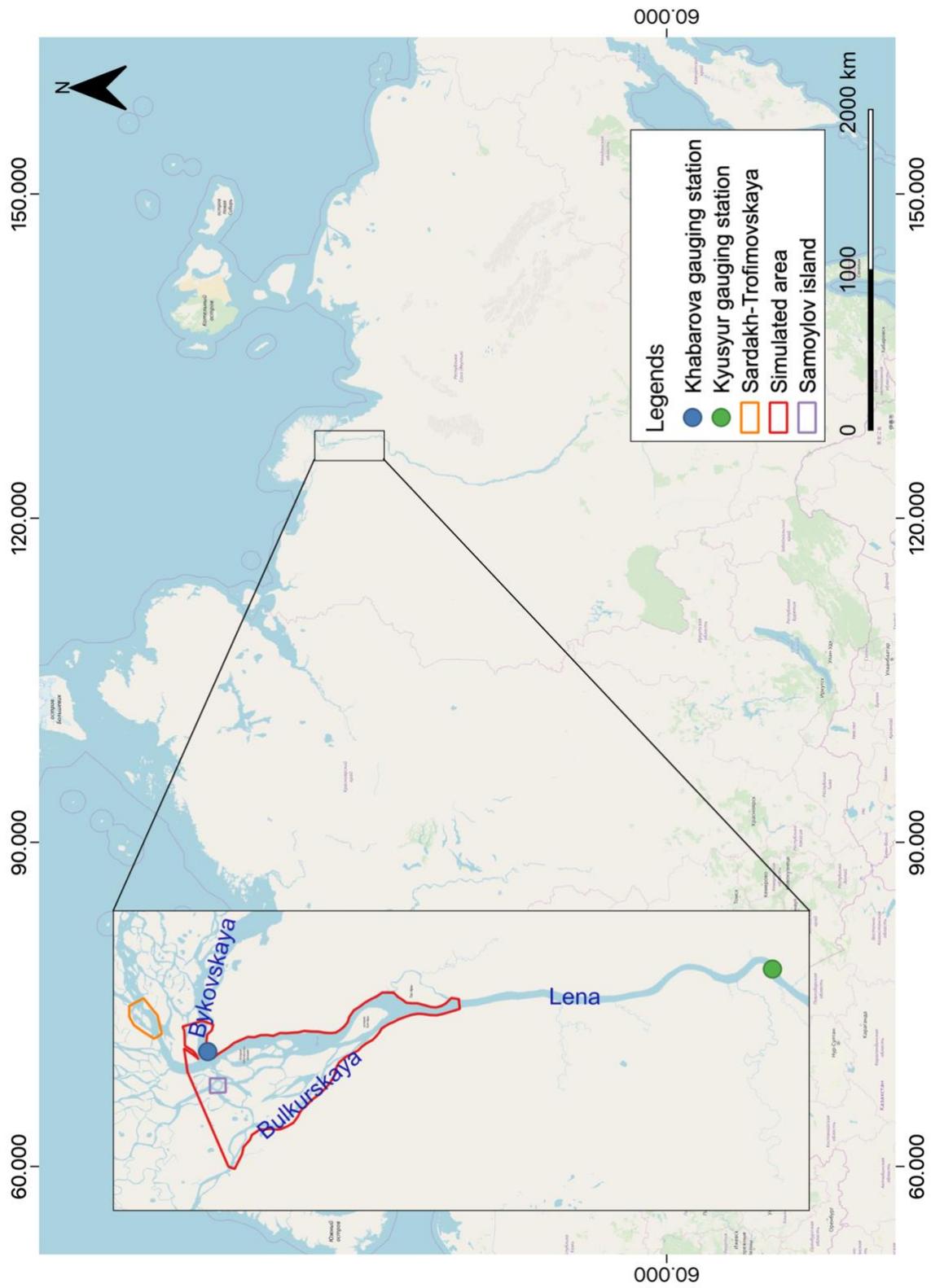


Figure 1 The Lena delta study area (OpenStreetMap, Pertiwi)

**Table 2 Summary of the common information of Lena delta**

<b>Area of Lena delta</b>	32,000 km <sup>2</sup> (ESA 2019a)
<b>Area of the simulated study area</b>	1,561.15 km <sup>2</sup>
<b>Length of Lena river</b>	4,500 km (ESA 2019a)
<b>Length of investigated section</b>	93.14 km
<b>Average summer rainfall</b>	145.2 mm (Boike et al. 2019)
<b>Spring flood period on the Lena delta</b>	May – June (ArcticGRO 2003, Minpirody 2019)

The floodplain of the study area on Lena delta is mainly covered in sand, rock, and low vegetation (Holocene and Pleistocene terrace) (Heim 2019). The low-elevation floodplain between the Lena and Bulkurskaya reaches was mostly flooded during the peak of the Spring flood. An aerial photo of part of Lena delta is shown in Figure 2. An observation center is located on Samoylov island (purple polygon in Figure 1). This observation center is used for permafrost studies. Through a camera attached to this building, the flood on this island can be monitored. A captured photo of before and after the Spring flood event on the Samoylov island in 2014 is shown in Figure 3.



**Figure 2 An aerial photo of the Lena delta**



Figure 3 The Spring flood event 2014 on the Samoylov island, before and during the peak flood, recorded by the camera

### 2.3. Simulation Period

Extreme Spring flood events take place annually in Lena. According to the gauged datasets (ArcticGRO 2003, Minprirody 2019) and remotely-sensed image visual interpretation, the runoff usually rises around late May to early June, then recedes abruptly until November on the delta. The hydraulic simulation was performed on the separate Spring flood events that occur starting from 2013 to 2019. The selected timeframe of the simulation was between May and July each year. Three conditions were simulated from each event: low runoff (before the flood), high runoff (during the peak flood), and medium runoff (when the flood was receding), totaling in 21 simulated cases. The selection of the event day depended on the availability and the quality of the remotely-sensed images, as well as the runoff value of the respective day

## 2.4. Study Limitations and Assumptions

Limitations in thesis were mainly caused by the limited availability of field data. Another reason that contributed to this was the limitations of the selected HEC-RAS simulation module. Due to these limitations, assumptions had to be made. Table 3 shows the full list of the limitation and the corresponding assumptions.

**Table 3 The study limitations and assumptions**

<b>Limitation</b>	<b>Assumption</b>
No bathymetry information	<ul style="list-style-type: none"> <li>- Trapezoidal-shaped channels</li> <li>- The bathymetry is digitized and interpolated</li> <li>- The average channel bed slope is assumed to be very small, according to the flat floodplain</li> </ul>
The river flow direction was unknown and could not be identified from the terrain, since the terrain is almost flat	The anastomosing river completely flows from the north to the south (1D flow direction)
Unknown coordinates of the WL gauging station	Estimated to be located at 72.403876°N, 126.812948°E
The discharge gauging station is 1km away from the upstream of the simulated area	The time of the discharge was adjusted according to the multi-temporal TSX interpretation
Limited coverage of TSX imagery on the study area	Use Landsat 8 and RapidEye imagery
Limited ground-truth land cover information of the study area	Use clustering method to group similar land cover
No ice modelling	During the simulation period, the river doesn't freeze
No weather data	Weather effects were neglected
Sediment transport not simulated	Channel bed and banks are fixed throughout the simulation period
Limited HEC-RAS geometry points might be lower than the dataset resolution	The geometry points were reduced in such a way to minimize the changes of the shape
The lakes on the floodplains were not modeled as storage area due to the unavailability of the corresponding data	The surface that was detected by the TDX was regarded as terrains
1D simulation is simplifying the simulation	<ul style="list-style-type: none"> <li>- 1D flow direction</li> <li>- Horizontal water level across one cross section</li> </ul>

The lack of available field data came to the expense of constraining limitations in this thesis. The channel bathymetry, which is a crucial parameter for hydraulic simulation was not known. Due to this reason, the river flow direction was also unknown. The ice and weather data was also unavailable. Therefore, the hydrological processes (weather effects and ice thawing) could not be incorporated in the simulation. The provided TSX product coverage didn't cover the whole simulated area. RapidEye (RE) and Landsat 8 (LS8) imageries were also utilized to compensate for the uncovered area. The available flow data wasn't exactly measured on the required location for the simulation, adjustments had to be made.

The last set of constraints came from the limitations of the selected modeling method. The 1D simulation which was chosen for this study simplified the geometries and the flow computations. The reasons behind the selection of this module will be elaborated in more details in 4.6.

## 2.5. Required Tools and Licenses

This thesis incorporated Remote Sensing methodology and hydraulic modeling. Remote sensing methods were performed to process remotely-sensed datasets into useful environmental parameters for the simulation. Geographic information system (GIS) methodology was useful to read, modify, write, and process the remotely-sensed datasets. Furthermore, GIS methodology was also utilized to extract hydraulic parameters from the remotely-sensed datasets, store them into geodatabases, and write the stored information into the preferred input of the hydraulic modeling tool requirements. These processes were performed either using the GDAL and OGR libraries of the Python programming environment or with GIS tools (QGIS, ArcGIS). Hydraulic modeling tool (HEC-RAS) was used to produce the desired hydraulic parameters by means of the input prepared from the remotely-sensed datasets. The full list of the tools and the licenses used in this thesis is provided in Table 4.

**Table 4 The used tools for this thesis and the license source of each tool**

<b>Tools</b>	<b>Extensions/Library</b>	<b>Function</b>	<b>License</b>
Google Earth Pro		Visual identification	Free to use
DLR Multi-SAR processor		Pre-process the SAR raw images	DLR
IDL Kennaugh to dB converter		To convert the Kennaugh number into Sigma naught	DLR
ArcGIS	<ul style="list-style-type: none"> <li>- Spatial Analyst</li> <li>- 3D Analyst</li> <li>- Sampling Design Tool</li> </ul>	<ul style="list-style-type: none"> <li>- Mean-shift segmentation</li> <li>- IDW interpolation</li> <li>- Sample point generation</li> </ul>	DLR
pgAdmin 4	PostGIS	PostgreSQL GUI to manage geodatabases	Open source
QGIS	<ul style="list-style-type: none"> <li>- GRASS</li> <li>- SAGA</li> <li>- RiverGIS</li> </ul>	To process and modify GIS data	Open source
Python 3.7.0 programming environment	<ul style="list-style-type: none"> <li>- Pandas</li> <li>- Geopandas</li> <li>- GDAL, OGR</li> <li>- Scikit-learn</li> <li>- Rasterio</li> </ul>	Land cover classification Flood map generation	Open source
HEC-RAS 5.0.7		Hydraulic simulation	Open source

### 3. Materials

This study was heavily dependent on the remotely-sensed data due to the very limited gauged data. Moreover, the study area was remote and of large extent. Fortunately, nowadays satellites have the ability to provide information with moderate detail due to their high spatial and temporal resolution. Additionally, available gauged datasets were also used to compensate for the information the satellites could not obtain. This subchapter comprises descriptions of both remotely-sensed and ground-truth datasets.

#### 3.1. Remotely-sensed Datasets

This study utilized remotely-sensed datasets from four different satellite platforms. These platforms are TerraSAR-X (TSX), TanDEM-X (TDX), RapidEye (RE), and Landsat 8 (LS8). The TDX DEM and TSX Stripmaps were acquired through the DLR database with the science exploitation permit. The available RE imagery was provided by the Alfred-Wegener-Institut (AWI) through their corporation with the DLR in cryosphere studies. Landsat 8 datasets were downloaded from U.S. Geological Survey (USGS)'s portal free of charge (USGS 2019a).

These platforms operate on different types of instruments. TDX and TSX, which are twin satellites, are active radar-based platforms. RE and LS8, on the other hand, are optical satellites. TDX and TSX receive backscatter, whereas RE and LS8 receive reflectance. Other than the transmission difference, their products differ in spatial resolution, temporal resolution, and area coverage. Table 5 shows the advantages and disadvantages of each satellite product.

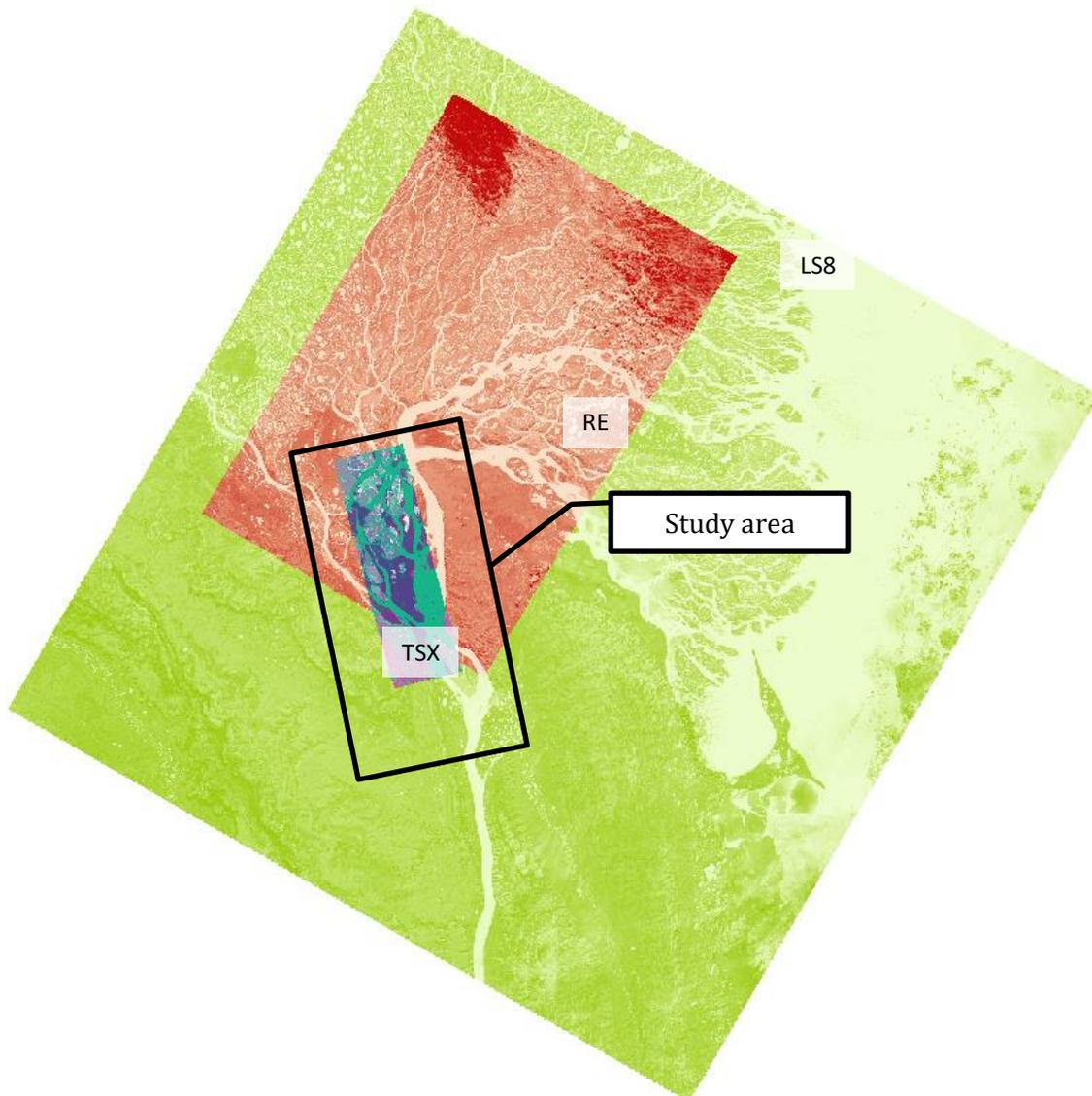
**Table 5 The advantages and disadvantages of the utilized TerraSAR-X, RapidEye, and Landsat 8 imageries**

	<b>TSX Stripmap</b>	<b>RE</b>	<b>LS8 L2 Surface Reflectance</b>
<b>Advantages</b>	<ul style="list-style-type: none"> <li>- High spatial resolution</li> <li>- High temporal resolution (almost daily with some gaps)</li> <li>- The microwave signal penetrates clouds</li> </ul>	<ul style="list-style-type: none"> <li>- High spatial resolution</li> <li>- Bigger swath width than that of TSX</li> </ul>	<ul style="list-style-type: none"> <li>- Big swath width</li> <li>- Available for free</li> </ul>
<b>Disadvantages</b>	<ul style="list-style-type: none"> <li>- Relatively small swath width</li> <li>- Various polarizations and incidence angles resulting in different sensitivity to different types of surface</li> </ul>	<ul style="list-style-type: none"> <li>- Optical sensor depends heavily on the weather, might have cloud covers</li> </ul>	<ul style="list-style-type: none"> <li>- Low spatial resolution</li> <li>- Optical sensor depends heavily on the weather, might have cloud covers</li> </ul>

Table 6 provides the acquisition dates, equipped sensors, georeferences, imaging modes, spatial resolutions, and swath widths of these products. TSX Stripmap products were used for multi-temporal analyses. 3 TSX Stripmaps were selected for each Spring flood event between 2013 and 2019, in regards to the runoff value of the acquisition day. As RE and LS8 image products were utilized to map the soil cover distribution, only a single acquisition of each RE and LS8 were selected, given that there was almost no significant change of soil cover on the study area during the simulation period (according to the visual multi-temporal TSX interpretation). Figure 4 shows the comparison of a single-tile area coverage between TSX, RE, and LS8 image products.

**Table 6 Summary of the specifications of the TerraSAR-X, RapidEye, and Landsat 8 datasets used in this study**

<b>Specifications</b>	<b>TSX Stripmap</b>	<b>RE Ortho Level 3A</b>	<b>LS8 L2 Surface Reflectance</b>
<b>Acquisition date(s)</b>	3 different acquisition days between 2013-2019 (Table 7)	9 July 2014	9 July 2014
<b>Type</b>	Radar	Optical	Optical
<b>Sensor</b>	Synthetic-Aperture Radar (SAR)	Jena-Optronik	Operational Land Imager (OLI)
<b>Transmission</b>	Active	Passive	Passive
<b>Band(s)</b>	X-band (9.65 GHz)	Blue, green, red, red edge, near-infrared (Table 8)	Coastal aerosol, blue, green, red, near-infrared, short-wave infrared (Table 9)
<b>Polarization</b>	HHVV or HHHV	-	-
<b>Georeference</b>	EPSG:32652	EPSG:4326	EPSG:4326
<b>Imaging Mode</b>	Stripmap	-	-
<b>Revisit cycle</b>	11 days	5.5 days	16 days
<b>Spatial Resolution</b>	5m	5m	30m
<b>Swath Width</b>	17 km	77 km	185 km



**Figure 4** An illustration of the swath width comparison between TerraSAR-X (purple and turquoise), RapidEye (red), and Landsat 8 (light green)

### 3.1.1. TerraSAR-X and TanDEM-X Mission Datasets

TSX and TDX are commercial twin satellites operating in the partnership between German Ministry of Education and Science (BMBF) through the German Aerospace Center (DLR) and Airbus Defence and Space (AIRBUS 2015). TSX was launched on 15 June 2007. TDX is an add-on to TSX for digital elevation measurement, launched on 21 June 2010 (Wessel 2016). TSX and TDX both have exceeded their designed operational lifetime (5 years) and are still working reliably. Both satellites are equipped with Synthetic Aperture Radar (SAR) sensor, operating on the X band, with the center frequency of 9.65 GHz, using Active Phases Array Antenna on single and dual polarization (AIRBUS 2015). Operating on the microwave band with active

sensors allows TDX and TSX to acquire backscatters regardless of the cloud cover, weather conditions, or the absence of daylight. Figure 5 illustrates the TDX-TSX platform over the earth.

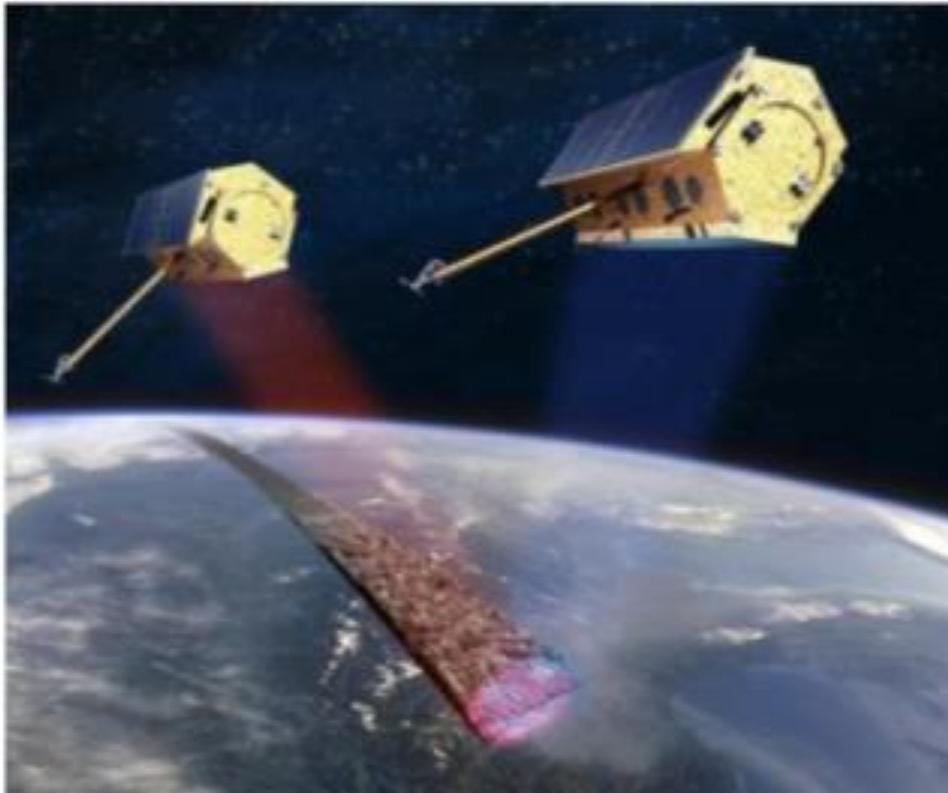


Figure 5 Artist view on the TSX and TDX mission (AIRBUS 2015)

TSX image products provide detailed soil cover and surface formation analyses (DLR 2019). Various types of surface coverage reflect the transmitted signal differently, making TSX image products good for land cover studies. TDX provides the global Digital Elevation Model (DEM) as a Digital Surface Model (DSM) instead of Digital Terrain Model (DTM), meaning that the product doesn't represent the elevation of the bare earth surface but includes the covers on top of the terrain, e.g. vegetation, water body, and man-made objects (EOC 2019).

#### **3.1.1.1. TanDEM-X DEM**

The TDX DEM covering the study area was generated from 2 TSX/TDX coverages in 2011 during which the surface water level of the water body was low (29 January, 28 September, and 31 October 2011). These acquisition days were marked in black lines in Figure 6 over the recorded surface water level at Khabarova Station. The relative and absolute vertical accuracy values were 2m and 10m, respectively. The 12.5m DEM resolution was sampled into 5m spacing. Edge preserving filter was applied.

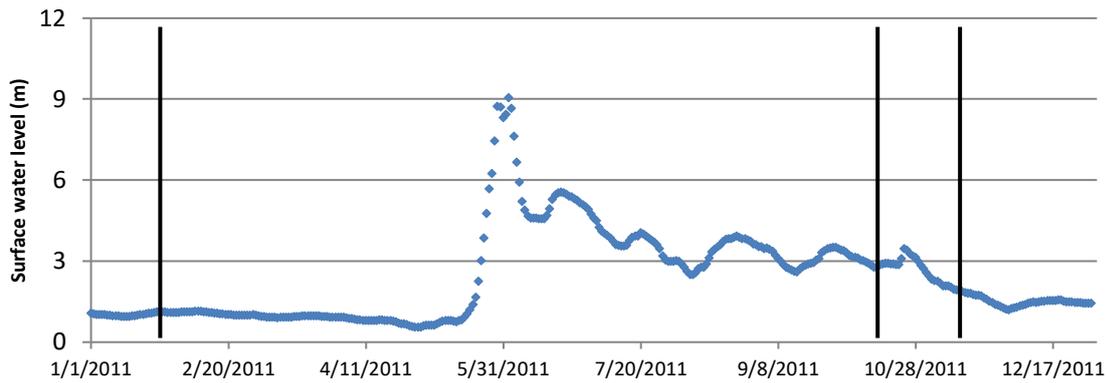


Figure 6 The TDX was generated from TSX/TDX acquisitions (marked in black vertical lines) during low surface water level (recorded at Khabarova station)

### 3.1.1.2. TerraSAR-X Stripmap

Although having the revisit cycle of 11 days (Table 6), the provided TSX Stripmaps for this thesis covered the study area almost on daily basis, with gaps of 2-5 days, depending on the satellite orbit and incidence angle. The availability of a great collection of TSX Stripmaps made it possible to choose specific ones according to need. All of the available TSX Stripmap acquisitions between 2013 to 2019 were observed first for a visual temporal analysis, then 21 TSX Stripmaps were selected according to the intended simulation cases.

Stripmap is a TSX imaging mode in the form of image strips (as shown in Figure 7, among other modes). In this imaging mode, the TSX antenna beam is fixed at a certain elevation and azimuth, illuminating the ground surface with a continuous sequence of pulses. All these TSX image strips were used for multi-temporal analysis, flood mapping, adjusting the gauging station datum, and land cover classification. Depending on the satellite trajectory, these strip images of different surface area may present different intensity values regarding polarizations, incidence angles, and pass direction in result. Table 7 provides an overview of these discrepancies. These discrepancies are important for the multi-temporal analysis and the flood map generations. More details of the effects from these discrepancies will be discussed in 6.5.

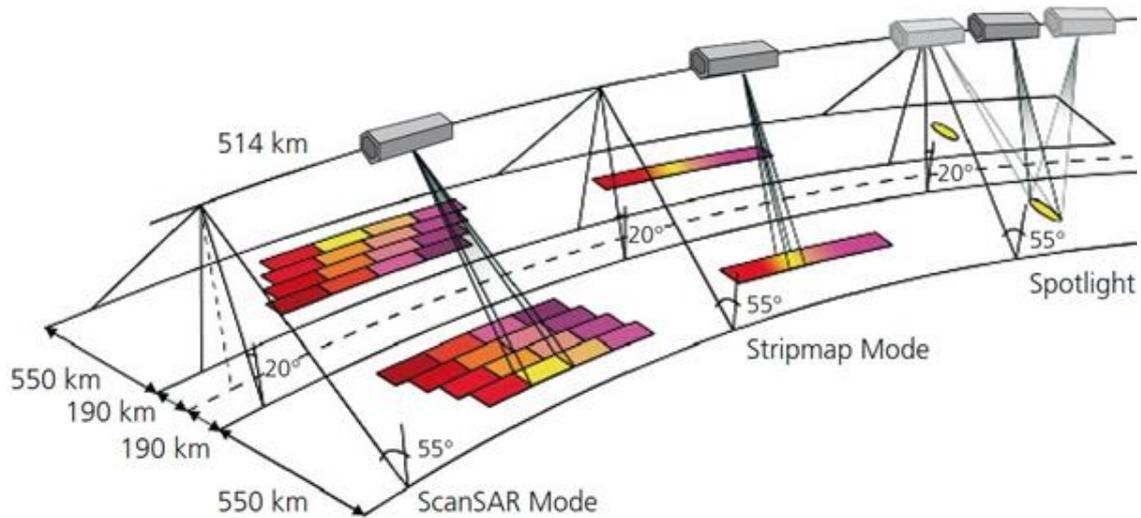


Figure 7 TerraSAR-X image acquisition modes (DLR 2009)

From Table 7, it can be concluded that the polarization, incidence angle, and pass direction were related to its corresponding orbit of a certain acquisition period. Satellite orbits are the trajectory paths that allow the satellites to cover different surface area over a certain period of time, in conjunction with the planet's rotation. Satellite orbits operate along a north-south trajectory whereas the earth rotates from west to east. Polarization is the oscillation direction of the pulse from and into the platform (Figure 8a). Incidence angle is the angle between the incident radar beam to the vertical of the intercepting surface (Figure 8b). Pass direction is the direction of the movement of the satellite. Ascending pass means the satellite travels northwards, whereas descending pass travels towards the southern pole.

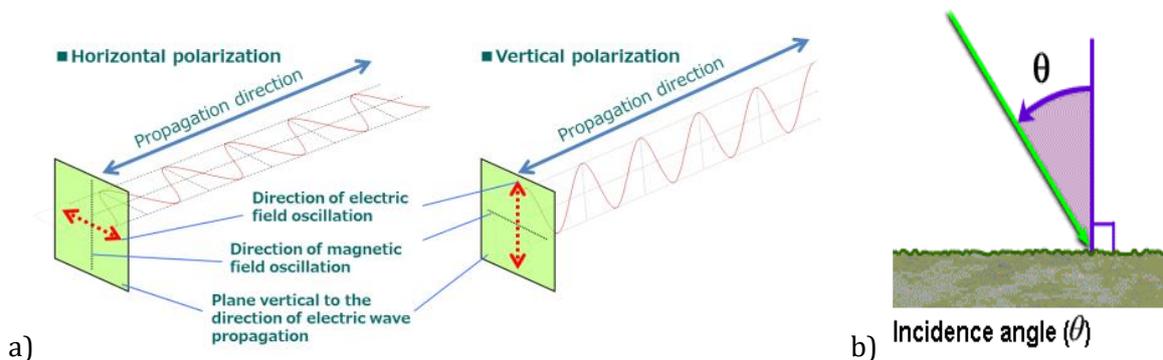
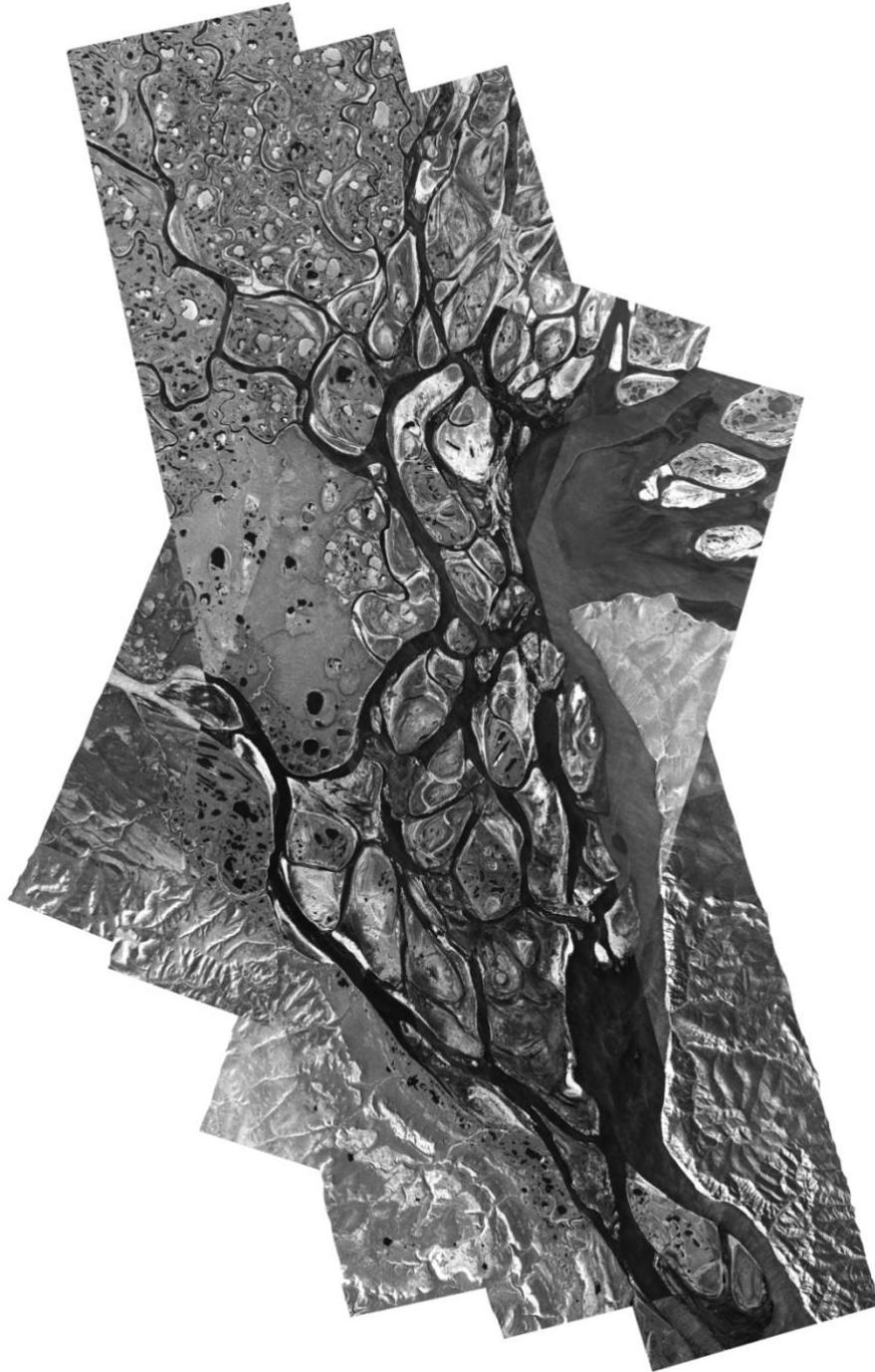


Figure 8 a) Horizontal and vertical polarization (JMA 2016); b) Satellite incidence angle (ESA 2014)

**Table 7 The specification of the utilized TerraSAR images**

No. Event	Date	Orbit	Polarization	Incidence Angle	Pass Direction
1	2013-05-14	65	HHVV	38 - 39	Ascending
2	2013-06-02	28	HHHV	36 - 37	Descending
3	2013-06-11	156	HHVV	33 - 34	Ascending
4	2014-05-09	28	HHHV	36 - 37	Descending
5	2014-05-31	28	HHHV	36 - 37	Descending
6	2014-06-11	28	HHHV	36 - 37	Descending
7	2015-05-07	28	HHHV	36 - 37	Descending
8	2015-06-04	119	HHHV	40 - 42	Descending
9	2015-06-23	65	HHVV	38 - 39	Ascending
10	2016-05-10	119	HHHV	40 - 42	Descending
11	2016-06-01	119	HHHV	40 - 42	Descending
12	2016-06-26	156	HHVV	33 - 34	Ascending
13	2017-05-19	119	HHHV	40 - 42	Descending
14	2017-06-04	28	HHHV	36 - 37	Descending
15	2017-06-29	65	HHVV	38 - 39	Ascending
16	2018-05-17	119	HHHV	40 - 42	Descending
17	2018-06-06	80	HHVV	27 - 29	Ascending
18	2018-06-13	28	HHHV	36 - 37	Descending
19	2019-05-04	119	HHHV	40 - 42	Descending
20	2019-06-04	80	HHVV	27 - 29	Ascending
21	2019-07-01	156	HHVV	33 - 34	Ascending

Figure 9 illustrates a collection of different TSX orbits covering the study area (not necessarily the ones chosen for the study). The Stripmaps leaning to the left are of the ascending pass direction, and the ones leaning to the right are of the descending pass direction. The revisit cycle of the exact same orbit is 11 days, however the study area was covered almost on daily basis due to the other available orbit coverage. This enabled the multi-temporal analysis to observe the development of the flood and the change of the land cover.



**Figure 9** Different coverage of TSX Stripmap orbits on the study area

### **3.1.2. RapidEye Ortho - Level 3A**

RapidEye (RE) is a five-satellite constellation (illustrated in Figure 10) with 5m-resolution imagery (DLR 2008). The single launch of the RE constellation took place on 29 August 2008 (ESA 2019b). RE products were previously operated by the RapidEye AG, a German Geographic Information System (GIS) provider. Now, they are commercialized by the Planet Labs. RE

platforms are equipped with the Jena-Optronik multi-spectral imager. This sensor captures five spectral bands of the electromagnetic spectrum: red, green, blue, red edge, and near-infrared (Table 8). Only a limited RE imageries were provided by AWI, each with different coverage and cloud cover percentage. The image acquired on 9 July 2014 was selected due to the good quality and the least amount of cloud cover percentage it has compared to the others. Since the RE imagery was used for land cover classification and considering that there was little to no change of land cover on the study area, only a single image was used.



**Figure 10 The RapidEye system consists of five satellites (DLR 2008)**

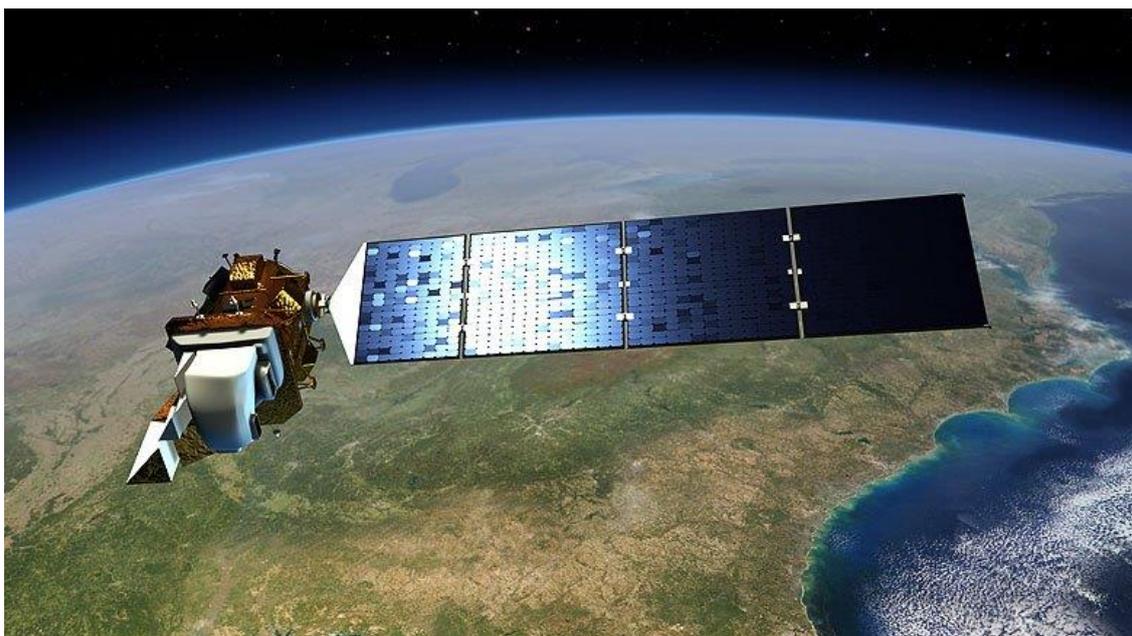
**Table 8 RapidEye bands**

<b>Band</b>	<b>Wavelength (nm)</b>
1 – Blue	440 – 510
2 – Green	520 – 590
3 – Red	630 – 685
4 – Red Edge	690 – 730
5 – Near-Infrared	760 – 850

**3.1.3. Landsat 8 Level 2 – Surface Reflectance**

Landsat 8 (illustrated in Figure 11) is a U.S. Geological Survey (USGS)’s observation satellite equipped with two push-broom instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) (USGS 2019b). LS8 was launched on 11 February 2013. The Level-2 surface reflectance datasets can be requested through the USGS Earth Explorer website

(USGS 2019a) for free and are processed automatically in ~5 hours. This product consists of 7 bands (Table 9). Since LS8 products can be downloaded for free, there were a collection of imageries to choose from. However, following that of the RE imagery selection, the search for the LS8 was narrowed down to the timeframe of around the Summer of 2014. The LS8 acquired on 9 July 2014 was then selected, as it also had a good quality of image and the least cloud cover percentage amongst others.



**Figure 11 Landsat 8 satellite platform (EOS 2019)**

**Table 9 The Landsat 8 L2 surface reflectance bands (Barsi et al. 2014)**

<b>Band</b>	<b>Wavelength (<math>\mu\text{m}</math>)</b>	<b>Useful for mapping</b>
1 – Coastal Aerosol	0.435 - 0.451	Coastal and aerosol studies
2 – Blue	0.452 - 0.512	Bathymetric mapping, distinguishing soil from vegetation, and deciduous from coniferous vegetation
3 – Green	0.533 - 0.590	Emphasizes peak vegetation, which is useful for assessing plant vigor
4 – Red	0.636 - 0.673	Discriminates vegetation slopes
5 – Near-Infrared	0.851 - 0.879	Emphasizes biomass content and shorelines
6 – Short-wave Infrared 1	1.566 - 1.651	Discriminates moisture content of soil and vegetation; penetrates thin clouds
7 – Short-wave Infrared 2	2.107 - 2.294	Improved moisture content of soil and vegetation and thin cloud penetration

## 3.2. Ground-Truth Materials

The available field datasets were gauged surface water level, upstream discharge, and land cover information. In this subsection, each dataset will be described and visualized.

### 3.2.1. Upstream Discharge at Kyusyur

The water runoff dataset was acquired through the Arctic Great Rivers Observatory (ArcticGRO)'s website (Holmes et al. 2018). The average daily water runoff was recorded at Kyusyur station (70.68°N, 127.39°E). The gauged water runoff between 2013 and 2019 is provided in Figure 12.

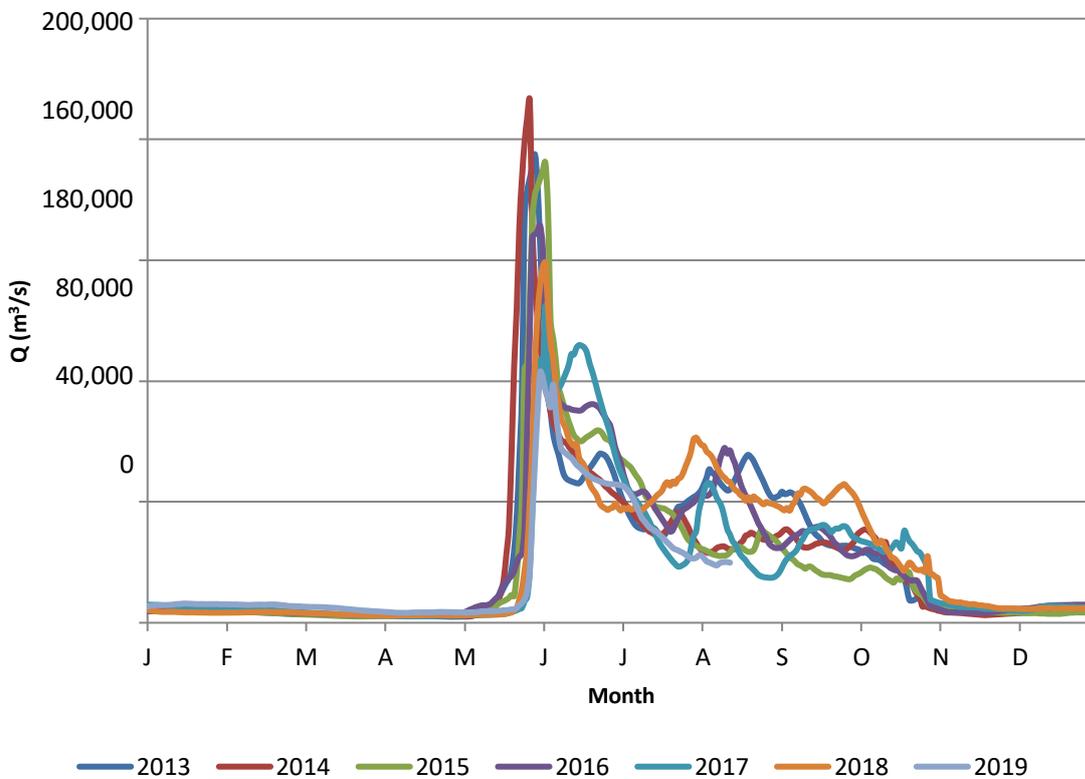


Figure 12 Daily gauged river discharge at the Kussyur station in 2014

### 3.2.2. Surface Water Level at Khabarova

The gauged water level dataset was acquired through the Russian Ministry of Natural Resources and Environment's website (Minprirody 2019). The water level was gauged at Khabarova station, located on the Bykovskaya reach, of which exact coordinate was not specified on the data file. The daily gauged surface water level is provided in Figure 13.

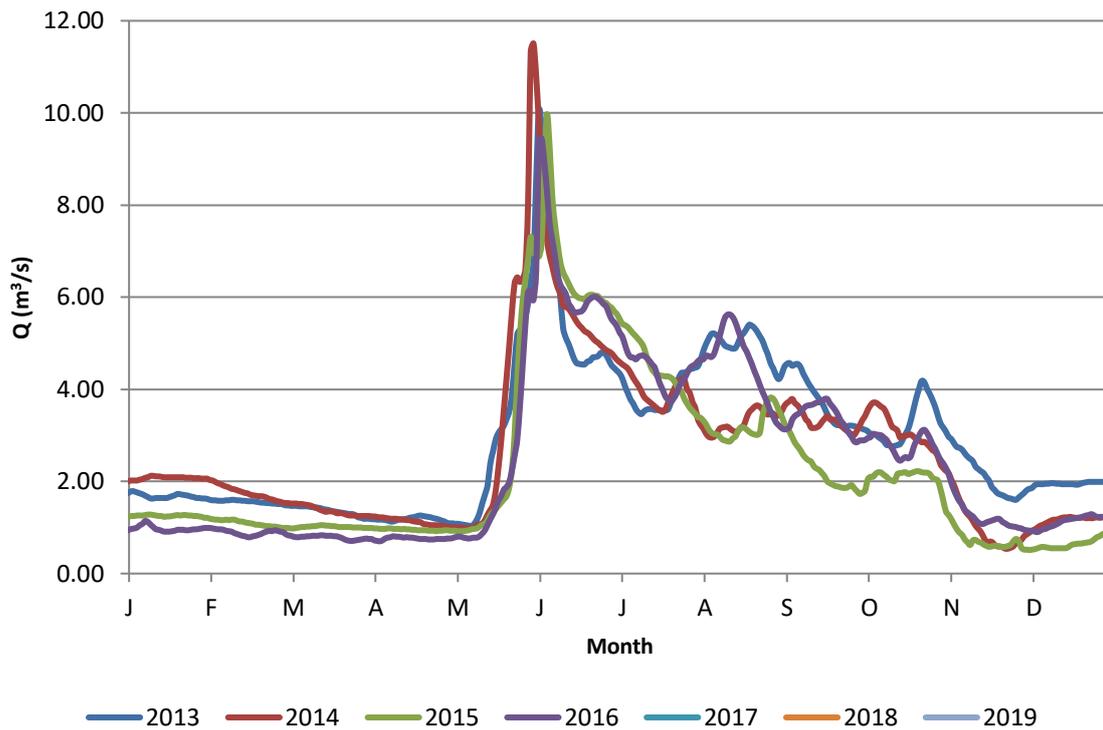


Figure 13 Daily gauged water level at the Khabarova station in 2014

The Khabarova and Kyusyur stations were located ~200,000 km away from each other. Therefore, the flood peak on the 2 stations was achieved on different time, with the average temporal discrepancy of 2 days.

### 3.2.3. Land Cover

The land cover information was provided by the Alfred-Wegener-Institut (AWI) from site visits (Heim 2019). Only a subset of the study area was observed, precisely surrounding the Samoylov island. The land cover information is provided in Figure 14. There are 7 types of land cover with different roughness height, as provided in Table 10.



Figure 14 The land cover information of center-north part of the study area (Heim 2019)

Table 10 Land covers and the corresponding roughness height (Heim 2019)

No.	Land Cover Type	Roughness height
1	Sandy floodplain	50mm
2	Holocene terraces (polygonal tundra)	100-300mm
3	Pleistocene terraces (wet sedge moss tundra)	30-60mm
4	Rock cliffs	N/A
5	Vegetated floodplains	200-400mm
6	Densely vegetated floodplains	800mm
7	Mountain tundra (on the top of the rock cliffs)	N/A

## 4. Remote Sensing Methodology

This chapter provides the step-by-step methodology that was performed to process the remotely-sensed datasets into the useful HEC-RAS parameter input. Information from the TDX Stripmaps, TDX DEM, RE Ortho Level 3A, and LS8 L2 Surface Reflectance were to be extracted and converted into a file format to be loaded into the HEC-RAS GUI. All these steps can be considered as the Remote Sensing methodology part of this thesis. These procedures include:

- Satellite dataset pre-processing
- DEM correction
- Bathymetry approximation
- Flood map generation
- Land cover classification to obtain Manning's surface roughness coefficients
- Converting the GIS data into HEC-RAS input

Each of these processes will be described in more details in the following subchapters.

### 4.1. TerraSAR-X, RapidEye, and Landsat 8 Dataset Pre-processing

In this thesis, the raw TSX data files were not used directly as there were corrections and calibrations to be applied beforehand. The raw TSX datasets were ordered according to need and then undergone three pre-processing steps:

- deriving Kennaugh elements from the raw SAR files
- deriving Sigma Naught (dB)

After the TSX imageries were pre-processed, mean-shift segmentation was applied. Mean-shift segmentation was also applied to the RE and LS8 datasets to reduce the heterogeneity.

#### **Deriving Kennaugh elements from the raw TSX Stripmap files**

This step might be of importance for future work using TerraSAR-X raw files. However, if pre-processed TerraSAR-X imageries in the form of Kennaugh element multiband rasters were already provided, this step is not needed.

Kennaugh elements describe the polarimetric information of the physical scattering mechanism of SAR images, consisting of 10 elements (described in Table 11). TSX Kennaugh elements were derived using the DLR Multi-SAR processor. By applying the Kennaugh matrix decomposition, Kennaugh elements from the TSX dataset are acquired as raster layers.

In this study, the available TSX Stripmaps were of two different polarization: 1) dual-co polarization (HHVV) and 2) dual-cross polarization (HHHV). After converted into Kennaugh elements, each of these dual-co and dual-cross polarization imageries would consist of four bands; total intensity, absorption, diattenuation, and retardance. However, the Kennaugh element of absorption, diattenuation, and retardance of these two polarization configurations varied. The Kennaugh element of the absorption, diattenuation, and retardance for the HHVV imageries were K3, K4, and K7, respectively. The Kennaugh element of the absorption, diattenuation, and retardance for the HHHV imageries were K1, K5, and K8, respectively. This is summarized in Table 12. These 4-band raster outputs were geocoded and radiometrically enhanced. (Schmitt, Wendleder, and Hinz 2015)

**Table 11 TerraSAR-X Kennaugh elements**

<b>Kennaugh Element</b>	<b>Description for</b>	<b>Meaning</b>
K0	Total intensity	Sum of both intensities
K1, K2, K3	Absorption	Intensity ratio (difference) between double bounce and surface intensity
K4, K5, K6	Diattenuation	Change of relation between two amplitude values
K7, K8, K9	Retardance	Phase change between double-bounce and surface scattering

**Table 12 Kennaugh element combinations for HHVV and HHHV polarization**

<b>Type of SAR Polarization</b>	<b>Polarization</b>	<b>Kennaugh elements</b>
Dual-co	HHVV	K0, K3, K4, K7
Dual-cross	HHHV	K0, K1, K5, K8

The needed inputs to process these raw files were the raw SAR file in an archived folder and the setup information as follows: the image polarization input, the desired Kennaugh element output, selected radiometric calibration method, pixel spacing (in UTM meter), and preferred image enhancement. The processor then converted the digital number (DN) into radar brightness ( $\beta^0$ ) by applying the calibration factor ( $k_s$ ), performed Polarimetric decomposition using Kennaugh matrix, performed multilooking/image detection, geocoded the image (ortho-rectification with Range-Doppler geocoding) using the TDX DEM (2011-2014), applied radiometric calibration, removed the incidence angle influences, enhanced the image enhancement by reducing multiplicative noise, and applied Multi-scale Multilooking & Schmittlet image enhancement.

### **Deriving Sigma Naught from the Pre-processed TSX Stripmap**

This process was applied to the 4-band TSX Stripmap raster files that were produced from the previous step. Sigma naught ( $\sigma_0$ ) has the scale of the backscatter values of -1 to 1, derived by normalizing the values of each cell of the raster bands. There were three scaling options: linear, logarithmic, or hyperbolic tangent. This process was performed in order to compress the data without loss and produce comparable data ranges (which is useful for the multi-temporal analysis). The SAR Kennaugh elements were converted into Sigma Naught (dB) in two forms:

- HH & VV (as separate layers) in dB
- Kennaugh elements (as separate layers) in dB

### **Image Segmentation**

In this study, the chosen TSX Stripmaps, RE Ortho Level 3A, and LS8 L2 Surface Reflectance have the resolution of 5 m, 5 m, and 30 m, respectively. High spatial imageries provide information of the heterogeneities of soil at a really fine scale (Aubert et al. 2011). For further processes in this thesis (e.g. LULC classification and flood mapping), the heterogeneities can be regarded as noise. For that reason, the image segmentation was necessary to be applied to these images.

The selected method to perform the segmentation is the mean-shift segmentation. Mean-shift segmentation is a technique that can damp shading or tone differences on localized objects. There are several other segmentation methods, but mean-shift was chosen because it considers the neighboring cells. This works better on reducing heterogeneity in mixed pixel values. ArcGIS starting from the version 10.3 is already equipped with this tool.

In order to perform mean-shift segmentation on ArcGIS, the raster statistics had to be calculated beforehand. Afterwards, the Spatial Analysis' Segment Mean Shift tool was selected. There are two important parameters to be defined to perform this segmentation: (1) spatial detail and (2) spectral detail, both ranging from 1 to 20. The combination of higher spectral detail and lower spatial detail produces the best resulting image. An illustration of combinations of spatial and spectral details is given in Figure 15. The blue box in the upper right shows the high-quality zone. The green box requires the fastest processing time, but the resulting images are only good enough to some extent. The area below the yellow line is best to be avoided as it only produces poor quality images. Therefore, in this study, the selected spatial detail is 5 and spectral detail 20.

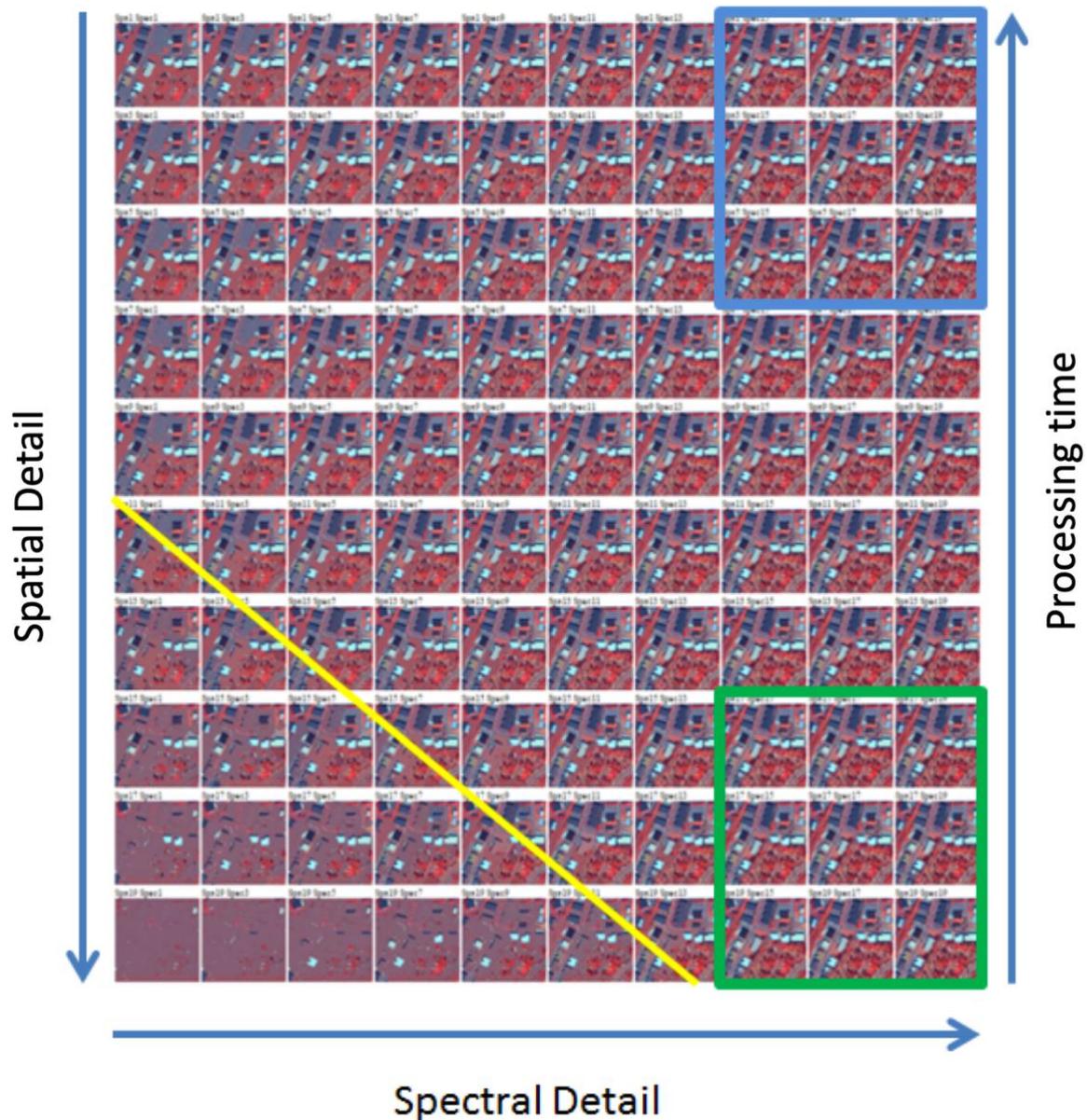


Figure 15 The combinations of defined spatial details and spectral details for mean-shift segmentation (ESRI 2015)

#### 4.2. TanDEM-X DEM Correction

TanDEM-X originally has the spatial resolution of 12.5 m. In order to make it pixel-compatible with TSX products, the DEM was sampled into 5 m. Kriging interpolation was performed to fill the gaps between the raster pixels. The TDX DEM that was provided by the DLR was already filtered and corrected, however the surface elevation still looks very rough at some area. Phase noises were produced while transforming the TDX signal to elevations. These noises with the range of 1-2 m were already filtered out. However, the rough surface values might be the results from the vegetation, rocks, or frozen sand on top of the land surface. Therefore, these values were taken as they were for the simulation.

With 5m-resolution, TDX produces a detailed and relatively accurate (post-processed) DEM. However, while observing the topography information retrieved from the DEM, an error in the acquisition was found. In the DEM, a rock island on the north part of the study area has a deep crater up to the depth of -60m with the direct link to the water body. Using the Google Earth Elevation Profile, it was confirmed that the island has a rather flat and high terrain. This error might affect the hydraulic modeling, as there are possibilities that water will go to the crater. The crater had to be filled. An Inverse Distance Weighted (IDW) interpolation was performed using point data of the max to mid elevation of the island in order to fill the inexistent crater. The before and after of this correction is illustrated in Figure 16.

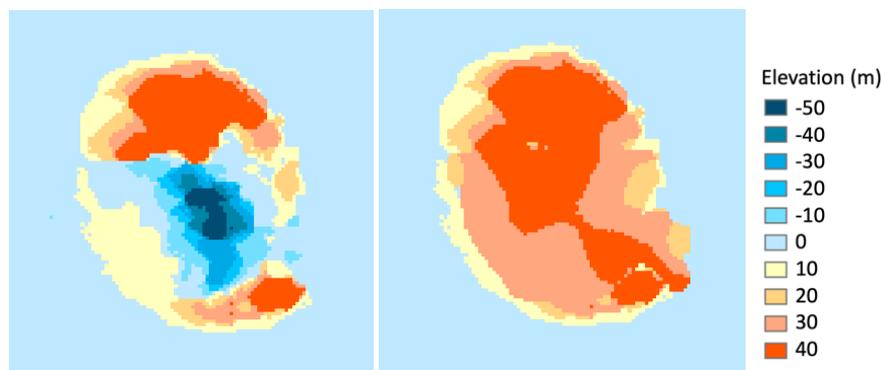


Figure 16 The rock island before and after the interpolation

### 4.3. Bathymetry Approximation

There are two well-known methods to approximate ungauged river bathymetry:

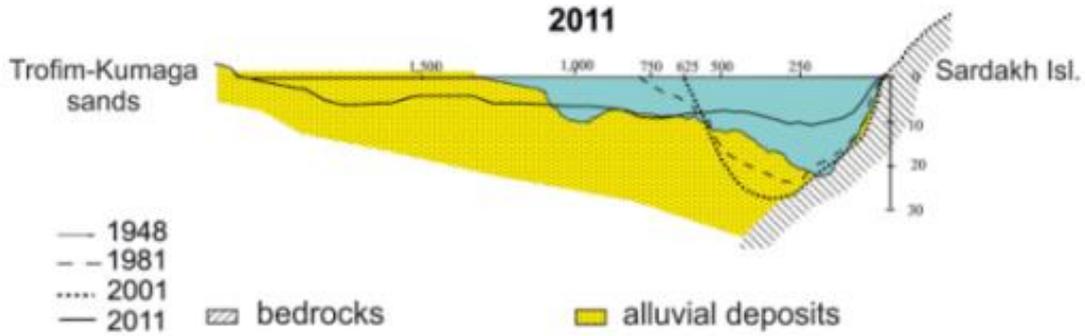
- Empirical approaches using Channel Bankfull and Slope Break approaches (Mersel et al. 2013) (Domeneghetti 2016) based on (Leopold, Maddock, and Thomas 1953)
- Remote sensing approach with the Satellite Derived Bathymetry (SDB) method, involving optical satellite imagery (Stumpf and Holderied 2003)

However, these methods were not fully applicable to this study due to these circumstances:

- Both methods require several cross-section channel profiles in order to calibrate the empirical equation
- SDB method is only applicable on clear water body

The bathymetry information was not available. There is one cross-section information from a publication by (Fedorova et al. 2015) close to the the study area, between the Sardakh and Trofimovskaya island (Figure 17). However, the exact coordinates of this cross-section were not specified and one cross-section profile is not sufficient to represent the bathymetry of all

channels. The reach between these two islands were noted in orange polygon in Figure 1, outside of the simulated area.



**Figure 17** Riverbed changes of between Sardakh and Trofimovskaya islands between 1948-2011. Units are in meters (Fedorova et al. 2015)

Due to the availability of a high-resolution optical imagery (RE), an attempt to approximate the channel bathymetry using the SDB method was performed. The relative depth  $\left(\frac{\ln(nR_w(\lambda_i))}{\ln(nR_w(\lambda_j))}\right)$  was derived using the green and blue band of the RapidEye datasets according to the method developed by Stumpf and Holderied (Stumpf and Holderied 2003). The empirical formula is as follows:

$$Z = m_1 \frac{\ln(nR_w(\lambda_i))}{\ln(nR_w(\lambda_j))} - m_0 \quad (4.1)$$

Where,

$Z$  : the actual satellite derived water depth

$m_1$  : a tunable constant to scale the ratio to depth

$n$  : a fixed constant for all areas

$R_w(\lambda_i)$  : the reflectance of the water of the band with lower absorption (blue band)

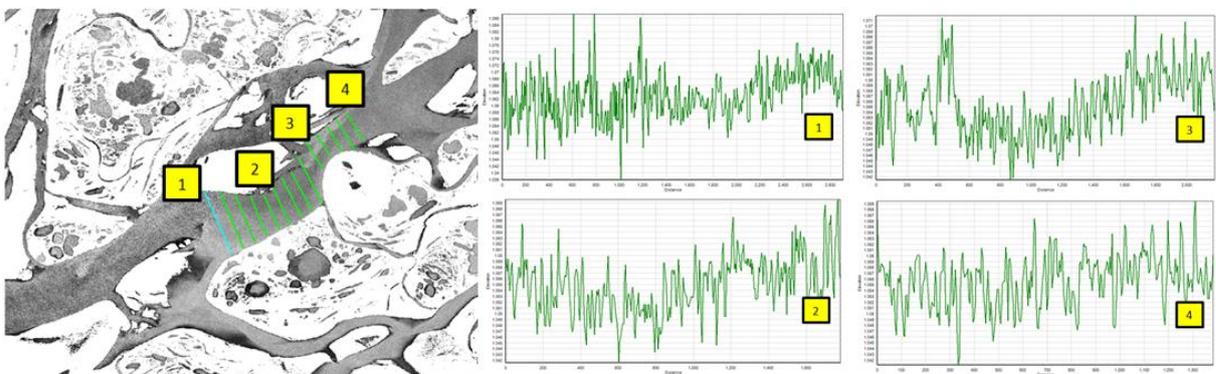
$R_w(\lambda_j)$  : the reflectance of the water of the band with higher absorption (green band)

$m_0$  : the offset for a depth of 0 m ( $Z = 0$ )



**Figure 18 Lena river near Kyusur by Max Holmes (ArcticGRO 2003)**

The results were not satisfying, as the resulting relative bathymetry was noisy (Figure 19 is showing four cross-section channel profile between the Sardakh and Trofimovskaya island) and didn't resemble the shape of the same reach channel provided in Figure 17. SDB method depends on the reflectance values from optical satellite imagery (green/yellow and blue bands), making it only applicable on clear water body. Based on the site visit by AWI, the surface water on the study area is highly turbid (Figure 18). This explains the noisy result of the relative bathymetry. The reflectance values represent the reflected wave by the particles on the water instead of the riverbed.



**Figure 19 The noisy cross-section profiles derived from an attempt to approximate the channel profile using SDB method**

Therefore, the channel bathymetry was approximated by using an assumption that the channels have a trapezoidal shape. The real topography of river channels generally has a complex anisotropy. Oftentimes, the anisotropy is mainly caused by the hydrodynamic and transport mechanisms by the river flow (Zhang et al. 2016). There are several ways to interpolate bathymetry. Four of which are mentioned and assessed in Curtarelli's study (Curtarelli et al. 2015). Inverse Distance Weighing (IDW) that is available in ArcGIS was selected for this study.

From the known cross-section (Figure 17), it can be concluded that the most recent maximum depth of the channels was around -20 m. Most of the obtained river elevation on the DEM was around -8 m, with some parts ranging to -20 m. Given the uncertainties of this approach, 5 bathymetry raster files with different maximum depth (-30, -25, -20, -15, -10 m) were created. These bathymetry files then were used as HEC-RAS geometry input and the simulation was run to see which one resulted the most fitting result.

The smaller channels on the floodplain in the middle of the main channels were kept as it is, assuming that the elevations of these small channels were already correct. In the winter the water in this area freezes over, which is apparent through the TSX data. Due to the shallow depth of these channels, TDX signal might have penetrated through the ice and then scattered back by the riverbed. The resulting bathymetry will be mentioned as *the synthetic bathymetry* in order to avoid ambiguity.

The IDW tool that is available in ArcGIS requires the input of point features with a field containing the information of the channel depth. Given that this method depends on the distances between those points, the points have to be generated continuously on a close distance on the same contour line. In order to create this, the contour lines of different channel depth (-2 m, -3 m, -4 m, -5 m, and the maximum depth) were digitized manually along the channel in a way that will result in trapezoidal-shaped channel (Figure 20). The depth of each contour line had to be declared in a field.

After these contour lines were digitized, they were converted into point features using the QGIS QChainage tool with the increment of 5 m. In order to make the interpolated channel depth connected to the channels that were not interpolated, the depth of the edges of the interpolation boundary was extracted from the DEM into vector points. After these feature points are merged, then these points were interpolated using the ArcGIS IDW tool in segments. The interpolation had to be performed in segments in order to reduce the computing time. Since the tool interpolates the feature points into a rectangular boundary, the segments had to

be made as lean as possible. After the interpolation processes, these interpolated segments were mosaicked together and clipped to the channel mask, and then were stitched onto the DEM.

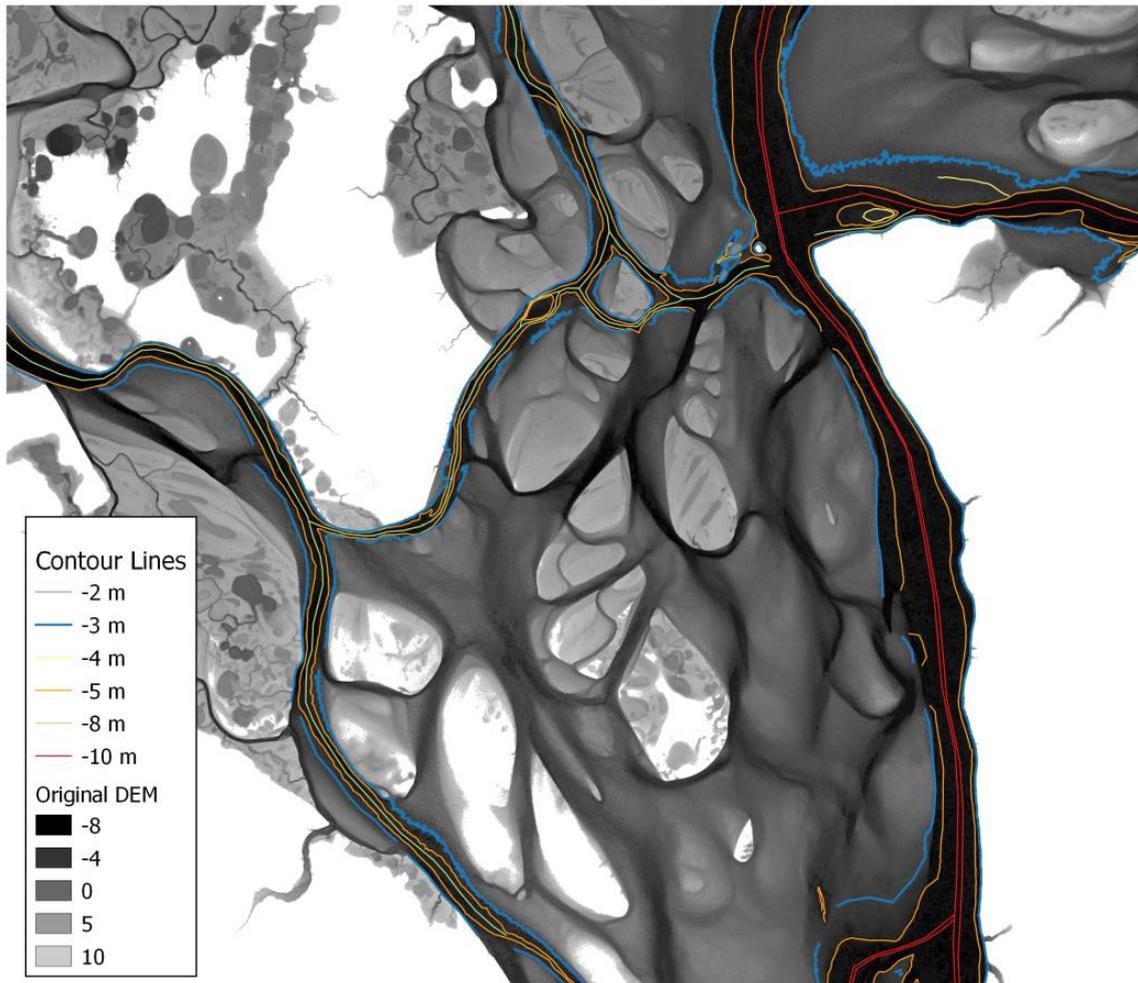


Figure 20 The generated bathymetry contour lines

#### 4.4. Flood Mapping

The aim of flood mapping is to produce the shape of the inundation boundary. The inundation boundaries were produced as binary raster files that would be used to validate the resulting inundation boundary from the hydraulic simulation. The raster files consisted of 0 and 1 value, in which 1 indicates the water body and the flooded area, whereas 0 the land surface that is not flooded.

Flood maps were commonly created by classifying the imagery into two classes: water and non-water. However, the pixel-based land cover classification for this purpose were not suitable for the selected TSX Stripmaps. The water and non-water cover of some of the images

were not easily numerically distinguishable. On some of the images, the water and the land looked similar, or the water had ripples that created a pattern. These impurities were caused by:

- **Wind;** wind induced ripples on the water surface, making a striped pattern on the water body
- **Low incidental angle;** images retrieved on low incidental angle were more sensitive to surface roughness, and other weather disturbances
- **Floating ice;** floating ice and ice cover on the land might have similar backscatter, making them undistinguishable
- **Different polarizations;** HHHV polarization had higher sensitivity to the moisture content; land surface with high moisture content might look similar to the water body

The approach in this study was to generate the flood maps from the TDX DEM based on eyesight and image-fitting. The DEM was flooded starting from the elevation of -6 m to 6 m with the increment of 25 cm. This range of elevation was all the needed threshold to cover all the observed flood event. By using this approach, these assumptions were taken:

- Flat water surface along the channel from the upstream to the downstream
- The DEM doesn't change during the whole simulation period

The flood maps were generated in Python programming language. The first step was to define the flooding threshold (the flat surface water elevation of the flood). The selected threshold was between -6 m to 6 m with the increment of 0.25 cm. Then, the GDAL library was used to open the DEM. After the DEM was open, the DEM was converted into an array. Sieve filtering then was applied into the array, to filter the number that is below-and-equal and above the threshold. The array cells with values of below-and-equal to the threshold were copied into a new array with the cell values of 1. The array cells with values of above the threshold were given the cell values of 0 on the copied array. The binary array then was converted into a raster file.

After a collection of flood maps with varying threshold were produced, these flood maps were fitted into a single acquisition TSX image. The flood map that visually fitted a TSX image of an acquisition day was selected as the inundation boundary of the corresponding day. At the end of the process, flood maps of each of the observed events were acquired. These flood maps would later be used to assess the accuracy of the hydraulic simulation. Figure 21 shows a

snippet of a) the DEM, b) a TSX Stripmap on top of the DEM, and c) the flood map that fits the TSX Stripmap.

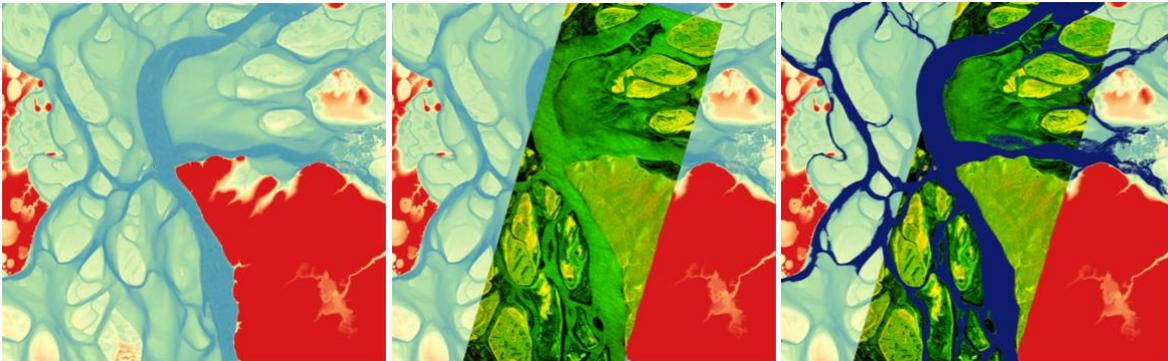


Figure 21 The flood map fitting process: a) the DEM, b) a TSX Stripmap on top of the DEM, c) fitting the DEM flood map from the collections of flood map into the TSX Stripmap

#### 4.5. Manning's Surface Roughness Coefficient

It is difficult to monitor and model soil surface characteristic because of their temporal and spatial variation. Therefore, remote-sensed data can be used for this purpose as well.

Two ways to derive the Manning's roughness coefficient from TSX products:

- Land cover classification (Bachofer et al. 2015)
- Directly retrieving the roughness height with empirical equation (Aubert et al. 2011, Baghdadi et al. 2011, Sadeh et al. 2018)

The latter requires in-situ soil moisture measurements, which wasn't possible to be carried out during the period of this study. Therefore, the surface roughness coefficients were based on the corresponding land cover, which were to be derived through classification methods.

2 consecutive Machine Learning methods were performed:

- **Iso-clustering** to define the classes for training datasets  
This is an unsupervised Machine Learning method. Iso-clustering was performed to separate the classes. The resulting separate classes were used as the training input for the next supervisor Machine Learning method. An iso-clustering tool with GUI was already available in ArcGIS. Iso-clustering is a modified iterative optimization clustering procedure. This method is also known as the migrating means technique. With this method, the cells are separated into the user-specified number of groups in the multidimensional space of the input bands (in image classification, bands function as features)

(ESRI 2016). This allows the “classifier” to separate the classes based on the values of the raster cells without training datasets. An image illustrating iso-clustering with only two features is provided in Figure 22. For this thesis, the features are more than two, as the bands of each satellite imageries also exceed that number.

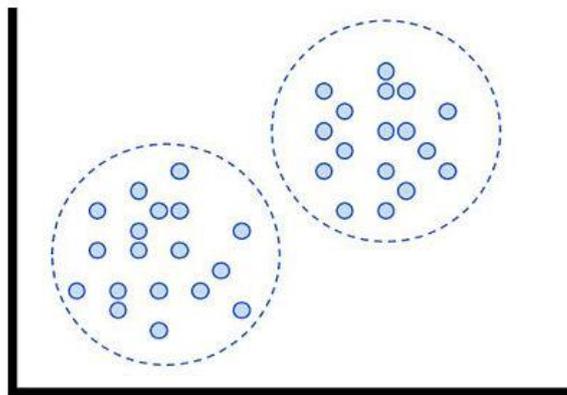


Figure 22 Iso-clustering of point data with two features (each on separate axis) (Klie 2017)

- **Random Forest** to classify the whole raster bands

Random Forest is a supervised Machine Learning method. The input to train this supervised classification method was generated from the previous unsupervised method (iso-clustering). Random Forest is an ensemble of decision trees of which the final class result will be based on majority voting. Random Forest classification was performed using the Python `scikit-learn` library, which is a well-known Python Machine Learning library. The supervised method requires training datasets. The classes for the training datasets were already defined from the previous iso-clustering classification. Figure 23 shows the simplified Random Forest classification scheme.

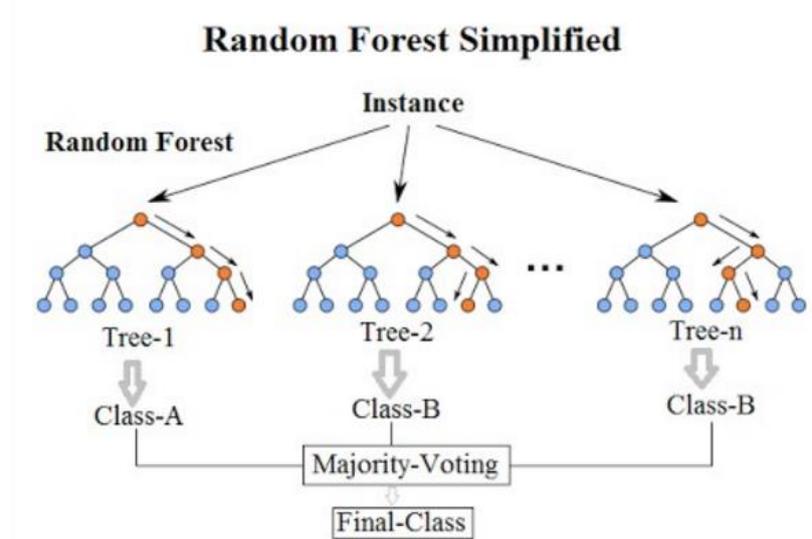


Figure 23 Random Forest classification scheme (Koehrsen 2017)

Land Use/Land Cover (LULC) classifications can be built by two imaging methods, i.e. optical and microwave remote sensing. Optical sensors rely on the surface reflectance on cloud-free conditions, whereas microwave sensors only capture the surface backscatter in a specific range of wavelength (Hütt et al. 2016). SAR techniques are useful for the land cover classification. SAR images provide information regarding the soil parameters under any weather conditions. Therefore, in this study, both optical satellite imagery (RE, LS8) and SAR imagery (TSX) were used for the land cover classification.

On this study, the training datasets were built based on the RE datasets, of which patterns fit to the land cover information provided by (Heim 2019). Then the feature importance (Variable Importance Measure, VIM) of both the TSX and RE datasets was analyzed in order to see which bands are more useful for the Random Forest classifier, and also to see whether or not the TSX and RE dataset were on agreement.

Images acquired on different time have different backscatter/reflectance values and value distribution, therefore the land cover classification is performed on each TSX acquisition tile separately. There are 2 raster band combinations of which the land cover classification was applied:

- Composite TSX-RE bands
- LS8 bands for a bigger extent that is not covered by the available TSX and RE datasets

The land cover classification in this study was performed through these steps:

- Iso-clustering the RapidEye image to create the training datasets

- Sorting and labeling the iso-cluster classes
- Extracting stratified sample points from the iso-cluster classes
- Dividing the sample points into training sample points (70%) and test sample points (30%) (Figure 24)

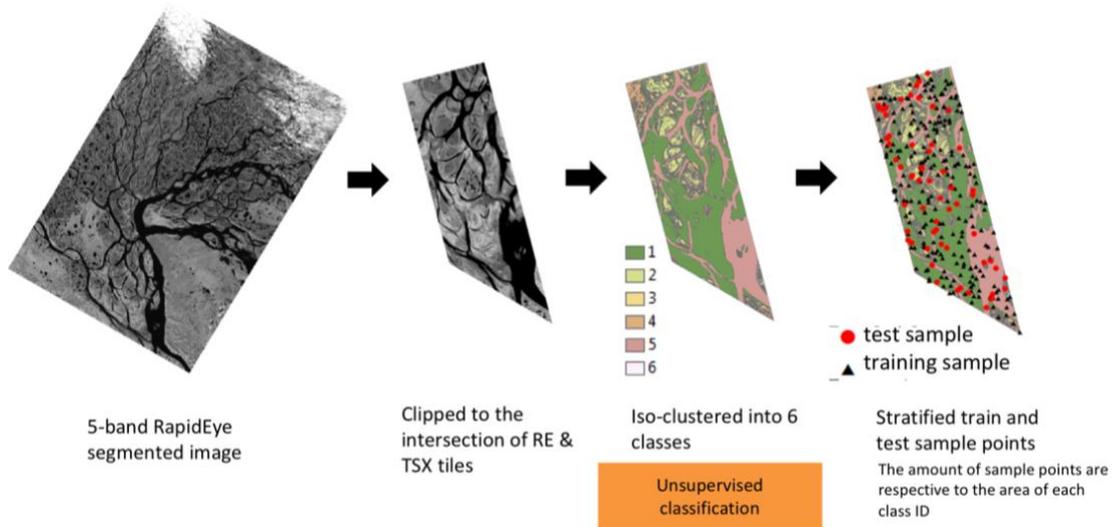


Figure 24 Creating training and test samples

- Preparing the band composites
- Reading the raster band composites as arrays
- Extracting the training classes and raster values (Figure 25)
- Extracting the test values (Figure 25)

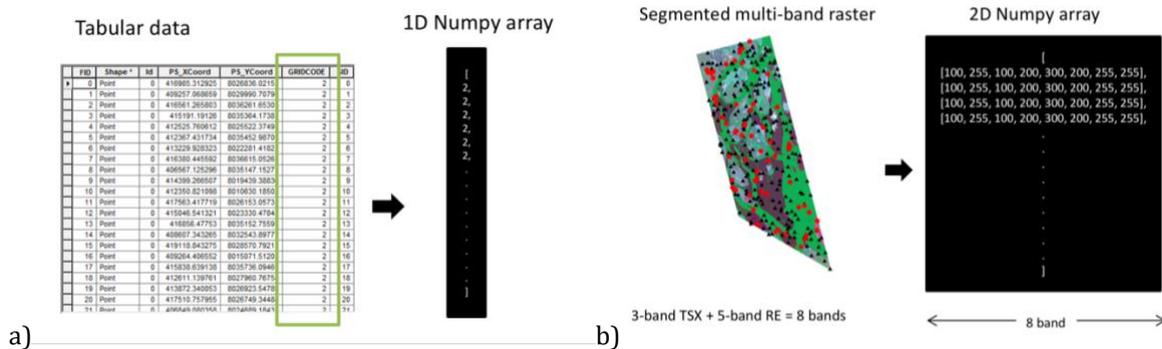


Figure 25 Extracting the classes and raster values for the training and test datasets, a) reading the class ID from sample point; b) taking out raster values on the sample points

- Training the Random Forest and calculating the VIM
- Classifying test samples and analyzing the confusion matrix (Figure 26)

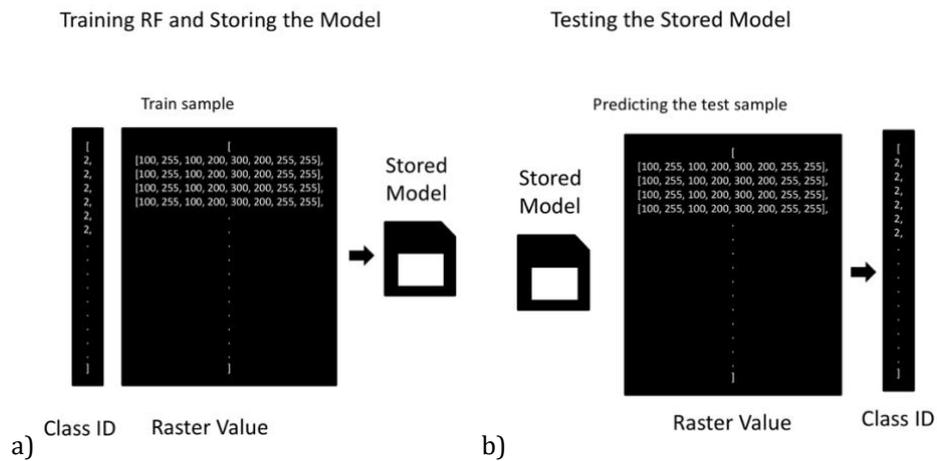


Figure 26 a) training the Random Forest classifier; b) testing the Random Forest classifier

- Classifying the whole raster image (Figure 27)
- Converting the classified arrays into raster files
- Image post-processing

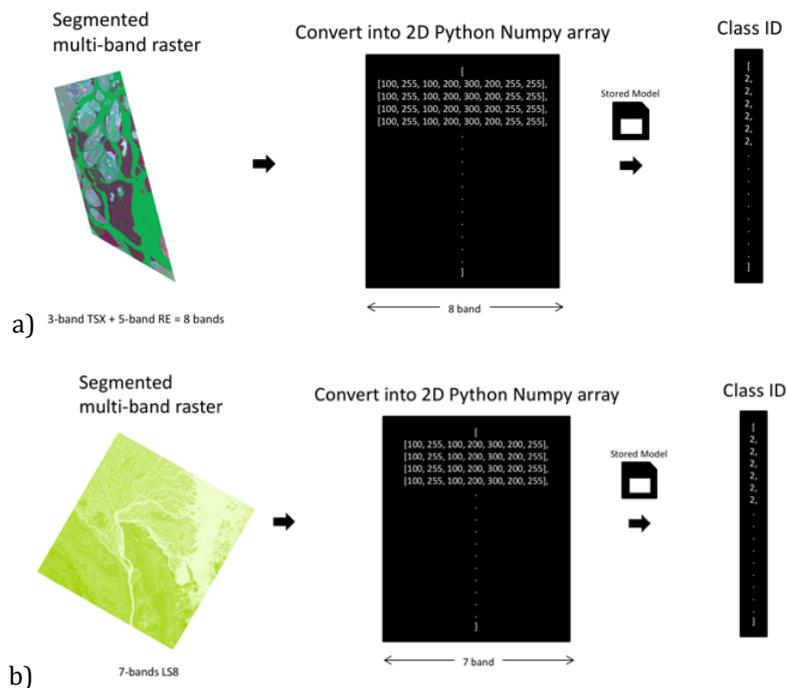


Figure 27 Classifying the a) TSX-RE Composite and b) LS 8 bands

The number of Manning's value input for 1D simulation is limited. Each cross-section line can only have up to 20 n-value. Therefore, the resulting land cover map were simplified by rounding up the nearby land covers. In order to simplify the land cover map, QGIS snakes vector generalization and polygon cleaning tools were used.

## 4.6. Converting the GIS Data into HEC-RAS Input

This subchapter consists of the method that bridges the Remote Sensing methodology part and the Hydraulic Simulation part of this thesis. This methodology is important as all the processed remotely-sensed datasets cannot be inputted into HEC-RAS directly. All the information from the processed remotely-sensed datasets had to be converted into a format that is recognized by HEC-RAS.

The main idea of incorporating the remote sensing data for HEC-RAS flow simulation is to extract the values of geometry and other required modeling input into geometry lines that are compatible with HEC-RAS. The official tool developed by HEC-RAS team to create this input is the HEC-GeoRAS ArcGIS extension. There are other approaches developed based on how HEC-GeoRAS, one of which is RiverGIS QGIS extension. The RiverGIS QGIS extension was selected for this study for its versatility and compatibility with the PostgreSQL/PostGIS database, making it easier to modify and work with. Moreover, the official tested HEC-GeoRAS was only compatible with ArcGIS up to the 10.2 version, which is an old version released in 2013. During working on this thesis, many unsolved bugs were found within the aforementioned version of ArcGIS. On the other hand, RiverGIS is compatible with new QGIS versions (3.x).

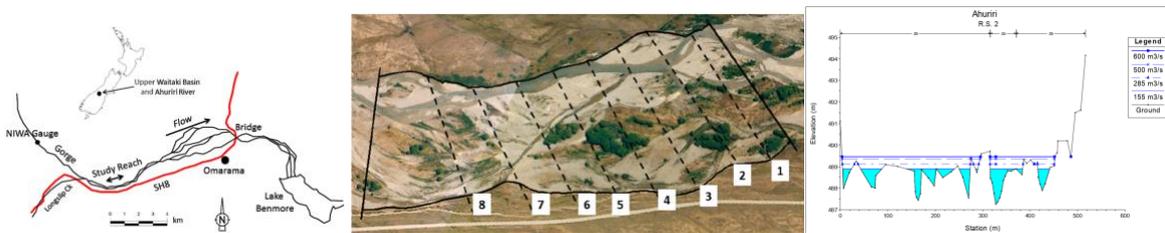
In order to use the RiverGIS QGIS extension, pgAdmin 4 (PostgreSQL GUI) with PostGIS bundle and QGIS and RiverGIS extension had to be installed. The pgAdmin 4 PostgreSQL GUI was utilized to create the database that would be where the extracted geometry values were to be stored in. The database then was connected through QGIS. In QGIS, the geometry lines were digitized manually and then used to extract the geometry values from the remote sensing datasets using the RiverGIS extension. The extracted geometry values then were stored in the database that was set up using pgAdmin 4. The last step is to create the HEC-RAS GIS geometry file (.SDF file) that contains the geometry values from the database using RiverGIS. These processes are elaborated in details in the following subsections.

### **The Geometrical Approach**

The Lena delta consists of a anastomosing river system. By observing the multi-temporal TSX images from 2013 to 2019 and running preliminary simulations, it was concluded that some parts of the reaches other than the main Lena reach were dry when there was not enough water runoff to fill them. HEC-RAS cannot simulate dry channels. In the preliminary simulations, an approach to model each of these small reaches were already tested. These setups resulted in model instability when some of these reaches were dried.

Another geometrical constraint for this simulation is that cross-sectional lines should start and end at the highest level of the terrain. The topography of the floodplain between these reaches was low. For this reason, it would be difficult to place the end of the cross-section lines.

The solution to these constraints is by applying an approach that was used by Caruso et al in their publication (Caruso et al. 2013). Working on a braided river system as well, Caruso et al regarded all the reaches as one big channel. Caruso et al made wide cross-section lines that include all the braided-river reaches perpendicular to the flow direction. Their approach can be seen in Figure 28.



**Figure 28 A 1D HEC-RAS geometry of a braided river system (Caruso et al. 2013): a) the shape of the braided river system, the red line is the stream center line; b) the cross-section lines; c) the profile of a cross-section line**

The same approach thus was implemented in generating the HEC-RAS geometry lines in this study. However, the anastomosing river system of the Lena delta indeed has much more complexity the braided river system than what is shown in Figure 28. However, since an attempt to draw the reaches on by one resulting in instability, Caruso et al. (2013)'s approach was applied to this thesis as well. Three reaches were digitized, connected with a junction. These reaches are the upstream Lena, downstream Lena, and Bykovskaya. Bykovskaya has the flow direction that is horizontally perpendicular to Lena, therefore were modeled as its own reach. The digitized geometry lines are shown in Figure 29 on the next subsection.

### Setting Up the PostgreSQL Geodatabase

A geodatabase is needed to store the hydraulic input parameters in the form for GIS datasets. With a geodatabase, it is easier to extract, update, erase, and modify the input parameters. RiverGIS, the tool that would be utilized to extract these parameters, is already integrated with PostgreSQL with PostGIS extension. These are the steps to set up a local geodatabase for RiverGIS:

- Installing PostgreSQL 10.9.2 x64 with PostGIS pgAdmin 10x64 bundle
- Setting up the superuser password and port
- Creating "rivergis" local database
- Creating database schema

Given that one database can nest many schemas, it is possible to have many different projects in the same rivergis local database. This makes working with the geometry more flexible, as it is possible to create different geometry scenarios stored on different database schema as well.

### **Digitizing the Geometry Lines**

Up to the time at which these processes were performed, there was no automatic approach to generate geometry lines that would follow the HEC-RAS 1D flow analysis prerequisites. The only automated cross-section profile generator does not work well for rivers that are not straight, as the resulting lines would overlap each other on meandering rivers and also would only be perpendicular to the center flow path, not to the overbank flow paths. The cross-section lines have to be perpendicular to any given flow path (Goodell 2012). The geometry lines then were digitized manually in order to meet these requirements.

These following geometry lines were digitized:

- Three river stream centerlines, digitized at the deepest part of the river reaches (Lena upstream, Lena downstream, and Bykovskaya)
- The bank lines, digitized at the depth of -5 m (in order to create precise lines, the digitization was snapped onto the -5 m contour lines)
- The center flow path lines, duplicated from the river stream centerlines
- The left and right overbank flow path lines, digitized near to the border of the simulated channel
- The cross-section profile lines, digitized along the floodplain, bent (dog-legged) in such a way to be kept perpendicular to the center, left overbank, and right overbank flow path lines (Goodell 2012, Brunner and CEIWR-HEC 2016)

This geometry setup consists of 2 main reaches (upstream and downstream Lena) and 1 tributary reach (Bykovskaya). These reaches were connected with one junction. The nodes (including the junction) and endpoints were generated automatically when the GIS data was imported to the geodatabase.

Figure 29 shows the digitized geometry lines. The calculation of the spacing between cross-sectional lines will be discussed in 5.3. The nodes and endpoints were to be generated automatically by the RiverGIS extensions after the geometry was imported into the database.

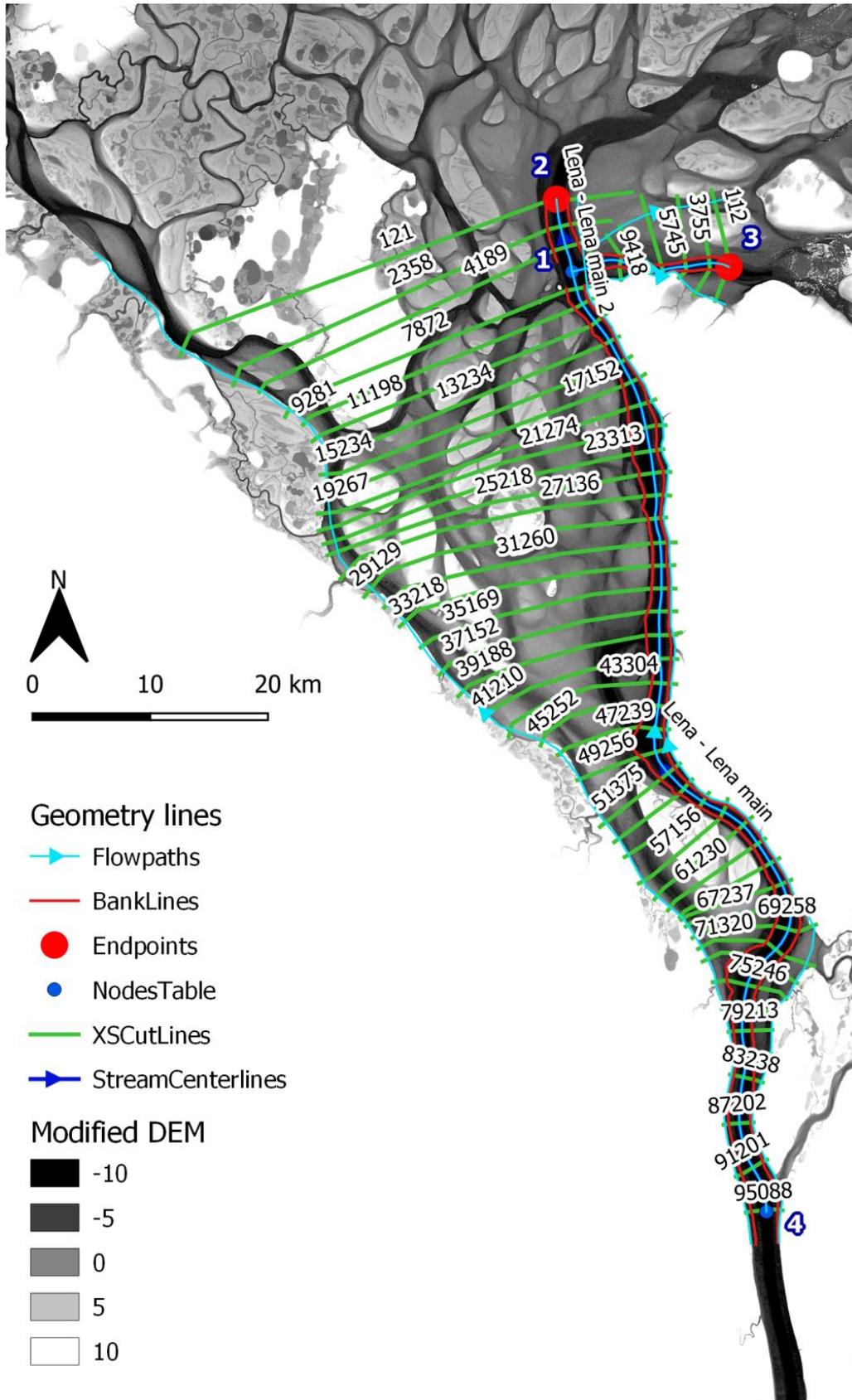


Figure 29 The generated HEC-RAS geometry lines in QGIS

## **Importing the GIS data to the Geodatabase**

Given the geometry and surface roughness values for the 1D simulation can only be useful for the hydraulic simulation in the form of geometry lines, the information than had to be extracted to those lines. QGIS RiverGIS extension was utilized for that purpose. The steps are:

- Connecting QGIS to the rivergis local database with credentials
- Creating the database tables
- Populating the tables using the digitized geometry lines (shapefiles)
- Filling the user-defined (noted in [UD] in the Table 13) attributes
- Extracting the geometry attributes onto the tables

Table 13 shows all the attributes that were extracted from the GIS datasets. It is important to double-check the values of each attributes on the geodatabase tables to make sure no information is missing. Missing values can be caused by mistakes made when generating the geometry lines, such as disconnected lines, bank lines not intersecting with the cross-sectional lines, flowpaths not generated one-directional, and other possible mistakes.

When extracting the DEM elevation into the cross-sectional lines, RiverGIS will convert those lines into points with the distance of the resolution of the DEM, each. Therefore, the number of extracted elevation depends on the length of the cross-sectional line and the resolution of the DEM. The bank lines will be also only extracted in the form of points, on the intersection of the bank lines and the cross-sectional lines. RiverGIS will automatically calculate the distance between these points to fill in the downstream reach length information.

**Table 13 Geodatabase HEC-RAS geometry table attributes**

No.	Table	Geometry	Type	Attribute
1	StreamCenterlines	The river stream centerlines	Polyline	<ul style="list-style-type: none"> <li>- ReachID [UD]</li> <li>- RiverCode [UD]</li> <li>- ReachCode</li> <li>- FromNode</li> <li>- ToNode</li> <li>- ReachLength</li> <li>- FromSta</li> <li>- ToSta</li> </ul>
2	XSCutlines	The cross-section profile lines	Polyline	<ul style="list-style-type: none"> <li>- ReachID</li> <li>- Station</li> <li>- RiverCode</li> <li>- ReachCode</li> <li>- LeftBank</li> <li>- RightBank</li> <li>- LLength</li> <li>- ChLength</li> <li>- RLength</li> <li>- DtmID</li> </ul>
3	BankLines	The channel left and right bank lines	Polyline	BankID
4	Flowpaths	The channel flow direction and the left and right overbank flow direction	Polyline	<ul style="list-style-type: none"> <li>- FpID (flowpath ID)</li> <li>- LineType (Channel, Left, or Right) [UD]</li> </ul>
5	LanduseAreas	The land use (/cover) map	Polygon	<ul style="list-style-type: none"> <li>- LUID (the land use ID)</li> <li>- LUCode (the land use name) [UD]</li> <li>- N_Value (the Manning's n value) [UD]</li> </ul>
6	NodesTable	The river nodes (upstream and joints)	Point	<ul style="list-style-type: none"> <li>- NodeID</li> <li>- X and Y Coordinates</li> </ul>
7	Endpoints	The downstream river node	Point	<ul style="list-style-type: none"> <li>- EndID</li> <li>- RiverCode</li> <li>- ReachCode</li> <li>- NodeID</li> </ul>
8	XSSurface	The XSCutlines with extracted elevations from the DEM	3D Polyline	<ul style="list-style-type: none"> <li>- PtID</li> <li>- XsecID</li> <li>- Station</li> <li>- Elevation</li> </ul>

## Generating the HEC-RAS SDF Input File

This last step was carried out to write the information stored into the geodatabase tables (Table 13) into a GIS-based file format that is recognized by HEC-RAS. This file format is SDF. SDF file consists of all the geocoded information from the previously mentioned geodatabase tables in a text format (as shown in Figure 30). The SDF file was generated with the RiverGIS extension, using the “Create RAS GIS Import File” tool.

```
#This file is generated by RiverGIS, a QGIS plugin (http://rivergis.com)
BEGIN HEADER:
  DTM TYPE: GRID
  DTM:
  STREAM LAYER: rivergis@localhost/mainchannel2km/StreamCenterlines
  NUMBER OF REACHES: 3
  CROSS-SECTION LAYER: rivergis@localhost/mainchannel2km/XSCutLines
  NUMBER OF CROSS-SECTIONS: 55
  MAP PROJECTION:
  PROJECTION_ZONE:
  DATUM:
  VERTICAL DATUM:
  BEGIN SPATIAL EXTENT:
    XMIN: 388750.509403801
    YMIN: 7955412.78387658
    XMAX: 442771.645517239
    YMAX: 8042954.57709528
  END SPATIAL EXTENT:
  UNITS: METERS
  HEADERS:
    439944.954403096, 7955601.47004146, 29.63
    439945.572365929, 7955601.34201729, 29.63
  END:
```

Figure 30 A snippet of the beginning and end of the SDF file

## 5. Hydraulic Simulation

After the simulation input was processed from the remotely-sensed datasets, the methodology continues to the Hydraulic Simulation part of this thesis. In this chapter, the step-by-step methodology to set up the HEC-RAS simulation is described. The description includes the Flow Data (discharge data) preparation, simulation setup, and stabilizing the Unsteady simulation.

The approach that was chosen for the hydraulic simulation is the 1-dimensional (1D) unsteady flow routing. This approach was selected because:

- HEC-RAS cannot simulate dry channels. Some reaches were dried during lower discharge (based on preliminary simulations and TSX Stripmap observation). Dry channels cause simulation to be unstable
- Due to the lack of bathymetry information, the flow direction of each reach could not be estimated
- The study area was big (~5,000 km<sup>2</sup>)
- A study by Caruso et al (2013) used a single streamline center for a braided river system using HEC-RAS 1D unsteady flow routing module

1D unsteady flow (dynamic wave) routing comes to the expense of simplifying the input parameters and the flow computation. However, this method would be useful as the initial approach to this remote study area, given the limited field information at hand. With this simplified method, the sensitivity of each approximated parameter (e.g. the channel bathymetry and the Manning's value) can be assessed. The hydraulic simulation that was performed in this study resembles that of manual iterative process, given there are many simulation cases to be run and calibrated. The computational time of 1D unsteady flow simulation is beneficial for that matter. The theoretical background in this chapter is based on the HEC-RAS Hydraulic Reference Manual (Brunner 2016).

The 1D unsteady flow in open channels was originally mathematically expressed by Barre de Saint-Venant in 1871 (Chow, Maidment, and Mays 1988). The mathematical expressions were thus acknowledged as the Saint-Venant (SV) equation. This equation is based on two conservation equations, i.e. mass and momentum. The basic concept of the conservation is a control volume (CV). CV is an imaginary and enclosed region within a flowing liquid. The shape of the CV is generated as to suit the analyzed condition. The further description of these two equations is as follows:

### 1) The law of conservation of mass

This term is also widely known as the continuity equation. The law states that the inflow into the CV equals to the outflow of the CV (Figure 31). In other words, the net rate of the flow into the CV equals to the rate of change of storage inside of the CV. This applies under the condition that no lateral in-/outflow occurs along the CV.

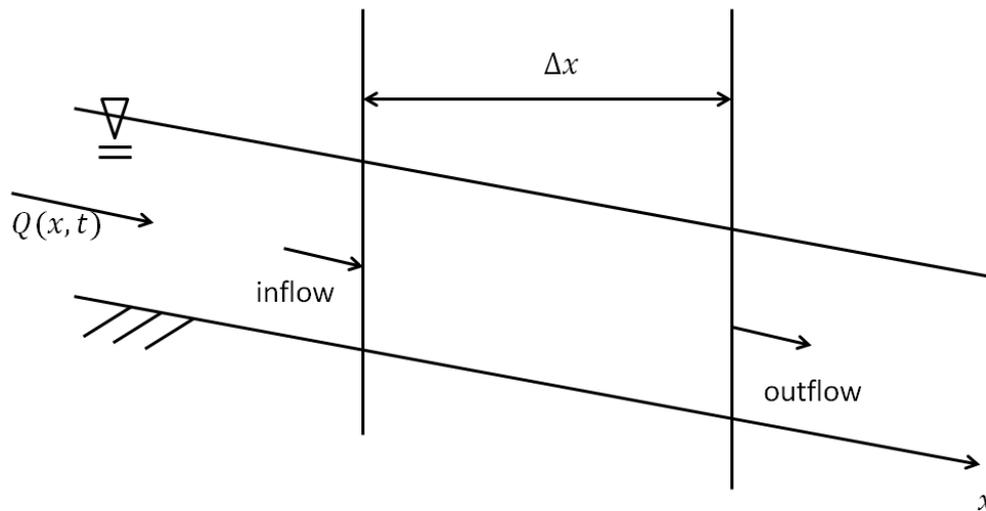


Figure 31 A Control Volume

The equation is expressed as follows:

$$\frac{\partial(AV)}{\partial x} + \frac{\partial A}{\partial t} - q = 0 \quad (5.1)$$

Where,

$A$  : cross-sectional flow area

$V$  : flow velocity

$q$  : lateral in-/outflow per unit length

$Q(x, t)$  : flow at the midpoint of the CV

$\Delta x$  : the distance between cross sections

### 2) The law of conservation of momentum

The law of conservation of momentum states that the net rate of momentum that enters a CV and the sum of all external forces that act on the CV equal to the rate of accumulation of momentum. The external forces are pressure, gravity, and frictions. The momentum equation is a vector equation that is applied on the longitudinal direction of the water flow.

$$\frac{\partial Q}{\partial t} + \frac{\partial QV}{\partial x} + gA \left( \frac{\partial h}{\partial x} + S_f \right) = 0 \quad (5.2)$$

Where,

$\frac{\partial Q}{\partial t}$  : the local acceleration term, that is the change in momentum due to the change in flow velocity over time

$\frac{\partial QV}{\partial x}$  : the convective acceleration term, that is the change in momentum due to the change in flow velocity along the channel

$g$  : the gravity force term, that is proportional to the bed slope,  $S_0$

$\frac{\partial h}{\partial x}$  : the pressure force term, that is proportional to the change of depth along the channel

$h$  : surface water level

$S_f$  : the friction force term, that is proportional to the friction slope

Looking at those SV equations, there are 2 dependent parameters; 1) the flow velocity and 2) the surface water level, that are varying in 2 independent parameters ( $x, t$ ). There is no analytical solution for these equation for practical application. In order to solve this equation, SV equation is derived based on these assumptions:

- The flow is one-dimensional; the velocity components on the  $y$  and  $z$  directions are neglected
- The water length is much bigger than the water depth, therefore the flow velocity on the  $z$  direction is neglected and the pressure is assumed to be hydrostatic
- The water level across one cross section is horizontal
- Channel bed and banks are fixed
- Average channel bed slope is small ( $<0.1$ )

Finite difference scheme is the method used by HEC-RAS to approximate the solutions, as the exact solution of the equations is not feasible for complex river system. The implicit solution scheme (as known as the 4-point scheme) is used to approximate the unknowns. The flow equations are expressed in finite difference form for all computational reaches with the length of  $\Delta x$  between the first and the last cross section for the unknown  $Q$  and  $t$  for each time step ( $\Delta t$ ). The information from the entire reach can influence the solution at any point.

## 5.1. Discharge Data Preparation

In HEC-RAS, the discharge data can be inputted as Hydrograph of the Boundary Condition in the Flow Data menu. This step was performed to provide the observed discharge values as a Hydrograph with a constant discharge value to fill the channel and stabilize the simulation. Three cases of each annual Spring flood event were observed: 1) low discharge, 2) medium discharge, and 3) high discharge.

### Discharge Data Adjustment

The Kyusyur discharge gauging station is located 124.65 km upstream from the starting point of the simulated area. Given that the channel shape and width seemingly doesn't change much along the way (Figure 1), it was assumed that there was also no change on the discharge values. However, by observing the TSX Stripmaps of the annual flood events, a discrepancy of the occurrence of the peak of the flood was found.

On the TSX Stripmaps, the peak of the flood was two days behind the gauged peak of the flood in Kyusyur station. This two-day discrepancy is consistent throughout the annual flood event from 2013 – 2019, based on the observation. Given the terrain was quite flat, it is possible it might take 2 days for the water flow from Kyusyur to disperse on the Lena delta.

In order to comply with this discrepancy, the gauged discharge datasets were adjusted by shifting the values by 2 days. Figure 32 shows the adjustment that was applied for the flood event in 2014. The availability of TSX Stripmap acquisition was indicated in grey lines. These TSX Stripmap acquisitions were used to make sure of the 2-day discrepancy.

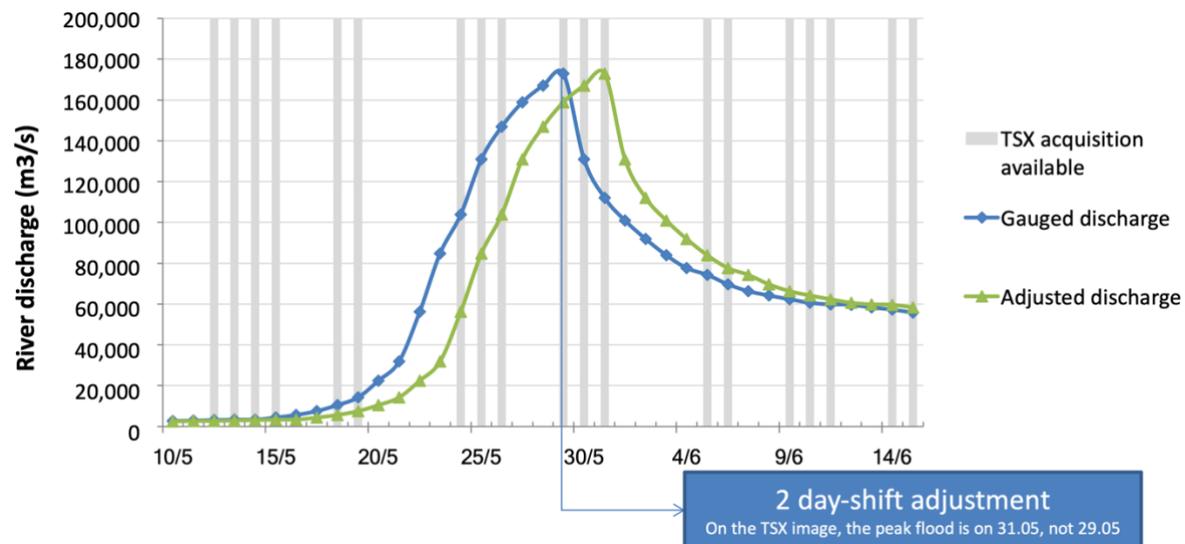


Figure 32 Adjusting the gauged discharge datasets

### Extracting the discharge value into a time-series constant discharge datasets

3 conditions that were observed in this study: 1) low discharge, 2) medium discharge, and 3) high discharge. The low discharge values were taken before the flood developed in early May. The high discharge values were taken from the peak of the flood. The medium discharge values were taken when the flood was receding. The time from which these values were extracted was selected accordingly to the availability of a good quality TSX Stripmap to match. Table 14 shows the selected days and the corresponding discharge values for the simulation.

**Table 14 The low, high, and medium discharge values for the simulation**

<b>No.</b>	<b>Date</b>	<b>Label</b>	<b>Q (m<sup>3</sup>/s)</b>
1	2013-05-14	Low discharge 2013	3,020
2	2013-06-02	High discharge 2013	155,000
3	2013-06-11	Medium discharge 2013	56,000
4	2014-05-09	Low discharge 2014	2,280
5	2014-05-31	High discharge 2014	173,000
6	2014-06-11	Medium discharge 2014	62,400
7	2015-05-07	Low discharge 2015	2,989
8	2015-06-04	High discharge 2015	147,029
9	2015-06-23	Medium discharge 2015	62,000
10	2016-05-10	Low discharge 2016	5,366
11	2016-06-01	High discharge 2016	128,593
12	2016-06-26	Medium discharge 2016	70,900
13	2017-05-19	Low discharge 2017	2,720
14	2017-06-04	High discharge 2017	87,120
15	2017-06-29	Medium discharge 2017	67,980
16	2018-05-17	Low discharge 2018	2,810
17	2018-06-06	High discharge 2018	119,200
18	2018-06-13	Medium discharge 2018	65,870
19	2019-05-04	Low discharge 2019	3,420
20	2019-06-04	High discharge 2019	83,000
21	2019-07-01	Medium discharge 2019	46,200

## 5.2. Simulation Setup

### Importing the Geometry to HEC-RAS

The geometry data for simulation was already generated from the remotely-sensed datasets. The geometry data in the form of SDF file (generated in 4.6) was imported to HEC-RAS. Figure 33 shows the default HEC-RAS geometric data view.

However, adjustments needed to be done to the geometry file in order to make the simulation work. These adjustments are: 1) joint data adjustment, 2) reducing the geometry data, and 3) simulated increment adjustment.

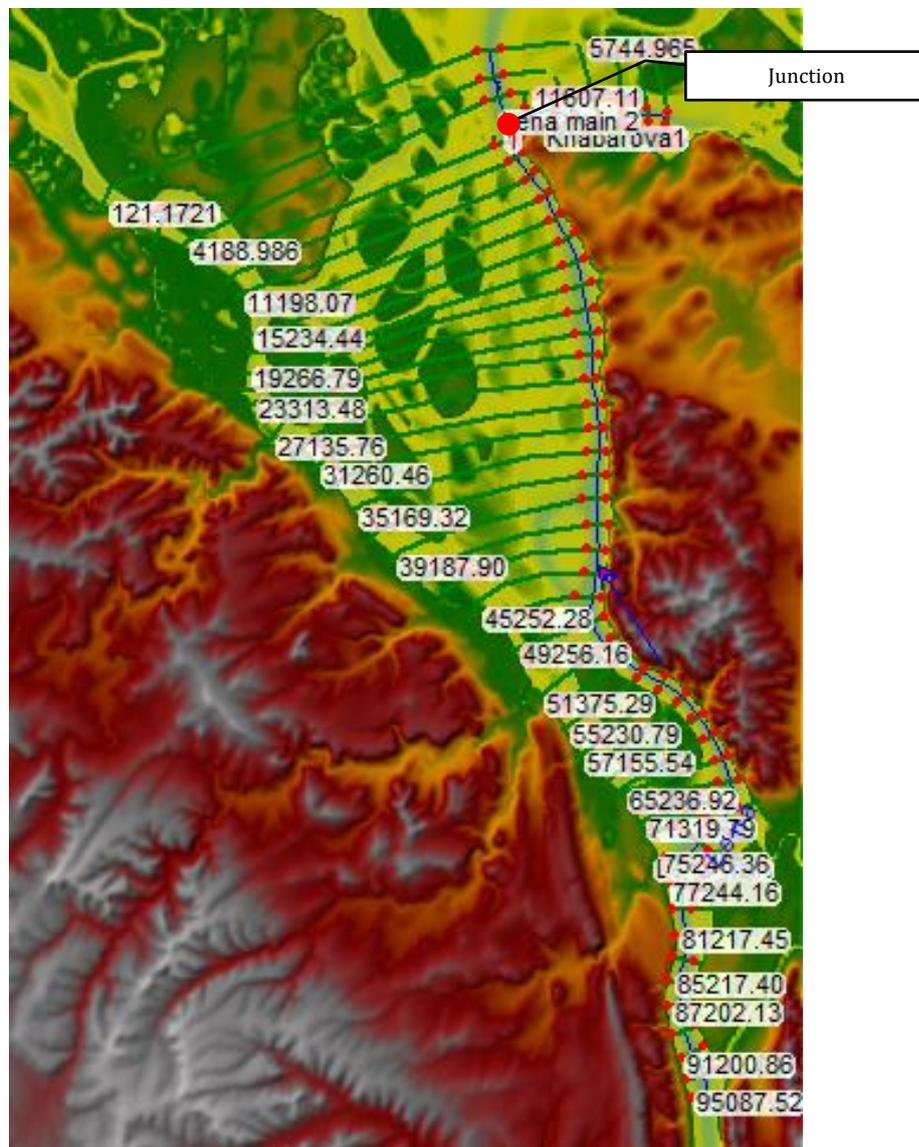


Figure 33 The HEC-RAS geometric data view

## **Junction Data**

The cross-sectional line upstream to the junction had to be set with the Downstream Reach Lengths of 0 (Brunner and CEIWR-HEC 2016). RiverGIS didn't define the junction that way. Therefore, the Downstream Reach Lengths for the cross-sectional line had to be set to 0 manually. Another important parameter that had to be added into the junction data is the junction length. The junction length is defined as the average length between the stream centerlines of the connecting rivers. The junction length to the downstream Lena reach and to the Khabarova reach is 3,679 m and 2,669 m, respectively. The Energy Balance Method was selected for the junction computation. The Energy Balance (Energy) method solves the WS elevations by using the energy equation across the junction.

## **Reduce Geometry Data**

Due to the nature of the high-resolution satellite datasets, the exported geometry was highly heterogeneous. Although this would promote a better accuracy, the heterogeneity cannot always be handled well by the HEC-RAS 1D simulation. HEC-RAS has a limitation in processing geometry data as follows:

- The maximum number of station-elevation points, blocked obstructions, and levees is 500 points
- There is a limit of 20 Manning's value per cross section

The cross-section points extracted from the modified TDX DEM varied according the channel width, from 300 to 2000 points. In order to fit the HEC-RAS' prerequisites, the elevation points of each cross-section were reduced to 499 points. The number of points were reduced with the XS Points Filter Tool. There are two options to filter the cross-section points:

- Near and Colinear Filter
- Minimize Area Change

The latter was chosen as the method preserves the natural shape of the channel (Figure 34), whereas the first method relies on a user-defined tolerance value.

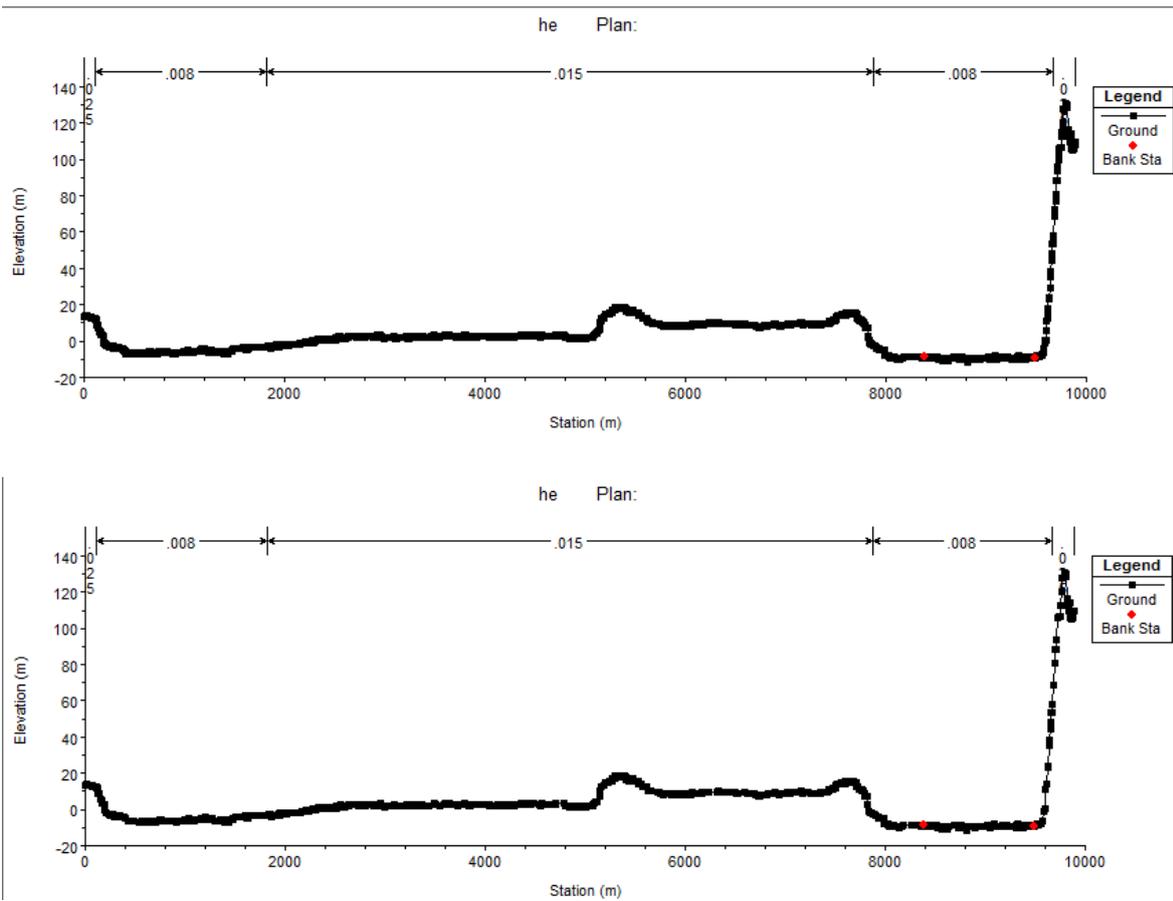


Figure 34 Cross-section 63216.78 before and after the elevation points were reduced to 500

### Steady Flow Data Input

The final results were to be obtained from the unsteady flow simulation, but the steady flow analysis was performed beforehand to an important parameter for the unsteady flow analysis. The steady flow analysis was performed with the aim to obtain the average flow velocity. The average velocity was used as the empirical base to approximate the computational time step, which is important for the stability of the unsteady flow analysis.

Steady flow analysis simulates a state where the fluid boundary condition does not change over time. Therefore, the steady flow rate input for this setup was the discharge values of each observed case presented in Table 14. On the cross-sectional lines after the junction, the discharge values were split into the two reaches (Lena downstream and Khabarova) respective to the cross-sectional wetted area of each reach (Table 15).

Table 15 Steady flow data input

	Date	Lena upstream reach runoff (m3/s)	Lena downstream reach runoff (m3/s)	Khabarova reach runoff (m3/s)
<b>2013</b>	2013-05-14	3,020.00	1,980.18	1,039.82
	2013-06-02	155,000.00	101,631.66	53,368.34
	2013-06-11	56,000.00	36,718.54	19,281.46
<b>2014</b>	2014-05-09	2,280.00	1,494.97	785.03
	2014-05-31	173,000.00	113,434.05	59,565.95
	2014-06-11	62,400.00	40,914.94	21,485.06
<b>2015</b>	2015-05-07	2,989.00	1,959.85	1,029.15
	2015-06-04	147,029.00	96,405.17	50,623.83
	2015-06-23	62,000.00	40,652.67	21,347.33
<b>2016</b>	2016-05-10	5,366.00	3,518.42	1,847.58
	2016-06-01	128,593.00	84,316.91	44,276.09
	2016-06-26	70,900.00	46,488.29	24,411.71
<b>2017</b>	2017-05-19	2,720.00	1,783.47	936.53
	2017-06-04	87,120.00	57,123.55	29,996.45
	2017-06-29	67,980.00	44,573.68	23,406.32
<b>2018</b>	2018-05-17	2,810.00	1,842.48	967.52
	2018-06-06	119,200.00	78,158.03	41,041.97
	2018-06-13	65,870.00	43,190.18	22,679.82
<b>2019</b>	2019-05-04	3,420.00	2,242.45	1,177.55
	2019-06-04	83,000.00	54,422.12	28,577.88
	2019-06-29	46,200.00	30,292.79	15,907.21

### Unsteady Flow Data Input

Unlike the steady flow analysis, the unsteady flow analysis simulates the changing boundary conditions in time-series. In this study, constant values of upstream runoff were used as the upstream hydrograph input. The runoff values are the same as displayed in Table 14. The runoffs are kept constant for 5 days in order to stabilize and fill the channels. The initial condition is set as the initial runoff upstream, and split runoff after the joint, similar to the values provided in Table 15.

### 5.3. Stabilizing the Unsteady Simulation

An unstable numerical model is identified by one or both of these characteristics:

- numerical errors that are growing to the extent of oscillating solutions
- error that is becoming large making the computations stop

In order to make HEC-RAS unsteady simulation steady, the temporal and spatial change has to be kept gradual. The ideal way is to try all the parameter combination of a big range of parameter values until the simulation works. But with the limitation of time, the range of the values was shortened with the help of empirical equations and references from similar studies.

In this study, the parameters that were tuned to stabilize the simulation are the cross-sectional line spacing, computational time step, cross-section table properties, contraction and expansion coefficient, and downstream slope.

### **Cross-sectional line spacing ( $\Delta x$ )**

A common rule of thumb in approximating the cross-sectional line spacing is Samuel equation (Samuels 1989, Goodell 2008). This empirical equation estimated the spacing by the channel bankfull depth and the overall slope.

$$\Delta x \leq \frac{0.15 D}{S} \quad (5.3)$$

$$\Delta x \leq \frac{0.15 \times 5}{2.72 \times 10^{-4}}$$

$$\Delta x \leq 2,720 \text{ m}$$

where,

- $\Delta x$  : distance between cross-section lines
- $D$  : bankfull depth
- $S$  : slope

The selected spacing is 2 km. More details of this selection will be discussed in the Sensitivity Analysis in 6.4.

### **Time step ( $\Delta t$ )**

The implicit solution scheme requires a certain computation time step to compute the unknown parameters. The time step is crucial to the simulation in order to keep the simulation stable. In this study, the time step is approximated with the following equation:

$$\Delta t \leq \frac{\Delta x}{V_w} \quad (\text{Brunner 2011}) \quad (5.4)$$

where,

- $\Delta x$  : distance between cross-section lines
- $V_w$  : flood wave speed

Given the used hydrographs for the simulation is of constant discharge values, the flood wave speed is taken as the average flow velocity of the channel. The average velocity input to

calculate the time step is taken from the result from the steady flow analysis. The time step values are provided in Table 16.

**Table 16 The average velocity and the computational time step of each simulation setup**

<b>No. Event</b>	<b>Date</b>	$\bar{V}$ (m/s)	$\Delta t \leq$ (mins)	$\Delta t$ (mins)
1	2013-05-14	0.845	39.469	30
2	2013-06-02	3.635	9.170	6
3	2013-06-11	2.483	13.422	12
4	2014-05-09	0.769	43.362	40
5	2014-05-31	3.924	8.496	6
6	2014-06-11	2.585	12.894	12
7	2015-05-07	0.841	39.614	30
8	2015-06-04	3.561	9.361	9
9	2015-06-23	2.578	12.929	12
10	2016-05-10	1.019	32.697	30
11	2016-06-01	3.378	9.867	9
12	2016-06-26	2.708	12.309	12
13	2017-05-19	0.816	40.850	30
14	2017-06-04	2.917	11.429	10
15	2017-06-29	2.666	12.502	12
16	2018-05-17	0.825	40.417	30
17	2018-06-06	3.282	10.156	10
18	2018-06-13	2.636	12.646	12
19	2019-05-04	0.878	37.965	30
20	2019-06-04	2.865	11.634	10
21	2019-07-01	2.314	14.404	12

### **The Cross-section Table Properties (HTab)**

A cross-section table properties (HTab) that has solid definitions help construct a stable unsteady flow model in HEC-RAS. HEC-RAS converts the cross-sectional geometry into a curve explaining the connection between the hydraulic parameters and the stage. This curve is used during the computation. Since this curve is defined in discrete points, the values between the points are linearly interpolated.

The interpolation increment defines the smoothness of the interpolated values. The increments that are too coarse might lead to error or instability. Therefore, it is important to define the starting point, increment, and the number of increment points that cover the wetted area of each cross sections. If the wetted area is not covered by the increment points, the simulation might go instable.

Figure 35 shows the HTab parameter table. On the right side, the interpolation increment lines of the cross section 63216.78 is shown. The lines were placed close together, hence it looks like a continuous grey color. The starting elevation was set to the lowest point of each cross section. The interpolation increment was set to 0.1. The increment points were set to 250 to cover the wetted area.

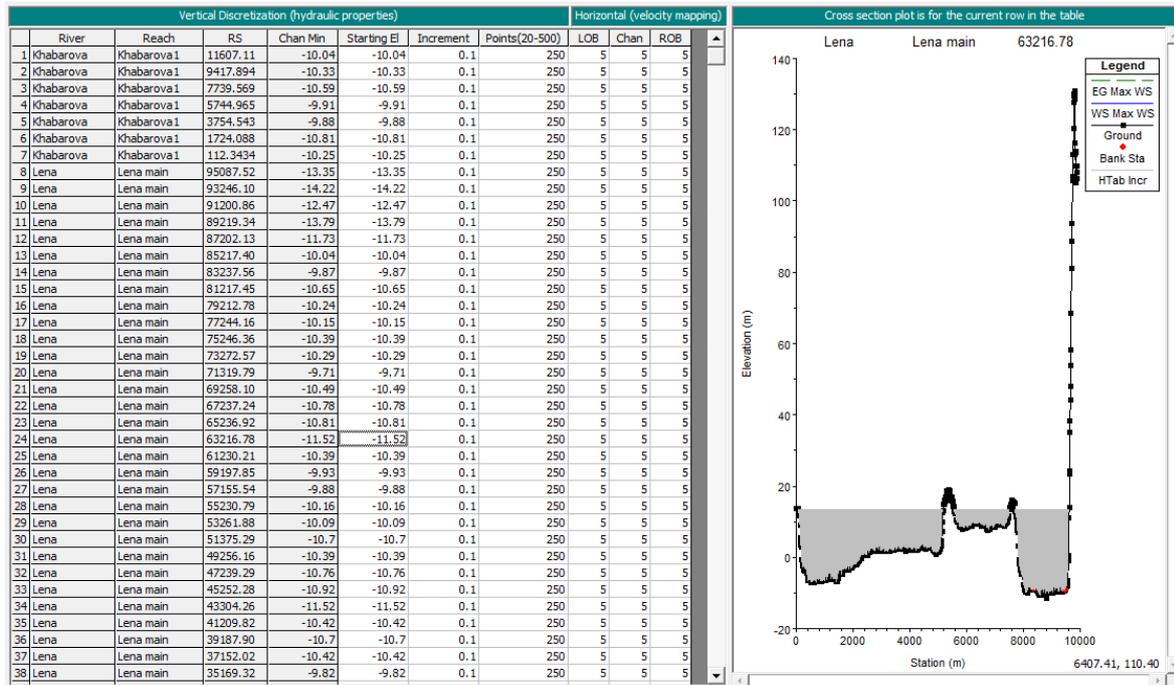


Figure 35 The HTab parameters, with cross section 63216.78 shown on the right

### Contraction and expansion coefficient

The 1D momentum equation does not always take into account the forces on the flow field at sharp contraction/expansion zones. Empirical contraction and expansion coefficients can be entered to better approximate the forces acting on the water on those zones. Natural channels are different to man-made channels, as the changes along the channel can be more abrupt. Therefore, this simulation used the contraction and expansion coefficients that were defined for natural channels, 0.05 and 0.2, respectively (Brunner and CEIWR-HEC 2016).

### Downstream slope

The downstream Boundary Condition of this simulation were set to the friction slope. The friction slope can be approximated to the riverbed slope. In order to keep the slope changes along the cross sections constant, the slope value of 0.0001 was used. This is discussed in more details in 6.4.3.

Table 17 summarizes the simulation setup parameters.

**Table 17 The simulation parameters**

<b>Boundary Condition</b>	<b>Upstream</b>	Hydrograph (Table 14)
	<b>Downstream</b>	Friction slope of 0.0001
<b>Initial Condition</b>	<b>Upstream</b>	Initial Discharge (Table 14)
	<b>Joints</b>	Discharge values according to the wetted area (Table 15)
<b>Cross-sectional line spacing</b>	2 km	
<b>Computational time step</b>	Ranging from 6 – 40 mins (Table 16)	
<b>Simulated area</b>	1,561.15 km <sup>2</sup>	

## 6. Results and Discussions

In this chapter, the results of the remote sensing method and the hydraulic simulation are provided and discussed. The results from the remote sensing method includes 1) bathymetry, 2) land cover, and 3) Manning's surface roughness coefficient. The results from the HEC-RAS hydraulic simulation includes 1) inundation boundary, 2) flood depth, and 3) flood velocity.

The sensitivity analyses of the hydraulic simulation parameters are also described. The sensitivity analyses include the model sensitivity to the cross-sectional line spacing, computational time step, and the downstream slope. This section explains why those values were selected from previous iterative processes by running the HEC-RAS simulation with different parameter values.

The model accuracy is assessed by comparing the modeled inundation boundaries to the flood maps. Model accuracy refers to the degree of closeness of the numerical solution to the true solution. The model accuracy depends highly on 1) assumptions and limitations of the model, 2) accuracy of the geometric data, 3) accuracy of the flow data and BC, and 4) numerical accuracy of the solution scheme. Numerical accuracy is different to the model accuracy, as it refers to the approximation by the method to solve the problem.

### 6.1. Bathymetry

The main reaches of Lena delta were approximated as trapezoidal channels, whereas the smaller reaches were taken as it is, given that with shallower depth the riverbed might have been detected when the water on these channels were completely frozen. Due to the fact that the floodplain on the study area was quite flat, the channels were assumed to be flat as well from the upstream to the downstream of the modeled area.

The bathymetry approximation procedures were performed 5 times, each with different maximum depth: -30 m, -25 m, -20 m, -15 m, and -10 m. The selected set was determined through an iterative process. Each of these *synthetic bathymetry* sets were inputted into HEC-RAS geometry data and the *synthetic bathymetry* set with resulting inundation boundaries most fitting to the flood map was selected.

The resulting inundation boundaries of the simulations using the maximum depth of -15 m to -30 m were underestimated. The reaches other than the main Lena reach were dried during the lower runoff, unlike that of the corresponding TSX-generated flood map. The results of the

simulation using the maximum depth of -10 m showed the best agreement to the TSX-generated flood maps. Figure 36 shows the resulting *synthetic bathymetry*.

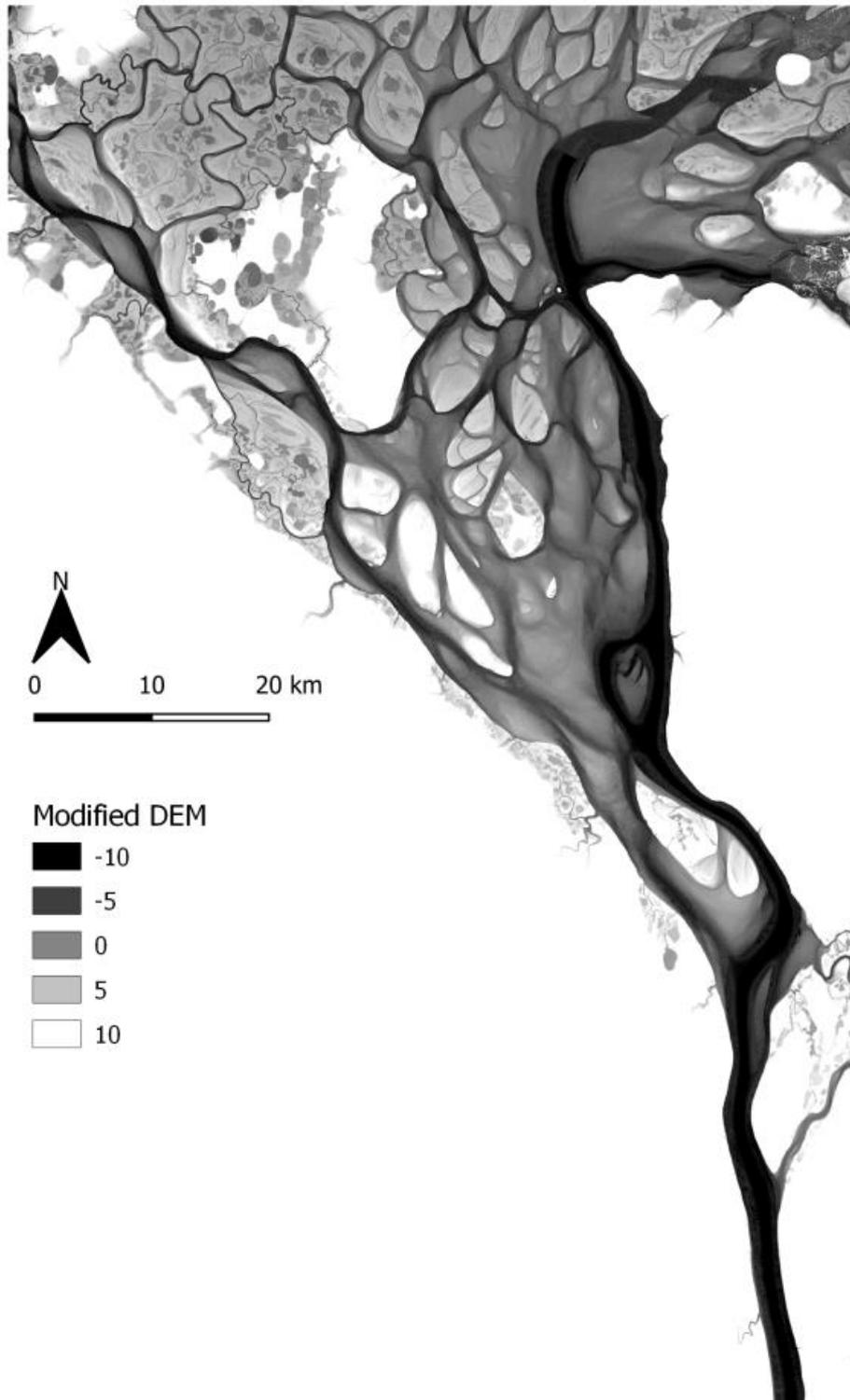


Figure 36 The DEM corrected with generated trapezoidal channel with the maximum depth of -10 m. The terrains higher than 10 m were not shown as those were irrelevant to the simulation

## 6.2. Land Cover

The land cover classification was performed twice: 1) with the TSX-RE datasets and 2) with the LS8 datasets. The final result was mosaicked together by prioritizing the TSX-RE land cover on the overlapping area, as the TSX-RE provided much better spatial resolution (5 m instead of 30 m). There are 6 land cover classes: 1) sand, 2) Holocene terraces, 3) vegetated floodplain and Pleistocene terraces, 4) densely-vegetated floodplain, 5) water body, and 6) rock.

In this subchapter, the land cover classification results are discussed by analyzing the confusion matrix between the predicted and actual classes, user's accuracy, producer's accuracy, and the overall accuracy.

### 6.2.1. TerraSAR-X and RapidEye Stack



**Figure 37** The land cover classification map obtained from the stack of TSX and RapidEye datasets

The TSX-RE land cover classification was performed separately on 3 different TSX tiles: west, center, and east tiles. The overall accuracy of the west, center, and east tiles are 93%, 95%, and

93%, respectively. The resulting land cover map is shown in Figure 37. Through the confusion matrices (Table 18 - Table 20), the most reoccurring of the misclassification was of the classes vegetated floodplain and densely vegetated floodplain. The backscatter intensity and surface reflectance of these classes are quite close at some area. Other than that, it is possible that the vegetation variants are mixed. In order to achieve better accuracy on these classes, the vegetation can be classed by species, as different species might reflect signal and light differently.

**Table 18 The confusion matrix of the TSX-RE west tile land cover classification**

		Predicted class						User's accuracy (%)
		Sand	Holocene terraces	Vegetated floodplain & pleistocene terraces	Densely vegetated floodplain	Water body	Rock	
Actual class	Sand	11	0	0	0	0	0	100.0
	Holocene terraces	0	7	0	0	0	0	100.0
	Vegetated floodplain & pleistocene terraces	0	1	9	0	0	0	90.0
	Densely vegetated floodplain	0	0	1	15	0	1	88.2
	Water body	0	0	0	0	10	0	100.0
	Rock	0	0	0	1	0	4	80.0
Producer's accuracy		100.0	87.5	90.0	93.8	100.0	80.0	

Table 19 The confusion matrix of the TSX-RE center tile land cover classification

		Predicted class					User's accuracy (%)	
		Sand	Holocene terraces	Vegetated floodplain & pleistocene terraces	Densely vegetated floodplain	Water body		Rock
Actual class	Sand	27	0	0	0	0	0	100.0
	Holocene terraces	0	6	0	0	0	0	100.0
	Vegetated floodplain & pleistocene terraces	0	0	4	1	0	0	80.0
	Densely vegetated floodplain	0	0	1	2	0	0	66.7
	Water body	0	0	0	0	18	0	100.0
	Rock	0	0	0	0	0	1	100.0
Producer's accuracy		100.0	100.0	80.0	66.7	100.0	100.0	

Table 20 he confusion matrix of the TSX-RE east tile land cover classification

		Predicted class						User's accuracy (%)
		Sand	Holocene terraces	Vegetated floodplain & pleistocene terraces	Densely vegetated floodplain	Water body	Rock	
Actual class	Sand	6	0	0	0	0	0	100.0
	Holocene terraces	0	4	0	0	0	0	100.0
	Vegetated floodplain & pleistocene terraces	0	0	14	3	0	0	82.4
	Densely vegetated floodplain	0	0	1	12	0	0	92.3
	Water body	0	0	0	0	13	0	100.0
	Rock	0	0	0	0	0	6	100.0
Producer's accuracy		100.0	100.0	93.3	80.0	100.0	100.0	

The highest user's accuracy, 100%, is achieved by the sand class, holocene terraces, and water body class for all tiles. The lowest user's accuracy, 66.7%, is achieved by the densely-vegetated floodplain on the center tile. The highest producer's accuracy, 100%, was achieved by the sand and waterbody classes on all tiles. The lowest producer's accuracy, 66.7%, was achieved by the densely-vegetated floodplain on the center tile. Through this result, it can be concluded that sand, holocene terraces, and water body are highly distinguishable, whereas densely-vegetated floodplain is not.

The variable importance of the Random Forest classifier is the measure of the ranking of important features, in this case, the raster bands. The Variable Importance Measure (VIM) of the Random Forest classifier for the west, center, and east tiles is presented in Figure 38. Given that the training datasets were based on the RE bands (blue, green, red, red edge, and near infrared), the VIM of these bands are much higher than that of the TSX bands (K0, K3, and K4).

This shows that the spatial distribution pattern of the TSX intensity values has low agreement with the spatial distribution pattern of the RE reflectance values. The band with the highest variable importance is near the infrared band. The band with the lowest variable importance is the K3 band. The near-infrared spectrum consists of a large set of overtones and combination bands, making it more varied according to the surface cover. K3, the absorption band, does not distinguish different surface cover that well.

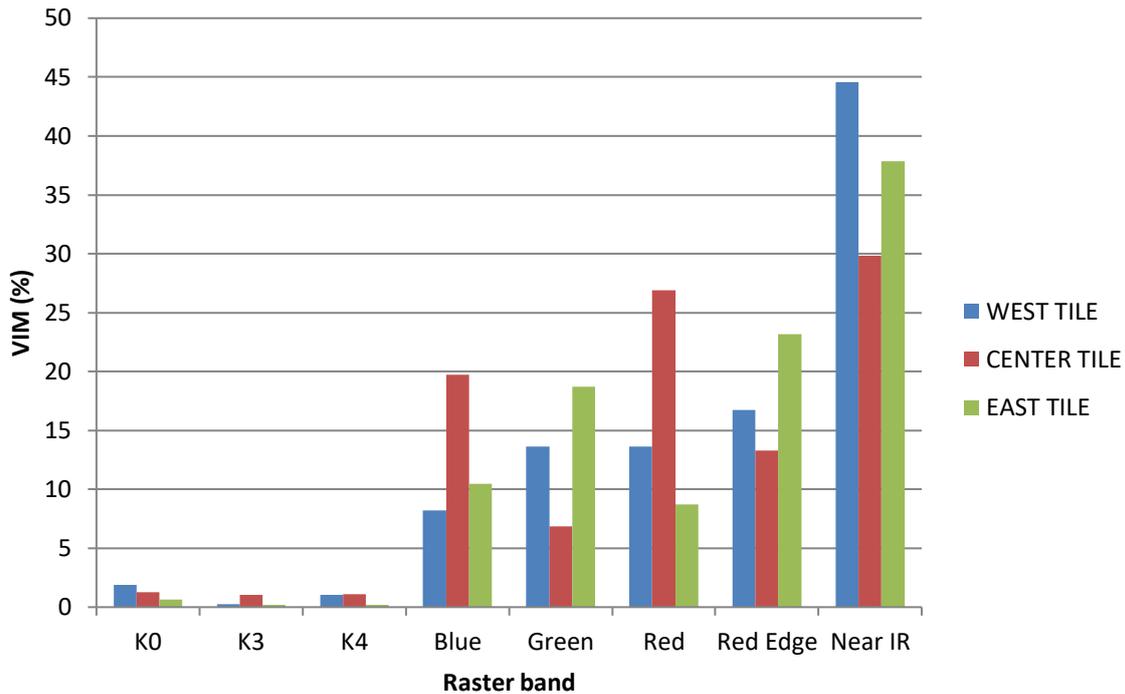


Figure 38 The Variable Importance Measures (VIM) of TSX bands and RE bands for the Random Forest-based land cover classification

### 6.2.2. Landsat 8

Figure 39 shows the land cover map obtained from the LS8 classification. Although the pattern looks similar to Figure 37, the result differs in details. This is due to the spatial resolution difference. LS8 pixel is 36 times bigger than that of TSX and RE. Therefore, those pixels are rounded up as each pixel in LS8 image represents 36 TSX/RE pixels. The overall accuracy of this classification is also way lower, 75%. The confusion matrix (Table 21) shows that the misclassifications are more distributed than that of TSX-RE classification. The most misclassified classes are between the Holocene terraces, vegetated floodplain, and densely vegetated floodplain.



- 1 = sand
- 2 = holocene terraces
- 3 = vegetated floodplain & pleistocene terraces
- 4 = densely vegetated floodplain
- 5 = water body
- 6 = rock

**Figure 39 The land cover classification map obtained from the LS8 datasets**

The highest user's accuracy, 93.2%, is achieved by the sand class. The lowest user's accuracy, 53.8%, is achieved by the rock class. The highest producer's accuracy, 93.2%, was achieved by the sand class. The lowest producer's accuracy, 50%, was achieved by the rock class. This showcases that the sand is highly distinguishable, whereas rock is the least distinguishable of all the classes through the LS8 bands.

**Table 21 The confusion matrix of the LS8 land cover classification**

		Predicted class						User's accuracy (%)
		Sand	Holocene terraces	Vegetated floodplain & pleistocene terraces	Densely vegetated floodplain	Water body	Rock	
Actual class	Sand	41	0	0	0	2	1	93.2
	Holocene terraces	0	8	8	0	1	0	47.1
	Vegetated floodplain & pleistocene terraces	0	4	20	6	0	2	62.5
	Densely vegetated floodplain	0	0	5	23	0	4	71.9
	Water body	2	1	2	0	36	0	87.8
	Rock	1	0	0	5	0	7	53.8
Producer's accuracy		93.2	61.5	57.1	67.6	92.3	50.0	

Similar to the Random Forest VIM of the TSX-RE classification, the variable with the highest importance is the near-infrared band (shown in Figure 40). The least important band for this classification is the coastal aerosol band. The near-infrared spectrum consists of a large set of overtones and combination bands, making it more varied according to the surface cover. The coastal aerosol band, on the other hand, is useful for coastal and aerosol studies, but does not represent different surface covers clearly.

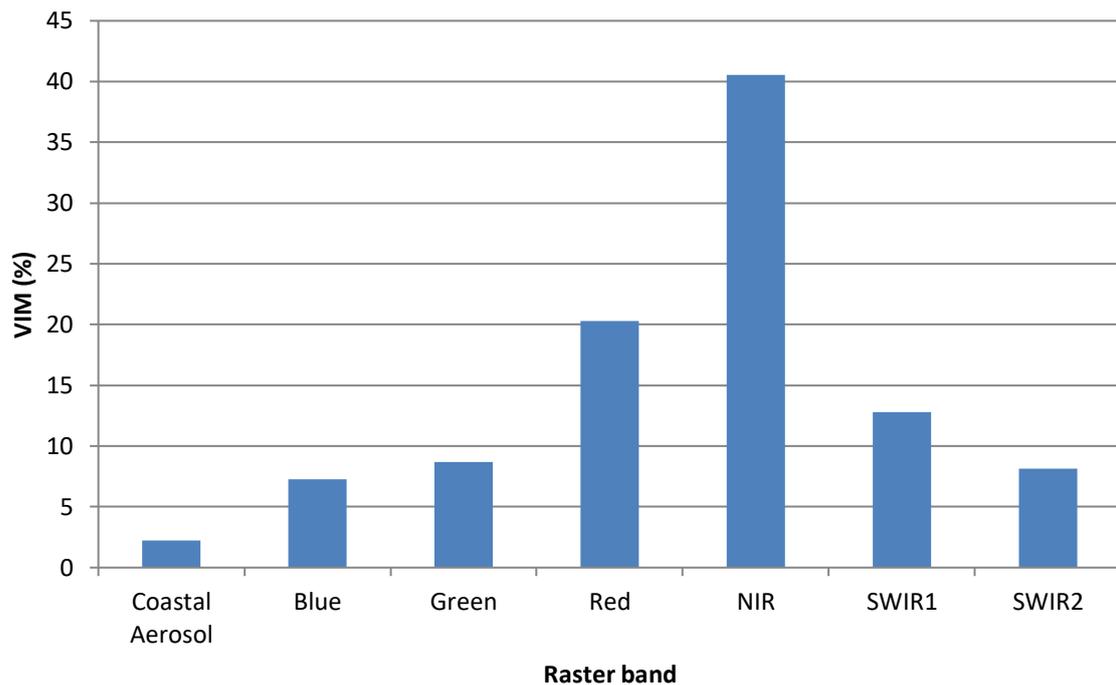


Figure 40 The Variable Importance Measures (VIM) of LS8 bands for the Random Forest-based land cover classification

### 6.3. Manning's Surface Roughness Coefficient

The Manning's surface roughness coefficients were assigned to each land cover class. Due to the lack of field surface roughness measurement, a well-known Manning's value reference (Chow 1959) was used to approximate the Manning's surface roughness coefficients based on the surface cover classes. The middle (normal) coefficient values were selected (Table 22). Based on a preliminary study, it was concluded that most of the riverbed in the study area is made of sand. Therefore, the water body class was considered as sandy riverbed.

Table 22 The Manning's roughness value of the channel and floodplain

No	Land cover class	Manning's roughness value
1	Sandy floodplain	0.048
2	Holocene terraces	0.060
3	Vegetated floodplain and Pleistocene terraces	0.070
4	Densely vegetated floodplain	0.150
5	Sandy riverbed	0.030
6	Rock	0.040

## 6.4. HEC-RAS Sensitivity Analysis

### 6.4.1. The model sensitivity to the cross-sectional line spacing ( $\Delta x$ )

Based on the empirical Samuel equation, the cross-sectional line spacing ( $\Delta x$ ) had to be smaller than 2.72 km. Spacing that is placed too far will cause numerical damping of the flood wave and model instability (Gary W. Brunner 2011).

Given that 2.72 km is quite a distance, the initial approach to the spacing was by trying to run the simulation with a smaller distance first. The earlier simulation with the  $\Delta x$  of 20m, 50m, 100m, 200m, 400m, 500m, 1 km, up to 1.5km were all unstable, resulting in spiking energy grade lines and rising flood waves at some cross-section (Figure 41).

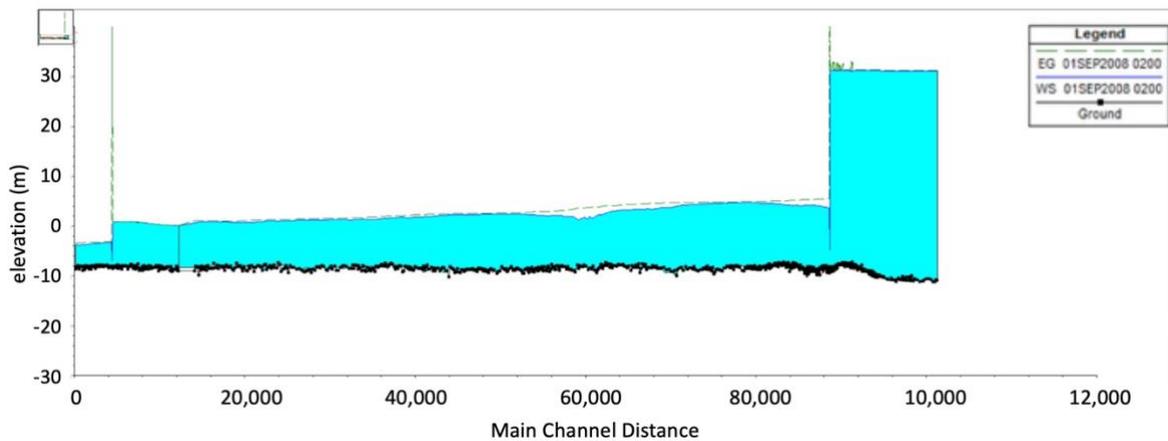


Figure 41 Unstable result due to cross-section lines being too close together (in the picture,  $\Delta x = 50\text{m}$ )

According to Gary W. Brunner (2011), spacing that is placed too close will cause overestimated energy loss, precipitated significant error, wave steepening, and model instability on the rising side of the flood wave. According to that description, it was clear that the spacing between 20 m to 1.5 km were too small for this simulation, as the wave steepening and the instability on the rising side of the flood wave occurred.

At the end, it was found out that the  $\Delta x$  of 2 km works to stabilize the simulation. The distance of 2 km is quite big and might result in the loss of the channel bathymetry details. However, given that the riverbed of this study area was approximated to be quite flat, the loss of details was not significant. The generated geometry lines in Figure 29 has shown that the islands that divide the reaches were all represented by the cross-section lines.

#### **6.4.2. The model sensitivity to the time step ( $\Delta t$ )**

There are 21 simulation setups and the time step  $\Delta t$  of each setup depends on the average flood wave velocity which was acquired from the steady state simulation. The  $\Delta t$  ranges from 6 to 40 minutes. The higher the upstream runoff, the higher the acquired average flood wave velocity. The  $\Delta t$ , which is inversely proportional to the average flood wave velocity, gets smaller as the upstream runoff gets higher. The  $\Delta t$  of the low, medium, and high upstream runoff ranges between 30-40 minutes, 6-10 minutes, and 12 minutes, respectively. The  $\Delta t$  that is set too small may cause model instability, long computation time, and steepening flood wave that might oscillates. On the other hand, the  $\Delta t$  that is set too big will cause numerical diffusion and model instability as well (Gary W. Brunner 2011).

#### **6.4.3. The model sensitivity to the downstream slope**

The channel was assumed to be flat from the upstream to the downstream of the modeled area. However, the downstream friction slope was required as the Boundary Condition of all the downstream reaches. The friction slope was adjusted to stabilize the Energy Grade line through an iterative process. In order to stabilize the simulation, the slope changes between cross-sections should not change abruptly. Hence, a small value of friction slope was required to make the simulation work. The friction slope of 0.0001, meaning there is 1 m elevation change every 10 km, fits the simulation well. The friction slope that is too big (0.01 and 0.001 were also tested before) caused wave steepening and unstable simulation.

### **6.5. Inundation Boundary**

The accuracy of the simulation was assessed by comparing the resulting inundation boundaries of each simulation setup were compared to the flood map created from the TDX-TSX imageries. The resulting HEC-RAS inundation boundaries were rendered into polygon shapefiles, therefore they were converted into raster files consisting of binary numbers; 1 notating the flooded pixel and 0 the non-flooded pixel. The comparison is quantified with 2-class confusion matrix (Table 23), where the YES column is designated for flooded pixel and the NO column for non-flooded pixel.

Table 23 Two-class confusion matrix

		<b>predicted</b>	
		<b>YES</b>	<b>NO</b>
<b>actual</b>	<b>YES</b>	<b>Hits</b> = True Positives	<b>Misses</b> = False Negatives
	<b>NO</b>	<b>False Alarms</b> = False Positives	<b>Correct Negative</b> = True Negatives

Hits, i.e. True Positives, are cases in which the classifier predicted YES (pixel is flooded) and it is flooded in real life. Correct Negatives, i.e. True Negatives, are cases in which the classifier predicted NO (pixel is not flooded) and it is not flooded in real life. Misses, i.e. False Negatives, are cases in which the classifier predicted NO (pixel is not flooded), but it is not flooded in real life. False Alarms, i.e. False Positives, are cases in which the classifier predicted YES (pixel is flooded), but it is not flooded in real life.

Given that the binary raster files are rectangular-shaped, the non-flooded pixels obviously outnumbered the flooded pixels. Looking into these non-flooded pixels would not be representative to the result accuracy, since HEC-RAS wasn't used to model the rectangular area. Therefore, the accuracy is rated by the True Positive (Hit) rate, calculated by the following equation:

$$True\ Positive\ rate = \frac{True\ Positives}{Actual\ flooded\ pixels}$$

The confusion matrix component values with the True Positive rates are presented in Table 24.

**Table 24 The values of the Hits, False Alarms, Misses, Correct Negatives, and the True Positive Rate of the modeled inundation boundaries and the TSX-TDX-generated flood maps**

Upstream Condition	Year	Date	Q (m <sup>3</sup> /s)	Hits	False Alarms	Misses	Correct Negatives	Actual Flooded Pixels	TP rate (%)
<b>Low runoff</b>	<b>2013</b>	2013-05-14	3,020	9,949,673	1,273	4,869,880	127,513,439	14,819,553	67.14
	<b>2014</b>	2014-05-09	2,280	9,488,014	1,422	4,403,787	128,441,042	13,891,801	68.30
	<b>2015</b>	2015-05-07	2,989	9,854,936	77,595	835,407	131,566,327	10,690,343	92.19
	<b>2016</b>	2016-05-10	5,366	10,768,918	3,211	3,122,883	128,439,253	13,891,801	77.52
	<b>2017</b>	2017-05-19	2,720	9,776,365	1,226	5,043,188	127,513,486	14,819,553	65.97
	<b>2018</b>	2018-05-17	2,810	9,827,743	1,552	4,064,058	128,440,912	13,891,801	70.74
	<b>2019</b>	2019-05-04	3,420	10,153,633	1,835	3,738,168	128,440,629	13,891,801	73.09
<b>Medium runoff</b>	<b>2013</b>	2013-06-11	56,000	23,982,605	17,788	11,391,349	106,942,523	35,373,954	67.80
	<b>2014</b>	2014-06-11	62,400	25,480,671	385,585	3,635,446	112,832,563	29,116,117	87.51
	<b>2015</b>	2015-06-23	62,000	25,720,984	30,855	8,534,481	108,047,945	34,255,465	75.09
	<b>2016</b>	2016-06-26	70,900	28,114,601	85,190	6,140,864	107,993,610	34,255,465	82.07
	<b>2017</b>	2017-06-29	67,980	27,353,963	60,497	6,901,502	108,018,303	34,255,465	79.85
	<b>2018</b>	2018-06-13	65,870	26,750,819	87,833	6,287,220	109,208,393	33,038,039	80.97
	<b>2019</b>	2019-07-01	46,200	20,767,281	265,736	4,401,529	116,899,719	25,168,810	82.51
<b>High runoff</b>	<b>2013</b>	2013-06-02	155,000	40,544,939	10,470	3,477,519	98,301,337	44,022,458	92.10
	<b>2014</b>	2014-05-31	173,000	41,486,381	9,863	3,563,180	97,274,841	45,049,561	92.09
	<b>2015</b>	2015-06-04	147,029	40,045,741	15,998	3,556,419	98,716,107	43,602,160	91.84
	<b>2016</b>	2016-06-01	128,593	38,457,361	200,678	1,750,369	101,925,857	40,207,730	95.65
	<b>2017</b>	2017-06-04	87,120	32,435,780	5,977	9,467,597	100,424,911	41,903,377	77.41
	<b>2018</b>	2018-06-06	119,200	37,712,240	7,028	5,084,650	99,530,347	42,796,890	88.12
	<b>2019</b>	2019-06-04	83,000	31,435,567	11,323	7,586,661	103,300,714	39,022,228	80.56

The relation between the True Positive rate and the upstream runoff shows the trend that the accuracy of the modeled inundation boundaries is higher when the flood reaches the peak, and lower the lower the upstream runoff is. This trend is presented in Figure 42.

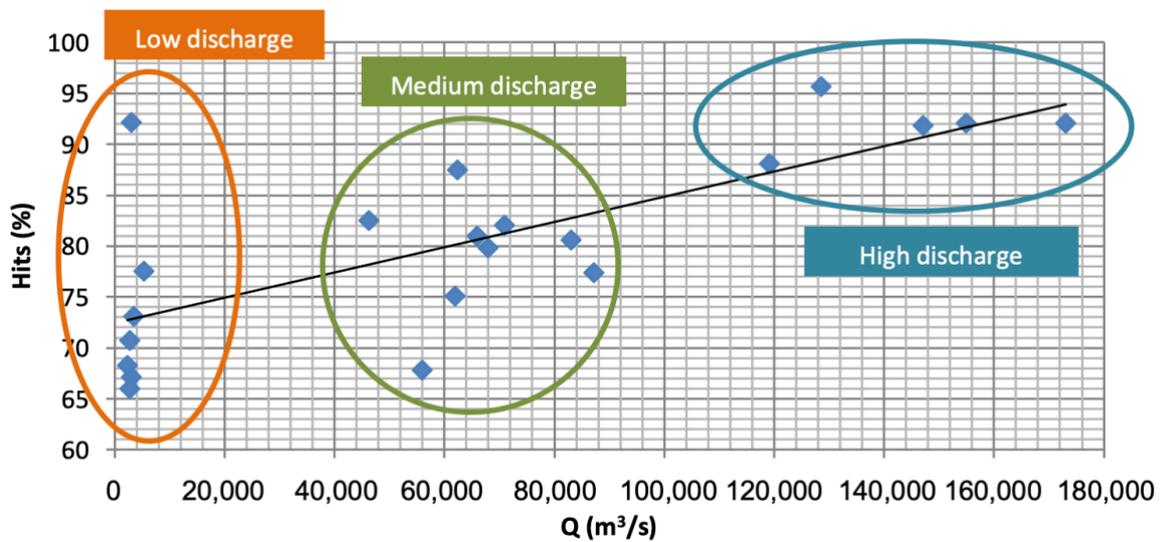
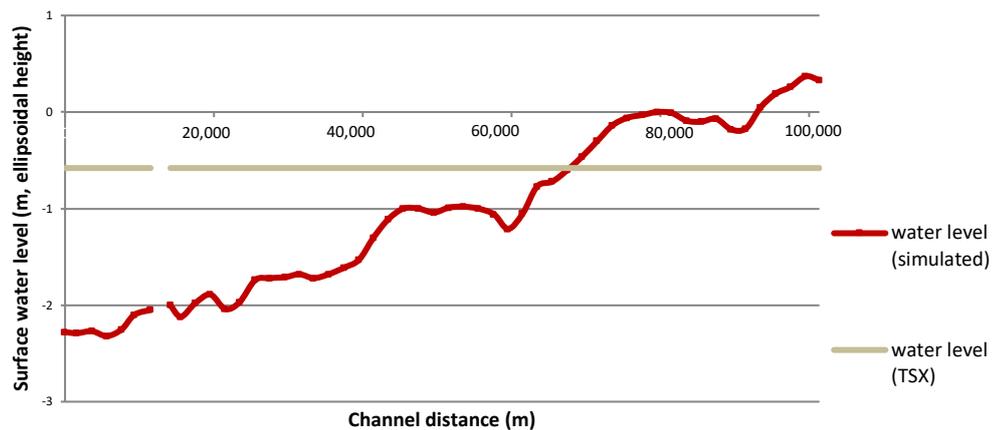


Figure 42 The upstream runoff plotted against the True Positive (Hits) rate (%)

The True Positive rate ranges from 65.97 to 95.65%. As previously mentioned, the model accuracy depends highly on 1) assumptions and limitations of the model, 2) accuracy of the geometric data, 3) accuracy of the flow data and BC, and 4) numerical accuracy of the solution scheme. This range can be explained with the set of limitations on modeling the flood, as follows:

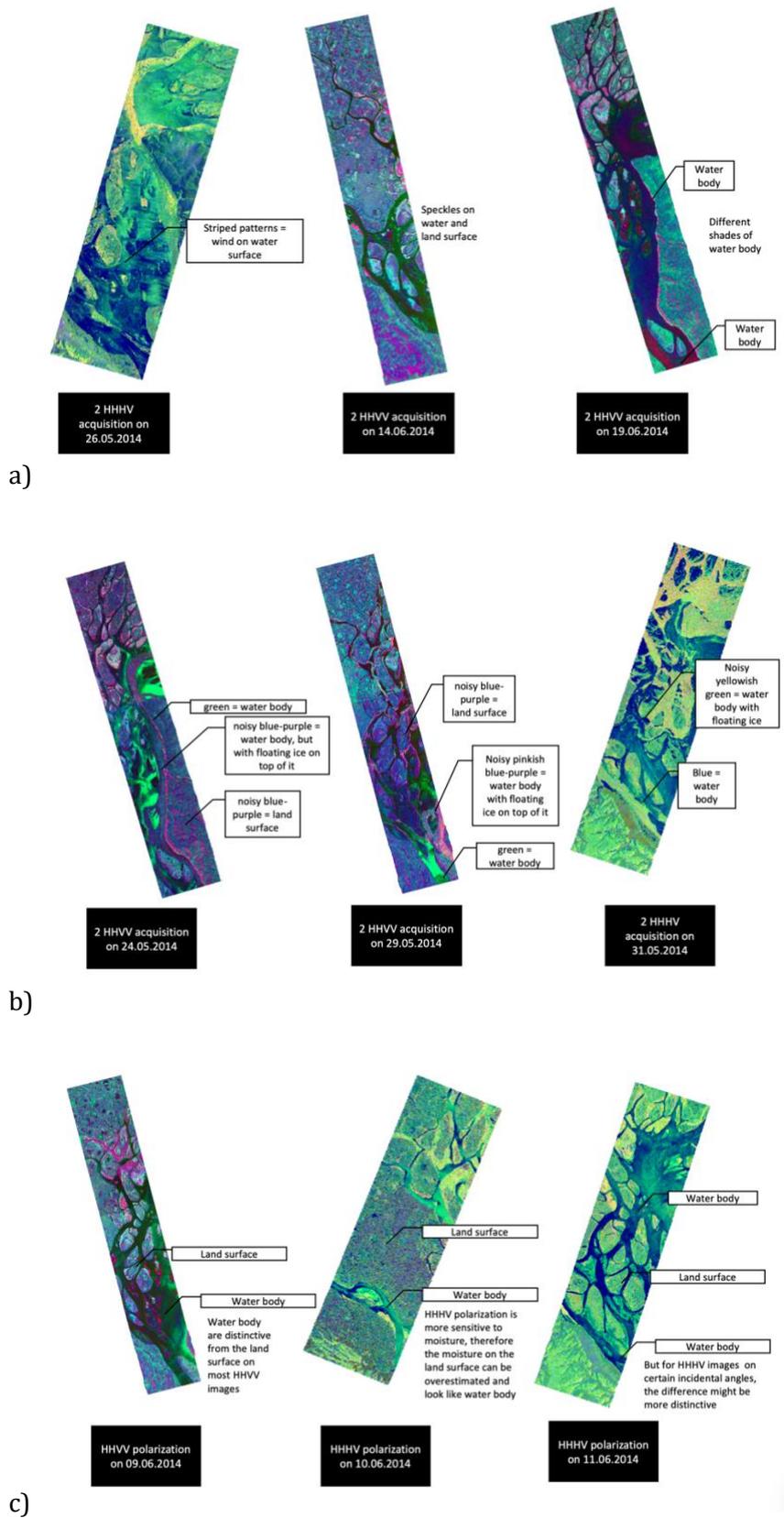
- the modeled inundation boundary was validated with the flood map that was generated by assuming that the surface water level along the channel was flat
- the flood extent from TSX Stripmaps was easier to observe during high flood
- ice was not modeled
- neglected weather impacts
- the assumption on the flow data
- simplifying the geometry and the flow direction
- the approximated Manning's roughness coefficient
- the inexistence of the real bathymetry data

The modeled inundation boundary was validated with the flood map that was generated by assuming that the surface water level along the channel was flat from the upstream to the downstream. Meanwhile, the modeled surface water level was not flat. Figure 43 shows an example of the comparison between the surface water profile acquired from the simulation and from the TSX-TDX-generated flood map of the medium upstream runoff event in 2014 (2014-06-11).



**Figure 43** The stream centerline surface water profile plot of the simulation result compared to that of the TSX-TDX-generated flood map of the medium upstream runoff 2014. Upstream of the channel is located on the right

The flood extent during the peak flood was easier to be observed than that of during the non-peak flood. The flood map depended highly on eyesight-based observation of the TSX Stripmaps. The Stripmaps quality differ from day to day depending on the wind condition, the existence of snow/ice cover on the surface, the incidental angle, and the polarization (previously described in 3.1.1.2). Strong wind created ripples on the water body, making the surface to have striped patterns on the image. The striped patterns blur out the flood boundary on the image. The wind creating ripples on the TSX acquisition can be seen in Figure 44a. Snow/ice covers reflects differently to the sensor. The existence of snow/ice cover on top of the water makes it difficult to see the flood extent as the backscatter looks different to the water body, but might look similar to the snow/ice-covered land surface. This situation can be seen in Figure 44b. Ice/snow cover still exists in early May, during the low upstream runoff. The flood map fitting accuracy might have been affected by this. During the high upstream runoff, the ice has already melted, making the water surface easier to be identified. Lower incidental angles are more sensitive to surface roughness height, which might lead to more speckles on the image. The speckles might mislead the observation given to the noise both on the water and land surface. The speckles on the image can be seen in Figure 44a. HHHV polarization is more sensitive to moisture than HHVV polarization, therefore the moisture on the land surface can be overestimated and look like water body on the HHHV Stripmaps. An HHHV acquisition is presented in Figure 44c.



**Figure 44 Challenges in visual interpretation for flood map generation using TSX Stripmaps: a) varying image quality due to speckles from wind or other factors, b) ice/snow cover, and c) HHHV polarization**

The ice and snow were neglected in the simulation. The accumulated snow and ice in the previous winter melt and contribute to a certain amount of water volume to the floodplain and the river. The flow input of the simulation only depended on the upstream runoff. During the flow from the upstream to the downstream, the ice and snow on the top of the river and floodplain melts and add water volume to the flow. Table 24 shows that the True Positive rates are always below 100%, meaning the inundation boundaries were mostly underestimated. The loss of water volume from the ice and snow in the model can affect this underestimation.

The modeled study area has a massive extent, however the weather effect is neglected in the simulation. Through the camera on the observation center, it can be seen that this area rains a lot throughout the year. The accumulated precipitation might also add water volume to the river and floodplain, which is missing from the simulation.

The upstream hydrograph for the model input was taken from the gauged runoff in Kyusyur, which was located 124.65 km upstream of the starting point of the simulated area (Figure 32). Although the channel shape doesn't seem to change between these two points, there might be weather or geographical effects that might take effects on the water runoff.

Due to the massive extent of the study area, the geometry was simplified into 1D geometrical lines. This results in the loss of details of such complex braided river system. The complex reaches were modeled as 1 channel. The flow is modeled to be one-directional, whereas in real life the flow direction varies on each reach. This, too, leads to different water distribution in real life compared to the model.

As there was no real data of the surface roughness values, the values were approximated based on the land cover classes. A common process to this is to approximate the surface roughness iteratively by running the simulation several times until the best fitting result is achieved. However, given the many other uncertainties in this study, the approximated Manning's roughness values were taken as it is. The input values that were taken through an iterative process are the channel bathymetry.

The last factor is the most important of all in regards to the accuracy. The inexistence of bathymetry data made it difficult to reach a better accuracy rate. The modeled area is big, therefore the heterogeneity of the channel shape must differ greatly. These details are missing from the simulation. The shape of the channel defines the distribution of the water flow from upstream to downstream.

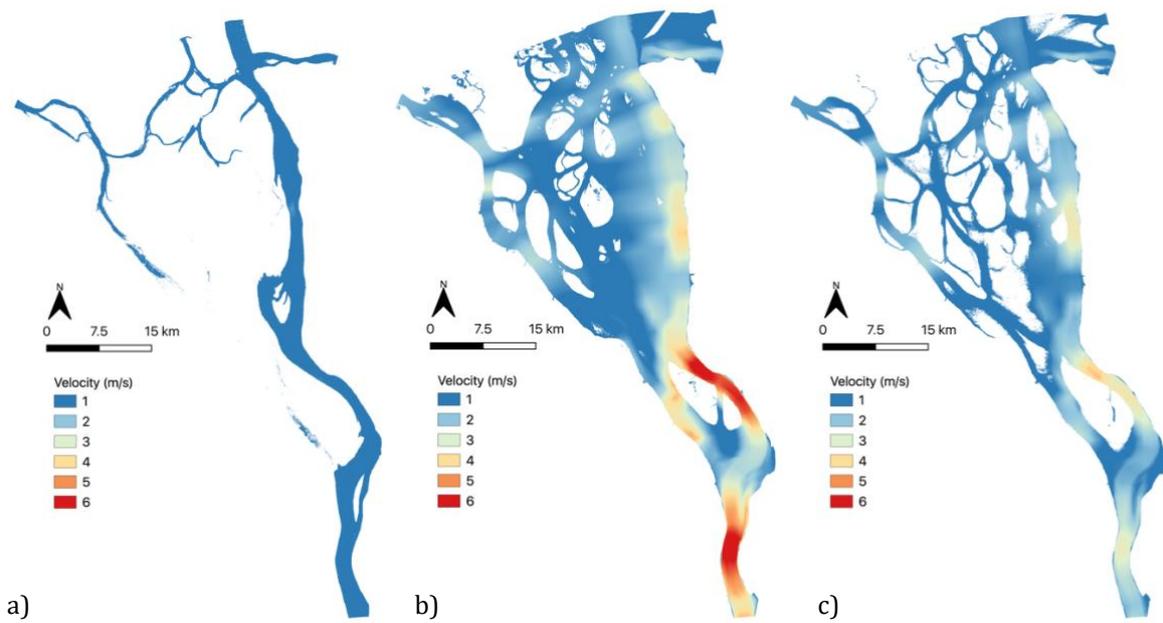
However, given the lack of data and information, the True Positive rate of 65-95% was good enough to represent that the methodology is applicable for such study. It indicates that with a

better data acquisition and more precise detailing, the simulation based on this approach can be improved.

## 6.6. Flow Velocity

The HEC-RAS 1D unsteady flow module compute the flow velocity on the main, left overbank, and right overbank channel separately with 1D flow direction. The water is only computed to flow from the upstream to the downstream in one direction (x), no velocity in the y- and z-direction is resulted. Figure 45 shows the resulting flow velocity computed during the 2014 flood event. Figure 45a shows the flow velocity in early May, before the thawing process started when the upstream runoff is still low. Figure 45b shows the flow velocity during the peak of the flood in late May. Figure 45c shows the flow velocity when the flood is receding, with the medium upstream runoff. Through this comparison, it can be seen that the flow velocity rose at some parts of the channel during the peak of the flood, and decreased again when the flood is receding.

The modeled flow velocity before the flood event was lower than 1.5 m/s. Then the velocity built up to 6 m/s on some parts of the channel during the peak of the flood. The velocity went down again as the flood was receding. The change in velocity from low to high might be linked to the sediment transport that was reported to often occur during the Spring flood. However, in order to quantify and analyze it thoroughly, a further sediment modeling is needed.



**Figure 45 The flow velocity of the 2014 Spring flood event during a) low upstream runoff (early May), b) high upstream runoff (late May), and c) medium upstream runoff (mid June)**

## 6.7. Flood Depth

Figure 46 shows the simulated water depth on the peak flood event in 2014. The flood depth on the floodplain varied from 0.001 to 2.5 m. Presented on the subset is the flood depth on the Samoylov island, where the observation center exists. Most part of the island are not flooded, only the East to Northeast part was slightly flooded to the depth of 2 m.

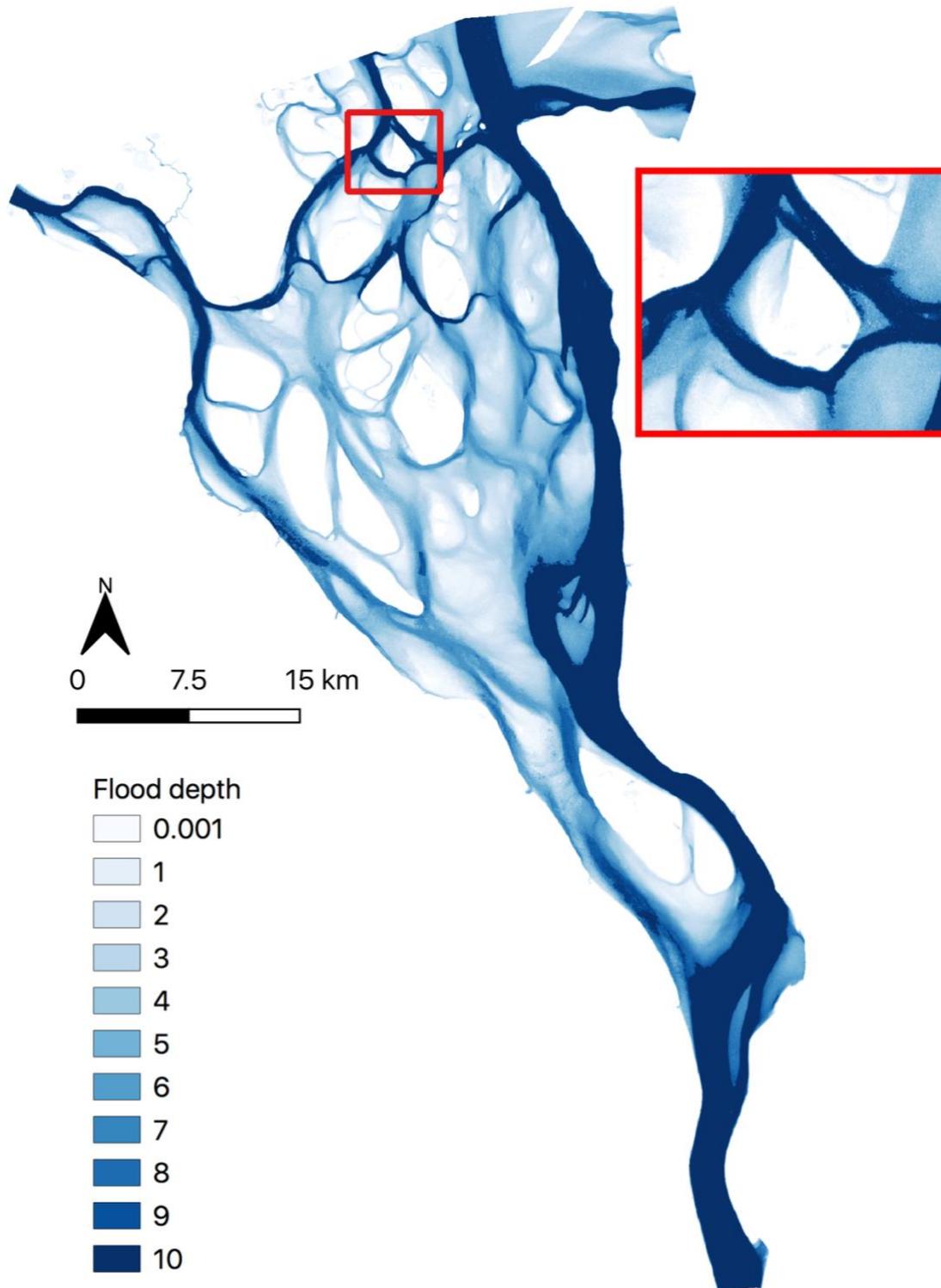


Figure 46 The flood depth of the peak flood in 2014 (the flood with the highest recorded upstream runoff during the simulation period)

## 7. Conclusions and Future Works

### 7.1. Conclusions

As part of the ongoing permafrost river delta study in the DLR DFD, this thesis is a feasibility study for incorporating hydraulic modeling into remote sensing method that is solely based on TerraSAR-X (TSX) and TanDEM-X (TDX) products. The remote study area, Lena delta in Siberia, is captured by TSX and TDX almost on daily basis, with the gap of 2-5 days. Through these products, the flood map, topography, and land cover were derived. However, the available field data was not sufficient to complement the simulation parameter input. The available field datasets were the gauged upstream runoff, downstream surface water level, and land cover information. The river bathymetry, surface roughness, and friction slope had to be approximated.

The parameter approximation came to the expense of varying model accuracy, ranging from 65 to 95%. The bathymetry and friction slope were approximated iteratively in regards to the simulation result. The values with the best fitting result were taken. The main channels were assumed to be trapezoidal-shaped with the maximum depth of -10 m, with the maximum depth almost flat from the upstream to the downstream. The friction slope was set to 0.0001 to match the overall riverbed slope, in order to stabilize the unsteady simulation. The Manning's surface roughness coefficients were assigned to each classified land cover class. Given that the bathymetry and the friction slope were iteratively approximated, the Manning's surface roughness coefficients were taken as it is. However, considering these results were achieved with many approximations of parameters with high uncertainties, the approach was considered applicable. This means, with better field data acquisition, the accuracy of the model can be pushed even higher, with better calibration and validation processes.

The land cover classification was performed with two consecutive Machine Learning methods: 1) iso-clustering to define the classes and 2) Random Forest to classify the whole raster image. The classified classes are sandy floodplain, sandy riverbed, Holocene terraces, vegetated floodplain (Pleistocene terraces), densely-vegetated floodplain, and rock. The most misclassified classes are between the vegetated floodplain and densely floodplain. The overall accuracy ranged between 93-95% for the classification based on TSX-RE bands, and was precisely 75% for the classification based on LS8 bands. The band with the highest importance for the classification is the Near-infrared (NIR) band (both of RE and LS8). The bands with the lowest importance were the Kennaugh element K3 (the absorption band of the TSX) and coastal aerosol band (of LS8).

The 1D unsteady flow simulation was highly sensitive to the cross-sectional line spacing ( $\Delta x$ ), computational time-step ( $\Delta t$ ), and the downstream slope. The simulation was stabilized with the  $\Delta x$  of 2 km. With shorter distance, the energy grade line spiked and resulted in rising oscillating flood wave upstream, making the simulation unstable. The  $\Delta t$  was easier to be approximated, based on the average velocity obtained from the steady state simulation, ranging from 6-40 minutes. The downstream slope of 0.0001 worked to keep the simulation stable. Bigger slopes lead to the rise of oscillating flood wave and instability.

The model accuracy was assessed by comparing the modeled inundation boundaries to the flood maps. The accuracy was quantified with the True Positive (Hit) rate ( $\frac{\text{True Positives}}{\text{Actual flooded pixels}}$ ). The accuracy ranged from 65-95%. The relation between the accuracy and the upstream runoff shows the trend that the accuracy is higher when the flood reaches the peak, and lower the lower the upstream runoff is. The modeled inundation boundaries tend to be underestimated. The varied accuracy rate was caused by the 1) bias in generating the flood map, 2) neglecting ice in the simulation, 3) neglecting weather impacts, 4) flow data not recorded exactly at the simulated upstream, 5) simplifying the geometry, 5) Manning's roughness coefficient not calibrated, and 6) the inexistence of the real bathymetry data.

Flood depth and flow velocity were extracted from the simulation result. During the flood event 2014, which was the most severe Spring flood during the simulation period, the flood depth on the floodplain varied from 0.001 to 2.5 m. The flow velocity rose at some parts of the channel during the peak of the flood (from 1 m/s up until 6 m/s), and decreased again when the flood is receding. This might be linked to the high sediment transport that often occur during the Spring flood. However, sediment modeling is needed in order to quantify and analyze the transport thoroughly.

The challenges in this thesis were: 1) processing big data, 2) utilizing many different tools that were not always well-integrated with each other, 3) approximations without real data for validation, 4) complicated method and long computing time for the bathymetry interpolation, and 5) post-processing the simulation result.

The TSX and TDX products have high spatial resolution (5 m). Given the massive extent of the study area, the file sizes of these raster files are also massive. Loading and processing these products take time, and some processes need to be redone in order to acquire different result, adding up to a long image pre-processing time. Same applies to the RE imagery.

GIS tools were needed to process the remotely-sensed and geospatial datasets. ArcGIS and QGIS were used interchangeably due to the different features these software offer. However,

oftentimes the offered features do not work on specific datasets and alternative solutions had to be applied. Specific dataset details were not always interpreted the same across these two platforms as well. A lot of adjustments had to be taken in order to make it work. Python batch processing worked really well on ArcGIS, but was quite buggy in QGIS due to many constraints in the library.

Without data for validation, the unknown parameters had to be approximated blindly. A long process from the approximation to the geospatial data processing and hydraulic simulation had to be performed in order to prove whether or not those parameters could be used.

The most time- and resource-consuming work in this thesis was the bathymetry approximation. Digitizing the contour lines, converting the contour lines into points, modifying the feature fields, segmenting the interpolation area, performing IDW interpolation, mosaicking the segmented interpolation result, and then testing the bathymetry for the simulation took a lot of time. These processes were repeated for different contour line configuration in order to get the best-fitting simulation result.

The computational time for the HEC-RAS unsteady simulation was short due to the simplified 1D geometry, however the post-processing of the resulting geospatial datasets took a much longer time, especially for the inundation boundary shapefiles. The results were rendered according to the DEM spatial resolution (5 m), resulting in raster and shapefiles of large sizes. One shapefile took roughly 1 hour to be rendered. The raster files took a shorter time to be rendered, around 15-30 minutes each. However, loading and post-processing these geospatial datasets for analysis purposes took long time, CPU load, and memory pressure.

By analyzing the processes and results of this feasibility study, the strengths and weakness of this approach can be concluded. The strengths are: 1) the potential to study the hydraulic processes of a remote area with limited gauged data, 2) the potential to perform hydraulic modeling of a massive study area extent, 3) high spatial resolution of TSX and TDX products giving a very detailed information of the study area, and 4) a proposed method to interpolate the bathymetry elevation in order to create a 2D bathymetric information. The weaknesses are: 1) the important data (bathymetry and surface roughness) were not available, 2) no data for calibration and validation of the hydraulic modeling, 3) high uncertainties due to many approximations of the important parameter inputs, and 4) the flood maps that were used to assess the model accuracy were generated based on visual interpretation. The solutions to improve and extend this study in the future will be described in the next subchapter.

## 7.2. Future Works

Both the remote sensing method and the hydraulic simulation can be improved with 1) better field data acquisition, 2) more precise flood mask generation, 3) coupling the hydraulic modeling with the hydrological modeling, and 4) performing 2D/3D simulation. As a starting point, this thesis already highlighted what need to be performed in order to obtain the sought-after hydraulic parameters. Improvements on the simulation input, calibration, and validation will significantly rise the model accuracy.

The most important key factor to improving the simulation is by having complete input datasets. As previously mentioned, the channel shape and bathymetry affects greatly how the flow is distributed from the upstream to the downstream of the modeled area. Previous studies utilizing TDX DEM incorporate LIDAR-sensed bathymetry (Heimhuber 2013, Krötzinger 2015). Other than the bathymetry, a real field-acquired surface roughness data will also benefit the simulation. A grain size can be a good approach into estimating the Manning's surface roughness values (Ab Ghani et al. 2007). Another alternative would be to measure the soil moisture, that can be used as limiting parameters for TSX-derived roughness height (Aubert et al. 2011, Baghdadi et al. 2011, Sadeh et al. 2018). Adding the weather data into the simulation input can also add the missing water volume from precipitation and ice thawing. This can cover for the water volume loss that cause the underestimation.

One of the uncertainties in this thesis was the datasets that were used to assess the accuracy; the flood maps. The multi-temporal flood maps were generated by fitting the flooding increment into the TSX based on the eyesight. The observed flood extent was highly biased and depended on the quality of the TSX Stripmap acquisition. The flood map generation can be improved with other methods such as object-based land cover classification. With object-based classification, the features such as texture, shape, among others can be extracted therefore the classification is not only based on the intensity values of the TSX, which are highly sensitive to the wind, incidence angle, and polarization.

With the availability of the weather and ice/snow data, the hydraulic simulation can be coupled with hydrological modeling. Hydrological modeling can be used to process the precipitation and river-ice interaction. Given the massive extent of the study are, the accumulated water volume from the precipitation and ice thawing might be of significant amount.

A more precise and detailed hydraulic simulation method can be performed with better field data availability. 2-dimensional or even 3-dimensional simulation can be performed. With 2D simulation, the flow in the x- and y-direction can be computed. With 3D simulation, the z-

direction is included into the equation. The complicated braided river system would benefit from this better-detailed simulation. However, with such large area, the computation time and effort would rise exponentially. Better computing resources, funding, and longer study period will be required.

Acquiring flood depth and flow velocity from the Spring flood events in Lena delta is a starting point for the permafrost river delta study. In the future, this simulation can be extended to further studies, such as sediment transport, water quality, freshwater-seawater interaction, and ecological modeling, among others. The study can also be integrated with climate-related researches, given the big impacts the arctic rivers impose on the ocean conveyor belt.

## List of Figures

Figure 1 The Lena delta study area (OpenStreetMap, Pertiwi) .....	6
Figure 2 An aerial photo of the Lena delta.....	7
Figure 3 The Spring flood event 2014 on the Samoylov island, before and during the peak flood, recorded by the camera.....	8
Figure 4 An illustration of the swath width comparison between TerraSAR-X (purple and turquoise), RapidEye (red), and Landsat 8 (light green) .....	14
Figure 5 Artist view on the TSX and TDX mission (AIRBUS 2015).....	15
Figure 6 The TDX was generated from TSX/TDX acquisitions (marked in black vertical lines) during low surface water level (recorded at Khabarova station) .....	16
Figure 7 TerraSAR-X image acquisition modes (DLR 2009).....	17
Figure 8 a) Horizontal and vertical polarization (JMA 2016); b) Satellite incidence angle (ESA 2014) .....	17
Figure 9 Different coverage of TSX Stripmap orbits on the study area.....	19
Figure 10 The RapidEye system consists of five satellites (DLR 2008).....	20
Figure 11 Landsat 8 satellite platform (EOS 2019) .....	21
Figure 12 Daily gauged river discharge at the Kusyur station in 2014.....	22
Figure 13 Daily gauged water level at the Khabarova station in 2014 .....	23
Figure 14 The land cover information of center-north part of the study area (Heim 2019) ....	24
Figure 15 The combinations of defined spatial details and spectral details for mean-shift segmentation (ESRI 2015) .....	28
Figure 16 The rock island before and after the interpolation .....	29
Figure 17 Riverbed changes of between Sardakh and Trofimovskaya islands between 1948-2011. Units are in meters (Fedorova et al. 2015) .....	30
Figure 18 Lena river near Kyusur by Max Holmes (ArcticGRO 2003).....	31
Figure 19 The noisy cross-section profiles derived from an attempt to approximate the channel profile using SDB method .....	31
Figure 20 The generated bathymetry contour lines.....	33
Figure 21 The flood map fitting process: a) the DEM, b) a TSX Stripmap on top of the DEM, c) fitting the DEM flood map from the collections of flood map into the TSX Stripmap .....	35
Figure 22 Iso-clustering of point data with two features (each on separate axis) (Klie 2017)	36
Figure 23 Random Forest classification scheme (Koehrsen 2017) .....	37
Figure 24 Creating training and test samples.....	38
Figure 25 Extracting the classes and raster values for the training and test datasets, a) reading the class ID from sample point; b) taking out raster values on the sample points .....	38

Figure 26 a) training the Random Forest classifier; b) testing the Random Forest classifier ..	39
Figure 27 Classifying the a) TSX-RE Composite and b) LS 8 bands .....	39
Figure 28 A 1D HEC-RAS geometry of a braided river system (Caruso et al. 2013): a) the shape of the braided river system, the red line is the stream center line; b) the cross-section lines; c) the profile of a cross-section line.....	41
Figure 29 The generated HEC-RAS geometry lines in QGIS.....	43
Figure 30 A snippet of the beginning and end of the SDF file .....	46
Figure 31 A Control Volume.....	48
Figure 32 Adjusting the gauged discharge datasets.....	50
Figure 33 The HEC-RAS geometric data view .....	52
Figure 34 Cross-section 63216.78 before and after the elevation points were reduced to 500 .....	54
Figure 35 The HTab parameters, with cross section 63216.78 shown on the right.....	58
Figure 36 The DEM corrected with generated trapezoidal channel with the maximum depth of -10 m. The terrains higher than 10 m were not shown as those were irrelevant to the simulation.....	61
Figure 37 The land cover classification map obtained from the stack of TSX and RapidEye datasets .....	62
Figure 38 The Variable Importance Measures (VIM) of TSX bands and RE bands for the Random Forest-based land cover classification.....	66
Figure 39 The land cover classification map obtained from the LS8 datasets .....	67
Figure 40 The Variable Importance Measures (VIM) of LS8 bands for the Random Forest-based land cover classification .....	69
Figure 41 Unstable result due to cross-section lines being too close together (in the picture, $\Delta x = 50\text{m}$ ).....	70
Figure 42 The upstream runoff plotted against the True Positive (Hits) rate (%).....	74
Figure 43 The stream centerline surface water profile plot of the simulation result compared to that of the TSX-TDX-generated flood map of the medium upstream runoff 2014. Upstream of the channel is located on the right.....	75
Figure 44 Challenges in visual interpretation for flood map generation sing TSX Stripmaps: a) varying image quality due to speckles from wind or other factors, b) ice/snow cover, and d) HHHV polarization .....	76
Figure 46 The flow velocity of the 2014 Spring flood event during a) low upstream runoff (early May), b) high upstream runoff (late May), and c) medium upstream runoff (mid June) .....	79

Figure 45 The flood depth of the peak flood in 2014 (the flood with the highest recorded upstream runoff during the simulation period) .....80

## List of Tables

Table 1 The hydraulic parameters derived from the TerraSAR-X and TanDEM-X products, the hydraulic simulation, and the unknown parameters in this thesis .....	5
Table 2 Summary of the common information of Lena delta .....	7
Table 3 The study limitations and assumptions .....	9
Table 4 The used tools for this thesis and the license source of each tool .....	11
Table 5 The advantages and disadvantages of the utilized TerraSAR-X, RapidEye, and Landsat 8 imageries.....	12
Table 6 Summary of the specifications of the TerraSAR-X, RapidEye, and Landsat 8 datasets used in this study .....	13
Table 7 The specification of the utilized TerraSAR images .....	18
Table 8 RapidEye bands.....	20
Table 9 The Landsat 8 L2 surface reflectance bands (Barsi et al. 2014) .....	21
Table 10 Land covers and the corresponding roughness height (Heim 2019) .....	24
Table 11 TerraSAR-X Kennaugh elements.....	26
Table 12 Kennaugh element combinations for HHVV and HHHV polarization.....	26
Table 13 Geodatabase HEC-RAS geometry table attributes .....	45
Table 14 The low, high, and medium discharge values for the simulation .....	51
Table 15 Steady flow data input.....	55
Table 16 The average velocity and the computational time step of each simulation setup .....	57
Table 17 The simulation parameters .....	59
Table 18 The confusion matrix of the TSX-RE west tile land cover classification .....	63
Table 19 The confusion matrix of the TSX-RE center tile land cover classification .....	64
Table 20 The confusion matrix of the TSX-RE east tile land cover classification .....	65
Table 21 The confusion matrix of the LS8 land cover classification .....	68
Table 22 The Manning's roughness value of the channel and floodplain .....	69
Table 23 Two-class confusion matrix .....	72
Table 24 The values of the Hits, False Alarms, Misses, Correct Negatives, and the True Positive Rate of the modeled inundation boundaries and the TSX-TDX-generated flood maps.....	73

# List of Abbreviations

<b>Abbreviation</b>	<b>Definition</b>
ArcticGRO	Arctic Great Rivers Observatory
AWI	Alfred-Wegener-Institut
BMBF	German Ministry of Education and Science
CEIWR-HEC	US Army Corps of Engineers Hydrologic Engineering Center
CV	Control Volume
DEM	Digital Elevation Model
DFD	German Remote Sensing Data Center Institute
DLR	German Aerospace Center
DSM	Digital Surface Model
DTM	Digital Terrain Model
EOC	Earth Observation Center
GIS	Geographic information system
Htab	Cross-section Table Properties
IDW	Inverse Distance Weighted
LS8	Landsat 8
LULC	Land Use/Land Cover
Minprirody	Russian Ministry of Natural Resources and Environment
NIR	Near-infrared
OLI	Operational Land Imager
RE	RapidEye
RF	Random Forest
SAR	Synthetic Aperture Radar
SDB	Satellite-derived Bathymetry
SV	Saint-Venant
TDX	TanDEM-X
TIRS	Thermal Infrared Sensor
TP	True Positive
TSX	TerraSAR-X
TUM	Technical University of Munich
USGS	US Geological Survey
VIM	Variable Importance Measure

## List of Symbols

$DN$	Digital Number
$\beta^0$	Radar brightness
$k_s$	Calibration factor
$\sigma^0$	Sigma naught
$\gamma^0$	Gamma naught
$Z$	The actual satellite derived water depth
$m_1$	A tunable constant to scale the ratio to depth
$n$	A fixed constant for all areas
$R_w(\lambda_i)$	The reflectance of the water of the band with lower absorption (blue band)
$R_w(\lambda_j)$	The reflectance of the water of the band with higher absorption (green band)
$m_0$	The offset for a depth of 0 m ( $Z = 0$ )
$Q$	Discharge
$A$	Cross-sectional flow area
$V$	Flow velocity
$q$	Lateral in-/outflow per unit length
$Q(x, t)$	Flow at the midpoint of the CV
$\frac{\partial Q}{\partial t}$	The local acceleration term, that is the change in momentum due to the change in flow velocity over time
$\frac{\partial QV}{\partial x}$	The convective acceleration term, that is the change in momentum due to the change in flow velocity along the channel
$g$	The gravity force term, that is proportional to the bed slope
$S_0$	The bed slope
$\frac{\partial h}{\partial x}$	The pressure force term, that is proportional to the change of depth along the channel
$h$	Surface water level
$S_f$	The friction force term, that is proportional to the friction slope
$\Delta x$	Distance between cross-sectional lines
$D$	Bankfull depth
$S$	Slope
$V_w$	Flood wave speed
$\Delta t$	Computational time-step

## References

- Ab Ghani, Aminuddin, Nor Zakaria, Chun Kiat Chang, Junaidah Ariffin, Zorkeflee Abu Hasan, and Ahmad Ghaffar. 2007. "Revised equations for Manning's coefficient for Sand-Bed Rivers." *International Journal of River Basin Management* 5:329-346. doi: 10.1080/15715124.2007.9635331.
- AIRBUS. 2015. *TerraSAR-X Image Product Guide: Basic and Enhanced Radar Satellite Imagery*: Airbus Defence and Space.
- ArcticGRO. 2003. "Arctic Great Rivers." <https://arcticgreatrivers.org/rivers/>.
- Aubert, M., N. Baghdadi, M. Zribi, A. Douaoui, C. Loumagne, F. Baup, M. El Hajj, and S. Garrigues. 2011. "Analysis of TerraSAR-X data sensitivity to bare soil moisture, roughness, composition and soil crust." *Remote Sensing of Environment* 115 (8):1801-1810. doi: <https://doi.org/10.1016/j.rse.2011.02.021>.
- Bachofer, Felix, Geraldine Quénéhervé, Volker Hochschild, and Michael Maerker. 2015. "Multisensoral Topsoil Mapping in the Semiarid Lake Manyara Region, Northern Tanzania." *Journal of Remote Sensing* 7:9563-9586. doi: 10.3390/rs70809563.
- Baghdadi, Nicolas, Pauline Camus, Nicolas Beaugendre, Oumarou Issa, Mehrez Zribi, Jean Franç Ois Desprats, Jean Rajot, Chadi Abdallah, and Christophe Sannier. 2011. *Estimating Surface Soil Moisture from TerraSAR-X Data over Two Small Catchments in the Sahelian Part of Western Niger*. Vol. 3.
- Barsi, J.A., K. Lee, G. Kvaran, B. L. Markham, and J. A. Pedelty. 2014. "The Spectral Response of the Landsat 8 Operational Land Imager." *MDPI - Remote Sensing* 6 (10):10232-10251.
- Boike, Julia, Jan Nitzbon, Katharina Anders, Mikhail N Grigoriev, Dimitry Yu Bolshiyarov, Moritz Langer, Stephan Lange, Niko Bornemann, Anne Morgenstern, Peter Schreiber, Christian Wille, Sarah Chadburn, Isabelle Gouttevin, and Lars Kutzbach. 2019. *Measurements in soil and air at Samoylov Station (2002-2018), version 201908*.: Alfred Wegener Institute - Research Unit Potsdam, PANGAEA.
- Brunner, Gary W. 2011. *Common Model Stability Problems When Performing an Unsteady Flow Analysis*: Hydrologic Engineering Center.
- Brunner, Gary W. 2016. *HEC-RAS River Analysis System: Hydraulic Reference Manual*: US Army Corps of Engineers.
- Brunner, Gary W., and CEIWR-HEC. 2016. *HEC-RAS River Analysis System User's Manual Version 5.0*. Davis, CA: US Army Corps of Engineers.
- Caruso, Brian, Alex Ross, Claire Shuker, and Tim Davies. 2013. "Flood Hydraulics and Impacts on Invasive Vegetation in a Braided River Floodplain, New Zealand." *Environment and Natural Resources Research* 3 (1). doi: 10.5539/enrr.v3n1p92.
- CEIWR-HEC. 2019. "Hydrologic Engineering Center." accessed 23 May 2019. <https://www.hec.usace.army.mil/software/hecras/>.
- Chow, V. T., D. R. Maidment, and L. W. Mays. 1988. *Applied Hydrology*. New York: McGraw-Hill.
- Chow, V.T. 1959. *Open-channel hydraulics*. New York: McGraw-Hill Book Co.
- Curtarelli, Marcelo, Joaquim Leão, Igor Ogashawara, João Lorenzetti, and José Stech. 2015. "Assessment of Spatial Interpolation Methods to Map the Bathymetry of an Amazonian Hydroelectric Reservoir to Aid in Decision Making for Water Management." *ISPRS International Journal of Geo-Information* 4:220-235.

- DLR. 2008. "RapidEye." accessed 15 July. [https://www.dlr.de/rd/en/desktopdefault.aspx/tabid-2440/3586\\_read-5336/](https://www.dlr.de/rd/en/desktopdefault.aspx/tabid-2440/3586_read-5336/).
- DLR. 2009. *TerraSAR-X: The German Radar Eye in Space*. Bonn, Germany: DLR Raumfahrt-Agentur.
- DLR. 2019. "TerraSAR-X/TanDEM-X gallery." accessed 23 May 2019. <https://www.dlr.de/dlr/en/desktopdefault.aspx/tabid-10422/>.
- Domeneghetti, Alessio. 2016. "On the use of SRTM and altimetry data for flood modeling in data sparse regions." *WATER RESOURCES RESEARCH*. doi: 10.1002/2015WR017967.
- EOC, Geoservice. 2019. "The TanDEM-X 90m Digital Elevation Model." EOC, DLR, accessed 29 April 2019. <https://geoservice.dlr.de/web/dataguide/tdm90/>.
- EOS. 2019. "LANDSAT 8." accessed 13 November. <https://eos.com/landsat-8/>.
- ESA. 2014. "Geometry Glossary." accessed 3 September. <https://earth.esa.int/handbooks/asar/CNTR5-5.html>.
- ESA. 2019a. "Lena River Delta." accessed 13 November. [http://www.esa.int/ESA\\_Multimedia/Images/2019/06/Lena\\_River\\_Delta](http://www.esa.int/ESA_Multimedia/Images/2019/06/Lena_River_Delta).
- ESA. 2019b. "RapidEye Earth Observation Constellation." accessed 13 November. <https://earth.esa.int/web/eoportal/satellite-missions/r/rapideye>.
- ESRI. 2015. "Pass the classification but hold the salt and pepper!", accessed 12 July. <https://www.esri.com/arcgis-blog/products/product/imagery/pass-the-classification-but-hold-the-salt-and-pepper/>.
- ESRI. 2016. "How Iso Cluster works." accessed 6 October. <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-iso-cluster-works.htm>.
- Fedorova, Irina, A. Chetverova, D. Bolshiyarov, A. Makarov, J. Boike, B. Heim, A. Morgenstern, P. P. Overduin, C. Wegner, V. Kashina, A. Eulenburg, E. Dobrotina, and I. Sidorina. 2015. "Lena Delta hydrology and geochemistry: long-term hydrological data and recent field observations." *Biogeosciences* 12:345–363. doi: 10.5194/bg-12-345-2015.
- Gary W. Brunner, P.E. 2011. "Common Model Stability Problems When Performing an Unsteady Flow Analysis." accessed 1 August. <https://www.nws.noaa.gov/oh/hrl/modelcalibration/6.%20%20Hydraulic%20Model%20Calibration/4.1%20L-11%20CommonModelStabilityProblemsInUnsteady%20FlowAnalysis.pdf>.
- Goodell, Chris. 2008. "Samuel's Equation for Cross Section Spacing ", accessed 15 August. <http://hecrasmodel.blogspot.com/2008/12/samuels-equation-for-cross-section.html>.
- Goodell, Chris. 2012. "How to draw cross sections. ." accessed 1 August. <http://hecrasmodel.blogspot.com/2012/07/how-to-draw-cross-sections.html>.
- Heim, Birgit. 2019.
- Heimhuber, Valentin. 2013. "GIS Based Flood Modeling as Part of an Integrated Development Strategy for Informal Settlements." M.Sc, Chair of Urban Water Systems Engineering, TU Munich.
- Holmes, R.M., J.W. McClelland, S.E. Tank, R.G.M. Spencer, and A.I. Shiklomanov. 2018. "Arctic Great Rivers Observatory. Water Quality Dataset, Version 20190709." <https://www.arcticgreatrivers.org/data>.

- Hütt, Christoph, Wolfgang Koppe, Yuxin Miao, and Georg Bareth. 2016. "Best Accuracy Land Use/Land Cover (LULC) Classification to Derive Crop Types Using Multitemporal, Multisensor, and Multi-Polarization SAR Satellite Images." *Remote Sens.* 8. doi: 10.3390/rs8080684.
- JMA. 2016. "Transition work of the HimawariCast service." accessed 3 September. [https://www.data.jma.go.jp/mscweb/en/himawari89/himawari\\_cast/transition.html](https://www.data.jma.go.jp/mscweb/en/himawari89/himawari_cast/transition.html).
- Klie, Hector. 2017. "Going Beyond K-means For Effective Clustering Of Production Curves." DeepCast.ai, accessed 23 September. [https://deepcast.ai/article/2017/07/18/going\\_beyond\\_k\\_means\\_for\\_effective\\_clustering\\_of\\_production\\_curves](https://deepcast.ai/article/2017/07/18/going_beyond_k_means_for_effective_clustering_of_production_curves).
- Koehrsen, Will. 2017. "Random Forest Simple Explanation." accessed 23 September. <https://medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d>.
- Krötzinger, Wolfgang. 2015. "Flash Flood Modeling and Sediment Analysis for the Evaluation of Mitigation Measures at an Ephemeral Stream in Canaan, Haiti." M.Sc., Chair of Hydrology and River Basin Management, TU Munich.
- Leopold, Luna B., JR. Maddock, and Thomas. 1953. *The Hydraulic Geometry of Stream Channels and Some Physiographic Implications*. Washington: United States Government Printing Office.
- Ma, Xieyao, Tetsuzo Yasunari, Tetsuo Ohata, and Yoshihiro Fukushima. 2005. "The influence of river ice on spring runoff in the Lena river, Siberia." *Annals of Glaciology* 40.
- Marcus, W. Andrew, and Mark A. Fonstad. 2010. "Remote sensing of rivers: the emergence of a subdiscipline in the river sciences." *Earth Surface Processes and Landforms* 35 (15):1867-1872. doi: 10.1002/esp.2094.
- Mersel, Matthew K., Laurence C. Smith, Konstantinos M. Andreadis, and Michael T. Durand. 2013. "Estimation of river depth from remotely sensed hydraulic relationships." *WATER RESOURCES RESEARCH* 49:3165–3179. doi: 10.1002/wrcr.20176.
- Minprirody. 2019. "Automated Information System for State Monitoring of Water Bodies (translated from Russian)." accessed 23 May 2019. <https://gmvo.skniivh.ru/>.
- Papa, Fabrice, Catherine Prigent, and W. B. Rossow. 2008. "Monitoring Flood and Discharge Variations in the Large Siberian Rivers From a Multi-Satellite Technique." *Surveys in Geophysics*. doi: 10.1007/s10712-008-9036-0.
- Pavelsky, Tamlin Muir, and Laurence C. Smith. 2004. "Spatial and temporal patterns in Arctic river ice breakup observed with MODIS and AVHRR time series." *Remote Sensing of Environment* 93:328–338. doi: 10.1016/j.rse.2004.07.018.
- Peterson, Bruce J., Robert M. Holmes, James W. McClelland, Charles J. Vörösmarty, Richard B. Lammers, Alexander I. Shiklomanov, Igor A. Shiklomanov, and Stefan Rahmstorf. 2002. "Increasing River Discharge to the Arctic Ocean." *Science* 298:2171-2173.
- Sadeh, Yuval, Hai Cohen, Shimrit Maman, and Dan G. Blumberg. 2018. "Evaluation of Manning's n Roughness Coefficient in Arid Environments by Using SAR Backscatter." *Remote Sens.* 10. doi: 10.3390/rs10101505.
- Sakai, Toru, Shigemi Hatta, Makoto Okumura, Tetsuya Hiyama, Yasushi Yamaguchi, and Gen Inoue. 2015. "Use of Landsat TM/ETM+ to monitor the spatial and temporal

- extent of spring breakup floods in the Lena River, Siberia." *International Journal of Remote Sensing* 36 (3):719-733. doi: 10.1080/01431161.2014.995271.
- Samuels, P.G. 1989. "Backwater lengths in rivers." *Proceedings -- Institution of Civil Engineers* (87):571-582.
- Schmitt, Andreas, Anna Wendleder, and Stefan Hinz. 2015. "The Kennaugh element framework for multi-scale, multi-polarized, multi-temporal and multi-frequency SAR image preparation." *ISPRS Journal of Photogrammetry and Remote Sensing* 102:122-139. doi: <https://doi.org/10.1016/j.isprsjprs.2015.01.007>.
- Stumpf, Richard P., and Kristine Holderied. 2003. "Determination of water depth with high-resolution satellite imagery over variable bottom types." *American Society of Limnology and Oceanography* 48 (1):547-556.
- Sutter, A. 2008. "Fluvial Environments." accessed 12 November. [http://www.seddepeq.co.uk/DEPOSITIONAL\\_ENV/Fluvial/Fluvial.htm](http://www.seddepeq.co.uk/DEPOSITIONAL_ENV/Fluvial/Fluvial.htm).
- USGS. 2019a. "EarthExplorer - Home." accessed 20 May 2019. <https://earthexplorer.usgs.gov/>.
- USGS. 2019b. "Landsat Missions: Landsat 8." accessed 23 May 2019. <https://www.usgs.gov/land-resources/nli/landsat/landsat-8>.
- Vuglinsky, V. S. 2002. "Peculiarities of ice events in Russian Arctic rivers." *Hydrological Processes* 16 (4):905-913. doi: 10.1002/hyp.365.
- Wessel, B. 2016. *TanDEM-X Ground Segment – DEM Products Specification Document*. Oberpfaffenhoden, Germany: EOC, DLR. Public Document TD-GS-PS-0021.
- Zhang, Yanjun, Cuiling Xian, Huajin Chen, Michael L. Grieneisen, Jiaming Liu, and Minghua Zhang. 2016. "Spatial interpolation of river channel topography using the shortest temporal distance." *Journal of Hydrology* 542:450-462. doi: <https://doi.org/10.1016/j.jhydrol.2016.09.022>.