

Doctoral Dissertation

Standardized Automatic Flood Detection by Multi-Sensor Satellite Data Integration

(マルチセンサー衛星データ統合による標準化された洪水の自動検出)

March, 2022

Husniyah Binti Mahmud

Graduate School of Sciences and Technology for Innovation,
Yamaguchi University

ACKNOWLEDGEMENT

Throughout the writing of this dissertation I have received a great deal of support and assistance. I would like to express my deepest gratitude to all those who made my dissertation possible either directly or indirectly. I would like to express my deepest thanks to my supervisor Prof Masahiko Nagai for his guidance, caring and patience. His invaluable help suggesting, commenting throughout this experimental even during his busiest hours have contributed to the success of this research. Not forgotten my appreciation to Prof Takahiro Osawa, Prof Koji Asai, Dr Suzuki Motoyuki, and Dr Akamatsu Yoshihisa for guiding me and gives an idea to improve the dissertation.

My sincere thanks to Dr Vaibhav Katiyar and Dr Noppawan for advising and mentoring me during this project work. I am most grateful to staffs from government agency Department of Irrigation and Drainage, JAXA, European Space Agency, and the United States Geological Survey for their generosity providing all the data.

Last but not least, I would also like to thank my beloved mom, sister and brother. They were always supporting me and encouraging me with their best wishes. Finally, I could not have completed this dissertation without the support of my friends who provided happy distractions to rest my mind outside of my research. And above all, I thank Allah S.W.T for His blessing and giving me the strength to complete this project.

PREFACE

This dissertation is being submitted to Yamaguchi University's Graduate School of Science and Technology for Innovation in partial fulfillment of the requirements for the Doctor of Engineering degree. The following are the publications that the author wrote during his PhD studies:

[1] Peer-reviewed Conference Paper

a. Automated Extraction of Flood for Large Scale Area Using Weight Average Otsu's Method from Alos-2 Dual Polarization and MODIS

Husniyah Binti Mahmud, Masahiko Nagai

Proceeding: The 40th Asian Conference on Remote Sensing (ACRS 2019) October 14-18, 2019. Daejeon Convention Center (DCC), Daejeon, Korea, 2019.

Author Contributions: Husniyah Binti Mahmud carried out the analysis, acquisition of the data, analysis and interpretation of the data and drafting the manuscript. Masahiko Nagai participated in study design and critical revision of the manuscript. All authors read and approved the final manuscript.

Conflict of Interest: The authors declare no conflict of interest in preparing this publication.

Publication: ISBN 978-1-7138-0326-3, Volume 4, Page 2505.

b. Classification of Surface Water Area from Multispectral and MultiSAR Data Using Automatic and Robust System

Husniyah Binti Mahmud, Masahiko Nagai

Proceeding: The 41st Asian Conference on Remote Sensing (ACRS 2020) November 9-11, 2020. Deqing, China, 2020.

Author Contributions: Husniyah Binti Mahmud carried out the analysis, acquisition of the data, analysis and interpretation of the data and drafting the manuscript. Masahiko Nagai participated in study design and critical revision of the manuscript. All authors read and approved the final manuscript.

Conflict of Interest: The authors declare no conflict of interest in preparing this publication.

Publication: ISBN 978-1-7138-2908-9, Volume 1, Page 84.

[2] Peer-reviewed Journal Paper

Improved Consistency of an Automated Multisatellite Method for Extracting Temporal Changes in Flood Extent

Husniyah Binti Mahmud, Vaibhav Katiyar, and Masahiko Nagai

Journal: Mathematical Problems in Engineering (Hindawi)

Author Contributions: Husniyah Binti Mahmud carried out the analysis, acquisition of the data, analysis and interpretation of the data and drafting the manuscript. Vaibhav Katiyar and Masahiko Nagai participated in study design and critical revision of the manuscript. All authors read and approved the final manuscript.

Conflict of Interest: The authors declare no conflict of interest in preparing this publication.

Publication: Volume 2021, Article ID 6164161, 17 pages

<https://doi.org/10.1155/2021/6164161>

SUMMARY

Most of the area in Malaysia are prone to floods, and get affected almost every year. Flooding can be caused by heavy rainfall during the monsoon, geographical conditions, and other factors. Flood change studies are divided into 2 dimensions: temporal and spatial. A high number of flood observation is crucial so that timely measures can be taken. However, because each satellite data has its own characteristics, obtaining a consistent and robust water classification is difficult. Change detection, visual interpretation utilizing RGB composition, supervised classification, image texture algorithms, and active contour models are some available SAR and optical image data classification technique. Despite the abundance of near real-time data available, decision-makers in the disaster response phase appear to be underutilizing the data and information due to a number of constraints. Despite numerous remarkable efforts, existing satellite technology or any single data product has not been able to overcome the current decision-makers problem. As a result, rather than establishing a new system for a better flood operation, there is a requirement to develop a process for improving the end product for an efficient disaster response. This study focuses on using the same processing platform to standardize multi-source remote sensing data. In this context, standardizing with the meaning to eliminate major inconsistency and differences between the dataset.

The first main objective of this dissertation is to show how to use SAR polarization to automatically detect floods. Otsu thresholding method and using either single-polarization or total backscatter were assessed to extract surface water from image data. In comparison, total backscatter of polarization was chosen for automatic extraction due to the higher frequency of bimodal histogram compared to using single polarization. This chapter also is to utilize optical images to detect floods automatically. Spatial performance of each classifier, Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), and Automated Water Extraction Index No shadow (AWEInsh), and Automated Water Extraction Index with shadow (AWEIsh), stability of each classifier was compared for extraction of surface water from image data. MNDWI is chosen to represent surface water classification from the optical image data.

The dissertation's second objective is to increase the stability of surface water classification by integrating multispectral and multi SAR images. The flood event in Perlis in September 2017 was chosen as the case study. During this flood episode, Perlis was hit by Typhon Doksuri, which passed over Northern Vietnam on September 15th, resulting in continuous rainfall from September 15th until early October 2017. This study proposes to use the Modified Normalized Water Index (MNDWI) on one Landsat 7 image, two Landsat 8 images, five MODIS images, and total backscatter of polarization in four ALOS-2 images and ten Sentinel-1 images, then to employed Otsu image segmentation to distinguish water and non-water areas. The potential for image fusion to improve water area extraction consistency was studied. In this context, fusion is required to obtain a single image that retains essential features of original images, the simple and robust fusion of images with the same observation period has been proposed in this study. Similar image registration and preprocessing are used in the overall for fusion processing. The backscatter of SAR is not directly comparable to the MNDWI, from optical images. The MNDWI data was rescaled by inverting the minimum and maximum index values, so that pixels near to 1 represent water and pixels close to 0 represent non-water. The rescaled MNDWI optical image and SAR were then used to a 2D wavelet transform. The wavelet transform was utilized in this work to fuse two images using Python 3.6. Finally, the fused images were processed using Otsu thresholding. The results suggest that using a grid to incorporate a flood inundation model provides a useful overview of the flood inundation process

and able to eliminate inconsistency, especially in areas where data is scarce. Unflooded areas can be ignored when utilizing a grid system, which reduces processing time.

The third objective is to discuss about the relation of the flood occurrence with the surface water map at a low-lying area. To validate our findings, we used global surface water (GSW), to be as referral map to associate flood occurrence with the surface water flood. This varied distribution suggests that the levels of risk of surface water flooding are determined by factors associated with topography and land use. The low-lying areas, which include the main paddy lands, residential areas, highways, rails, and sugarcane plantations, are identified as those that are destructible by floods. The other class in the reclassified map includes primarily forest and rubbers which refer as highland area, which are less likely to be inundated or disrupted by floods than low-lying areas.

Although there is still challenge to eliminate temporal inconsistency, standardization of flood map is critical step to classified water extent from multisensor satellite.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENT	2
PREFACE	3
SUMMARY	5
TABLE OF CONTENTS	7
LIST OF FIGURES	9
LIST OF TABLES	12
LIST OF ACRONYMS	13
CHAPTER I INTRODUCTION	15
1.1 Background	15
1.2 Research Problems	19
1.3 Research Objectives	19
1.4 Dissertation Structure and Outline	20
References	20
CHAPTER II LITERATURE REVIEW	22
2.1 Flood mapping: Radar and optical	22
2.2 L-band and C-band for flood detection	23
2.3 Flood mapping using optical	27
2.4 Principle of single image water mapping	28
2.5 Otsu method	29
References	30
CHAPTER III OVERVIEW OF THE WORK	36
3.1 Overview of the work	36
CHAPTER IV AUTOMATED CLASSIFICATION USING THRESHOLDING	37
4.1 Introduction	37
4.2 Study area	38
4.3 Satellite data	39
4.4 Research flowchart	39
4.5 Development of water references	40
4.6 Automatic thresholding using Otsu Thresholding in Python	41
4.7 Results and discussions	44
4.7.1 Automated thresholding with Otsu	44
4.7.2 Water classification using Otsu technique	48
4.7.3 Cloud masking of multispectral images	48
4.7.4 Automatic thresholding using Otsu Thresholding	50
4.7.5 Changes of water extent in the study area from Landsat 7, Landsat 8, and MODIS data using various water indexes.	52
4.8 Ground Station data	52
4.9 Multitemporal water spatial classification using Sentinel-1	56
4.10 Integration of Multi SAR and Multispectral	59
4.11 Discussion of chapter IV	60
References	61

CHAPTER V IMPROVED CONSISTENCY INTEGRATION OF SAR AND OPTICAL WATER EXTENT DATA	62
5.1 Introduction	62
5.2 Satellite data	62
5.3 Research flowchart	64
5.4 Grid resolution and image resampling	65
5.5 2D wavelet transform for image fusion	66
5.6 Performance evaluation for data-scarce areas	67
5.7 Grid-based image fusion	67
5.8 Evaluation results	71
5.9 Discussion of chapter V	72
References	73
CHAPTER VI ASSESS WATER EXTENT MAP WITH SURFACE WATER MAP FROM GSW	74
6.1 Introduction	74
6.2 Study area and Satellite data	75
6.3 Accuracy assessment Method	77
6.4 Results and Discussion	78
6.4.1 Region a)	78
6.4.2 Region b)	79
6.4.3 Region c)	80
6.5 Discussion for Chapter VII	81
References	82
CHAPTER VII CONCLUSION	
7.1 Dissertation Conclusions	83
7.2 Limitations and Future Work	84

LIST OF FIGURES

	Page
Figure 1.1: Regional wind flow and seasonal monsoon flows.	15
Figure 1.2: Flood-prone area in Malaysia mark in green. Source: Drainage and Irrigation Department Malaysia [Online] (2012).	16
Figure 1.3: Multi satellite earth observation data for flood spatial study.	17
Figure 1.4: Example of Multi satellite earth observation for Flood Mapping	17
Figure 1.5: Methods for flood mapping.	18
Figure 1.6: Common method for flood mapping.	19
Figure 2.1: Conceptual model of spatial and temporal variations for pre-flood, flood, and post-flood.	22
Figure 2.2: Different radar signal shows different penetration levels [39]	24
Figure 2.3: Example of the horizontal and vertical radar signal. [39]	25
Figure 2.4: Example of the horizontal and vertical radar signal. [39]	25
Figure 2.5: Example of the incidence angle. [39]	26
Figure 2.6: Interaction of L-band and C-band with different types of surface during and no flood condition.	27
Figure 2.7: SAR backscattering histogram thresholding.	29
Figure 2.8: Optical satellite histogram thresholding.	29
Figure 3.1: Overview of the work.	36
Figure 4.1: Photographs from the Perlis flood disaster on September 22nd and 23rd, 2017.	37
Figure 4.2: Study area.	38
Figure 4.3: Flowchart of spatial-temporal water movement detection	40
Figure 4.4: Water references selection criteria	41
Figure 4.5: Determination of foreground and background using water references with bimodal.	42
Figure 4.6: Calculation within class variance	43
Figure 4.7: Same calculation needs to be performed for all the possible threshold values.	43
Figure 4.8: Automatic thresholding using OTSU Thresholding in Python	44
Figure 4.9: The weigh average of HH applied on the whole area	45
Figure 4.10: The weigh average of HH+HV applied on the whole area	45
Figure 4.11: Extracted result of weight average from the image acquired on October 02 2014 a) raw ALOS-2 HH, b) OTSU ALOS-2 HH, c) raw ALOS-2 HH+HV d) OTSU ALOS-2 HH+HV.	46
Figure 4.12: Close-ups of the extracted result from the image HH+HV acquired on October 02 2014.	47
Figure 4.13: Visual of surface water extent using ALOS-2 6.25m for 2014 and 2017 flood.	48
Figure 4.14: Cloud masking of optical images.	49
Figure 4.15: Determination of foreground and background using water references with bimodal histogram.	50
Figure 4.16: Same calculation needs to be performed to all water indices.	50
Figure 4.17: Sample of water classification using different water index using Landsat 7 (2017/09/07).	51

Figure 4.18: Changes of water extent in the study area from Landsat 7, Landsat 8, and MODIS data using various water indexes.	52
Figure 4.19: The study area's elevation map. The map displays the location and flood information for some of the damaged residential areas, which are represented by red dots. Blue triangles show water level data from the ground station.	53
Figure 4.20: Study area contour map.	54
Figure 4.21: Downstream ground observation data	54
Figure 4.22: Upstream ground observation data.	55
Figure 4.23: Upstream and downstream maximum water level comparison.	55
Figure 4.24: Sample of water classification using different water index using Sentinel 1 from 06/09/2017 to 30/10/2017	56
Figure 4.25: Temporal coverage of ALOS-2 and Sentinel-1	56
Figure 4.26: Changes of water extent in the study area from Sentinel-1 and ALOS-2 data.	57
Figure 4.27: Extent of the September 2017 flood extracted from the Sentinel-1 satellite images.	57
Figure 4.28: Examples of water area detection in a) lake and b) river (bottom). The water area is shown in blue and the background image shown is the Google Earth image.	58
Figure 4.29: Temporal coverage of the satellite images used in this study.	59
Figure 4.30: Changes of surface water extent in the study area from Sentinel-1, ALOS-2, Sentinel-2, Landsat 7, Landsat 8, and MODIS data between 06/09/2017 to 30/10/2017.	59
Figure 5.1: Flowchart of the automatic flood mapping process.	65
Figure 5.2: A reference vector grid of 30 x 30-pixel (w x h) size.	66
Figure 5.3: 2D wavelet transform fusion process [5].	67
Figure 5.4: Landsat 7 optical image (left) and MODIS Terra optical image (center), both acquired on 7 September 2017, and a fused image created using these two images (right). The water area is shown in blue.	68
Figure 5.5: Scatterplots between (a) MNDWI of Landsat 7 optical images and that of MODIS optical image, (b) MNDWI of Landsat 7 optical image and that of fused image, and (c) MNDWI of MODIS optical image and that of the fused image.	68
Figure 5.6: Sentinel-1 (left) and MODIS Terra (center) images, both acquired on 12 September 2017, and a fused image created using these two images (right). The water area is shown in blue.	69
Figure 5.7: Scatterplots between (a) MNDWI of Sentinel-1 backscatter images and that of MODIS optical images, (b) MNDWI of Sentinel-1 backscatter images and that of fused images, and (c) MNDWI of MODIS optical images and that of fused images.	69
Figure 5.8: Time series of the surface water extent extracted from multiple satellite images for some of the flood-affected grids. (a) Grid 5, (b) Grid 20, (c) Grid 21, (d) Grid 29, and (e) Grid 39.	70
Figure 5.9: Time series of (a) the extracted surface water area in grid 21 and the water level measured at the Kg Repoh station, and (b) the extracted surface water area in grid 29 and the water level measured at the Kg Bakau station.	71
Figure 6.1: a) No flood view in Kedah area and b), c) and d) Photos of submerged houses and paddy field in torrential rain in Kedah source from social media. Kedah land use majorly comprise of paddy field.	74

Figure 6.2: a) ALOS-2 HH+HV b) GWS binary image. Dark color indicates water region while, white color indicates non-water area.	75
Figure 6.3: Closed up of study area a), b) and c) and the blue is the flood extent on 2014 report by DID.	76
Figure 6.4: a) ALOS-2 HH+HV b) GWS binary image. White color indicates water region while, dark color indicates non-water area.	78
Figure 6.5: a) ALOS-2 HH+HV b) GWS binary image. Dark color indicates water region while, white color indicates non-water area.	79
Figure 6.6: a) ALOS-2 HH+HV b) GWS binary image. Dark color indicates water region while, white color indicates non-water area.	80

LIST OF TABLES

	Page
Table 2.1 Summary of L band and C band characteristics.	26
Table 4.1: Utilized PALSAR-2 data.	39
Table 4.2: Automatic threshold values using the OTSU technique of water reference with circle areas from the HH and HH + HV sigma-naught values taken on the evening of 02 October 2014 5.30 PM until 01 March 2018 5.30 PM.	47
Table 5.1: Details of the (SAR) and optical satellite images used in this study.	63
Table 6.1: Result of Errors of Omission, Errors of commission and Kappa coefficient for Region a)	78
Table 6.2: Result of Errors of Omission, Errors of commission and Kappa coefficient for Region b).	80
Table 6.3: Result of Errors of Omission, Errors of commission and Kappa coefficient for Region c)	81

LIST OF ACRONYMS

ALOS-2:	Advanced Land Observing Satellite-2
AWEInsh:	Automated Water Extraction Index No shadow
AWEIsh:	Automated Water Extraction Index with shadow
DEM:	Digital Elevation Model
DID:	Department of Irrigation and Drainage
ENSO:	El-Niño Southern Oscillation
EO:	Earth Observation
HH:	Horizontal transmission and horizontal reception
HV:	Horizontal transmission and vertical reception
HH:	Containing high-frequency diagonal information
HL:	Containing high-frequency vertical information
IOD:	Indian Ocean Dipole
GEE:	Google Earth Engine
GSW:	Global Surface Water
LH:	Containing high-frequency horizontal information
LL:	Containing low-frequency approximation information
MBE:	Mean Bias Error
MJO:	Madden-Julian Oscillation
MODIS:	Moderate Resolution Imaging Spectroradiometer
MNDWI:	Modified Normalized Difference Water Index
NaFFWS:	National Flood Forecasting and Warning System
NDWI:	Normalized Difference Water Index
NIR:	Near-infrared

OLI:	Operational Land Imager
PALSAR-2:	Phased Array L-band Synthetic Aperture Radar-2
RGB:	Red, Green and Blue
RMSE:	Root Means Square Error
SAR:	Synthetic Aperture Radar
SWIR:	Shortwave infrared
SLC:	Scan Line Corrector
TIRS:	Thermal Infrared Sensor
UTM:	Universal Transverse Mercator
VV:	Vertical transmission and vertical reception
VH:	Vertical transmission and horizontal reception

Chapter 1

Introduction

1.1 Background

Peninsular of Malaysia is located at the equatorial zone extending in between northern latitude 1° and 6° N and the eastern longitude from 100° to 103° E. Malaysian climate is influenced by Northeast monsoon and southwest monsoon, which blow alternately by season. There is an increase in rainfall during the Northeast monsoon, which takes place from November and ends in February. While the Southwest monsoon starts from May to August, Malaysia is more likely in a dry season due to the weaker wind speed.

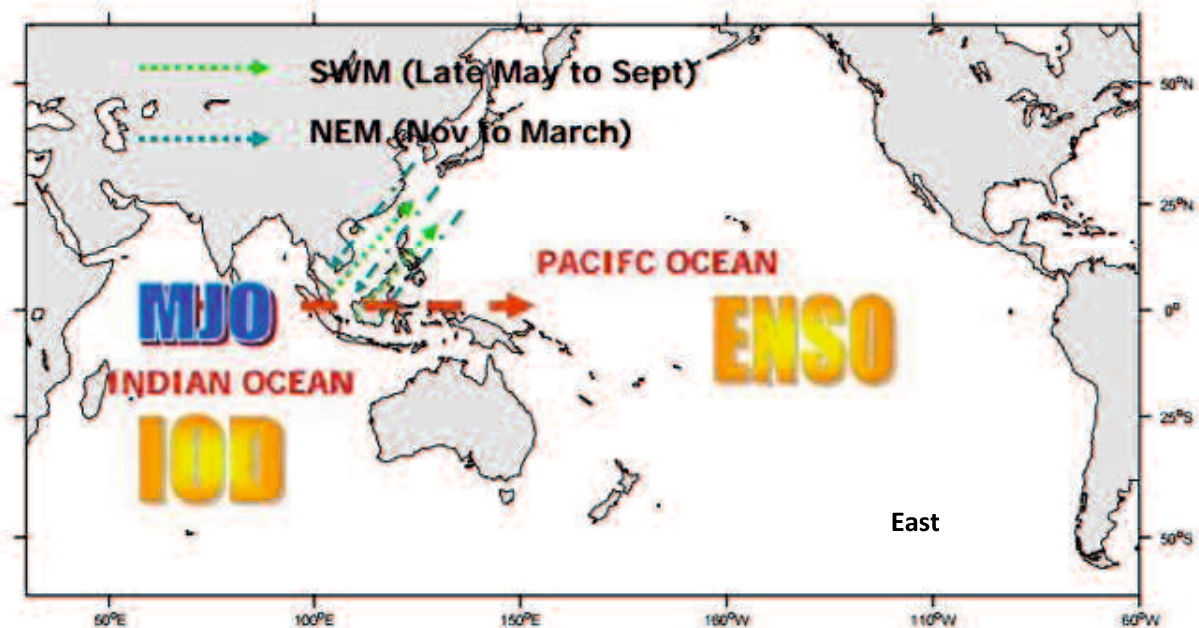


Figure 1.1: regional wind flow, seasonal monsoon flows, and other dominant modes of climate variability [12].

Figure 1.1 shows the regional wind flow and seasonal monsoon flows. The Northeast wind cross over the South China Sea during the Northeast monsoon, and East and South of Peninsular Malaysia are under heavier rainfall than usual [1]. During Southwest monsoon-influenced by southwesterly winds usually last for 3-4 months from May to August [2] [3] [4]. Since the southwest monsoon has a weaker wind speed than the northeast monsoon, the dry season is more likely to occur during this period. Dry seasons are more frequent than wet seasons due to the stable atmosphere, resulting in less intense convective development. West coastal area experienced thunderstorms, heavy rain, and strong gusting winds in the predawn and early morning.

Two shorter transitional periods occur between the monsoon period, usually in April and October [5]. The winds are generally light and variable.

Large-scale phenomena such as the El Niño-Southern Oscillation (ENSO) [7] [8], Indian Ocean Dipole (IOD) and the Madden-Julian oscillation (MJO) have a significant impact on the climate of Malaysia [9] [10] [11]. On an intraseasonal period, the Madden-Julian Oscillation (MJO), which was present throughout the northeast monsoon season, has a considerable impact on large-scale circulations over the Indian Ocean and Maritime Continent.

Tangang *et al.* (2008) [9] showed the MJO's relation to the Maritime Continent when they studied at the impact of atmospheric variables in causing extreme flooding in Malaysia's southern peninsula. They suggested that the MJO's influence over the Indian Ocean stimulates strong easterly winds over the Maritime Continent. The easterly wind increases rainfall in Malaysia's southern peninsula in 2 directions. First, it prevents the Borneo vortex from developing. When the Borneo vortex is present, the most of the rainfall usually falls in the western part of East Malaysia. Second, the 5 easterly winds increase the northeasterly wind's effect. The cold and dry northeasterly wind from the Siberian High warmed up as it passed over the warm tropical sea, merging with the warm and moist easterly from the Pacific Ocean, resulting in a more intense formation of convective clouds across southern peninsular Malaysia.

The development of sea surface temperature across the Indian Ocean, known as the Indian Ocean Dipole (IOD), influences convection during the northeast monsoon season. [13] stated that the anomalous colder sea surface temperature (SSTA) (negative IOD) across the southeastern Indian Ocean reduces convection over that region while enhancing convection over the Malaysian region during the NEM monsoon.

The ENSO is categorized into three phases: neutral, warm (El Niño), and cold (La Niña). In general, most studies agree that El Niño tends to reduce rainfall across the Maritime Continent, whereas La Niña enhances higher rainfall.

Subramaniam (2009), were discussed for NEM 2016/2017. Over the years, Malaysia has been prone to flood risk, with an area of about 29,000 km², as shown in Figure 1.2, affecting more than 4.82 million people (22% of the population) and suffering annual damage of almost USD 298.29 million. The previous large-scale floods have destroyed infrastructure, agriculture, fatality, interfered the human activities, and slowed down economies.



Figure 1.2: Flood-prone area in Malaysia mark in green. Source: Drainage and Irrigation Department Malaysia [Online] (2012).

Peninsular Malaysia are very vulnerable to floods and frequently occur during the monsoon season. Over the years, Malaysia has been prone to flood risk, with an area of about 29,000 km², as shown in Figure 1.2, affecting more than 4.82 million people (22% of the population) and suffering annual damage of almost USD 298.29 million. The previous large-scale floods have destroyed infrastructure, agriculture, fatality, interfered the human activities, and slowed down economies.

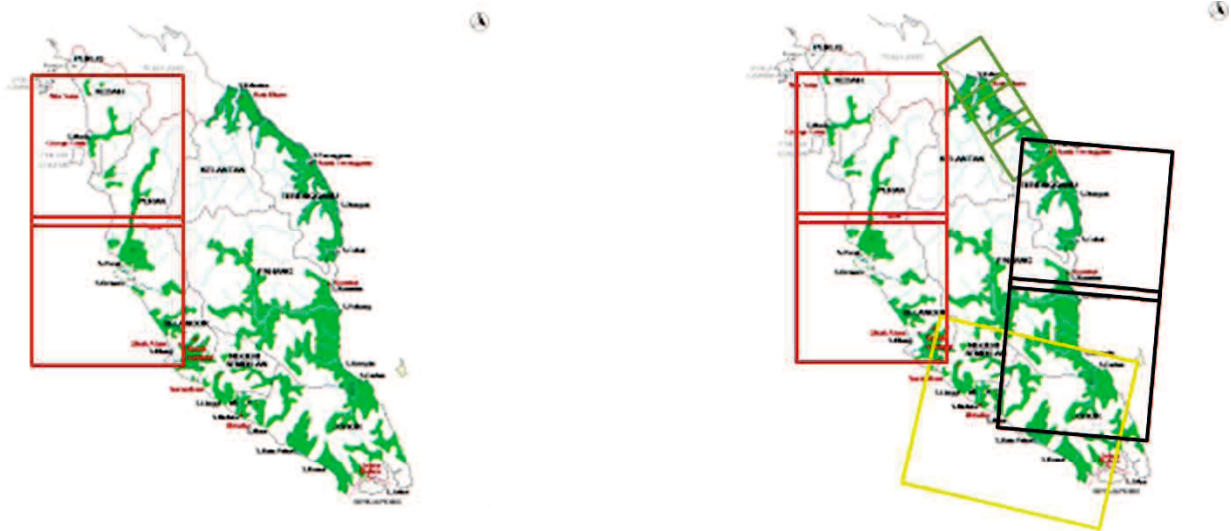


Figure 1.3: Multi satellite earth observation data for flood spatial study.

Every year, there are many parts of Malaysia affected with flood. and, it is difficult for a single earth observation satellite to cover those affected areas in a short time. the use of Multisatellite earth observation data to catch significant information of those flood affected areas was proposed in this study.

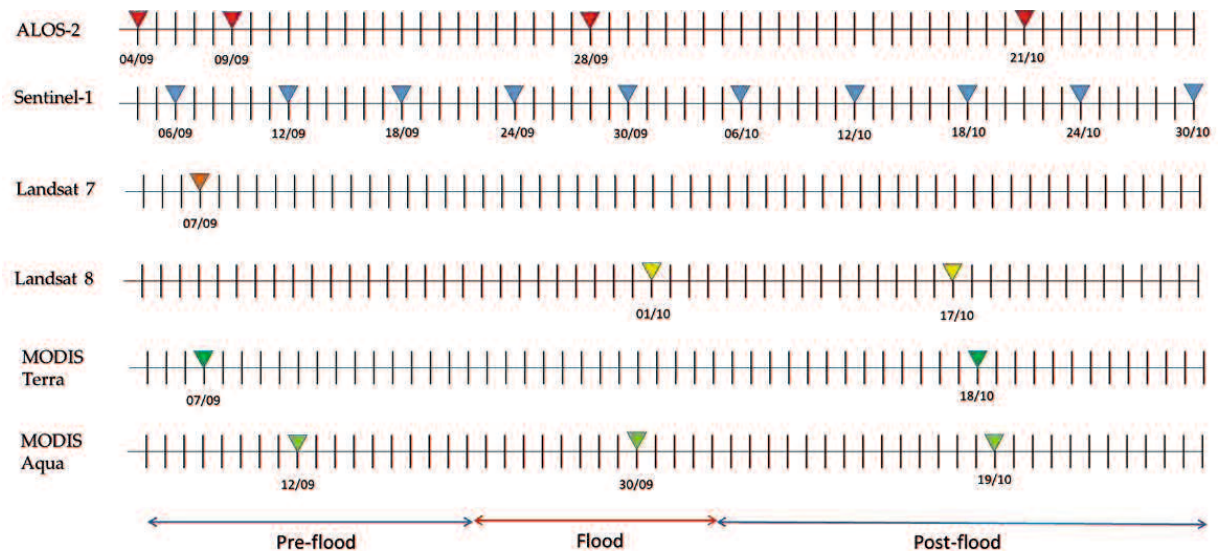


Figure 1.4: Example of Multi satellite earth observation for Flood Mapping

Figure 1.4 shows the example of Multi satellite earth observation near real time data based on a flood case in Perlis, Malaysia on September 2017. Another important aspect of flood study, is its temporal changes. And multisatellite have significant advantage on increasing the frequency of flood observations during the duration of flood. For example, in here we are showing multisatellite coverage before, during and after a flood case in Perlis on September to end of October 2017. in case we only use alos-2 images, we can only get 4 numbers of observation within those 2 months. but when we maximize the utilization of satellite images we can increase the number of observation frequency and reduce the interval times of observation.

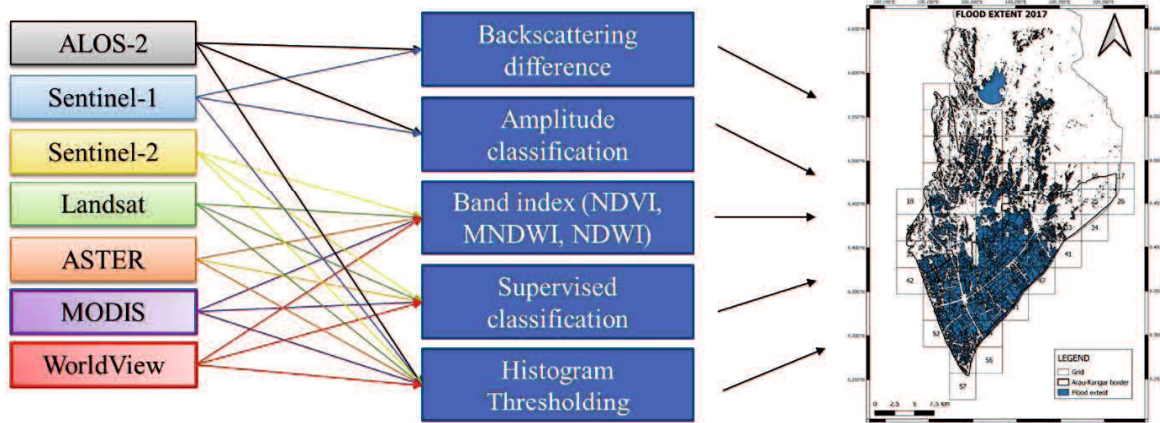


Figure 1.5: Methods for flood mapping.

Many techniques are used to extract water using optical satellite images and SAR images, from image processing techniques to image classification techniques: Backscattering difference use two SAR images (one post-flood /one co-flood or pre-flood) were used for change detection analyses (Davide et al., 2018). Amplitude classification where training data using water signature in SAR amplitude image. Signature value is highly variable depending on image acquisition (Liu et al., 2018). Band index (NDVI, MNDWI, NDWI) is a derivation of index using imaging spectral band and its limitation is selecting the optimal water related bands is vital. Supervised classification (Giordan et al., 2018) select water training data over composite band image however, it is time consuming (Campbell and Wynne, 2011). Histogram thresholding separate water and non-water pixels from histogram distribution of SAR backscatter or optical image. It is based on a single image (Vaibhav et al., 2021), effective pixel-based approach, common method (Fan and Lei, 2012; Nakmuenwai et al., 2017) and required less time-consuming (Chen et al., 2021).

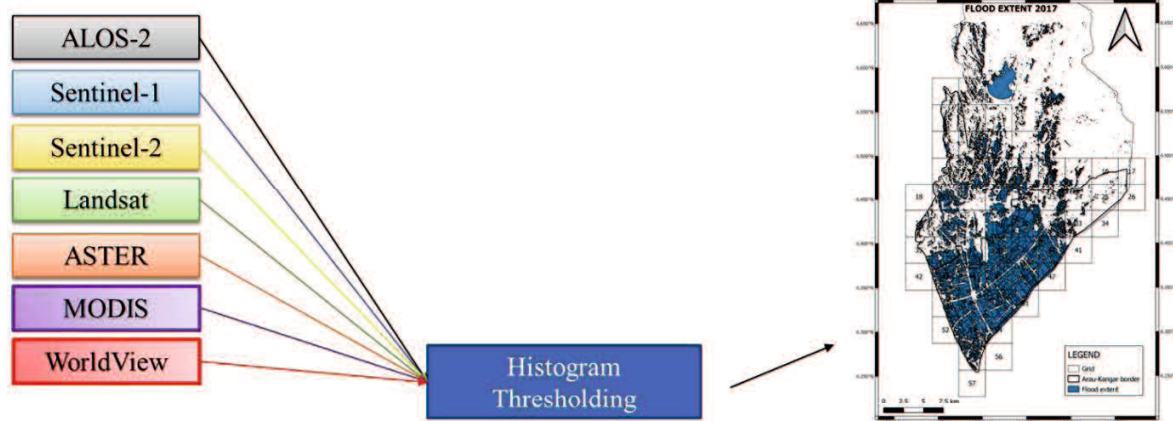


Figure 1.6: Common method for flood mapping

Presently, even though an abundance of near real-time data is available, the data and information still appear to be underutilized by decision-makers in the disaster response phase due to a few constraints, such as the limitation on (1) time and capacity of a person to process new geospatial datasets, (2) accessibility of near real-time data, (3) compatibility of user platforms and geospatial data formats, (4) knowledge of the data availability and inadequacy of data latency, and (5) understanding of the end-user demand on the product and timing needs [6]. Despite various notable efforts, to this end, existing satellite technology or any single data product could not solve the current challenge from the decision-making standpoint. Hence, to fulfill decision-makers' needs, there is a demand to develop a process for improving the end product for an effective disaster response rather than developing a new system to improve flood operation. The study motivation is to propose the use of a common method that is highly adaptable to any satellite input data in order to increase the frequency of flood extent observation during the duration of disaster.

1.2 Research Problem

1. There are many emerging satellite imageries, the data and information still appear to be underutilized by decision-makers in the disaster response phase.
2. Processing satellite images separately can lead to inconsistency when combining different datasets for time change analysis using a specific time.
3. Accuracy assessment of satellite imagery is needed to identify the problem of the input dataset.

1.3 Research Objectives

1. To establish an automated water classification by using SAR and optical images using a common method.
2. To standardize the flood extent output from multi-satellite in the context of elimination of major inconsistency and differences between datasets.

3. To discuss about the relation of the flood occurrence with the surface water map at a low-lying area.

1.4 Dissertation Structure and Outline

This dissertation consists of six chapters. This paper begins with an introduction of the research, given in Chapter 1. A short explanation of the basic theory behind SAR and related methods of water classification and thresholding using SAR and optical data are presented in Chapter 2. To summarize the research objectives, the overview of the work was shown in Chapter 3. The automatic flood detection using ALOS-2 images and for optical images, water surface detection is described in Chapter 4. Chapter 5 focuses on the applications of both automatic methods on both SAR and optical images and in order to improve consistency of multi satellite output by standardizing the automatic flood detection method and the results obtained. Discussion about the relation of the flood occurrence with the surface water map at a low-lying area is evaluated in Chapter 6. Finally, Chapter 7 discusses and summarizes the main results of the study.

REFERENCES

- [1] Wan-Zin, W.Z., Jemain, A.A., Ibrahim, K., Suhaila, J. and Sayang, M.D., 2009. A comparative study of extreme rainfall in Peninsular Malaysia: with reference to partial duration and annual extreme series, *Sains Malaysia*, 38(5): 751–760.
- [2] Camerlengo, Alejandro Livio and Monica Ines Demmler. “Wind-driven circulation of Peninsular Malaysia's eastern continental shelf.” *Scientia Marina* 61 (1997): 203-211.
- [3] Tangang, F.T. 2001. The quasi-biennial and low-frequency oscillation in the Malaysian precipitation anomaly. *International Journal of Climatology* 21(10): 1199-1210
- [4] Jamaludin, Shariffah & Deni, Sayang & Zin, Wan. (2010). Spatial patterns and trends of daily rainfall regime in Peninsular Malaysia during the southwest and northeast monsoons: 1975–2004. *Meteorology and Atmospheric Physics*. 110. 1-18. 10.1007/s00703-010-0108-6.
- [5] Jamaludin Suhaila and Abdul Aziz Jemain , 2007. Fitting Daily Rainfall Amount in Malaysia Using the Normal Transform Distribution. *Journal of Applied Sciences*, 7: 1880-1886.
- [6] G. J-P. Schumann, S. Frye, G. Wells et al., “Unlocking the full potential of Earth observation during the 2015 Texas flood disaster,” *Water Resources Research*, pp. 3288–3293, 2016.
- [7] Aldrian, E., and Y. S. Djamil, 2008: Spatio-temporal climatic change of rainfall in east Java Indonesia. *Int. J. Climatol.*, 28, 435-448, doi:10.1002/joc.1543.
- [8] Moron, V., A. W. Robertson, and J.-H. Qian, 2010: Local versus regional-scale characteristics of monsoon onset and post-onset rainfall over Indonesia. *Climate Dyn.*, 34, 281-299, doi:10.1007/s00382-009-0547-2
- [9] Tangang, F. T., Juneng, L., Salimun, E., Vinayachandran, P. N., Seng, Y. K., Reason, C. J. C., Behera, S. K., and Yasunari, T. (2008), On the roles of the northeast cold surge, the Borneo vortex, the Madden-Julian Oscillation, and the Indian Ocean Dipole during the extreme

2006/2007 flood in southern Peninsular Malaysia, *Geophys. Res. Lett.*, 35, L14S07, doi:10.1029/2008GL033429.

[10] Hidayat, R., and S. Kizu, 2010: Influence of the Madden-Julian Oscillation on Indonesian rainfall variability in austral summer. *Int. J. Climatol.*,30, 1816-1825, doi:10.1002/joc.2005.

[11] Salahuddin, A., and S. Curtis, 2011: Climate extremes in Malaysia and the equatorial South China Sea. *Global Planet. Change*, 78, 83-91, doi: 10.1016/j.gloplacha.2011. 05.001.

[12] Ahmad Fairudz Jamaluddin, 2016. *The Weather and Climate of Malaysia*, Malaysian Meteorological Department, Ministry of Science, Technology and Innovation.

[13] Chang, C.-P., P.A. Harr, and H.-J. Chen, 2005: Synoptic disturbances over the equatorial South China Sea and western Maritime Continent during boreal winter. *Mon. Wea. Rev.*, 133, 489-503, doi:10.1175/MWR-2868.1.

Chapter II Literature Review

2.1 Flood mapping: Radar and optical

To monitor surface water dynamics, two forms of Earth Observation (EO) satellite images with high geographical and temporal resolutions are used: multispectral and radar data. Earth observation satellites provide multitemporal coverage of huge areas, which is commonly used to study differences in land cover before, during, and after flood occurrences. [1,2,3,4]. Figure 2.1 shows the dynamic behavior of a flood in an inundated area. The conceptual model is consistent with prior research [5,6], which found that the flood extent peaked during the flood and then rapidly declined a few days later.

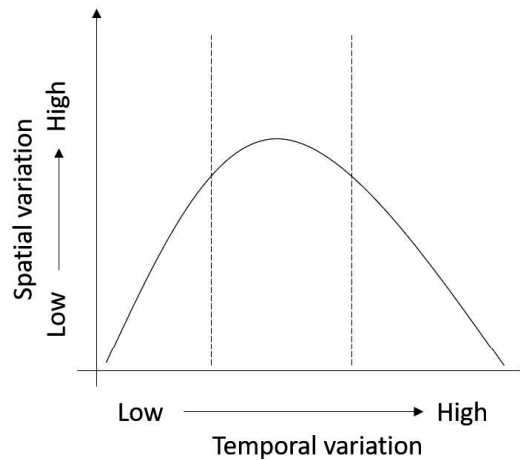


Figure 2.1: Conceptual model of spatial and temporal variations for pre-flood, flood, and post-flood.

Optical and radar satellite images can be used to extract large-scale flood information. Optical sensors useful for mapping inundation areas include the Moderate Resolution Imaging Spectrometer (MODIS) onboard the Aqua and Terra satellites [7,8], the Land Remote-Sensing Satellite (Landsat) [9,10], and Sentinel-2 [11]. Both Sentinel-1 [12,13] and ALOS-2 [14,15,16] are equipped with radar sensors and provide valuable data for flood mapping. Optical images provide extensive information on land cover, while radar data can be utilized to identify flooded regions quickly. The number of flood surface observations could be increased by combining optical and radar imagery [17].

The appearance of a cloud during a flood occurrence is quite often unavoidable, especially in the tropical zone, where cloud cover varies with time [18,19]. Because of its wavelength capability penetrating clouds, synthetic aperture radar (SAR) is effective for consistently monitoring flooded areas [17]. In general, the wavelength, polarization, resolution, look angle, surface roughness, dielectric characteristics, and slope of a SAR sensor determine how microwaves interact with the earth's surface. The smooth water surface behaves as a specular reflector, giving SAR images a dark appearance and low backscatter [20]. For high-accuracy surface water mapping, choosing the right SAR polarization is crucial. Horizontal transmission and horizontal reception (HH), horizontal transmission and vertical reception (HV), vertical transmission and vertical reception (VV), or vertical transmission and horizontal

reception (VV) are all examples of single-polarization SAR (VH). For the delineation of open water bodies, HH polarization has proven high classification accuracy [21,22,23]. Using the C-band Radar Satellite (RADARSAT-2) SAR data, Manjusree et al. [24] investigated the backscatter response in multi-polarizations and discovered a difference in the response of the flooded region discrimination depending on the polarization. HH's radiometric profile is less backscattered than HV and VV's; yet, HH's accuracy in heterogeneous regions is lower than HV and VV's [25]. The use of dual-polarization in flooded vegetation was studied by Henry et al. [26], Horritt et al. [27], and Irwin et al. [28], and their findings proved the benefits of dual-polarization observations in detecting inundated areas in vegetated regions. There are two channels of intensity and phase information in dual-polarization SAR, providing for the differentiation of scattering mechanisms [25]. The most reliable result for automatic classification mapping utilizing a large dataset was reached by combining HH and HV (HH + HV) [28]. Furthermore, if the histogram has a bimodal distribution, the Otsu thresholding [29] approach performs well because a threshold is chosen at the bottom of a steep valley linking two tops representing object and background. Furthermore, instead of adopting single-polarization, HH + HV maintains bimodality (HH or HV) [30,31].

Visual interpretation of optical images at visible and near-infrared (NIR) wavelengths can be utilized to extract water surfaces. Optical images, unlike SAR images, are affected by weather conditions. Sunlight and the inability to penetrate clouds have a significant impact on the quality of information in optical images. Furthermore, optical images lack information on surface structure and roughness when compared to SAR images. Using a water index derived using two or more bands is the quickest and most reliable method of identifying water areas from an optical image [32]. In recent years, a number of water indices have been established [33,34,35,36]. The Normalized Difference Water Index (NDWI) and the Modified Normalized Difference Water Index (MNDWI) are the most frequently used (MNDWI) [37]. The NDWI detect water characteristics in remote sensing images by using green and NIR bands [33]. The MNDWI, on the other hand, is a more dependable and stable water index since it enhances open water features and successfully removes noise in images of urban areas, vegetation, and soil using shortwave infrared (SWIR) rather than near-infrared (NIR) [34]. Based on optical images from Sentinel-2, Landsat-7, Landsat-8, and MODIS MOD09A1, Husniyah et al. [38] compared NDWI, MNDWI, Automated Water Extraction Index with no shadow (AWEInsh), and Automated Water Extraction Index with shadow (AWEIsh), and concluded that MNDWI was relatively stable and could reduce the inconsistency of flood extent detection when combined with multiSAR images.

2.2 L-band and C-band for flood detection

It leads to a rapid increase in the availability of near-real-time SAR data because SAR independence from sunlight and weather conditions. Longer wavelength SAR sensors may also penetrate deeper into vegetation cover and identify sub-canopy conditions; thus, radar remote sensing can be quite useful for monitoring flood change. The SAR systems onboard Earth Observation satellites for flood detection often use the X-, C-, or L-band frequencies. Varying X-, C-, and L-band sensors have different capabilities based on their sensor properties (wavelength, polarization, and incidence angle) and environmental circumstances aspects (for example, vegetative type, phenology of plants, soil moisture). Figure 2.2 demonstrates several radar signals and their penetration levels. The wavelength of the X-band is around 3.0 cm, the

C-band is around 5.6 cm, and the L-band is up to 24 cm. Because it is only effective for detecting flooding beneath sparse vegetation or forest after leaf-off, the X-band radar has lower reliability for flood classification beneath a forested canopy, as shown in Figure 2.2. The C-band can penetrate deeper into vegetation volume than X-band sensors, which scatter at the canopy at the trees' tips. L-band, on the other hand, is more reliable due to a larger range of backscattering enhancement and better penetration.

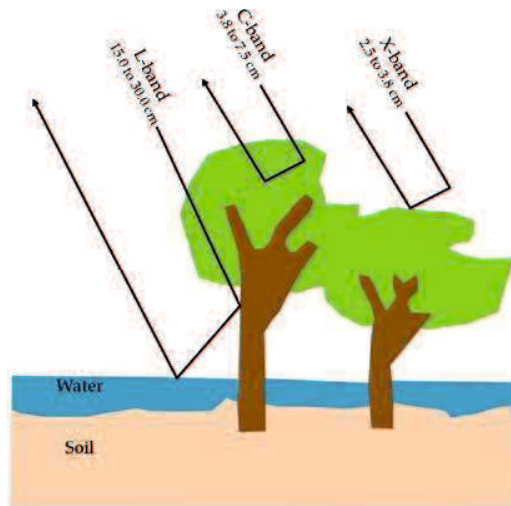


Figure 2.2: Different radar signal shows different penetration levels [39]

Multifrequency SAR is essential for flood analysis because it improves consistency and temporal coverage. For pre-flood, during, and post-flood analysis, this study uses SAR images acquired in the L and C bands for automatic and robust flood extent estimation. SAR data was obtained in a timely manner from European Space Agency's Copernicus for Sentinel-1 satellites' frequent six-day revisits and the Japan Aerospace Exploration Agency's for ALOS-2. Several previous studies have demonstrated successful Sentinel-1 and ALOS-2 time series data integration, with good agreement with validation sources such as aerial imaging, optical imagery, and other validation sources.

For a single channel, SAR polarisation can be horizontal (H) or vertical (V), and is determined by the orientation of the electromagnetic field vector with respect to its propagation direction. Figure 2.3 shows the horizontally and vertically polarized radar signal's direction. The multipolarized L-band can transmit and receive radar in a linear manner, such as polarized: horizontal transmit, horizontal receive (HH); vertical transmit, vertical receive (VV); and cross-polarized: horizontal transmit, vertical receive (HV); and vertical transmit, horizontal receive (VH). In comparison to single polarized radar data potential HH and HV dual-polarization evaluated in this study, a polarized radar system can influence backscatter strength and detect submerged cropped area. SAR images with single co polarization HH (horizontal transmit and horizontal receive) have been employed in many past researches to estimate water and flood areas because they improve the contrast between forest and flooded vegetation and maximize canopy penetration. Wind is less of an issue for cross polarized HV (horizontal transmit vertical receive) and VH (vertical transmit horizontal receive) than for co-polarized HH. As a result, dual-polarization has the potential to improve water extraction efficiency. Sentinel-1A images with dual-polarization vertical transmit, vertical receive VV,

vertical transmit, and horizontal receive VH were employed in this study. The ESA Copernicus Open Access Hub was used to download all of the Sentinel-1A images.

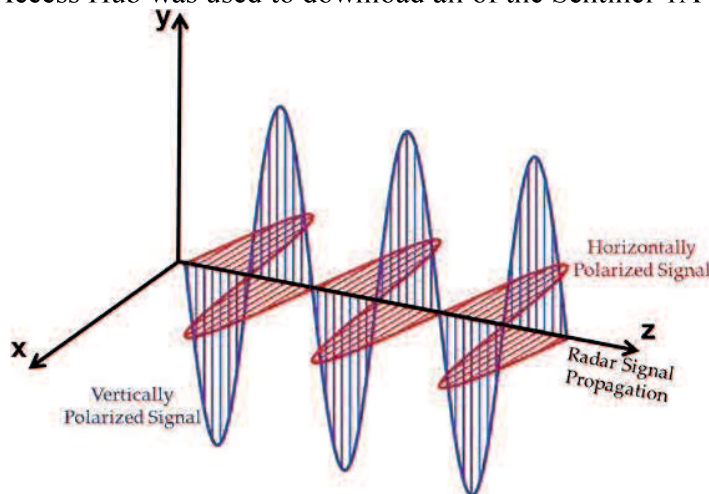


Figure 2.3: Example of the horizontal and vertical radar signal. [39]

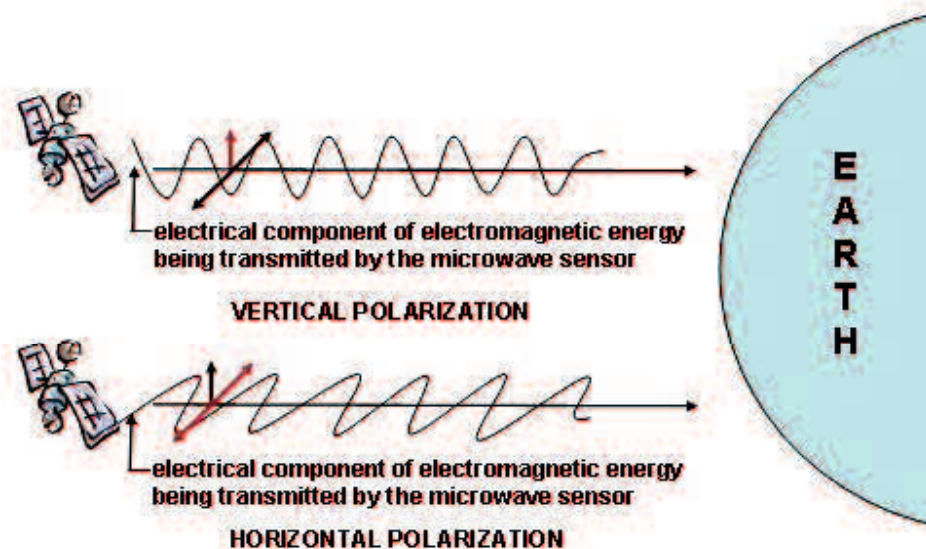


Figure 2.4: Example of the horizontal and vertical radar signal. [39]

The incidence angle, in addition to wavelength and polarization, can affect its ability to detect floods. The angle of the pulse of microwave energy transmitted by satellite perpendicular to the earth's surface is known as the incidence angle. The explanation of the incidence angle is shown in Figure 2.5. The larger the incidence angle, the shallower it is, and the smaller the incidence angle, the steeper it is. Because shallow incidence angles are more influenced by the top of the tree, creating more volume scattering, a steeper incidence angle is preferred for separating water from non-water in the flooded forest. The use of a combination of multi-incidence angles, on the other hand, is useful for monitoring floods in wetlands.

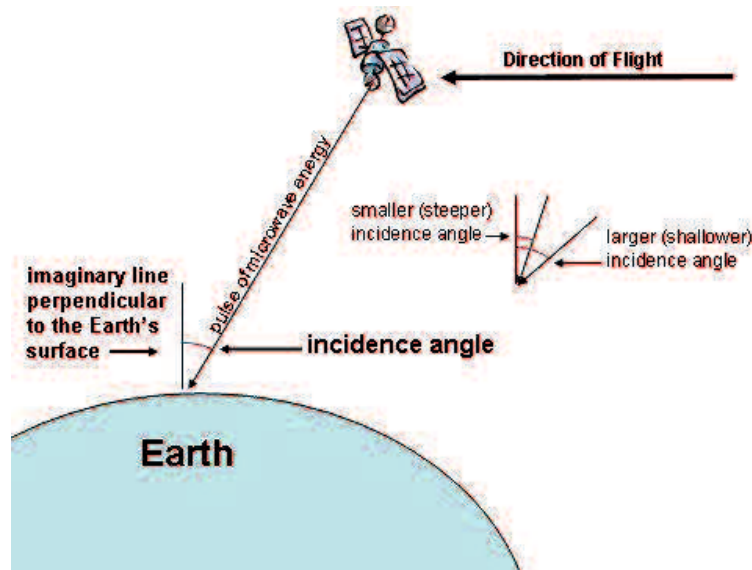


Figure 2.5: Example of the incidence angle. [39]

Table 2.1 Summary of L band and C band characteristics.

	ALOS-2	Sentinel-1
Band	L	C
Wavelength (cm)	22.9	5.6
Incidence angle (°)	8° - 70°	20° - 46°
Polarization	HH and HV	VV and VH
Revisit days	14 days	12 days
Spatial resolution (m)	6.25 m, 25 m	10 m

Figure 2.6 represents the microwave scattering interaction on high and low vegetation during dry and flood conditions. In general, forested wetlands should use longer wavelengths (L-band) to detect flooding, while herbaceous wetlands should use shorter wavelengths (C-band). L band SAR operates as a mirror, reflecting most energy away from the sensor, and reflects flat surfaces such as paddy fields to water. Backscattered radiation, on the other hand, will increase as it interacts with rougher surfaces. When the forest is not flooded and interacts with the forest canopy, volumetric scattering occurs. The volumetric scattering intensity is determined by the height of the vegetation layer underneath the forest's main canopy.

The interaction of incident energy with the tree trunk is followed by a shift in direction to a specular surface (usually bare earth, extremely short vegetation, or water), where energy is reflected back to the sensor; this process also occurs in the opposite direction. When the region is flooded, even if only temporarily, there is a high double-bounce reflection between the tree trunks and the water surface, which adds to the volumetric scattering within the canopy and greatly improves the return signal to the sensor. The main scattering processes are the same for C-band and L-band, but due to the shorter wavelength of C-band, the interactions between incident radiation and specific cover types differ (5.6 cm). The shorter wavelengths of the C-

band, for example, do not penetrate deep forest canopy. As a result, volumetric scattering within the canopy records for the majority of backscattering for this cover type; short vegetation, such as pasture, which may not be visible at longer wavelengths due to specular reflection, will provide a moderate backscattering return, also due to volumetric scattering.

The smooth water surface acts as a specular reflector, while the surrounding environment (soil and plant) acts as diffuse reflectors, which is how side-looking radar images are used to map water bodies. As a result, water surfaces have minimal backscatter, but soil and vegetation morphologies have more backscatter. The minimal size of detectable objects and the sensors' position in the satellite's range of view are both limitations in radar data. As a result of the latter, shadows or blind areas appear in synthetic aperture radar (SAR) satellite images, which are mistakenly perceived as water surfaces.

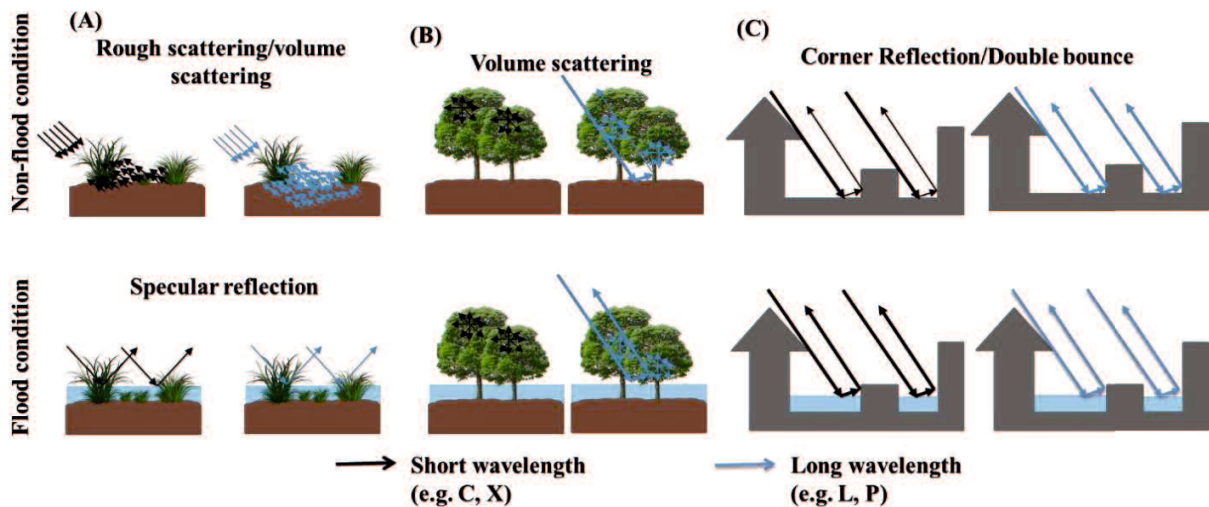


Figure 2.6: Interaction of L-band and C-band with different types of surface during and no flood condition.

In this, when the cover is “dry,” there is surface reflection at both short and long wavelengths due to the relative surface roughness in addition to volume backscatter from the grasses (short wavelength) and penetration into the soil itself (penetration will be greater at the long wavelengths). When flooded, the reflection becomes specular (assuming little surface roughness of the water) although emergent vegetation will create backscatter at the short wavelengths. For forest (B), short wavelength scatter will be dominated by volume scattering within the canopy and, if dense, the electromagnetic energy might not penetrate to the surface. Longer wavelengths will scatter from branches and tree structure in addition to “double-bounce” surface/trunk backscatter. When inundated the “double bounce” return will be highly amplified. In urban regions (C) the “double-bounce” effect can tend to dominate at both scales, although surfaces will appear “rougher” at short wavelengths, dulling this. When flooded, the “double” and indeed multiple-bounce returns will be heightened significantly.

2.3 Flood mapping using optical

Optical remote sensing uses visible, near infrared (NIR), and short-wave infrared (SWIR) sensors to detect sun radiation reflected from targets on the ground to produce images of the earth's surface. Because different materials reflect and absorb differently at different

wavelengths, the targets can be distinguished by their spectral reflectance characteristics in remotely sensed images. The number of spectral bands used in the imaging process is used to classify optical remote sensing.

However, optical sensors are limited by their dependency on both sunlight and clouds, which may limit the ability to collect data during a flood. Many factors allow atmospheric correction of satellite data over inland water difficult. The adjacency effect of neighboring land pixels, which significantly raises water reflectance, is one of the most significant challenges. Other significant challenges include terrestrial atmospheric pollution, high turbidity, floating objects, and the adjacency effect of neighboring land pixels, which significantly raises water reflectance.

The fact that clear water absorbs almost all near-infrared light, in contrast to highly reflecting adjacent soil or plants, has been utilized to detect surface-water locations using optical data. The technique of categorizing individual pixels in an image, usually based on spectral reflectance properties.

Thematic classification [40], linear unmixing [41], single-band thresholding [42] and two-band spectral water indices are four common water classification methods using optical remote sensing pictures [43]. [40], [41], [42]. Because thresholding from single and two-band indices takes less time to complete, it is the most often utilized approach for water extraction [44]. Manual classifications are time-consuming and prone to error due to false interpretation of satellite imagery.

2.4 Principle of single image water mapping

Various image processing approaches have been introduced that so far. Using satellite data, these techniques assist in the identification of water features. Thresholding-based methods have been the most extensively employed in the literature, in part because they are less computationally time-consuming while yet producing equivalent accuracy to more complicated segmentation approaches [45], [46]. The thresholding approach is the quickest way to classify an image into binary categories. The selection of an appropriate threshold has a significant impact on the classification results. Thresholds are commonly calculated by evaluating the SAR backscatter intensity histogram or water index from optical image and determining the probability distributions of water and non-water pixels and was shown in Figure 2.7 and Figure 2.8. The basic principle is that a SAR intensity image's histogram contains two partially overlapped distributions of water and non-water pixels. Their intersection produces the best threshold for distinguishing the two classes with the least amount of error. A smaller threshold usually identifies water extent with higher confidence, while a larger threshold tends to increase the confidence of identifying dry land [45], [47].

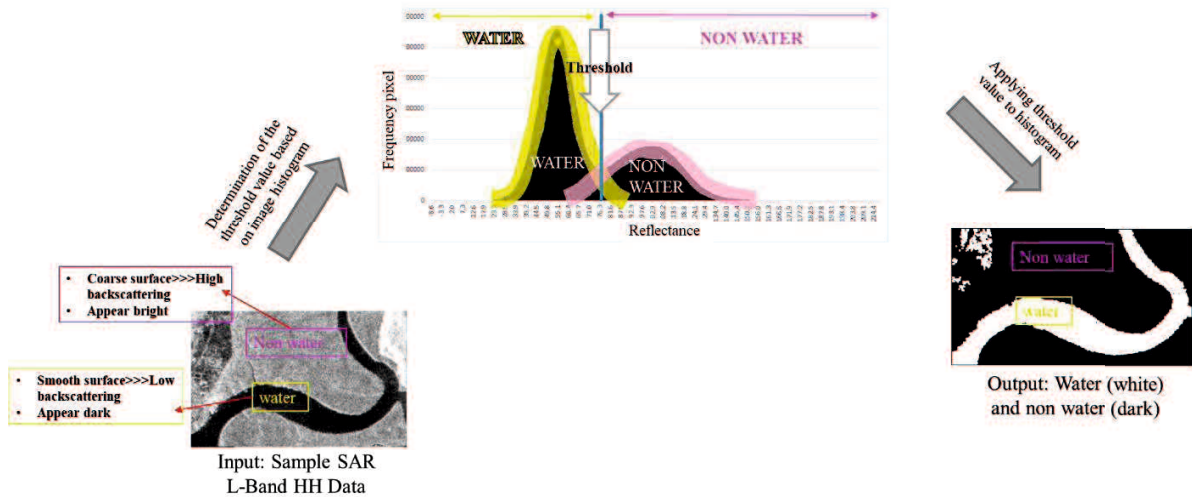


Figure 2.7: SAR backscattering histogram thresholding.

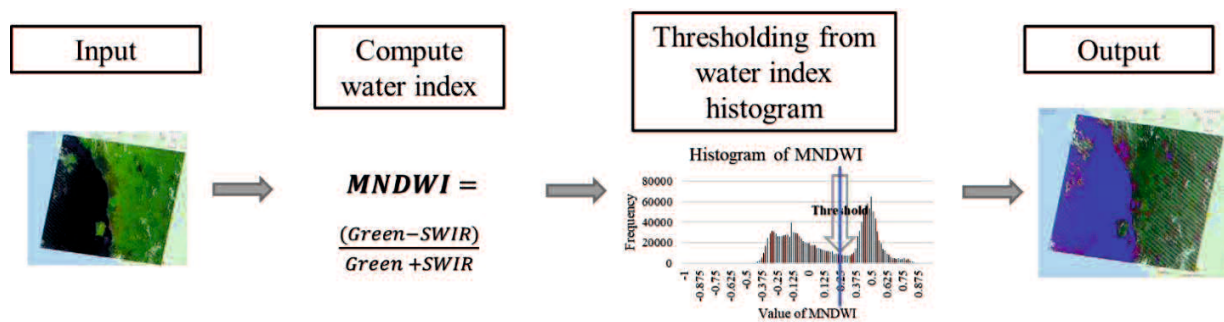


Figure 2.8: Optical satellite histogram thresholding.

2.5 Otsu method

A basic, simple option method could help with a specific task while reducing the computational cost for broader application. Because of their simplicity, robustness, and adaptability, this study focuses on Otsu thresholding approaches, which is one of the most often used thresholding algorithms. Nobuyuki Otsu proposed the Otsu thresholding technique. By increasing the between-class variation of the gray levels in the object and background sections, Otsu thresholding approaches select the best threshold. Image segmentation is a key step in extracting water features information. The threshold segmentation has been adopted in water index to separate the image into two classes: water features and background features. The threshold values for water index, NDWI and MNDWI were set to zero, but the adjustment of the threshold based on actual situations is necessary. And that could achieve a more accurate result for the water information delineation. Hence, dynamic thresholds are needed when different regions or different phases of remote sensing data are employed to detect water features information. Many methods can be used for image threshold segmentation. The Otsu

method is a dynamic threshold method that has been successfully used in delineating water bodies and monitoring water, area changes [49].

Furthermore, if noise removal and character identification are implemented appropriately, Otsu can be utilized for thresholding and performs satisfactorily. Because the calculation requires 1D intensity data, the advantage is that determining the threshold is simple, which helps to reduce computational processing time in real-world applications. Many techniques to improve the original Otsu method have been presented as a result of these advantages. However, if the images have noise, these methods are unlikely to give satisfactory segmentation results.

Eq. 1 shows the equation for within-class variance (iterative technique) (1). Let σ_w^2 represent the mathematical expression for within-class variance, σ_b^2 represent the variance of background pixels and σ_f^2 represent the variance of foreground pixels. Furthermore, W_b and W_f are mathematical symbols representing the background and foreground weights, respectively.

$$\sigma_w^2 = W_b\sigma_b^2 + W_f\sigma_f^2$$

Eq. 2 shows the equation for between-class variance (custom approach) (2). Let σ_B^2 represents the mathematical expression for between-class variance, with the background pixels' mean value representing the background, and the foreground pixels' mean value representing the foreground.

$$\sigma_B^2 = W_bW_f(\mu_b - \mu_f)^2$$

The Otsu thresholding approach is effective for image segmentation since it takes less time to compute than other methods. When identifying or calculating the best threshold value, Otsu thresholding uses a simple mathematical equation in its algorithm. Otsu's method is a finding an optimal threshold based on the observed distribution of pixel values (Otsu, 1979). Otsu is an acceptable method to achieve stable temporality classification for time series analysis (Wolfgang, 2012). However, if the histogram is unimodal or close to unimodal, the threshold is difficult to detect (Nicolas et al., 2010; Coudray et al., 2021). Applying automatic clustering-based thresholding to a large area or global thresholding may return unsatisfying results because its histogram sometimes is not bimodally distributed (Sezgin et al., 2004). Local thresholding methods try to overcome these problem considering thresholds only for a region with bimodally distributed and a value then, assigned to the large image (Vala et al., 2013).

References

- [1] Rättich, Michaela, Sandro Martinis, and Marc Wieland. "Automatic Flood Duration Estimation Based on Multi-Sensor Satellite Data." *Remote Sensing* 12, no. 4 (2020). <https://doi.org/10.3390/rs12040643>.
- [2] Senthilnath, J., S. N. Omkar, V. Mani, and P. G. Diwakar. "Multi-Temporal Satellite Imagery for Flood Damage Assessment." *Journal of the Indian Institute of Science* 93, no. 1 (2013): 105–16.
- [3] Tavus, B., S. Kocaman, H. A. Nefeslioglu, and C. Gokceoglu. "A Fusion Approach for Flood Mapping Using Sentinel-1 and Sentinel-2 Datasets." *International Archives of the*

- Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives 43, no. B3 (2020): 641–48. <https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-641-2020>.
- [4] Anusha, N., and B. Bharathi. “Flood Detection and Flood Mapping Using Multi-Temporal Synthetic Aperture Radar and Optical Data.” *Egyptian Journal of Remote Sensing and Space Science* 23, no. 2 (2020): 207–19. <https://doi.org/10.1016/j.ejrs.2019.01.001>.
- [5] Sakamoto, Toshihiro, Nhan Van Nguyen, Akihiko Kotera, Hiroyuki Ohno, Naoki Ishitsuka, and Masayuki Yokozawa. “Detecting Temporal Changes in the Extent of Annual Flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS Time-Series Imagery.” *Remote Sensing of Environment* 109, no. 3 (2007): 295–313. <https://doi.org/10.1016/j.rse.2007.01.011>.
- [6] Nigro, Joseph, Daniel Slayback, Frederick Policelli, and G Robert Brakenridge. “NASA/DFO MODIS Near Real-Time (NRT) Global Flood Mapping Product Evaluation of Flood and Permanent Water Detection.” *Evaluation*, Greenbelt, MD, 2014, 27.
- [7] Sivanpillai, Ramesh, Kevin M. Jacobs, Chloe M. Mattilio, and Ela V. Piskorski. “Rapid Flood Inundation Mapping by Differencing Water Indices from Pre- and Post-Flood Landsat Images.” *Frontiers of Earth Science* 15, no. 1 (2021). <https://doi.org/10.1007/s11707-020-0818-0>.
- [8] Yulianto, Fajar, NFn Suwarsono, Sayidah Sulma, and Muhammad Rokhis Khomarudin. “Observing the Inundated Area Using Landsat-8 Multitemporal Images and Determination of Flood-Prone Area in Bandung Basin.” *International Journal of Remote Sensing and Earth Sciences (IJReSES)* 15, no. 2 (2019): 131. <https://doi.org/10.30536/j.ijreses.2018.v15.a3074>.
- [9] Pena-Regueiro, Jesús, Maria Teresa Sebastia-Frasquet, Jesús A. Aguilar-Maldonado, Javier Estornell, José Andrés Sanchis-Blay, Sergio Morell-Monzó, and Vicent Altur-Grau. “Sentinel-2 Analysis of Flooded Areas: Applied Case Study - La Safor Wetland, Spain.” *WIT Transactions on Ecology and the Environment* 242, no. December (2020): 63–70. <https://doi.org/10.2495/WP200061>.
- [10] Cao, Han, Hong Zhang, Chao Wang, and Bo Zhang. “Operational Flood Detection Using Sentinel-1 SAR Data over Large Areas.” *Water (Switzerland)* 11, no. 4 (2019). <https://doi.org/10.3390/w11040786>.
- [11] Singha, Mrinal, Jinwei Dong, Sangeeta Sarmah, Nanshan You, Yan Zhou, Geli Zhang, Russell Doughty, and Xiangming Xiao. “Identifying Floods and Flood-Affected Paddy Rice Fields in Bangladesh Based on Sentinel-1 Imagery and Google Earth Engine.” *ISPRS Journal of Photogrammetry and Remote Sensing* 166, no. May (2020): 278–93. <https://doi.org/10.1016/j.isprsjprs.2020.06.011>.
- [12] Hki, Masato O, Manabu W Atanabe, Ryo N Atsuaki, Takeshi M Otohka, Hiroto N Agai, Takeo T Adono, Shinichi S Uzuki, et al. “Flood Area Detection Using ALOS-2 PALSAR-2 Data for the 2015 Heavy Rainfall Disaster in the Kanto and Tohoku Area,

- Japan.” *Journal of the Remote Sensing Society of Japan (Japan)* 36, no. 4 (2016): 348–58. <https://doi.org/10.11440/rssj.36.348>.
- [13] Vanama, V. S.K., Mohamed Musthafa, Unmesh Khati, R. Gowtham, Gulab Singh, and Y. S. Rao. “Inundation Mapping of Kerala Flood Event in 2018 Using ALOS-2 and Temporal Sentinel-1 SAR Images.” *Current Science* 120, no. 5 (2021): 915–25. <https://doi.org/10.18520/cs/v120/i5/915-925>.
- [14] Tay, Cheryl W.J., Sang Ho Yun, Shi Tong Chin, Alok Bhardwaj, Jungkyo Jung, and Emma M. Hill. “Rapid Flood and Damage Mapping Using Synthetic Aperture Radar in Response to Typhoon Hagibis, Japan.” *Scientific Data* 7, no. 1 (2020): 1–9. <https://doi.org/10.1038/s41597-020-0443-5>.
- [15] Schumann, Guy, Paul D. Bates, Matthew S. Horritt, Patrick Matgen, and Florian Pappenberger. “Progress in Integration of Remote Sensing-Derived Flood Extent and Stage Data and Hydraulic Models.” *Reviews of Geophysics* 47, no. 3 (2009): 1–20. <https://doi.org/10.1029/2008RG000274>.
- [16] Wonsick, Margaret M., Rachel T. Pinker, and Yves Govaerts. “Cloud Variability over the Indian Monsoon Region as Observed from Satellites.” *Journal of Applied Meteorology and Climatology* 48, no. 9 (2009): 1803–21. <https://doi.org/10.1175/2009JAMC2027.1>.
- [17] Hess, Laura L., John M. Melack, Adriana G. Affonso, Claudio Barbosa, Mary Gastil-Buhl, and Evlyn M.L.M. Novo. “Wetlands of the Lowland Amazon Basin: Extent, Vegetative Cover, and Dual-Season Inundated Area as Mapped with JERS-1 Synthetic Aperture Radar.” *Wetlands* 35, no. 4 (2015): 745–56. <https://doi.org/10.1007/s13157-015-0666-y>.
- [18] Conde, Francisco Carreño, and María De Mata Muñoz. “Flood Monitoring Based on the Study of Sentinel-1 SAR Images: The Ebro River Case Study.” *Water (Switzerland)* 11, no. 12 (2019): 1–25. <https://doi.org/10.3390/w11122454>.
- [19] Hess, Laura L., John M. Melack, John M. Melack, Solange Filoso, Yong Wang, and Yong Wang. “Delineation of Inundated Area and Vegetation Along the Amazon Floodplain with the SIR-C Synthetic Aperture Radar.” *IEEE Transactions on Geoscience and Remote Sensing* 33, no. 4 (1995): 896–904. <https://doi.org/10.1109/36.406675>.
- [20] Jung, Yoon Taek, Sang-Eun Park, Chang-Sun Baek, and Dong-Hwan Kim. “Evaluation of Polarimetric Parameters for Flood Detection Using PALSAR-2 Quad-Pol Data.” *Korean Journal of Remote Sensing* 34, no. 1 (2018): 117–26.
- [21] Ohki, Masato, Takeo Tadono, Takuya Itoh, Keiko Ishii, Tsutomu Yamanokuchi, Manabu Watanabe, and Masanobu Shimada. “Flood Area Detection Using PALSAR-2 Amplitude and Coherence Data: The Case of the 2015 Heavy Rainfall in Japan.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12, no. 7 (2019): 2288–98. <https://doi.org/10.1109/JSTARS.2019.2911596>.

- [22] Manjusree, Panchagnula, L. Prasanna Kumar, Chandra Mohan Bhatt, Goru Srinivasa Rao, and Veerubhotla Bhanumurthy. "Optimization of Threshold Ranges for Rapid Flood Inundation Mapping by Evaluating Backscatter Profiles of High Incidence Angle SAR Images." *International Journal of Disaster Risk Science* 3, no. 2 (2012): 113–22. <https://doi.org/10.1007/s13753-012-0011-5>.
- [23] Manavalan, Ramanuja. "Review of Synthetic Aperture Radar Frequency, Polarization, and Incidence Angle Data for Mapping the Inundated Regions." *Journal of Applied Remote Sensing* 12, no. 02 (2018): 1. <https://doi.org/10.1117/1.jrs.12.021501>.
- [24] Henry, J. B., P. Chastanet, K. Fellah, and Y. L. Desnos. "Envisat Multi-Polarized ASAR Data for Flood Mapping." *International Journal of Remote Sensing* 27, no. 10 (2006): 1921–29. <https://doi.org/10.1080/01431160500486724>.
- [25] Horritt, M. S., D. C. Mason, D. M. Cobby, I. J. Davenport, and P. D. Bates. "Waterline Mapping in Flooded Vegetation from Airborne SAR Imagery." *Remote Sensing of Environment* 85, no. 3 (2003): 271–81. [https://doi.org/10.1016/S0034-4257\(03\)00006-3](https://doi.org/10.1016/S0034-4257(03)00006-3).
- [26] Irwin, Katherine, Alexander Braun, Georgia Fotopoulos, Achim Roth, and Birgit Wessel. "Assessing Single-Polarization and Dual-Polarization TerraSAR-X Data for Surface Water Monitoring." *Remote Sensing* 10, no. 6 (2018). <https://doi.org/10.3390/rs10060949>.
- [27] Otsu, N. "OTSU Paper." *IEEE Transactions on Systems, Man and Cybernetics* 20, no. 1 (1979): 62–66.
- [28] Nakmuenwai, Pisut, Fumio Yamazaki, and Wen Liu. "Automated Extraction of Inundated Areas from Multi-Temporal Dual-Polarization Radarsat-2 Images of the 2011 Central Thailand Flood." *Remote Sensing* 9, no. 1 (2017). <https://doi.org/10.3390/rs9010078>.
- [29] Conference, Asian, Remote Sensing, Korea Thp-, Husniyah Binti Mahmud, Masahiko Nagai, and Yamaguchi Email. "Automated Extraction of Flood for Large Scale Area Using," no. Acrs (2019): 1–10.
- [30] Huang, Chang, Yun Chen, Shiqiang Zhang, and Jianping Wu. "Detecting, Extracting, and Monitoring Surface Water from Space Using Optical Sensors: A Review." *Reviews of Geophysics* 56, no. 2 (2018): 333–60. <https://doi.org/10.1029/2018RG000598>.
- [31] MCFEETERS, and SK. "The Use of Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features." *International Journal of Remote Sensing* 17 (1996): 1425–32.
- [32] Xu, Hanqiu. "Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery." *International Journal of Remote Sensing* 27, no. 14 (2006): 3025–33. <https://doi.org/10.1080/01431160600589179>.

- [33] Feyisa, Gudina L., Henrik Meilby, Rasmus Fensholt, and Simon R. Proud. “Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery.” *Remote Sensing of Environment* 140 (2014): 23–35. <https://doi.org/10.1016/j.rse.2013.08.029>.
- [34] Jiang, Wei, Yuan Ni, Zhiguo Pang, Xiaotao Li, Hongrun Ju, Guojin He, Juan Lv, Kun Yang, June Fu, and Xiangdong Qin. “An Effective Water Body Extraction Method with New Water Index for Sentinel-2 Imagery.” *Water (Switzerland)* 13, no. 12 (2021). <https://doi.org/10.3390/w13121647>.
- [35] Vu, Authors, Anh Tuan, L E Thi, Thu Hang, and Nguyen Hong Quang. “Monitoring Urban Surface Water Bodies Changes Using MNDWI Estimated from Pan-Sharpned Optical Satellite Images Monitoring Urban Surface Water Bodies Changes Using MNDWI Estimated from Pan-Sharpned Optical Satellite Images (9771) Vu Anh Tuan, Le Thi T,” no. 9771 (2019).
- [36] Husniyah. B. M. and M. Nagai. “Classification of Surface Water Area from Multispectral and Multisar Data Using Automatic and Robust System,” *The 41st Asian Conference on Remote Sensing (ACRS 2020) November 9-11, 2020. Deqing, China, 2020*.
- [37] Mohammed Daboor and Brian Brisco (November 5th 2018). *Wetland Monitoring and Mapping Using Synthetic Aperture Radar, Wetlands Management - Assessing Risk and Sustainable Solutions*, Didem Gökçe, IntechOpen, DOI: 10.5772/intechopen.80224. Available from: <https://www.intechopen.com/chapters/63701>
- [38] Lira, J. Segmentation and morphology of open water bodies from multispectral images. *Int. J. Remote Sens* 2006, 27, 4015–4038
- [39] Sethre, P.R.; Rundquist, B.C.; Todhunter, P.E. Remote detection of prairie pothole ponds in the Devils Lake basin, North Dakota. *GISci. Remote Sens.* 2005, 42, 277–296, doi:10.2747/1548-1603.42.4.277.
- [40] Jain, S.K., A.K. Saraf, A. Goswami, and T. Ahmad, 2006. Flood inundation mapping using NOAA AVHRR data, *Water Resources Management*, 20(6):949–959.
- [41] McFeeters, S. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens* 1996, 17, 1425–1432.
- [42] Rogers, A.S., and M.S. Kearney, 2004. Reducing signature variability in unmixing coastal marsh Thematic Mapper scenes using spectral indices, *International Journal of Remote Sensing*, 25(12):2317–2335.
- [43] Xu, H., Wang, X. & Zhang, X. Decreased vegetation growth in response to summer drought in Central Asia from 2000 to 2012. *International journal of applied earth observation and geoinformation* 52, 390–402 (2016).

- [44] Ryu J., Won J., and Min K.D. 2002. Waterline extraction from Landsat TM data in a tidal flat: a case study in Gomso bay, Korea. Department of Geography & Anthropology Louisiana State University. Xu, H., Wang, X. & Zhang, X. Decreased vegetation growth in response to summer drought in Central Asia from 2000 to 2012. *International journal of applied earth observation and geoinformation* 52, 390–402 (2016).
- [45] Gstaiger, V., Huth, J., Gebhardt, S., Wehrmann, T., Kuenzer, C. and multi-Skaliger, M.S., 2008, Vergleich Von Ueberflutungsflaechen Abgeleitet Aus TerraSAR-X and Envisat ASAR. In *Proceedings of the 20th AGIT Symposium*, 2–4 July 2008, Salzburg, Austria (Salzburg: Wichmann), pp. 364–369
- [46] Kuenzer, C., H. Guo, J. Huth, P. Leinenkugel, X. Li and S. Dech. 2013. Flood mapping and flood dynamics of the Mekong Delta: ENVISAT-ASAR-WSM based time series analyses. *Remote Sensing* 52: 687–715
- [47] Claudia Kuenzer & Kim Knauer (2013) Remote sensing of rice crop areas, *International Journal of Remote Sensing*, 34:6, 2101-2139, DOI: 10.1080/01431161.2012.738946
- [48] S Kalathiya, VP Patel. Implementation of Otsu Method with Two Different Approaches. *International Journal of Software & Hardware Research in Engineering*. 2014; 2(2).
- [49] Du, Zhiqiang, Bin Linghu, Feng Ling, Wenbo Li, Weidong Tian, Hailei Wang, Yuanmiao Gui, Bingyu Sun, and Xiaoming Zhang. “Estimating Surface Water Area Changes Using Time-Series Landsat Data in the Qingjiang River Basin, China.” *Journal of Applied Remote Sensing* 6, no. 1 (2012): 063609. <https://doi.org/10.1117/1.jrs.6.063609>.

Chapter III Overview of the work

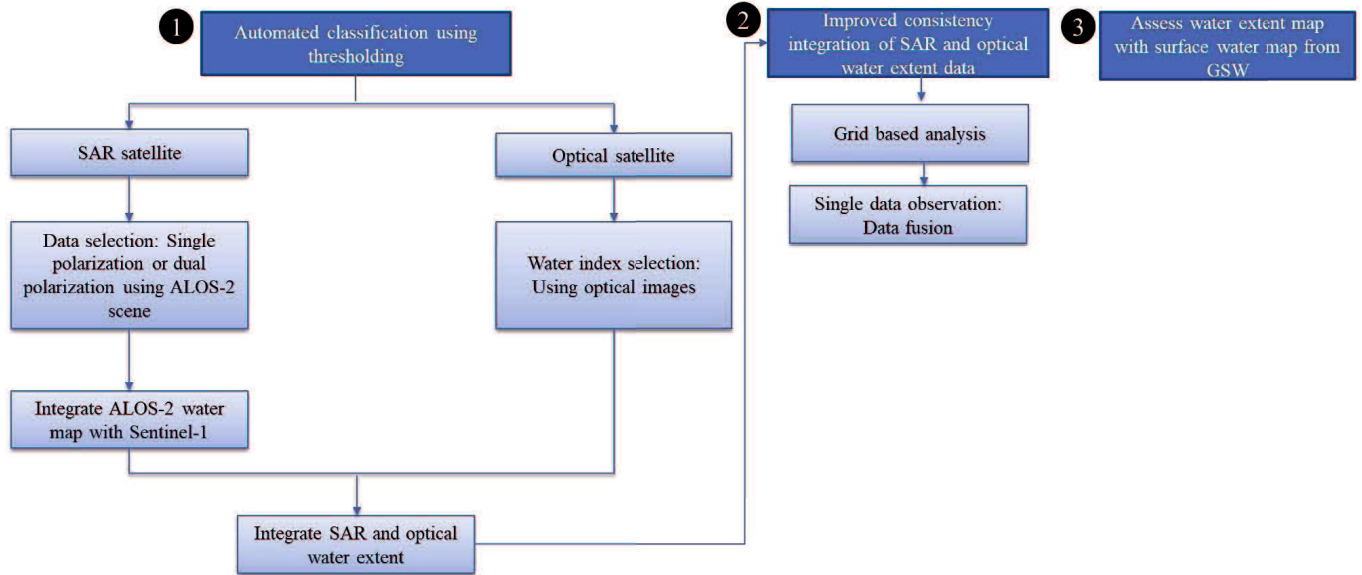


Figure 3.1: Overview of the work.

To summarize the research objectives, we create the overview of the work as shown in Figure 3.1. First of all, we start with assessing suitable data for automated classification using SAR data. Either to use single or dual polarization using ALOS-2 images. Next, we integrate ALOS-2 and Sentinel-1 extracted water area based on a temporal coverage of flood event. Then, we assess the stability of water index using various optical data and we integrate multiSAR and multispectral together. Second main work, is to improve the consistency of the integration result using grid based and fused of single data. In addition, we assess the water extent with surface water map from GSW.

Chapter IV

Automated classification using thresholding

4.1 Introduction

Flood is the most disastrous natural disaster in Malaysia. According to DID's annual flood report for 2016/2017, 404 flood occurrences were identified across Malaysia, resulting in a total economic loss of USD12,649,379.84 and 95,929 people being displaced from their homes [1]. In Malaysia, there are two types of floods: monsoon floods and flash floods. From November to March, the Northeast monsoon brought heavy rains to Peninsular Malaysia's east coast, the northern part of Sabah, and the southern part of Sarawak.



Figure 4.1: Photographs from the Perlis flood disaster on September 22nd and 23rd, 2017.

Delineation approaches, object or pixel-based classifications, and pre and post-flood comparisons are all common techniques for detecting floods using remote sensing data. The purpose of this research is to apply machine learning to identify floods automatically. The automatic thresholding method using radar data proposed was adopted in this study, Pisut et al., 2017 [2].

Using averages of local Otsu's threshold values and total backscatter of HH and HV of ALOS-2, this study aims to create an automated surface water extraction method. Sentinel-1, Landsat 7, Landsat 8 and MODIS, derived indices used for monitored spatial-temporal water movement during the start date and end date of the 2017 flood.

4.2 Study area

The study area is in the northwest of Peninsular Malaysia, which includes the states of Perlis and Kedah. The study area is approximately 5.35°N to 6.35°N longitude and 99°46'3.22"E to 99°59'25.07"E latitude, as illustrated in Fig. 4.2. The largest paddy growing area in Malaysia is in the northwest. The climate regime in Northwest Malaysia is characterized by a typical dry season that lasts 2 to 4 months from December to March. The rainy season occurs from September to November and April to May.

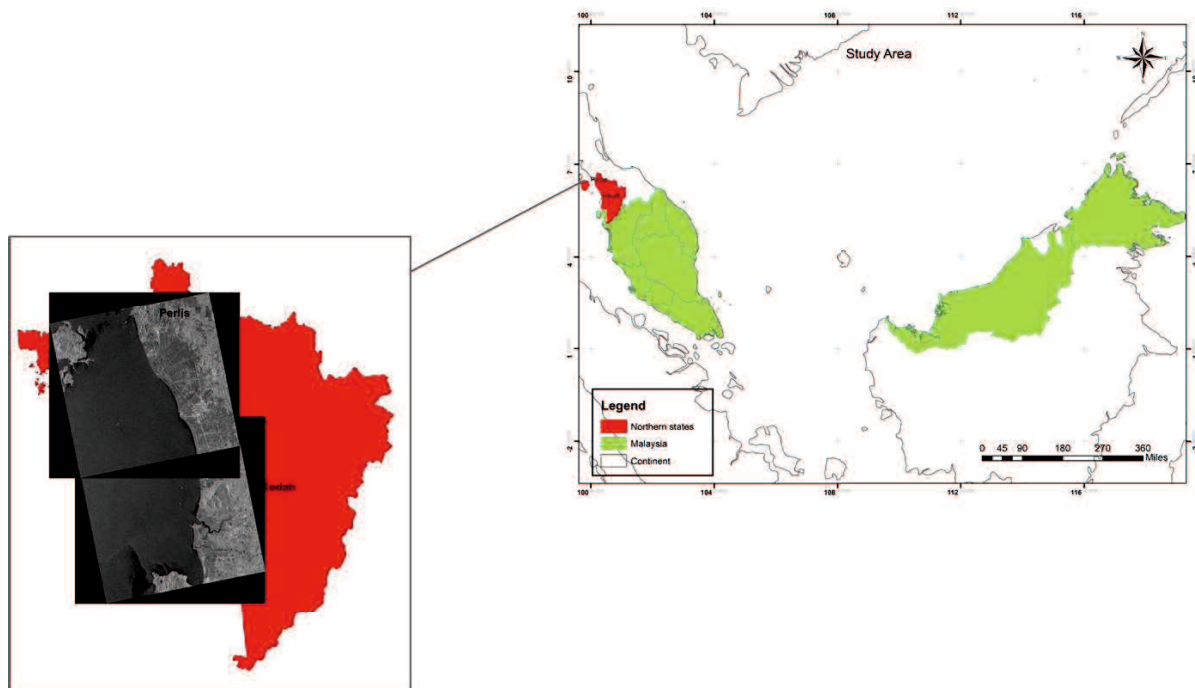


Figure 4.2: Study area

The state of Kedah receives an average annual rainfall of 2400 mm. In 2017, the average annual rainfall for Kedah was less than 7%, which was 2245mm. Although, the annual average rainfall is reduced but floods still occur in the state. A total of 57 flood incidents were recorded in 2017. The recorded flood incidents are a series of flash flood incidents and flood incidents caused by overflow from rivers. In September 2017, there were 17 reports of floods throughout the state of Kedah. The state of Kedah was in the northeast monsoon season where rain and strong winds hit almost every day. The water level in the main rivers increased and eventually

overflowed when the rain continued and high tide conditions that caused the outflow of river water to the sea is disrupted. The flood period was also long and involved a lot of evacuation, especially for victims who live close to the river. The topography of Perlis mostly flat with less than 61 meters above sea level. Perlis experience rainfall from April to May and August to October with average rainfall per year ranges between 1704 mm and 2005 mm. In this study, only Arau and Kangar chose a study area because of only these two areas reported in the Department of Irrigation and Drainage Flood Annual Report 2017/2018 facing flood twice in the year 2017/2018.

4.3 Satellite data

In this study, ALOS-2 data are utilized. Table 4.1 summarizes the description of ALOS-2. In total, there are 18 ALOS-2 images with ascending mode, with a constant off-nadir angle of 32.9°, a spatial resolution of 6.25 m, and in dual-polarization, HH and HV are selected. ALOS-2 data was ordered from <https://auig2.jaxa.jp>.

Table 4.1: Utilized PALSAR-2 data.

No.	Scene ID	Operation Mode	Date	Pass	Frame	Polarization mode
1	ALOS2019460100-141002	FBD	2014/10/02	Ascending	100	HH/HV
2	ALOS2019460110-141002	FBD	2014/10/02	Ascending	110	HH/HV
3	ALOS2040160100-150219	FBD	2015/02/19	Ascending	100	HH/HV
4	ALOS2040160110-150219	FBD	2015/02/19	Ascending	110	HH/HV
5	ALOS2073280100-151001	FBD	2015/10/01	Ascending	100	HH/HV
6	ALOS2073280110-151001	FBD	2015/10/01	Ascending	110	HH/HV
7	ALOS2093980100-160218	FBD	2016/02/18	Ascending	100	HH/HV
8	ALOS2093980110-160218	FBD	2016/02/18	Ascending	110	HH/HV
9	ALOS2127100100-160929	FBD	2016/09/29	Ascending	100	HH/HV
10	ALOS2127100110-160929	FBD	2016/09/29	Ascending	110	HH/HV
11	ALOS2147800100-170216	FBD	2017/02/16	Ascending	100	HH/HV
12	ALOS2147800110-170216	FBD	2017/02/16	Ascending	110	HH/HV
13	ALOS2180920100-170928	FBD	2017/09/28	Ascending	100	HH/HV
14	ALOS2180920110-170928	FBD	2017/09/28	Ascending	110	HH/HV
15	ALOS2199550100-180201	FBD	2018/02/01	Ascending	100	HH/HV
16	ALOS2199550110-180201	FBD	2018/02/01	Ascending	110	HH/HV
17	ALOS2203690100-180301	FBD	2018/03/01	Ascending	100	HH/HV
18	ALOS2203690110-180301	FBD	2018/03/01	Ascending	110	HH/HV

4.4 Research flowchart

Figure 4.3 shows the flowchart of spatial-temporal water movement detection. Initially, ALOS-2 image preprocessed using Sentinel Application Platform (SNAP) software. Preprocessed of ALOS-2 image involve merge of two ALOS-2 frames to cover study area, calculation of beta Nought, Lee speckle filtering with 5x5 windows and sigma Nought calculation. Afterward, each co polarized HH (horizontal transmit and horizontal receive) image and dual polarized which is total backscatter of HH and HV cut into small specific water

references. The criteria of water references explained in the next section. Histogram of Otsu of each water references determined fully automatic by developing Python script with software library like GDAL and Numpy.

Otsu threshold value of each water references with bimodal shape used for averaging to represent the whole region threshold while unimodal histogram rejected for next processed. Among the water extraction map from each average Otsu image threshold was compared and the most precise segmented image was selected for final representation of water movement map and integrated with Sentinel-1, Landsat 7, Landsat 8 and MODIS.

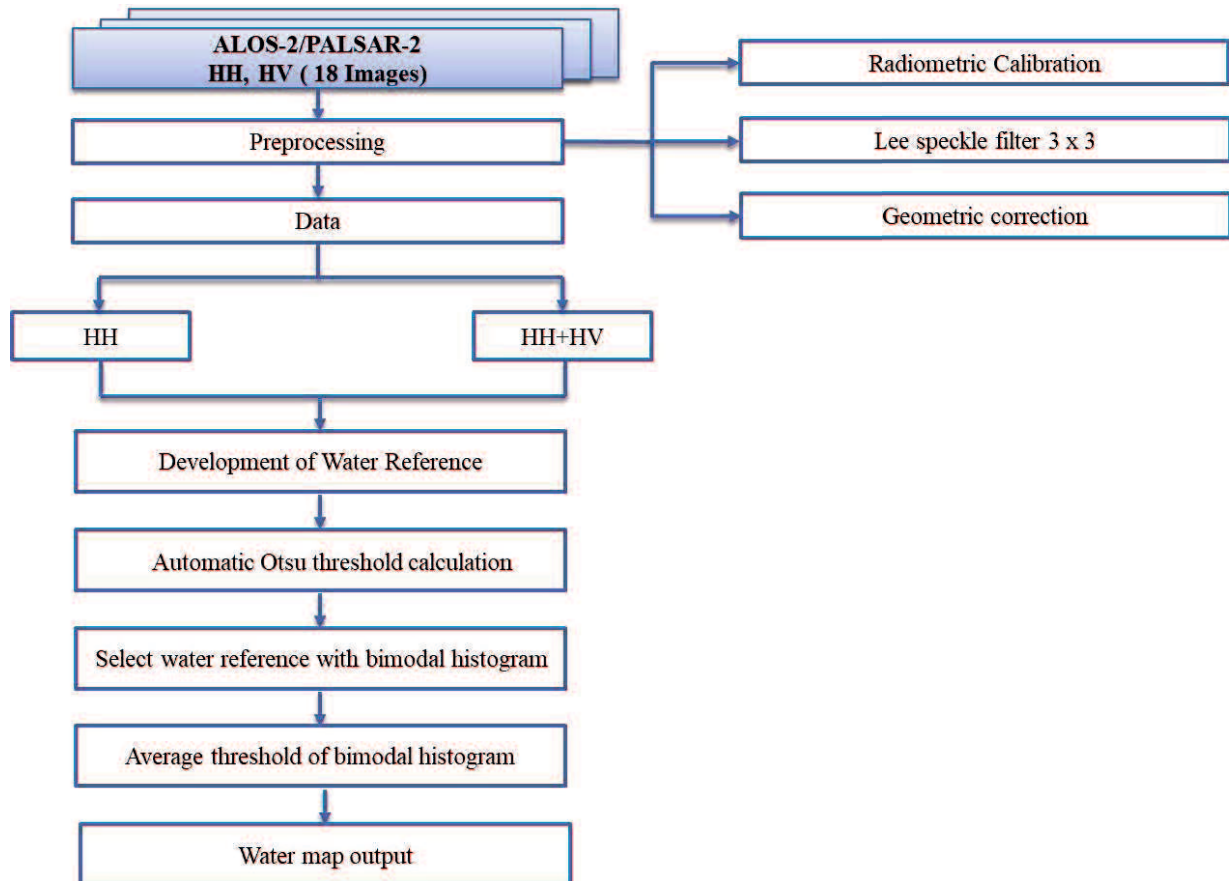


Figure 4.3: flowchart of spatial-temporal water movement detection

4.5 Development of water references

All the water references were selected using the following criteria proposed by Pisut et al. (2017): having an area larger than 320000m² (8192 pixels for a resolution 6.25m), containing water throughout the year, not facing a flood situation, located on flat ground as much as possible, and having water and non-water cover ratio of nearly 1:1. An irregular shape was allowed. In this study a circle shape is used to maintain ratio of water and non-water 1:1 ratio.

The 68 water references were selected from different types of water bodies, natural and man-made. The locations water references distribution and criteria are shown in Fig. 4.4.

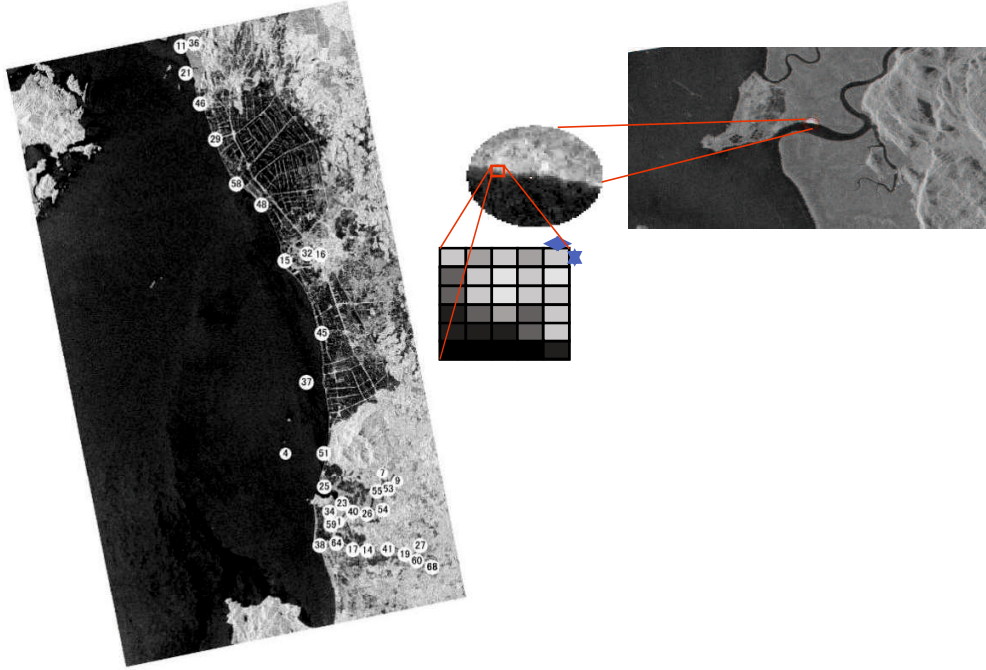


Figure 4.4: water references selection criteria.

4.6 Automatic thresholding using Otsu Thresholding in Python

Various thresholding algorithm have been proposed to optimally segment an image into two classes for example minimum error thresholding, IsoData thresholding and Yen thresholding. In this study, we decide to use Otsu because of its simplicity, robustness and adaptability on any input data as we are using multi source of satellite images. Among many threshold selection methods, Otsu is the optimum one in the sense that it maximizes the between class variance, a well-known measure used in statistical discriminant analysis.

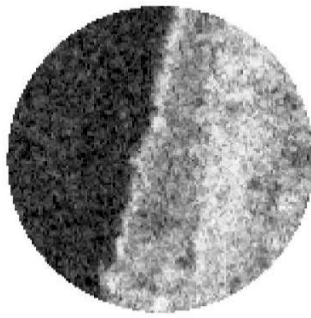
However, we did try to compare on using other modification of Otsu method. Three methods were tested to detect a threshold value for water presence from ALOS-2 automatically: Otsu, Otsu Village emphasis thresholding and Otsu neighbourhood emphasis, but the result from the experiment shows no significant difference between these three methods. Therefore, Otsu's is sufficient for further analysis. Image thresholding involves converting grayscale image ($f(x,y)$) to binary image ($g(x,y)$) where pixels value lower than threshold classified as zero while pixels higher than threshold classified as one. In general, this can be simplified as follows:

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) \geq T \\ 0 & \text{otherwise} \end{cases}, (1)$$

(1) First, to compute the Otsu threshold, the histogram of the pixels computed. Probability of threshold, T determined by dividing gray image pixels into the background, b and foreground, p . Then, the weight of background, w_b and foreground pixels, w_p calculated by dividing the total of background pixels frequency, f_b with the total number of entire pixels as follows:

$$W_b (\text{Weight background}) = \frac{f_b}{f_1 + f_2 + \dots + f_n} \text{ Total number of entire pixels}, (2)$$

$$W_p (\text{Weight foreground}) = \frac{f_p}{f_1 + f_2 + \dots + f_n} \text{ Total number of entire pixels}, (3)$$



Pixel value
 ■ 11.9626
 □ 535.438

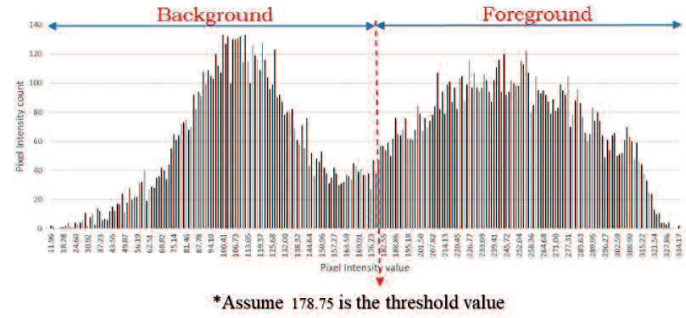


Figure 4.5: Determination of foreground and background using water references with bimodal

(2) the mean of background and foreground pixels calculated by adding multiplication of background pixels frequency, f_b with its pixels value V_b and divide with the sum of background pixels. The calculation is as follows:

$$\mu_b (\text{Mean background}) = (f_{b1} * v_{b1}) + (f_{b2} * v_{b2}) + \dots + (f_n * v_n) / \text{Total number of background pixels}, \quad (4)$$

$$\mu_p (\text{Mean foreground}) = (f_{p1} * v_{p1}) + (f_{p2} * v_{p2}) + \dots + (f_n * v_n) / \text{Total number of foreground pixels}, \quad (5)$$

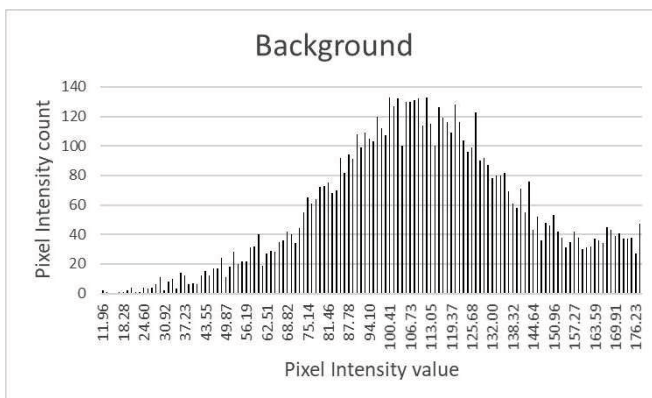
(3) The variance of background pixels and foreground pixels calculated by averaging square distance background frequency pixels with the mean background

$$V_b (\text{Variance background}) = ((f_{b1} - \mu_b)^2 * v_{b1}) + ((f_{b2} - \mu_b)^2 * v_{b2}) \dots ((f_n - \mu_b)^2 * v_n) + ((f_n - \mu_b)^2 * v_n) / \text{Total number of background pixels}, \quad (6)$$

$$V_p (\text{Variance foreground}) = ((f_{p1} - \mu_p)^2 * v_{p1}) + ((f_{p2} - \mu_p)^2 * v_{p2}) \dots ((f_n - \mu_p)^2 * v_n) + ((f_n - \mu_p)^2 * v_n) / \text{Total number of foreground pixels}, \quad (7)$$

(4) Afterward, the within-class variance was calculated by totaling background and foreground variance multiplied with their weights.

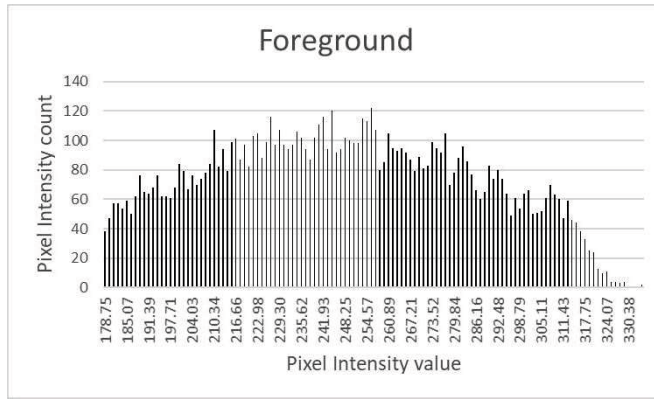
$$\text{Within class variance} = W_b V_b + W_f V_f, \quad (8)$$



$$\text{Weight } W_b = \frac{2 + 1 + 0 + 0 \dots}{16265} = 0.45$$

$$\begin{aligned} \text{Mean } \mu_b &= \frac{(11.96 \times 2) + (13.23 \times 1) + (14.49 \times 0) + (15.75 \times 0) \dots}{7239} \\ &= 110.79 \end{aligned}$$

$$\begin{aligned} \text{Variance } \sigma_b^2 &= \frac{((11.96 - 110.79)^2 \times 2) + ((13.23 - 110.79)^2 \times 1) + ((14.49 - 110.79)^2 \times 0) + ((15.75 - 110.79)^2 \times 0) \dots}{7239} \\ &= 959.72 \end{aligned}$$



$$\text{Weight } W_f = \frac{38 + 47 + 57 + 57 \dots}{16265} = 0.56$$

$$\begin{aligned} \text{Mean } \mu_f &= \frac{(178.75 \times 38) + (180.02 \times 47) + (181.28 \times 57) + (182.55 \times 57)}{9026} \\ &= 247.50 \end{aligned}$$

$$\begin{aligned} \text{Variance } \sigma_f^2 &= \frac{((178.75 - 247.50)^2 \times 6) + ((180.02 - 247.50)^2 \times 14) + ((181.28 - 247.50)^2 \times 14) + ((182.55 - 247.50)^2 \times 14)}{9026} \\ &= 1107.59 \end{aligned}$$

$$\text{Within class variance} = \sigma_w^2 = 1180.65$$

Figure 4.6: Calculation within class variance

Threshold	T = 154.75	T = 169.91	T = 176.23	T = 178.75
Weight, Background	$W_b = 0.40$	$W_b = 0.43$	$W_b = 0.44$	$W_b = 0.45$
Mean, Background	$\mu_b = 104.83$	$\mu_b = 108.39$	$\mu_b = 110.10$	$\mu_b = 110.79$
Variance, Background	$\sigma_b^2 = 690.68$	$\sigma_b^2 = 839.39$	$\sigma_b^2 = 923.84$	$\sigma_b^2 = 959.72$
Weight, Foreground	$W_f = 0.59$	$W_f = 0.57$	$W_f = 0.56$	$W_f = 0.55$
Mean, Foreground	$\mu_f = 241.67$	$\mu_f = 245.39$	$\mu_f = 246.93$	$\mu_f = 247.50$
Variance, Foreground	$\sigma_f^2 = 1702.71$	$\sigma_f^2 = 1470.68$	$\sigma_f^2 = 1386.86$	$\sigma_f^2 = 1357.84$
Within class variance	$\sigma_w^2 = 1295.85$	$\sigma_w^2 = 1200.04$	$\sigma_w^2 = 1182.89$	$\sigma_w^2 = 1180.65$

Figure 4.7: Same calculation needs to be performed for all the possible threshold values

Otsu's method is one of the best threshold selection methods for general gray-level images. This technique chooses the threshold value of the minimum within-class variance (σ^2_w) or the maximum between-class variance (σ^2_B) in Equation (1). Although this method can obtain satisfactory segmentation results in many cases, it is limited to images with background and foreground Gaussian distributions of equal variance. Therefore, images that do not meet this criterion may return unsatisfactory results, especially when the gray level histogram is unimodal or close to a unimodal distribution.

To address this weakness, many modifications of the Otsu method have been proposed. For example, the valley-emphasis method (VE), modified by weight σ^2_B with $p(t)$, the

complement of a probability at a threshold value t , causes the valley in the histogram to be more likely to be better determined. The neighborhood valley-emphasis (NE) improves the valley-emphasis method by weighting σ^2B with the neighborhood information in $n = 2m + 1$ intervals at the threshold value. The result is closer to the valley of the histogram because it considers the neighborhood around the threshold point in addition to the threshold point. Figure 4.8 shows the coding for automatic thresholding using Otsu thresholding and its modified in Python. The automatic process used the modified applicable Python script as the open-source Otsu algorithm is unsuitable for spatial data images.

```

Threshold.py - C:\Threshold\Threshold\Threshold.py (3.7.0)
File Edit Format Run Options Window Help
#print("%04d",hist_total)

hist_valley = hist_count.ravel()/hist_total

otsu_min = np.inf
otsu_key = -1

otsu_max = -np.inf
otsu_key2 = -1

valley_max = -np.inf
valley_key = -1

neighb_shift = int(math.ceil(neighb_count-1/2))
neighb_max = -np.inf
neighb_key = -1

for i in range(data_bin):
    nL,nR = np.hsplit(hist_count,[i+1]) #number of items
    sL,sR = np.hsplit(hist_series,[i+1]) #value of items
    wL,wR = nL.sum()/hist_total, nR.sum()/hist_total #weight of side
    mL,mR = (nL*sL).sum()/nL.sum(), (nR*sR).sum()/nR.sum() #mean of side
    vL,vR = (((sL-mL)**2)*nL).sum()/nL.sum(), (((sR-mR)**2)*nR).sum()/nR.sum() #variance of side

    otsu_wic = (wL*vL)+(wR*vR) #within-class

    if otsu_wic < otsu_min:
        otsu_min = otsu_wic
        otsu_key = i

    otsu_btc = wL*wR*(mL-mR)**2
    if otsu_btc > otsu_max:
        otsu_max = otsu_btc
        otsu_key2 = i

    """
    print("%-8s")
    print("M", mL, mR)
    print("V", vL, vR)
    print("WIC",i, otsu_key, otsu_wic, otsu_min)
    print("BTC",i, otsu_key2, otsu_btc, otsu_max)
    """

    #-----
    # calculate valley within-class
    valley_val = otsu_btc*(1-hist_valley[i])
    if valley_val > valley_max:
        valley_max = valley_val
        valley_key = i
    #print("Valley", valley_key, valley_val, valley_max)

    #-----
    # calculate neighborhood within-class
    neighb_left = i - neighb_shift
    neighb_right = i + neighb_shift + 1
    neighb_val = otsu_btc*(1-hist_valley[neighb_left:neighb_right].sum())
    if neighb_val > neighb_max:
        neighb_max = neighb_val
        neighb_key = i
    #print("Neighb",neighb_key, neighb_val, neighb_max)

hist_delta = int(round(((polyfit peakdet/100)*hist_count.sum()))

```

Figure 4.8: Automatic thresholding using Otsu Thresholding in Python

4.7 Results and discussions

4.7.1 Automated thresholding with OTSU

By considering the peaks, valleys, and curvatures of the smoothed histograms, the water references with unimodal distributions were rejected, and only those with bimodal distributions were considered. The weight average value of automated threshold for Otsu technique display on each whole HH PALSAR-2 image histogram in Fig. 4.9.



Figure 4.9: The weigh average of HH applied on the whole area

Fig. 4.10 shows the histogram of HH+HV PALSAR-2 with applied weight average of Otsu.



Figure 4.10: The weigh average of HH+HV applied on the whole area

Based on Fig. 4.9 by applying Otsu on the HH image the threshold value is not accurately located at the valley of the histogram compare with Fig. 4.10 where by using HH+HV threshold value mostly located at closest valley point of the histogram.

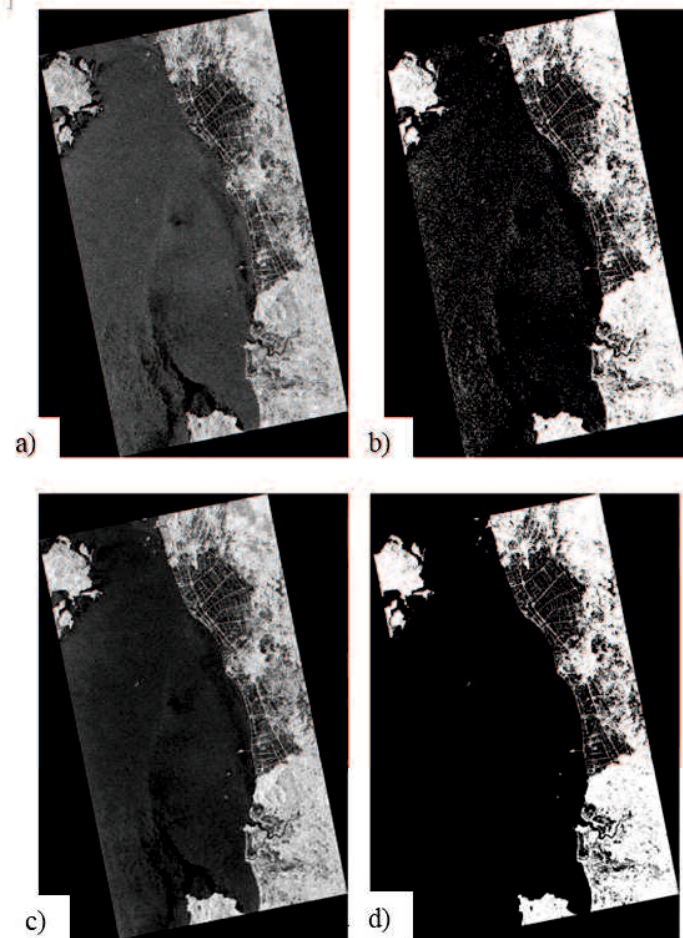


Figure 4.11: Extracted result of weight average from the image acquired on October 02 2014 a) raw ALOS-2 HH, b) Otsu ALOS-2 HH, c) raw ALOS-2 HH+HV d) Otsu ALOS-2 HH+HV.

The image segmentation by using Otsu was used to simply extract water surface by applying a global threshold value to each PALSAR-2 image. Fig. 4.11 shows October 02 2014 as the sample image of the extracted result. Otsu threshold of HH image = 48.4 while, Otsu threshold of HH+HV = 84.3. The result shows that HH+HV gives a better surface water extraction compared to HH by using VE technique, threshold of HH = 48.5 and threshold of HH+HV = 84.6. While, NE weight average threshold for HH image = 48.0 and HH+HV image = 83.4. The range of each threshold technique for each single image indistinct with each other.

Based on Fig. 4.11, HH+HV image separated the object clearly with more object details than HH image. Otsu, VE, and NE technique does not give significant impact on separating the object with the background. In this study, threshold from Otsu technique selected to represent water movement pre-flood, during and post flood event.

As seen in Table 4.2, the bimodal percentage number indicates the occurrence probability of the bimodal distribution for each image. Thus, HH + HV is more likely to have a bimodal distribution with 73%-85% and is more suitable for automatic classification compare to HH that have percentage 56%-72% to get bimodal distribution. In other words, the HV polarization can

improve the efficiency of water surface extraction. Thus, the extracted water areas presented in this study were derived from HH + HV.

Table 4.2: Automatic threshold values using the Otsu technique of water reference with circle areas from the HH and HH + HV sigma-naught values taken on the evening of 02 October 2014 5.30 PM until 01 March 2018 5.30 PM.

Local Time	Pass	Beam	Number of Water References	Bimodal Threshold HH		Bimodal Threshold HH+HV	
				Bimodal (%)	Weigh Average	Bimodal (%)	Weigh Average
2014/10/02	Ascending	FBD	68	67.64	48.4	80.88	84.3
2015/02/19	Ascending	FBD	68	66.18	49.4	73.53	85.3
2015/10/01	Ascending	FBD	68	72.06	50.9	77.94	87.2
2016/02/18	Ascending	FBD	68	57.35	49.1	80.88	85.1
2016/09/29	Ascending	FBD	68	67.65	51.4	80.88	88.7
2017/02/16	Ascending	FBD	68	55.88	51.4	83.82	88.7
2017/09/28	Ascending	FBD	68	69.12	137.3	80.88	146.9
2018/02/01	Ascending	FBD	68	83.82	43.5	85.29	79.0
2018/03/01	Ascending	FBD	68	67.65	52.5	79.41	86.7

Some of the close-ups of the extracted result from the HH+HV image acquired on October 02 2014 with Otsu threshold value are shown in Figure 4.12. The water boundaries obtained using this technique appear to be reasonable. The results of the water area are shown in white, and the non-water area shown in blue.

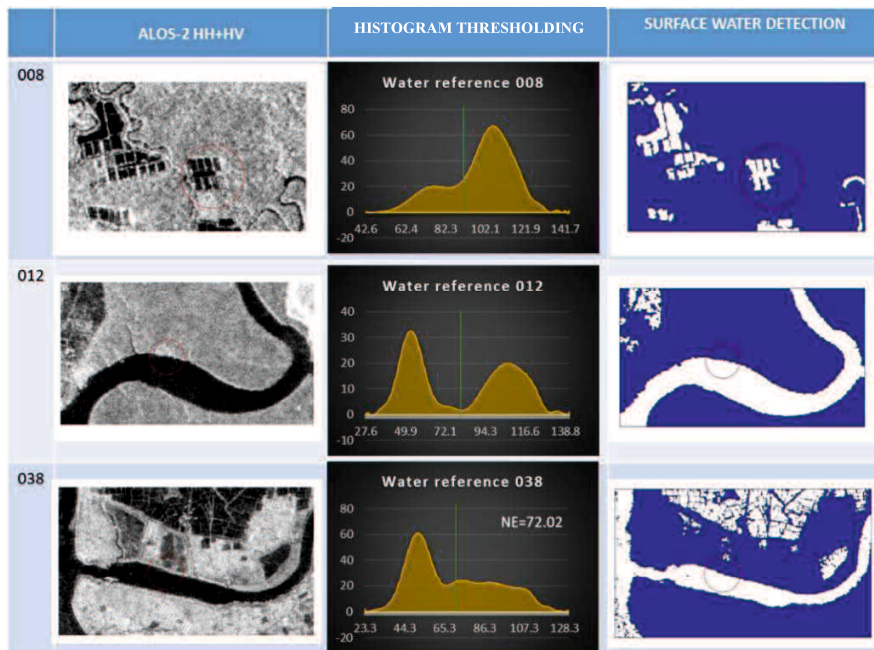


Figure 4.12: Close-ups of the extracted result from the image HH+HV acquired on October 02 2014

There are some errors in waterline area positions due to side-looking geometry of the SAR sensors cause shadowing, foreshortening and layover effects and underestimation of waterline.

4.7.2 Water classification using Otsu technique

Figure 4.13 shows a comparison of ALOS-2 for 2014 and 2017 flood.

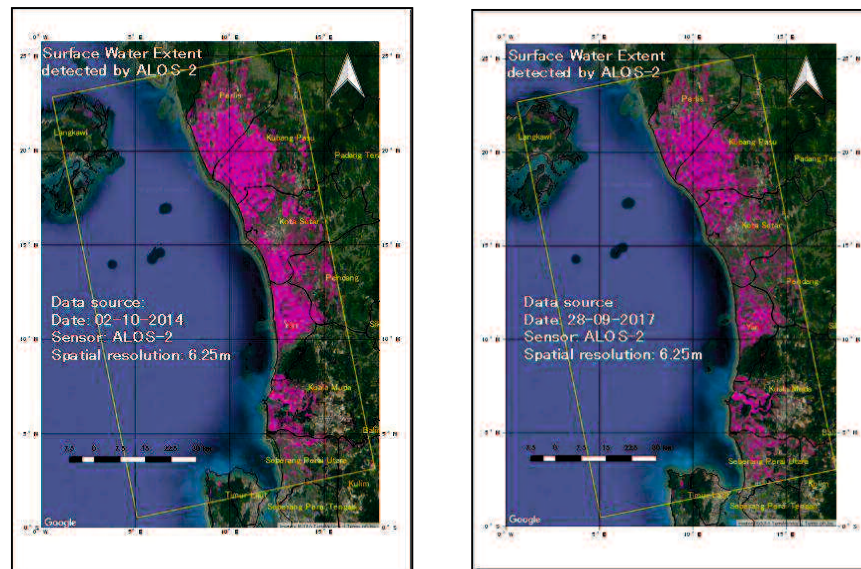


Figure 4.13: Visual of surface water extent using ALOS-2 6.25m for 2014 and 2017 flood.

Figure 4.13 shows the extracted surface water with basis from HH+HV data during 2014 and 2017 flood event. ALOS-2 gives detail and precise flood occurrence due to the high spatial resolution.

4.7.3 Cloud masking of multispectral images

Multispectral images are from Sentinel-2, Landsat-7, Landsat-8, and MODIS MOD09A1. Sentinel-2 revisit period is five days. Sentinel-2B was launched on March 7, 2017. The Sentinel-2 image has a different spatial resolution on each band. Due to its different spatial resolution size for the SWIR band, we decide to only apply NDWI on the Sentinel-2 image. We use only green band B3 (560 nm) and its Near-infrared is B8 (842 nm) with both 10m spatial resolution.

For Landsat 7 carries the Enhanced Thematic Mapper Plus (ETM+) sensor and has acquired and delivered data with data gaps caused by the Scan Line Corrector (SLC) failure. It has 30 m spatial resolution and the image we utilized in this study are Band 1 Visible (0.45 - 0.52 μm), Band 2 Visible (0.52 - 0.60 μm), Band 4 Near-Infrared (0.77 - 0.90 μm), Band 5 Short-wave Infrared (1.55 - 1.75 μm) and Band 7 Mid-Infrared (2.08 - 2.35 μm).

Meanwhile, Landsat 8 carries the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) instruments. Band 2 Visible (0.450 - 0.51 μm), Band 3 Visible (0.53 - 0.59 μm), Band 5 Near-Infrared (0.85 - 0.88 μm), Band 6 SWIR 1(1.57 - 1.65 μm) and Band 7 SWIR 2 (2.11 - 2.29 μm). All of these images are 30 m resolution.

The MODIS Terra MOD09A1 is an 8-day product (<http://LPDAAC.usgs.gov>) starting from the 6 September 2017 is used in this study. The layers used from the MOD09A1 product are the surface reflectance band 2 (841-876 nm), surface reflectance band 3 (459-479 nm), surface reflectance band 4 (545-565 nm), surface reflectance band 5 (1230-1250 nm), and surface reflectance band 6 (1628-1652 nm). All of these multispectral images data are available on GEE, which enables a quick and multi-scale analysis. For the quantitative analyses of the Landsat, MODIS Terra and Aqua images cloud masking was performed as shown in Figure 3.14 [48]. The pre-processing of optical images was fully implemented in Google Earth Engine (GEE). Landsat 7 ETM+ contain data gaps due to the failure of the Scan Line Corrector (SLC), which were filled in GEE. For Landsat 7 masked for “cloud shadows” is bit 3 while “clouds” are bit 5 respectively. Cloud mask for Landsat 8 is bit 8-9 and for MODIS state_1km was masked and assign according to this cloudBitMask = 1 << 2, cloudshadowBitMask = 1 << 9, cloudInternalBitMask = 1 << 10; adjacentBitMask = 1 << 13.

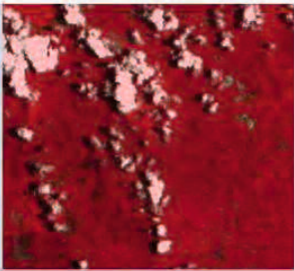
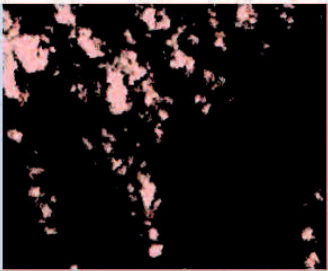
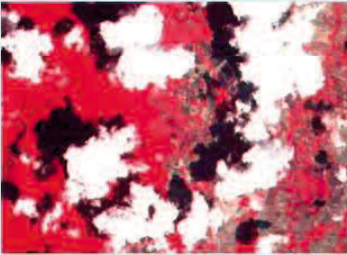

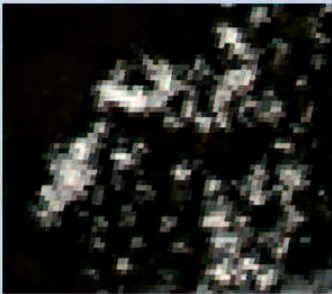
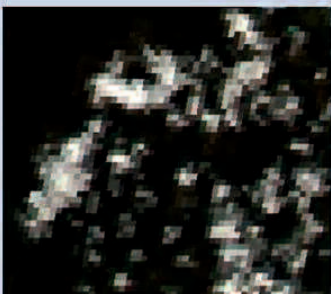
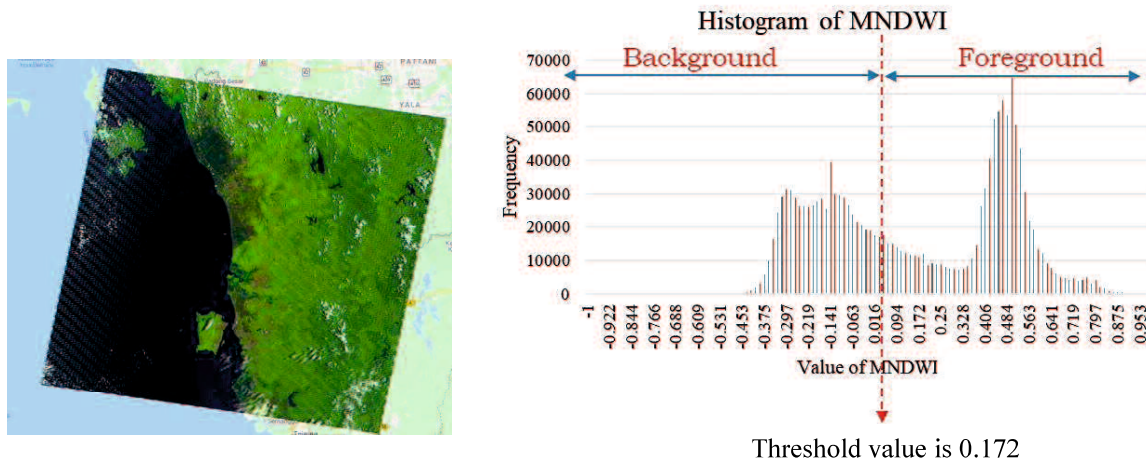
Satellite	Cloudy image	Cloud mask
Landsat 7		
Sentinel 2		
MODIS		

Figure 4.14: Cloud masking of optical images.

The equation of MNDWI, (Equation 8), was used to extract water features using the green and SWIR bands.

$$\text{MNDWI} = (\text{green} - \text{SWIR}) / (\text{green} + \text{SWIR}), \quad (8)$$

4.7.4 Automatic thresholding using OTSU Thresholding



Threshold value is 0.172

Figure 4.15: Determination of foreground and background using water references with bimodal histogram.

Threshold	T = -0.070	T = 0.172	T = -0.531	T = -0.078
Input				
Weight, Background	$W_b = 0.535$	$W_b = 0.501$	$W_b = 0.216$	$W_b = 3.682$
Mean, Background	$\mu_b = -0.336$	$\mu_b = -0.135$	$\mu_b = -0.765$	$\mu_b = -0.297$
Variance, Background	$\sigma_b^2 = 0.026$	$\sigma_b^2 = 0.019$	$\sigma_b^2 = 0.055$	$\sigma_b^2 = 0.016$
Weight, Foreground	$W_f = 0.465$	$W_f = 0.499$	$W_f = 0.784$	$W_f = 6.314$
Mean, Foreground	$\mu_f = 0.241$	$\mu_f = 0.469$	$\mu_f = -0.004$	$\mu_f = 0.163$
Variance, Foreground	$\sigma_f^2 = 0.019$	$\sigma_f^2 = 0.015$	$\sigma_f^2 = 0.059$	$\sigma_f^2 = 0.012$
Within class variance	$\sigma_w^2 = 0.023$	$\sigma_w^2 = 0.017$	$\sigma_w^2 = 0.059$	$\sigma_w^2 = 0.134$

Figure 4.16: Same calculation needs to be performed to all water indices.

Gennadii et al. (2016) proposed the unsupervised classification step based on the local adaptive threshold detection method. The segmentation of the water from the spectral index was build based on the extension of the Otsu method by a Canny edge filter where the number of input pixels only to those located near water-land edges. Using the morphological dilation, water and land pixels located near water are then computed applied to the detected edges. In the case of thin, single-pixel wide water bodies skewed distribution might be obtained. A buffer size (dilation) equal to half of the pixel is used to overcome the problem. A bimodal distribution is

expected so that a clear distinction of land and water can be acquired. Using the following parameters: $\sigma = 0.7$, the $\tau = 0.99$ for the Canny edge filter, and a structuring element with the size $15\text{ m} \times 15\text{ m}$ to dilate the edges and create a surrounding buffer region in a case of two classes in the grid tile, we were able to get an almost perfect detection of water pixels. The σ and τ parameters are used to define the standard deviation of the Gaussian smoothing kernel and the threshold used to define the sensitivity of the filter, respectively.

4.7.5 Changes of water extent in the study area from Landsat 7, Landsat 8, and MODIS data using various water indexes.

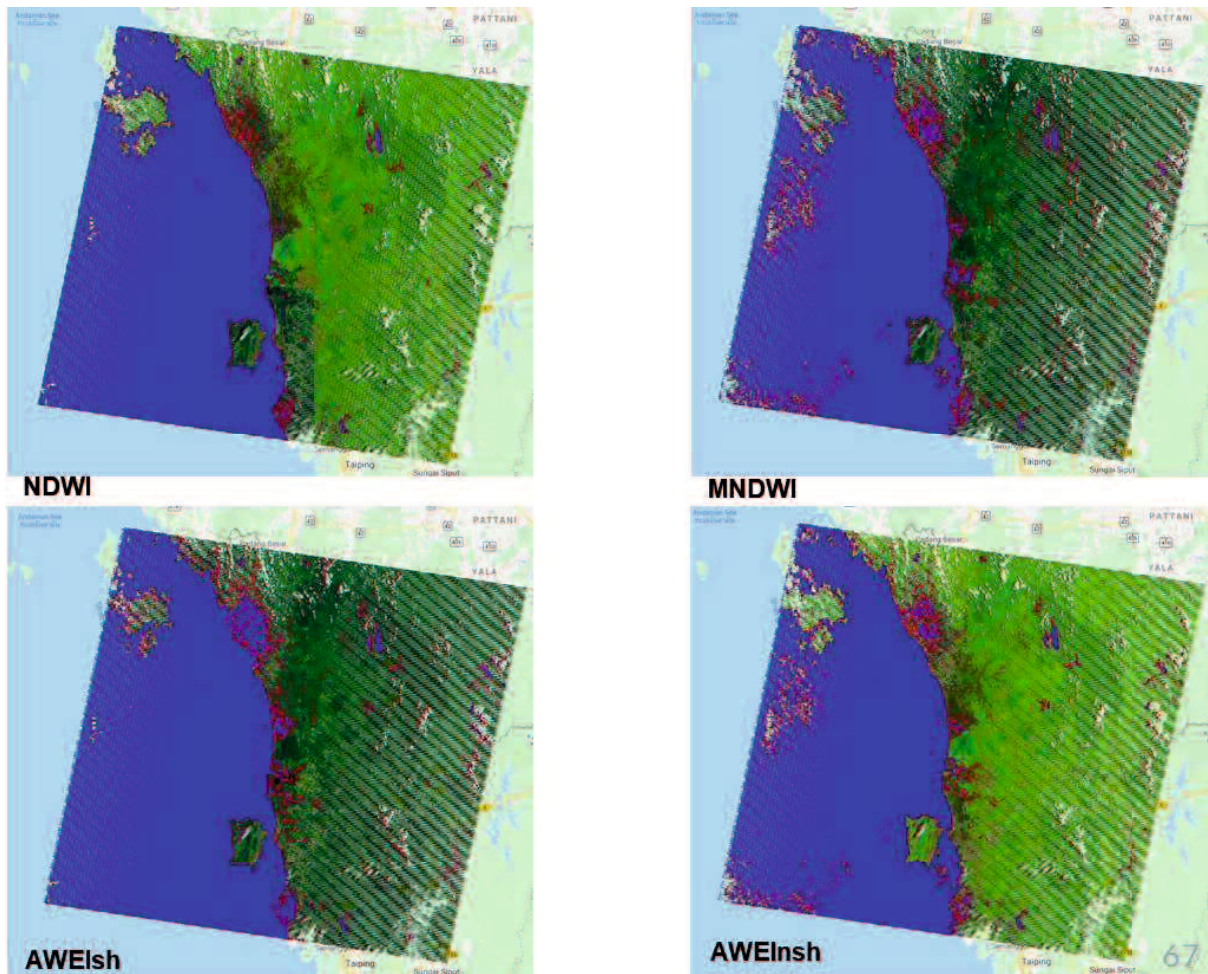
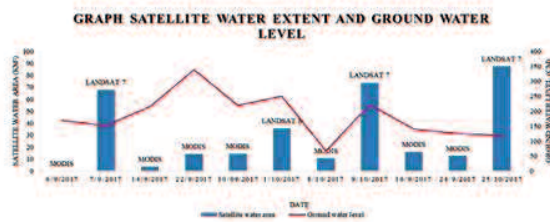
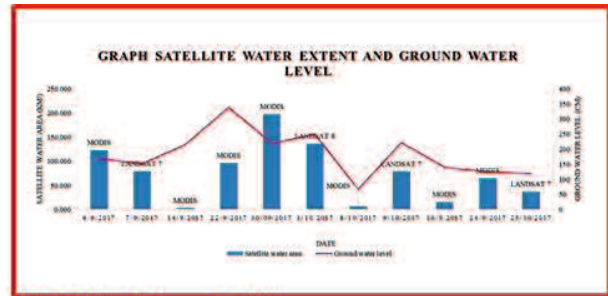


Figure 4.17: Sample of water classification using different water index using Landsat 7 (2017/09/07)

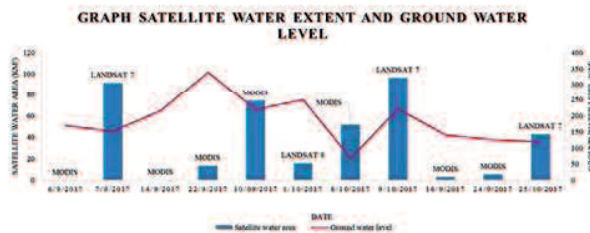
The extraction result shows that both can distinguish water area from the land area. Figure 4.17 shows the changes of classified surface water extent from Landsat 7 by the time during the flood event using the threshold of NDWI, MNDWI, AWEInsh, and AWEIsh. MNDWI.



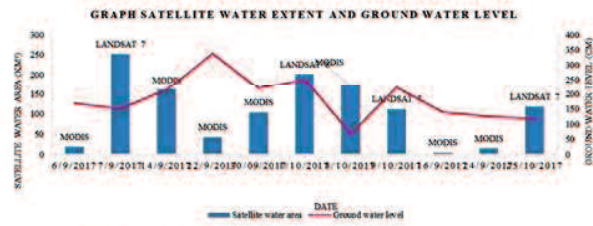
NDWI with $r = -0.099$



MNDWI with $r = 0.498$



AWEInsh with $r = -0.089$



AWEIsh with $r = -0.059$

68

Figure 4.18: Changes of water extent in the study area from Landsat 7, Landsat 8, and MODIS data using various water indexes.

Figure 4.18 shows the changes of classified surface water extent from Landsat 7, Landsat 8, and MODIS by the time during the flood event using the threshold of NDWI, MNDWI, AWEInsh, and AWEIsh. MNDWI shows the best development of flood progression. MNDWI has been selected to represent the water index for Landsat 7, Landsat 8, and MODIS. MNDWI shows more stability when integrating different optical data. MNDWI use green and SWIR able to enhanced positive value for water and negative for built-up land, soil and vegetation compared to NDWI. Eventhough AWEI developed to suppress shadow but it is not good as NDWI and MNDWI in the detection of water bodies in the non-built land.

4.8 Ground Station data

The area of study is in Malaysia's Perlis region, which was flooded in 2017. The study region extends from 6°43'19" N latitude to 100°07'59" E longitude, covering approximately 795 km². There are two distinct seasons: a rainy season with heavy rainfall and a hot season with prolonged drought are the two distinct seasons. The rainy season, which is related to the southwestern monsoon, begins in April or May and lasts until September or October. Meanwhile, the northeast monsoon influences the dry season, which lasts from December to March. The average annual rainfall ranges from 2000 to 3000 mm, and the average annual temperature is 27°C. Figure 4.19 shows an elevation map of the study area, which includes the locations of the affected villages as described in a report by the Department of Drainage and Irrigation (DID) [1].

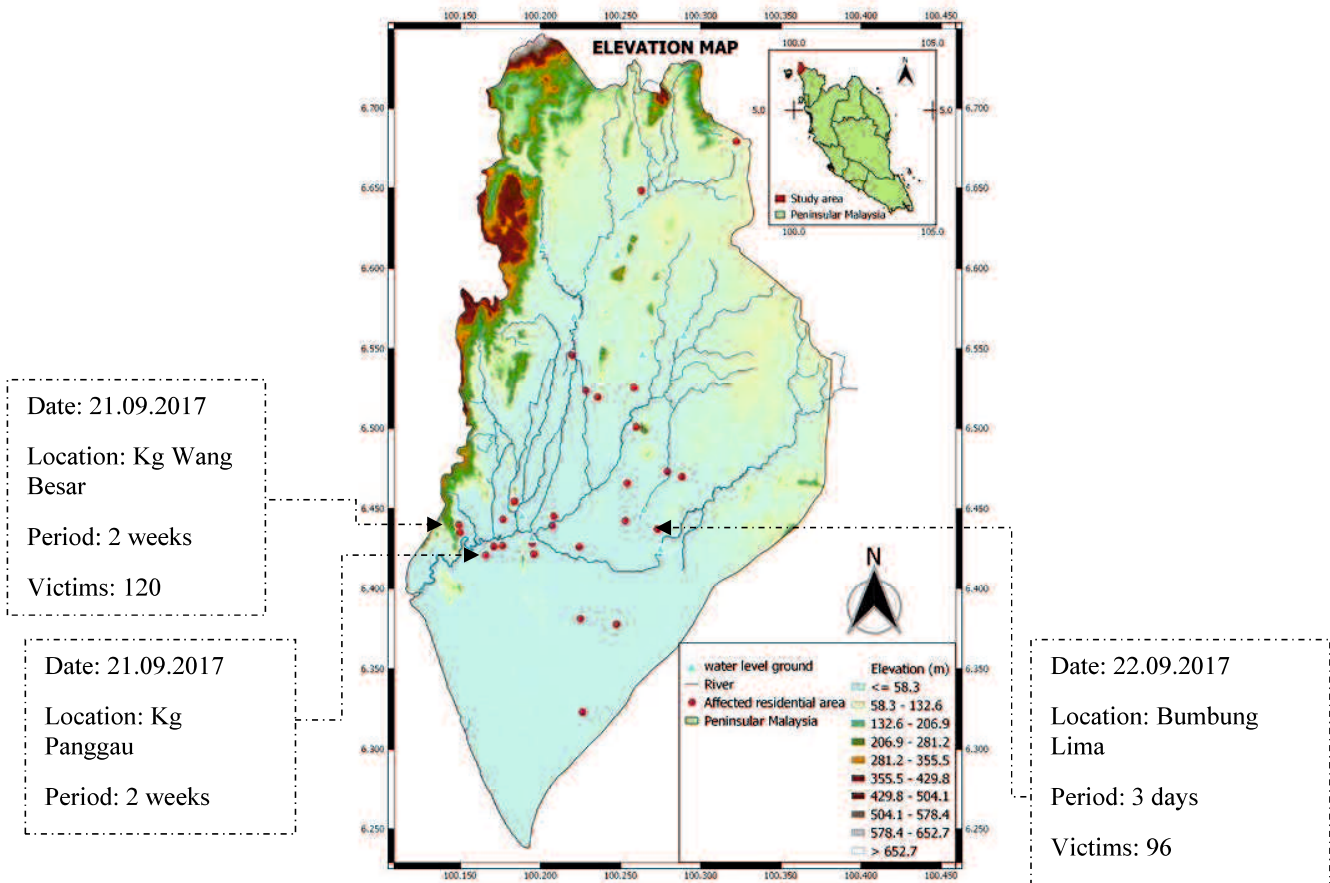


Figure 4.19: The study area's elevation map. The map displays the location and flood information for some of the damaged residential areas, which are represented by red dots. Blue triangles show water level data from the ground station.

Flooding is reported in two districts in Perlis, Kangar and Arau, by the DID in 2017 and 2018. Continuous rainfall was observed in several districts from September 15 until early October 2017. Rainfall was triggered by high atmospheric humidity around this time, as well as Typhoon Doksuri, which passed over Northern Vietnam on September 15th. In this monsoon's final phase, Malaysia was impacted by the relatively weak southern monsoon and the Northern Hemisphere summer monsoon. Wind gusts were caused by the tail effect of Typhoon Doksuri in Malaysia, particularly in northern Peninsular Malaysia [2]. This condition contributed to the significant rains that hit Perlis in September 2017, resulting in two-week-long flooding in some locations. Additionally, high tides worsened the floods.

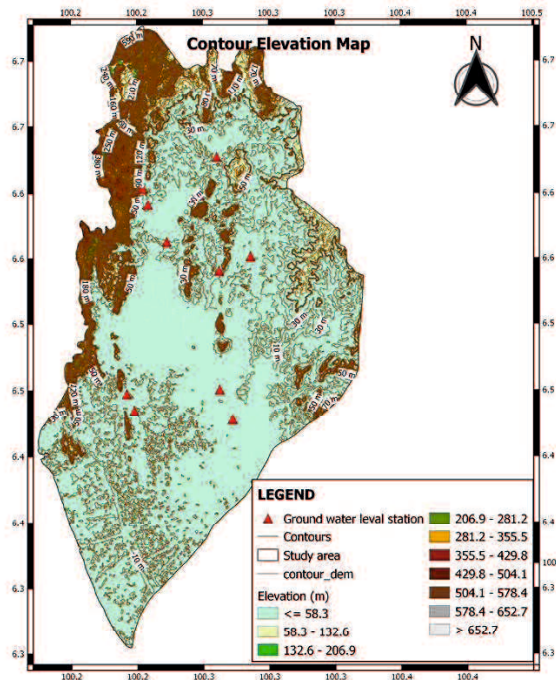


Figure 4.20: Available Ground Dataset for September 2017 flood

The river is the principal cause of flooding because it is constricted at its lower reaches. The capacity of the river at downstream area is less than $10,000\text{m}^3/\text{s}$, therefore flood that exceeds this capacity will overspill the banks and inundation flood water at land surface area and finally moving to the sea. The downstream and upstream ground station in this study was divided based on the boundary of the main river mouth (Sg.Perlis) as shown in Figure 4.21. The elevation of ground station at the downstream approximately 0 m (Figure 4.21) while, at the upstream the elevation is 20 m to 50 m (Figure 4.22).

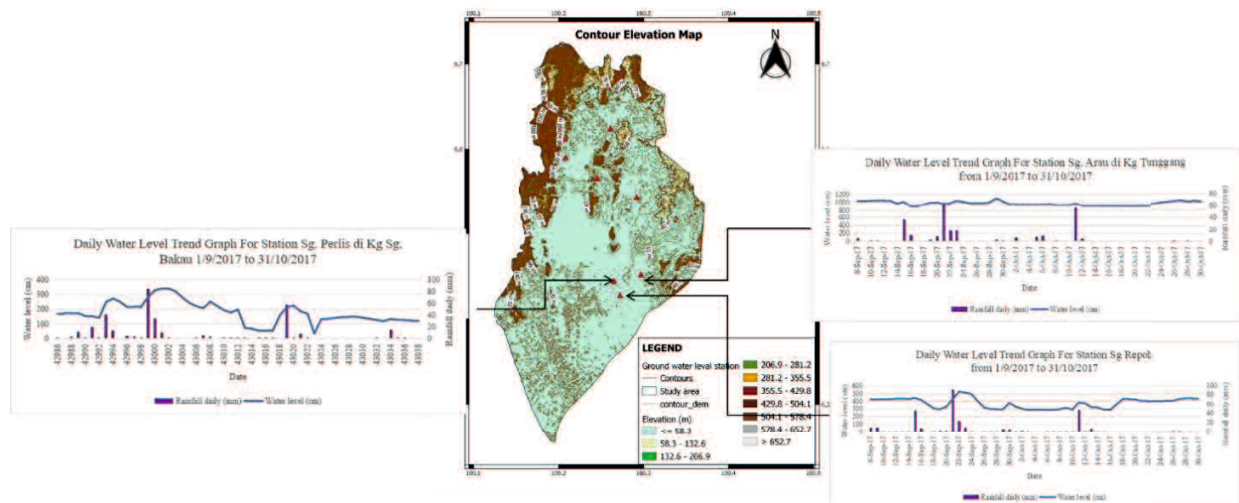


Figure 4.21: Downstream ground observation data

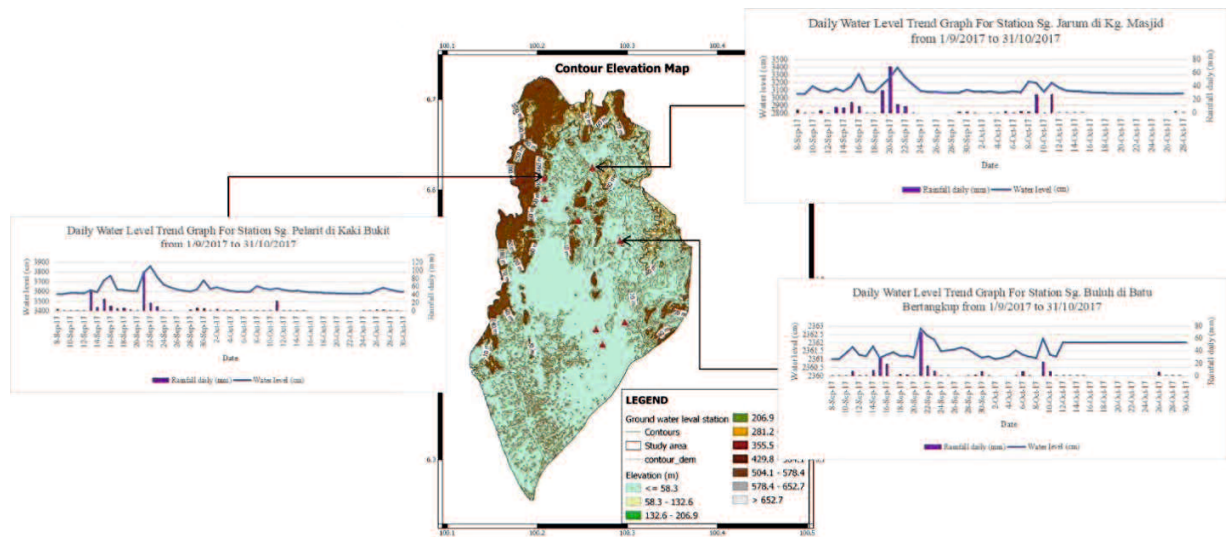


Figure 4.22: Upstream ground observation data.

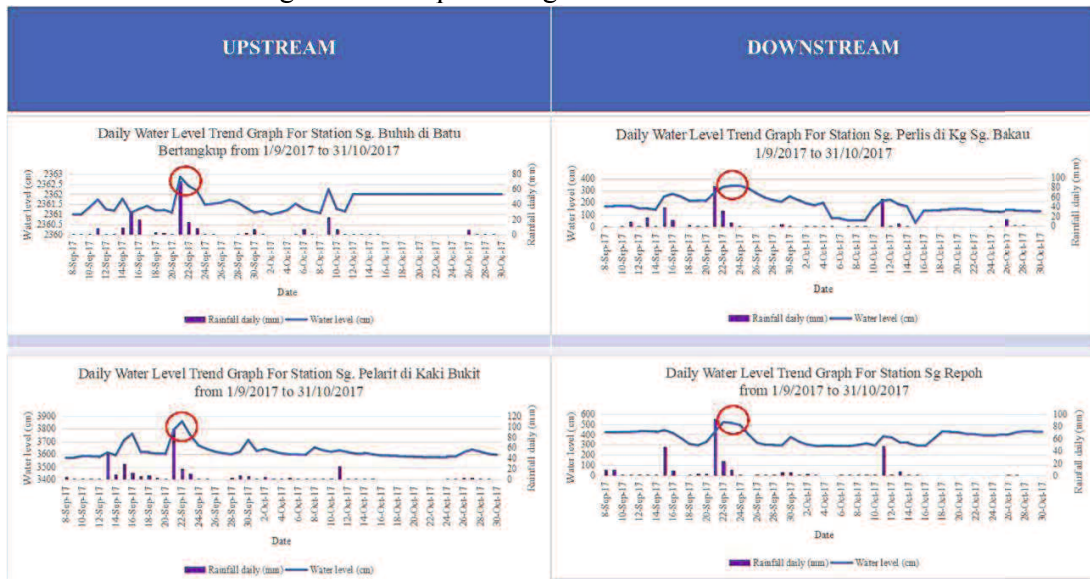


Figure 4.23: Upstream ground observation data.

Water level data from ground station were available at DID website (<https://publicinfobanjir.water.gov.my/aras-air/data-paras-air/?lang=en>) and opened for public access. The only provided information is 15 minutes resolution water level, monthly and daily rainfall data. Based on the elevation and peak of maximum water level, the study area was a pretty flat land and the maximum water of downstream and upstream has no interval difference (Figure 4.23), this study will only focus on extract surface water and inclusion of depth will be our future study.

4.9 Multitemporal water spatial classification using Sentinel-1

Sentinel-1 images of the Arau and Kangar areas on 24 September 2017 were used as flood stage images. The methods described in Section 4.3 were used to extract the surface water area, and an automated thresholding method was developed to differentiate water and non-water areas. Figure 4.24 shows an example of the water area extracted from the Sentinel-1 images.

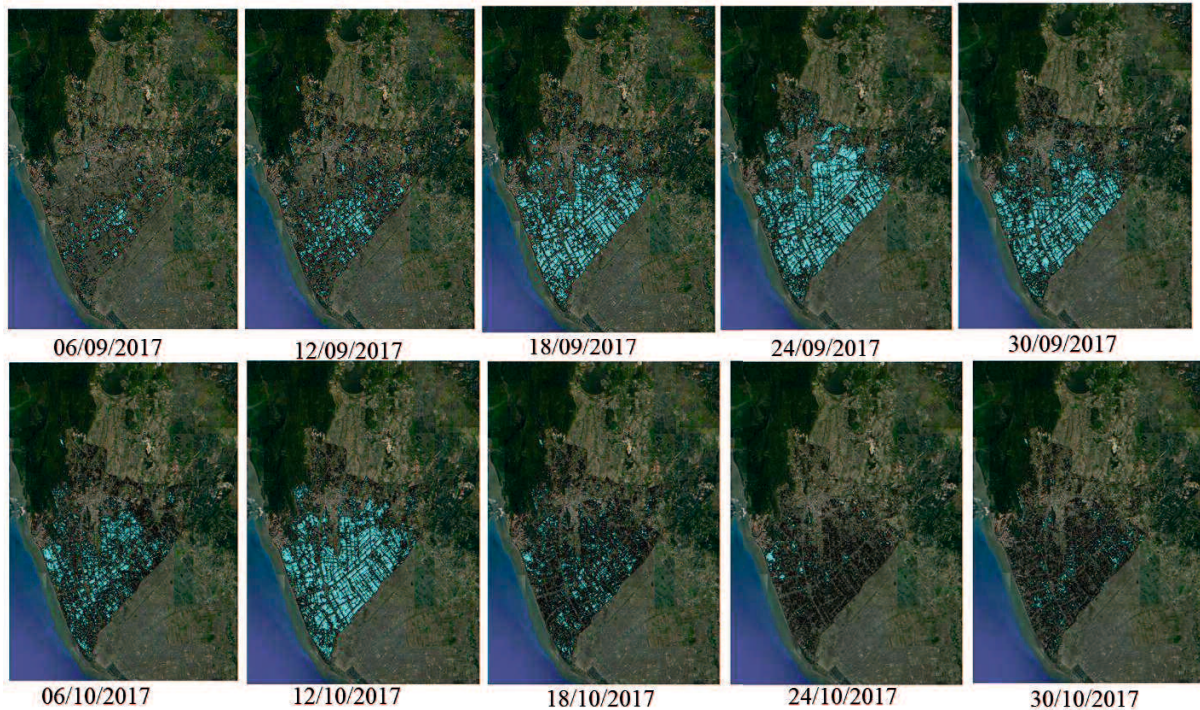


Figure 4.24: Sample of water classification using different water index using Sentinel 1 from 06/09/2017 to 30/10/2017

MultiSAR used in this study are images from ALOS-2 and Sentinel-1 satellites. The ALOS-2 image use in this study is a high sensitive 6 m resolution image with 14 days revisit time. We used both HH and HV of the 28 September 2017 image and it is in the ascending mode. Sentinel-1 is a C-band SAR data with a revisit time of 6 days and high spatial resolution with a resolution of 10 m. We use VV and VH polarisation from each image and both ascending descending images are included to achieve high temporal coverage.

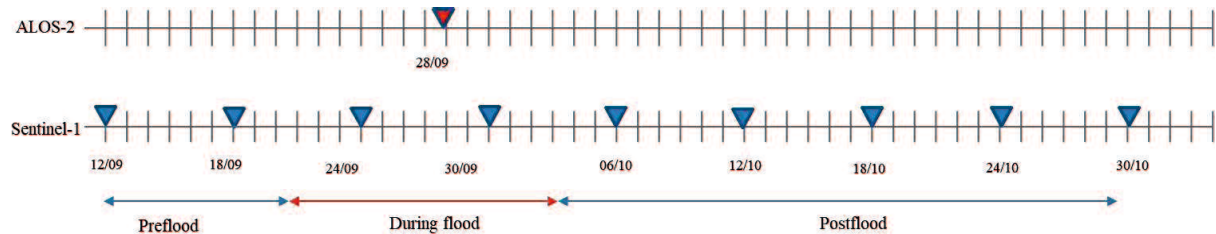


Figure 4.25: Temporal coverage of ALOS-2 and Sentinel-1

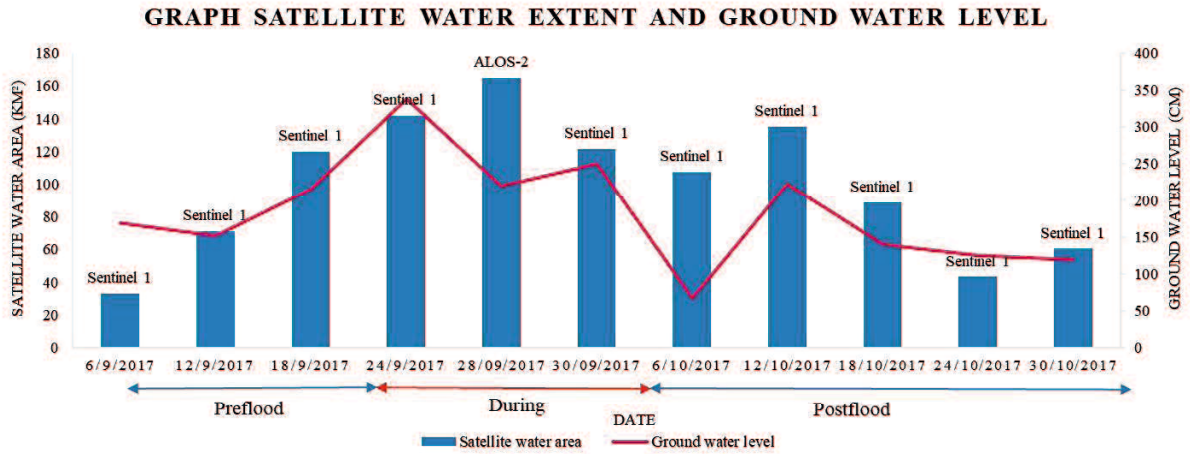


Figure 4.26: Changes of water extent in the study area from Sentinel-1 and ALOS-2 data.

Figure 4.26 shows that satellite water extent from SAR nearly in agreement with ground water level. The graph shows the relationship of satellite water extent with ground water level with correlation 0.609. The integration increase the frequency of observation.

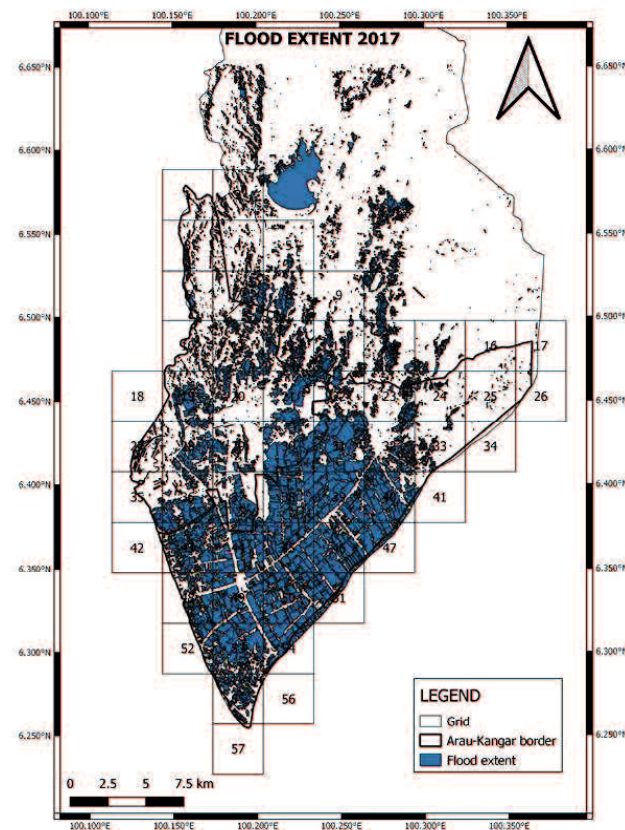


Figure 4.27: Extent of the September 2017 flood extracted from the Sentinel-1 satellite images.

The result shows that the flood extent information is consistent with the classification results of flood susceptibility in previous studies, in which major floods were found to be located

in the south of Perlis and paddy field areas [38,55]. As shown in Figure 4.27, the extracted flood extent (blue color) covered 41% of the total combined area (150.95 km²) of the Kangar and Arau districts. The high flood susceptibility in the south of Perlis is due to the presence of alluvial deposits, paddy fields, low-humidity clay soil, and flat terrain. The flood event in 2017, due to Typhoon Doksuri, severely damaged rice crop. Most paddy field areas were submerged in water for over three days, preventing seedlings from growing. Based on the result, this study shows agreement with the other previous study which indicated that flood mostly occur due to its terrain structure where low-lying areas like paddy fields located [56,57]. In general, it is challenging to perform flood classification using radar sensors when the land pixels have similar characteristics to water pixels. The results of this study show that the cause of overestimation of flooded areas is shadow due to terrain characteristics or the incidence angle of satellite images; the backscatter value of the shadow which is similar to that of water causes the misclassification of shadow as floodwater [58].

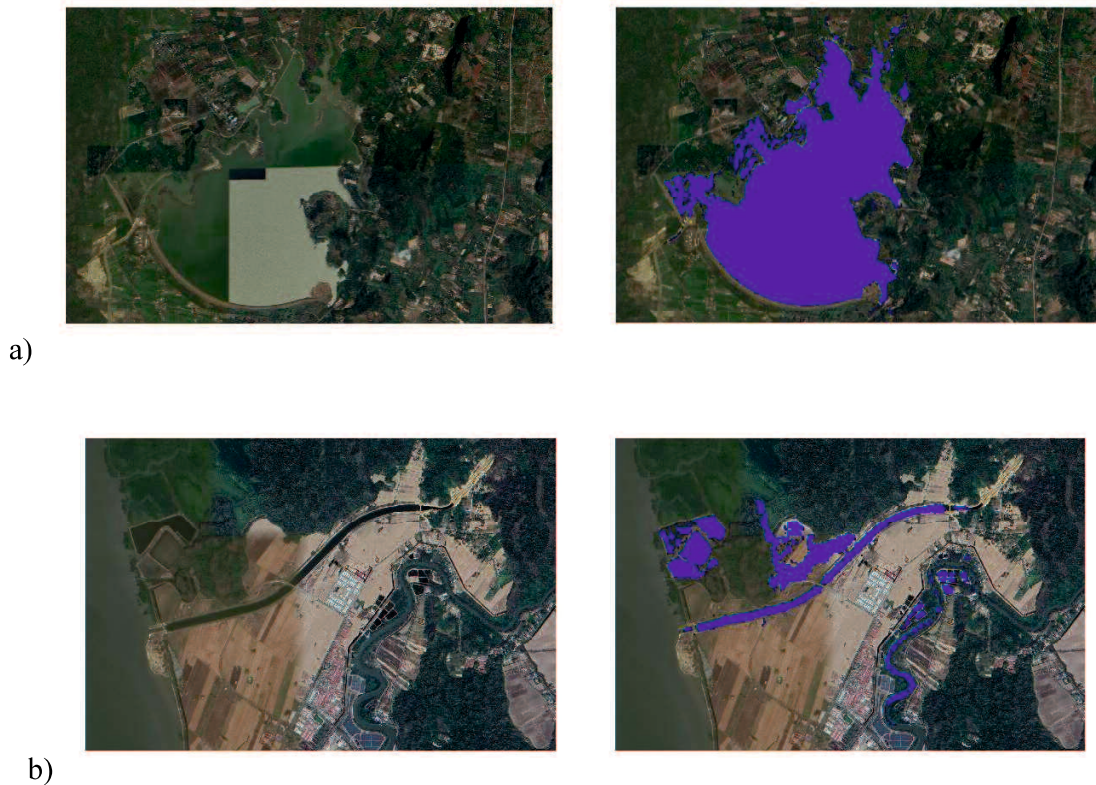


Figure 4.28: Examples of water area detection in a) lake and b) river (bottom). The water area is shown in blue and the background image shown is the Google Earth image.

4.10 Integration of MultiSAR and Multispectral

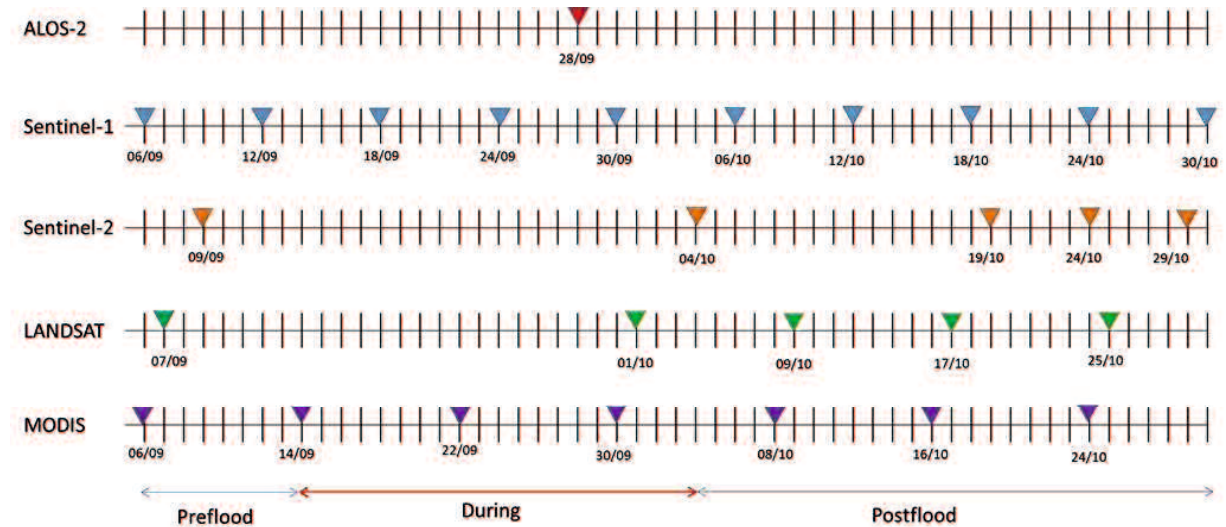


Figure 4.29: Temporal coverage of the satellite images used in this study.

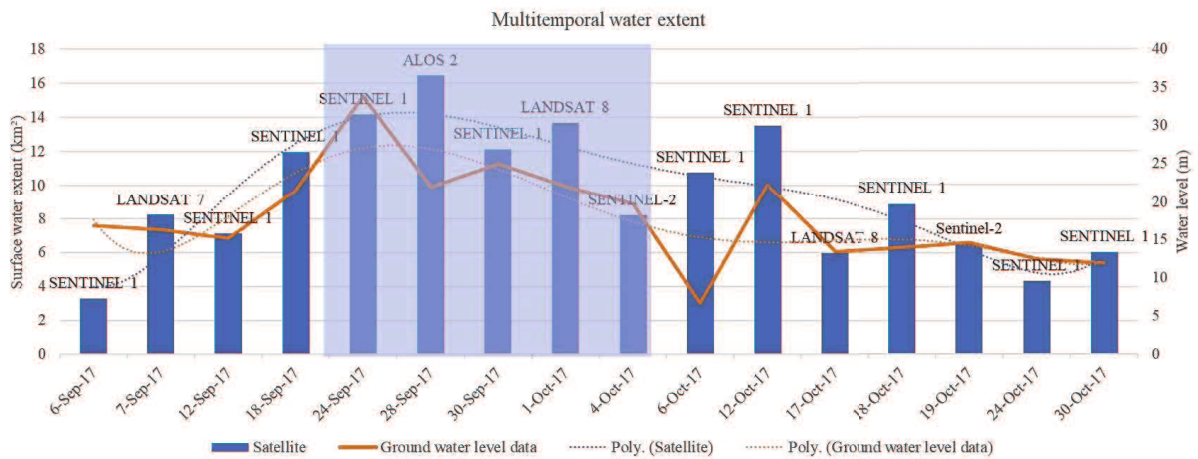


Figure 4.30: Changes of surface water extent in the study area from Sentinel-1, ALOS-2, Sentinel-2, Landsat 7, Landsat 8, and MODIS data between 06/09/2017 to 30/10/2017.

Figure 4.29 shows temporal coverage of the satellite images used in this study and Figure 4.30 shows the integrated surface water extent from various multi-satellite. The dataset still has a fluctuation of classified surface water however, it shows a distinct preflood, during, and post flood of surface water area. The lowest recorded surface water was suited at the beginning before the flood start and the highest peak recorded during flood occurrence and the total area gradually decrease after the flood ended. However, due to varying spatial resolution, MODIS overestimate the water extent compare to Sentinel-1 during flood on 30th September where flood area by Sentinel-1 around 60 km but MODIS estimate water area around 200 km. MODIS also highly affected with cloud and most of the pixels were removed and caused no water area detected on 14th September and 8th October. This bias and variance can be overcome by applying statistical procedure like polynomial regression to achieve low bias and low variance dataset. The model

can be very useful in quickly understanding the changes in water bodies by the time during flood events from multi-satellite.

4.11 Discussion of chapter IV

HH+HV grayscale image have a higher frequency of bimodal distribution with range 73.53 % to 85.53% compare to HH backscattering with 55.88 % to 83.82 % frequency bimodal. Thus, HH+HV chosen for automatic thresholding using OTSU. This technique provides a satisfactory water extraction result.

Next this chapter has discussed the technique for automatic extraction of the time series of the flooded area, calculating the surface water extent and integrate SAR and optical images. Based on the analysis, MNDWI shows the best and the most stable water index for the automatic thresholding of a multispectral image. After the optical and multiSAR integration, although there is still a fluctuation of classified surface water in the dataset, it does reveal a distinct pre-flood, during, and post-flood surface water area. The lowest recorded surface water level occurred before the flood began, and the highest peak occurred during the flood, with the total area progressively decreasing after the flood stopped. However, MODIS overestimates the water extent compared to Sentinel-1 during the flood on September 30th, where Sentinel-1 estimated the flood area to be around 60 km whereas MODIS estimated the water area to be over 200 km due to differing spatial resolution. MODIS was also heavily impacted by cloud, with the majority of pixels being deleted, resulting in no water area being recorded on September 14th and 15th. On the 14th and 8th of October, MODIS was also heavily impacted by cloud, with the majority of pixels being deleted and no water area being recorded. To achieve low bias and low variance datasets, statistical procedures such as polynomial regression can be used to overcome bias and variance. The model can help quickly comprehend how water bodies change over time during flood events using multi-satellite data. We'd like to execute a fusion on an overlapping satellite image, which includes MODIS and Sentinel-1 images, in the next chapter. On the 14th and 8th of October, MODIS was also heavily impacted by cloud, with the majority of pixels being deleted and no water area being recorded. To achieve low bias and low variance datasets, statistical procedures such as polynomial regression can be used to overcome bias and variance. The model can help quickly comprehend how water bodies change over time during flood events using multi-satellite data. Compare to multispectral, multiSAR show higher correlation with ground observation. This is due to the optical images affected by cloud occlusion, causing loss of data during the cloud masking process.

To summarise, the result also shows that the extracted surface water from multispectral images using the Otsu threshold from a stable water index can reduce the inconsistency with a set of multiSAR images when integrated. The agreement water extent from satellite with ground data slightly increase when number of observation increase. We'd like to execute a fusion on an overlapping satellite image, which includes MODIS and Sentinel-1 images, in the next chapter.

References

- [1] Department of Irrigation and Drainage (Jabatan Pengairandan Saliran), Malaysia, 2018. Laporan Banjir Tahunan 2017/2018. Available online: <https://www.water.gov.my/> (accessed on 02nd February 2019).
- [2] Nakmuenwai, P.; Yamazaki, F.; Liu, W. Automated Extraction of Inundated Areas from Multi-Temporal Dual Polarization RADARSAT-2 Images of the 2011 Central Thailand Flood. *Remote Sens.* 2017, 9, pp 78.
- [3] Donchyts, G.; Schellekens, J.; Winsemius, H.; Eisemann, E.; Van de Giesen, N. A 30 m Resolution Surface Water Mask Including Estimation of Positional and Thematic Differences Using Landsat 8, SRTM and OpenStreetMap: A Case Study in the Murray-Darling Basin, Australia. *Remote Sens.* 2016, 8, 386.
- [4] Department of Irrigation and Drainage (Jabatan Pengairandan Saliran), Malaysia, 2018. Laporan Banjir Tahunan 2017/2018. Available online: <https://www.water.gov.my/> (accessed on 02nd February 2019).
- [5] Committee, W M O Typhoon, Integrated Workshop, and Trcg Forum. “MEMBER REPORT ESCAP / WMO Typhoon Committee 8 Th Integrated Workshop / 2 Nd TRCG Forum (Philippines),” no. December (2013).
- [6] Husniyah. B. M. and M. Nagai. “Classification of Surface Water Area from Multispectral and Multisar Data Using Automatic and Robust System,” The 41st Asian Conference on Remote Sensing (ACRS 2020) November 9-11, 2020. Deqing, China, 2020.
- [7] Dano, Umar Lawal, Abdul Lateef Balogun, Abdul Nasir Matori, Khmaruzzaman Wan Yusouf, Ismaila Rimi Abubakar, Mohamed Ahmed Said Mohamed, Yusuf Adedoyin Aina, and Biswajeet Pradhan. “Flood Susceptibility Mapping Using GIS-Based Analytic Network Process: A Case Study of Perlis, Malaysia.” *Water (Switzerland)* 11, no. 3 (2019). <https://doi.org/10.3390/w11030615>.
- [8] Dutsenwai HS, Ahmad BB, Mijinyawa A, and Tanko AI (2016). Fusion of SAR images for flood extent mapping in northern peninsula Malaysia. *International Journal of Advanced and Applied Sciences*, 3(12): 37-48.
- [9] Alexandridis, Th., K. Perakis, and N. Silleos. “Flood Monitoring Using ERS-1 SAR Imagery and Digital Elevation Model.” *IFAC Proceedings Volumes* 31, no. 12 (1998): 67–73. [https://doi.org/10.1016/s1474-6670\(17\)36043-3](https://doi.org/10.1016/s1474-6670(17)36043-3).
- [10] Sherpa, Sonam Futi, Manoochehr Shirzaei, Chandrakanta Ojha, Susanna Werth, and Renaud Hostache. “Probabilistic Mapping of August 2018 Flood of Kerala, India, Using Space-Borne Synthetic Aperture Radar.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13, no. August 2018 (2020): 896–913. <https://doi.org/10.1109/JSTARS.2020.2970337>.

Chapter V

Improved consistency integration of SAR and optical water extent data

5.1 Introduction

Despite various notable efforts, to this end, existing satellite technology or any single data product could not solve the current challenge from the decision-making standpoint. Hence, to fulfill decision-makers needs, there is a demand to develop a process for improving the end product for an effective disaster response rather than developing a new system to improve flood operation. This study focuses on the standardization of multi-source remote sensing data by using the same processing platform. In this context, to obtain a single image that retains essential features of original images, the simple and robust fusion of images with the same observation period has been proposed in this study.

The main objectives of this study are i) to classify the water extent based on the average Otsu threshold values calculated for individual grids with a bimodal histogram, ii) then, the temporal changes of the flood extent are assessed by analyzing the grids that indicate flood-affected areas, iii) the analysis proceeds to applying 2D wavelet transform to fuse two images with the same observation period. Finally, the of temporal distribution of water classification was compared with the ground water level data. That the proposed model can be used to estimate flood duration as well as to estimate the flood-related losses, especially in ungauged or data-poor regions.

5.2 Satellite data

In this study, 22 SAR and optical images collected over the study region between 4 September 2017 and 30 October 2017 were acquired (Table 5.1). Four ALOS-2 SAR images in the HH and HV polarization modes were acquired from the Japan Aerospace Exploration Agency (JAXA) (<https://auig2.jaxa.jp/ips/home>). The acquired ALOS-2 images were in the L-band with a wavelength of 22.9 cm and spatial resolutions of 6.25 m and 25 m. The incidence angle of the ALOS-2 images varied between 8° and 70°. A total of 10 Sentinel-1 SAR images with VV and VH polarization with ascending and descending modes were used (<https://scihub.copernicus.eu/dhus/#/home>); these were Level-1 processed ground range detected (GRD) products collected in interferometric wide swath (IW) mode. The Sentinel-1 images were in the C-band with incidence angles between 29° and 46° and wavelength 5.6 cm.

The cloud-covered images less than 20% [1], were obtained from Landsat 7, Landsat 8, daily MODIS Terra MOD09GA, and MODIS Aqua MYD09GA. The Landsat 7 Enhanced Thematic Mapper (ETM+) and the Landsat 8 Operational Land Imager (OLI) have approximately 710 km sun-synchronous circular 98.2° inclined orbit with a revisit time of 16 days, but these overpass each other by eight days [1]. MOD09GA and MYD09GA version 6 provide bands 1–7 in a daily gridded L2G product in a sinusoidal projection, including reflectance values with a resolution of 500 m and observation and geolocation statistics with a resolution of 1 km.

Table 5.1: Details of the (SAR) and optical satellite images used in this study.

Satellite	Sensor	Acquisition date	Resolution (m)	Pass (Polarization)	Source	Projection	Central Wavelength (nm)
ALOS-2	PALSAR-2	4/9/2017	25	Descending (HH & HV)	JAXA	UTM	L band
ALOS-2	PALSAR-2	9/9/2017	25	Descending (HH & HV)	JAXA	UTM	L band
ALOS-2	PALSAR-2	28/9/2017	6.25	Ascending (HH & HV)	JAXA	UTM	L band
ALOS-2	PALSAR-2	21/10/2017	25	Descending (HH & HV)	JAXA	UTM	L band
Sentinel-1	C-SAR	6/9/2017	10	Descending (VV & VH)	ESA	Ellipsoid	L band
Sentinel-1	C-SAR	12/9/2017	10	Ascending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	18/9/2017	10	Descending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	24/9/2017	10	Ascending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	30/9/2017	10	Descending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	6/10/2017	10	Ascending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	12/10/2017	10	Descending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	18/10/2017	10	Ascending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	24/10/2017	10	Descending (VV & VH)	ESA	Ellipsoid	C band
Sentinel-1	C-SAR	30/10/2017	10	Ascending (VV & VH)	ESA	Ellipsoid	C band
Landsat 7	ETM+	7/9/2017	30		USGS	UTM	Green: 560 SWIR1: 1648
Landsat 8	OLI	1/10/2017	30		USGS	UTM	Green: 562 SWIR1: 1610
Landsat 8	OLI	17/10/2017	30		USGS	UTM	Green: 562

							SWIR1: 1610
MODIS Terra	Terra	7/9/2017	500	USGS	Sinusoidal	Green: 555	SWIR1: 1640
MODIS Aqua	Aqua	12/9/2017	500	USGS	Sinusoidal	Green: 555	SWIR1: 1640
MODIS Aqua	Aqua	30/9/2017	500	USGS	Sinusoidal	Green: 555	SWIR1: 1640
MODIS Terra	Terra	18/10/2017	500	USGS	Sinusoidal	Green: 555	SWIR1: 1640
MODIS Aqua	Aqua	19/10/2017	500	USGS	Sinusoidal	Green: 555	SWIR1: 1640

5.3 Research flowchart

The analysis involved the assessment of the temporal changes of floodwater area in the affected grid. The following four main steps were taken to achieve this objective: (1) pre-processing the images, (2) resampling and dividing the study area into an aligned grid, (3) identifying flood water extent time changes and (4) fusion by a grid.

Refer to Figure 5.1, the image with the same observation date was used for the image fusion process. There are two kinds of fusion based on the availability of data: optical with another optical image (example, Landsat 7 and MODIS Terra on 7 September 2017) and optical image with SAR image (example, Sentinel 1 and MODIS Terra on 12 September 2017). The overall fusion processing goes through image registration and pre-processing. MNDWI was calculated for optical images, however, the backscatter of SAR is not directly comparable with the MNDWI. Hence, the MNDWI information was rescaled by inverting the minimum and maximum index values so that the pixel close to the value of -1 represents water while close to 1 was classified as non-water [2]. Then 2D wavelet transform was applied to the rescaled MNDWI optical image and SAR. This study used wavelet transform to fuse two images using Python 3.6. Finally, Otsu thresholding was applied on the fused images.

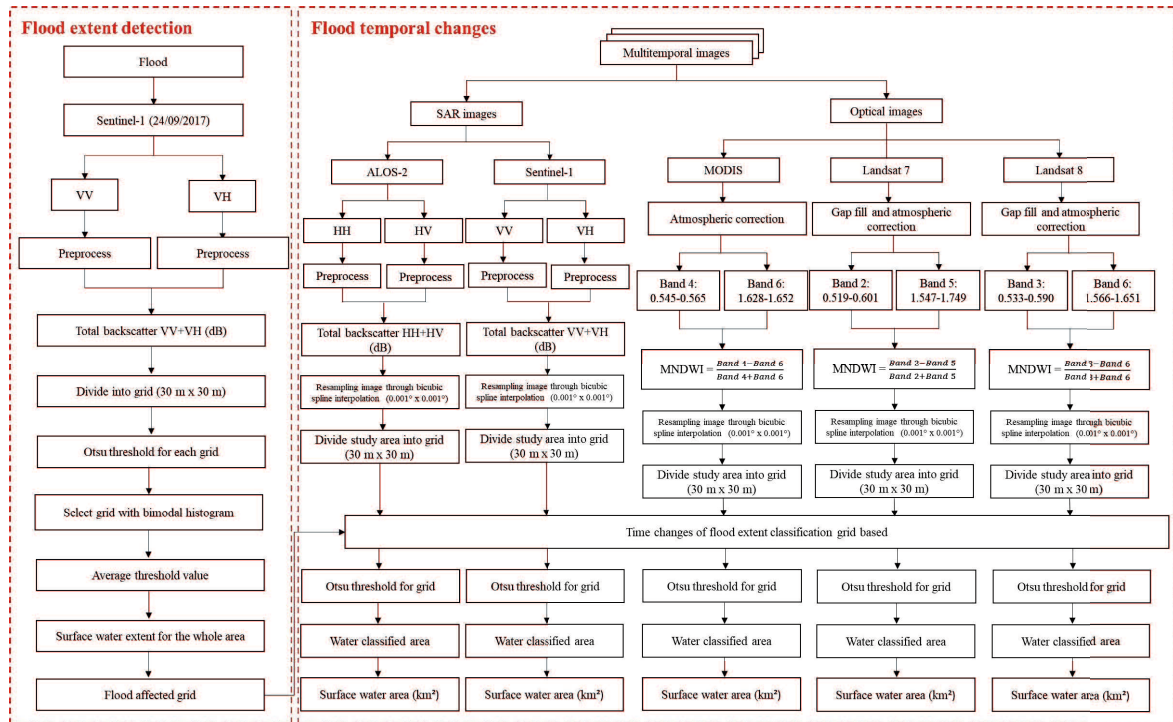


Figure 5.1: Flowchart of the automatic flood mapping process.

5.4 Grid resolution and image resampling

This study uses multi-satellite images with different orbital platforms. These satellites have different properties such as daily revisit time but lower spatial resolutions (500 m), such as MODIS Terra and Aqua. Other satellites with a better spatial resolution have a lower temporal resolution, such as Landsat 7, Landsat 8, Sentinel-1 and ALOS-2. Hence, image alignment and resampling is an essential step for this multitemporal and multi-satellite study. In this study, the input remote sensing images were resampled using bicubic spline interpolation with a pixel size of 0.001° . The size of the pixel was decided considering the spread of the ground data and convenience for the user since high spatial resolution might cause inconvenience during access [4]. A vector grid (Figure 5.2), was built as a reference for image projection and alignment. The vector grid was reprojected as a Universal Transverse Mercator (UTM) zone 47N coordinate system. The proposed method required overlapping images, and to do this the pixel of each image is aligned with the corresponding pixel of the referred vector grid. Then, each resampled and reprojected image was cropped to a 30×30 pixel ($w \times h$) following the reference vector grid for further processing.

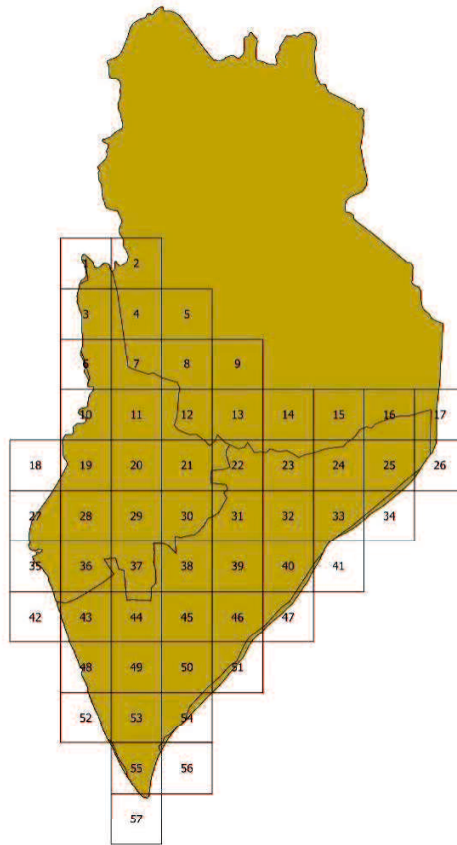


Figure 5.2: A reference vector grid of 30 x 30-pixel (w x h) size.

5.5 2D wavelet transform for image fusion

The fusion process begins by decomposing the image volume to frequency bands. Wavelet transform decomposes images into four quadrants of high- and low-frequency bands. The input images were transformed into four wavelet coefficients called HH (containing high-frequency diagonal information), LH (containing high-frequency horizontal information), HL (containing high-frequency vertical information), and LL (containing low-frequency approximation information). Then, the frequency bands of each image are analyzed by the set fusion rules to determine which one can be combined and which one has to be removed from the final volume of coefficients. To produce the fused and combined coefficient map of the input images, “the average” was used, and then an inverse wavelet transform was implemented to reconstruct the fused image [5]. Figure 5.3 shows the schematic diagram of 2D wavelet transform fusion process. The 2D wavelet transform was performed using PyWavelets which is wavelet transform software for Python.

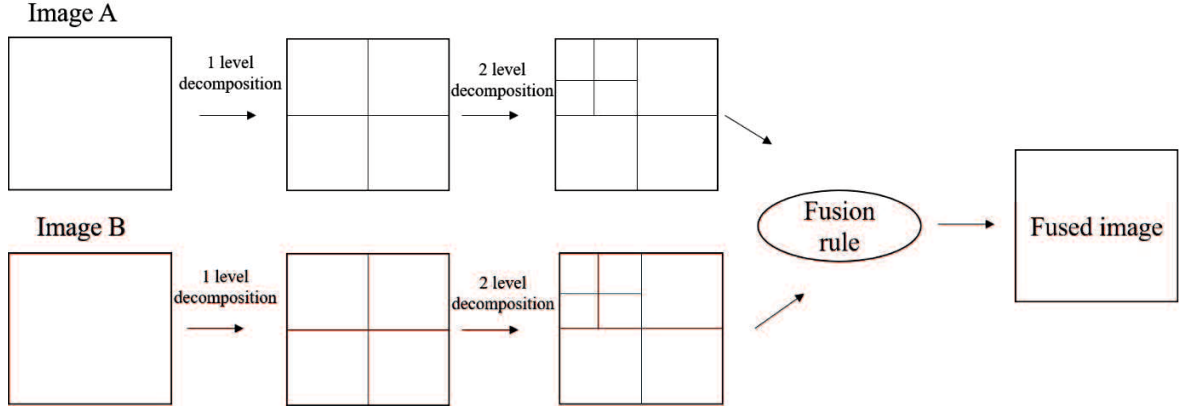


Figure 5.3: 2D wavelet transform fusion process [5].

5.6 Performance evaluation for data-scarce areas

To assess the accuracy of the surface water area derived from the multi-satellite time series, a correlation coefficient (Equation 1) was used to measure the correlation between the satellite-derived water area and the ground-measured water area.

$$\text{Correlation coefficient} = \frac{\sum_{i=1}^N (P_{S_i} - \bar{P}_S)(P_{O_i} - \bar{P}_O)}{\sqrt{\sum_{i=1}^N (P_{S_i} - \bar{P}_S)^2} \sqrt{\sum_{i=1}^N (P_{O_i} - \bar{P}_O)^2}} \quad (1)$$

Where, N is the number of samples, P_{S_i} is the estimated surface water extent value, \bar{P}_{S_i} is the mean of estimated surface water extent, P_{O_i} is the observed water level value, and \bar{P}_{O_i} is the mean of observed water level value. Additionally, the root means square error (RMSE) was used to measure the magnitude of the average error in the extracted surface water area. The RMSE, (Equation 5), is important for detecting undesirable large errors [6].

$$\text{Root Mean Square Error (RMSE)} = \left[\frac{\sum_{i=1}^N (P_{S_i} - P_{O_i})^2}{N} \right]^{1/2} \quad (2)$$

Finally, the mean bias error (MBE) (Equation 6), was used to measure the mean absolute value of the error. If the predicted value is larger than the observed value, the mean bias error will be positive, and vice versa.

$$\text{Mean Bias Error} = \frac{\sum_{i=1}^N (P_{S_i} - P_{O_i})}{N} \quad (3)$$

5.7 Grid-based image fusion

Subsequently, the method explained in Section 5.3 was applied to the MNDWI images since it is less complex than applying the fusion method to the green and SWIR bands, which are used to calculate the MNDWI. Figure 5.4 shows the extracted water area for grid 21 (Figure 5.2) based on two optical images acquired on 7 September 2017, one from Landsat 7 and one from MODIS Terra.

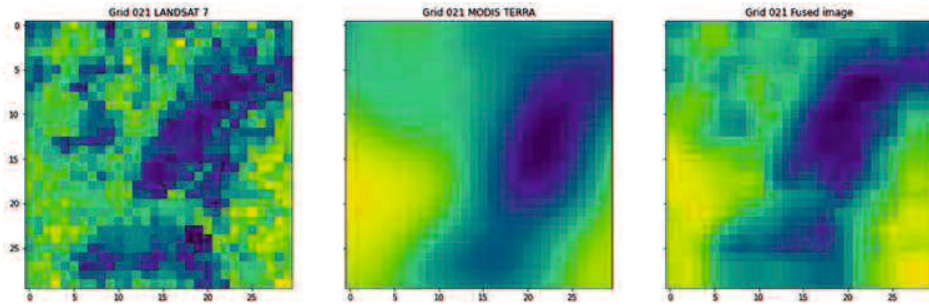


Figure 5.4: Landsat 7 optical image (left) and MODIS Terra optical image (center), both acquired on 7 September 2017, and a fused image created using these two images (right). The water area is shown in blue.

Figure 5.5 shows the scatterplots between the MNDWI of Landsat 7 optical images and that of MODIS optical images, the MNDWI of Landsat 7 optical images and that of fused images, and the MNDWI of MODIS optical images and that of fused images. The correlation coefficients were 0.31, 0.56, and 0.66, respectively. The amount of spectral content is preserved.

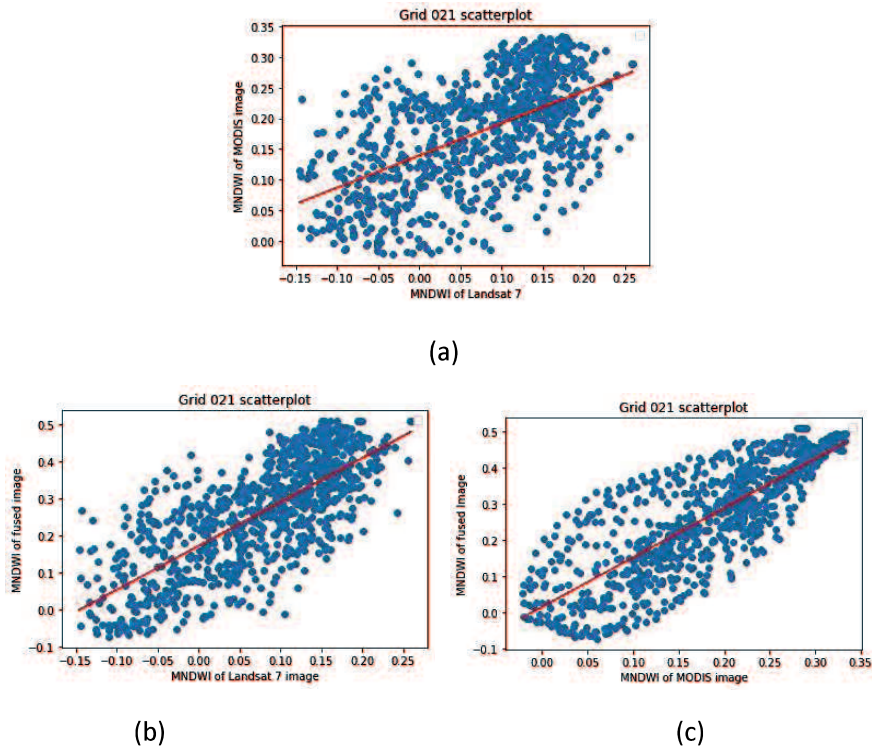


Figure 5.5: Scatterplots between (a) MNDWI of Landsat 7 optical images and that of MODIS optical image, (b) MNDWI of Landsat 7 optical image and that of fused image, and (c) MNDWI of MODIS optical image and that of the fused image.

Figure 5.6 shows the results of the fusion of a MODIS Terra optical image and a Sentinel-1 SAR image acquired on 12 September 2017 for grid 21. The results show that the fused image can preserve the information of the multispectral image while maintaining the structural and geometrical information of the SAR image. Although the Sentinel-1 image has

been filtered to remove noise, a few pixels of noise still appear, and the boundaries of objects are more apparent in the fused image than in the Sentinel-1 or MODIS Terra images.

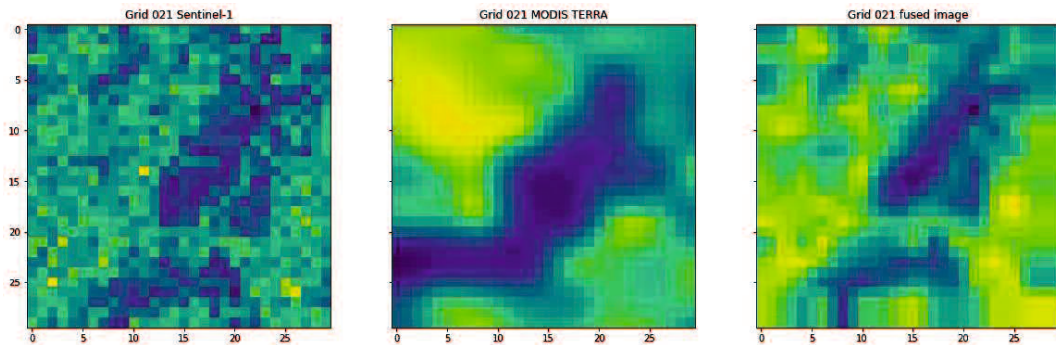


Figure 5.6: Sentinel-1 (left) and MODIS Terra (center) images, both acquired on 12 September 2017, and a fused image created using these two images (right). The water area is shown in blue.

Figure 5.7 shows the correlation between the backscatter of Sentinel-1 and that of MODIS optical images, the backscatter of Sentinel-1 and that of fused images, and the MNDWI of MODIS optical images and that of fused images. The correlation coefficients were 0.06, 0.33, and 0.18, respectively. Compared with the fused images created using two optical images, the fused images created using optical and SAR images have a lower correlation coefficient owing to the differences in data types, data structures, spatial resolution, and geometric characteristics.

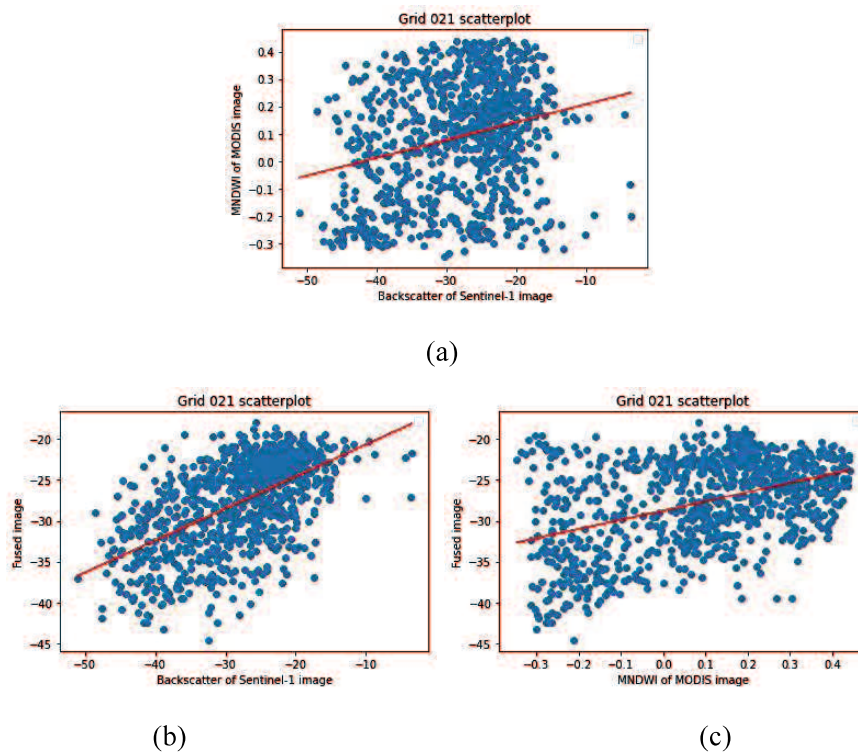
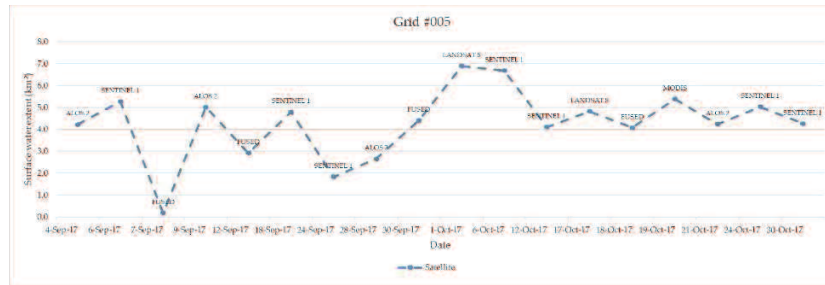
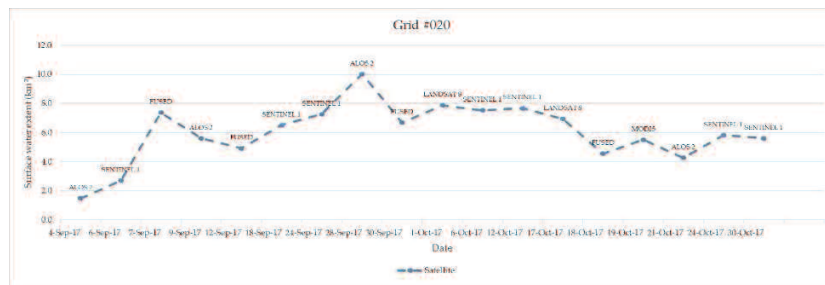


Figure 5.7: Scatterplots between (a) MNDWI of Sentinel-1 backscatter images and that of MODIS optical images, (b) MNDWI of Sentinel-1 backscatter images and that of fused images, and (c) MNDWI of MODIS optical images and that of fused images.

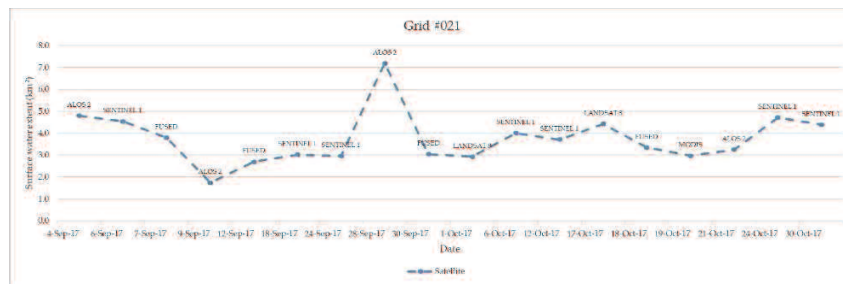
The classification of the surface water area using the integration of optical, SAR, and fused images from 4 September to 30 October 2017 for some affected grids is shown in Figure 5.8. In general, the water area in all of the grids is consistent with the ideal flood variation modelling curve (Figure 2.1). In grid 5, the maximum surface water extent occurred on 1 October, with a water area of 6.89 km², and abruptly declined after 11 days. The maximum surface water extent in grids 20, 21, 29, and 39 was observed on 28 September 2017, with water areas of 7.19, 8.07, 7.47, and 10.02 km², respectively. The surface water area abruptly declined in the next period, 30 September 2017, and subsequently decreased slowly until the end of the observation period.



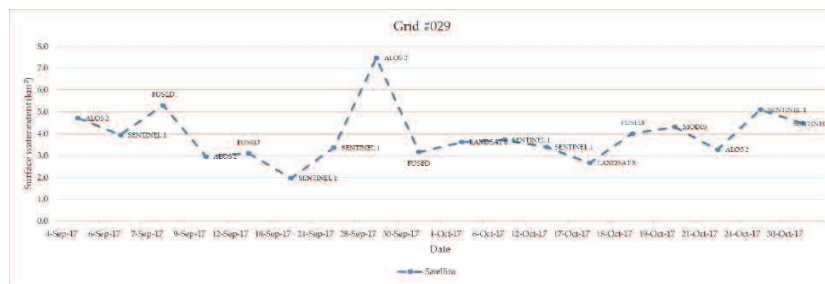
(a)



(b)



(c)



(d)



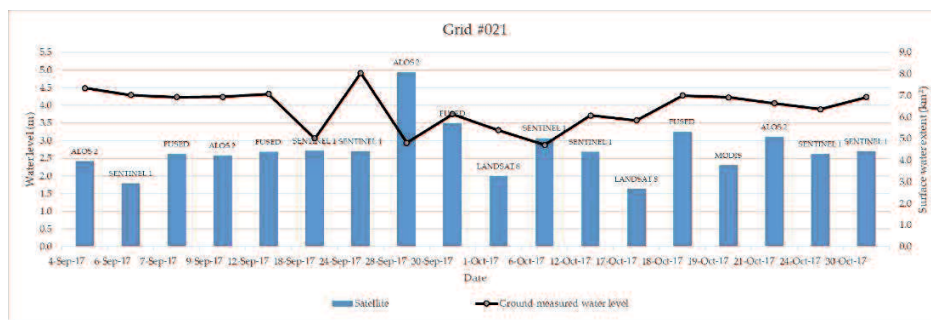
(e)

Figure 5.8: Time series of the surface water extent extracted from multiple satellite images for some of the flood-affected grids. (a) Grid 5, (b) Grid 20, (c) Grid 21, (d) Grid 29, and (e) Grid 39.

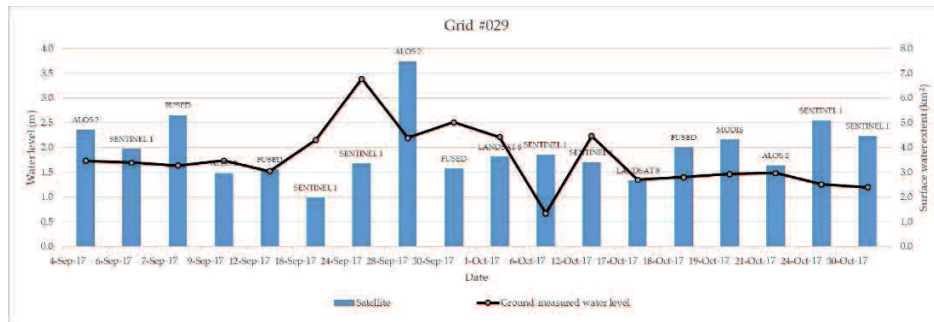
The fused image on 7 September 2017 for grid 5 showed a zero amount of surface water recorded. This might be due to the fused optical images seriously affected by cloud occlusion, causing loss of data during the cloud masking process. In such circumstances, alternatively a single image with lower contaminated data is a better option to represent data on that certain date. Discrepancies of each image can cause fluctuations in a certain period. Since this study aimed to use all the images regardless of image condition, standardization is required. For improvement, polynomial modelling could be utilized to achieve low bias and low variance time series data.

5.8 Evaluation results

Because of the scarcity of rain gauge data for the September 2017 flood event, ground observations of water level were only available for two grids. The correlation coefficient (r), RMSE, and MBE were calculated to evaluate the extraction results. Figure 5.9 shows the temporal variations of the extracted surface water area and the ground-measured water level for grids 21 and 29.



(a)



(b)

Figure 5.9: Time series of (a) the extracted surface water area in grid 21 and the water level measured at the Kg Repoh station, and (b) the extracted surface water area in grid 29 and the water level measured at the Kg Bakau station.

Based on the observed water level in grid 21, on 24 September, the water level was at a dangerous level (>3.5 m), with a recorded level of 4.91 m. Additionally, in grid 21, the peak of the ground-measured water level (24 September) occurred before the peak of the extracted water area (28 September). The RMSE between the extracted water area and the ground-observed water level is 1.55, and the MBE is -0.57 (underestimation). Similarly, for grid 29, the peak of the ground-measured water level also occurred before the peak of the extracted water area, and the values of RMSE and MBE for this grid are 2.56 and -2.15 (underestimation), respectively. Overall the two data did not have a good relationship, but there is an assumption that there is a delay of flooding event associated with the increase of river water level [59]. Therefore, the results are acceptable for further analysis.

5.9 Discussion of chapter V

In mid-September 2017, Typhoon Doksuri caused heavy rain and high winds in northern Peninsular Malaysia, which resulted in a flood event in Kangar and Arau districts. In this study, Landsat 7, Landsat 8, MODIS, ALOS-2, and Sentinel-1 images were integrated to estimate changes in the floodwater extent during the September 2017 event. A simple and operational flood monitoring methodology is presented. The integration of data from different sensors enables a higher observation frequency, which is favorable for flood response and increases the reliability of the estimation of flood duration. The degree of economic loss and structural damage is highly dependent on the flood duration.

This study found that the proposed automatic flood detection method is efficient for mapping the water surface area during flood events. The water area extraction results are acceptable for lakes that are large and homogenous. However, the classification showed a lower performance for rivers, especially when the rivers are narrow, close to urban structures, or follow a meandering path. Therefore, mapping flood occurrence in urban areas using SAR images remains challenging. It was impossible to extract the flood area in urban areas using Sentinel-1 data due to the high radar backscatter. Furthermore, the results showed that incorporating a flood variation spatial and temporal model gives a useful overview of the flood inundation process, especially in data-scarce areas. By using a grid system, unflooded areas can be ignored, thus reducing the processing time.

In the future, we plan to introduce a more reliable method to detect floods in urban areas and enhance the flood-detection ability in areas with small rivers. The implementation of a

well-defined method that considers spatial and temporal data about flood depth using digital elevation models (DEMs) and topographic information should be considered in future research.

References

- [1] Jung, Yoon Taek, Sang-Eun Park, Chang-Sun Baek, and Dong-Hwan Kim. “Evaluation of Polarimetric Parameters for Flood Detection Using PALSAR-2 Quad-Pol Data.” *Korean Journal of Remote Sensing* 34, no. 1 (2018): 117–26.
- [2] FEMA, and Analysis Grids. “Guidance for Flood Risk Analysis and Mapping Flood Depth and Analysis Grids,” no. May (2014).
- [3] Umar Lawal, Dano, Abdul-Nasir Matori, Khamaruzaman Wan Yusof, Ahmad Mustafa Hashim, Mansir Aminu, Soheil Sabri, Abdul-Lateef Balogun, Imtiaz Ahmed Chandio, and Munirah Radin Mohd Mokhtar. “Flood Susceptibility Modeling: A Geo-Spatial Technology Multi-Criteria Decision Analysis Approach.” *Research Journal of Applied Sciences, Engineering and Technology* 7, no. 22 (2014): 4638–44. <https://doi.org/10.19026/rjaset.7.845>.
- [4] Dutsenwai HS, Ahmad BB, Mijinyawa A, and Tanko AI (2016). Fusion of SAR images for flood extent mapping in northern peninsula Malaysia. *International Journal of Advanced and Applied Sciences*, 3(12): 37-48.
- [5] Mallat, Stephane G. “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation.” *Fundamental Papers in Wavelet Theory*, no. May (2009): 494–513. <https://doi.org/10.1515/9781400827268.494>.

Chapter VI

Assess water extent map with surface water map from GSW

6.1 Introduction

Serious floods have hit the State of Perlis and Kedah in 2005, 2010, 2011, 2014 as well as 2017. Frequent areas flooded are Kota Setar, Pokok Sena, Kubang Pasu, Padang Terap, Kuala Muda, Baling and Sik. There are many factors that influence the occurrence of unusual flood in Kedah and one of the causes is human activities that exploit an area unsustainably [1]. According to a study by Hafsah Saleh et al. [2], the low-lying areas, which include the main paddy lands, residential areas, highways, rails, and sugarcane plantations, are identified as those that are destructible by floods. The other class in the reclassified map includes primarily forest and rubbers which refer as highland area, which are less likely to be inundated or disrupted by floods than low-lying areas [2]. Flood on September 2017 in Kedah, overflow of river water in addition of high tide phenomenon during long duration of heavy rainfall contribute to flood [3]. The flood receding at a slow pace due to the high tide phenomenon.



Figure 6.1: a) No flood view in Kedah area and b), c) and d) Photos of submerged houses and paddy field in torrential rain in Kedah source from social media. Kedah land use majorly comprise of paddy field.

Surface water map can indicate where surface water flooding could occur as a result of local rainfall [4]. This chapter explores the spatial distribution of surface water flooding and land use that affect people's exposure to flooding in Kedah area based on proposed method to extract flood extent using ALOS-2. We used global surface water (GSW) which maps global surface water (GSW) bodies from 1984 to 2020 and provides statistics on their extent and changes were used to validate our result.

Jie Liu et. al [5], used global surface water (GSW) to validate their flood hazard zoning of the Angkor site using synthetic aperture radar (SAR) data. Based on the result the global surface

water recurrence and occurrence band dataset shows that there is correlation where high values of occurrence and recurrence shows higher flood risk in the high hazards zones and lower occurrence and recurrence value shows moderate flood risks in moderate hazard zones. Meanwhile, the zero occurrence and recurrence indicate absence of water in the region and low possibility of flood.

The GSW monthly water history datasets was referred to in order to show that the extent of the flood is higher in the lowest lying paddy field areas, thereby ensuring the degree of accuracy of the new flood extent that was formed from ALOS-2 HH+HV using an Otsu method. During the period of flooding and rice transplanting, there is a large proportion of surface water in a land surface consisting of water, vegetation and soils and with heavy rainfall it can cause surface water flooding. The GSW data was importantly referred to validate the outcome map because it shows specifically the association of the value of each pixel with the surface water flooding. The aim of the GWS monthly water history datasets in this section is to aid in justifying the flood extent map by showing the low-lying areas covered by paddy in high risk of flood.

6.2 Study area and Satellite data

Three area has been selected to be validated. This area consists of region a), b) and c) which shown in figure 6.2. Based on annual flood report on 2014 from DID [6], these three regions have flood recorded and shown in Figure 6.3. Region a) consist of residential area and paddy field and closely located with straits of Malacca and Merbok river, b) area is majorly classified as paddy field and c) consist of paddy field, several manmade which are airport and residential area.

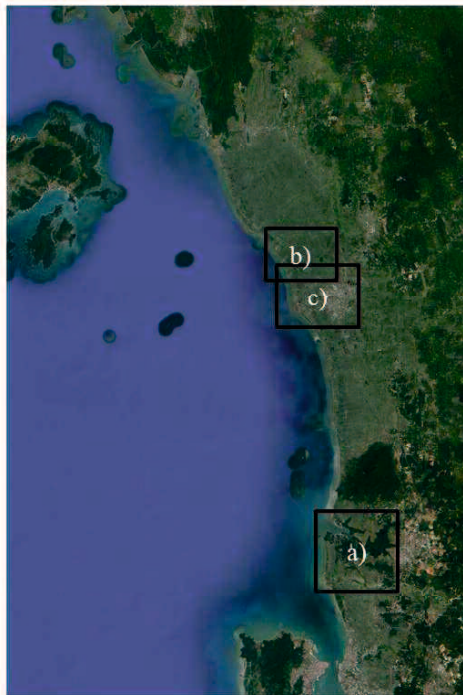


Figure 6.2: a) ALOS-2 HH+HV b) GSW binary image. Dark color indicates water region while, white color indicates non-water area.

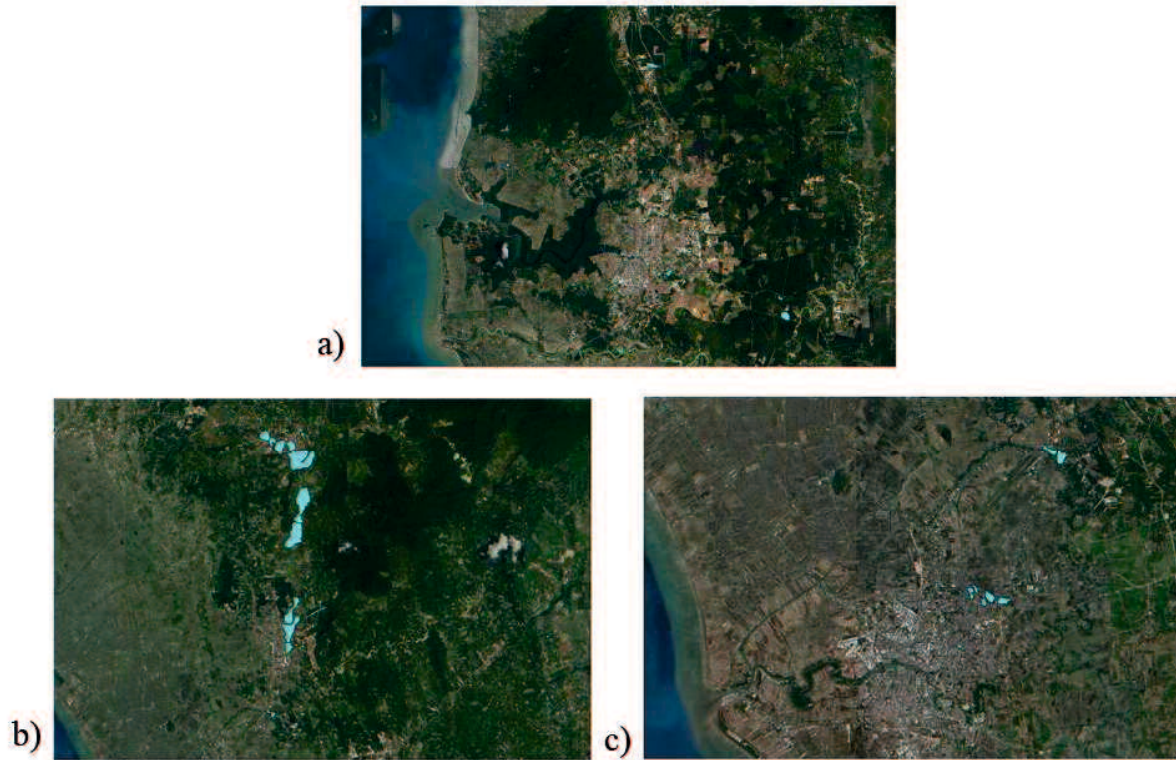


Figure 6.3: Closed up of study area a), b) and c) and the blue is the flood extent on 2014 report by DID.

Local scale flood observation may benefit from global geospatial datasets, especially in data-poor areas [7]. Pekel et al. [8] reported created the 1984–2015 global 30 m 30 m resolution surface water dataset in 2016, which maps global surface water (GSW) bodies from 1984 to 2020 and provides statistics on their extent and changes. GSW consists of information of the seven bands of global surface water dataset which are occurrence, absolute change in occurrence normalized change in occurrence, seasonality, recurrence, transition and maximum extent.

GSW monthly water history datasets were used to validate our results. The data set presented here freely available at <https://global-surface-water.appspot.com/>. Three million Landsat 5, 7, and 8 images with a 30 m spatial resolution were used to create this dataset. From 1984 to 2020, this dataset includes maps of the location and temporal distribution of global surface water, as well as statistics on the extent and variation of those water surfaces. Using an expert system, visual analytics, and evidential reasoning, each pixel in the dataset was identified as water or non-water. The dataset consists of three bands, with band 0 indicating "no data," band 1 represent "not water," and band 2 showing "water." The classifier makes less than 1% false water detections and misses less than 5% of the water [8].

Long-term water records were used to create thematic products that show many aspects of surface water dynamics. Water surfaces present during an entire year's observations are separately mapped from those that are seasonal, and the frequency with which water returns from year to year across the time-series is represented as recurrence. Between any two years of observation, transitions between permanent water, seasonal water, and land classes can be determined based on transitions between the first and last year of observation. Change

measurements at the global, continental, and country scales are obtained by merging these complimentary information layers, which capture water history per pixel, month, and year.

The GSW monthly dataset for September 2017 has scan line error so, was Classified water area from ALOS 2 HH+HV on 2014 was preferable to be validated. Classified water area from ALOS 2 HH+HV on 2014 (refer Chapter 3) has been used to validate with September 2014 GSW monthly water history dataset to justifying the flood prone area in low lying paddy field area. The GSW monthly dataset for September 2017 was

6.3 Accuracy assessment Method

The GWS map were compared with HH+HV ALOS-2 flood maps using three accuracy indicators, which are Errors of Omission, Errors of Commission and Kappa coefficient.

Reference sites that were left out (or excluded) from the correct class in the classified map are referred to as omission errors. On the classified map, the real land cover type was left out or omitted. Error of omission is sometime also referred to as a Type I error. In one area, an error of omission will be counted as an error of commission in another. By checking the reference sites for inaccurate categories, omission errors are calculated. This is accomplished by summing up the wrong classifications for each class and dividing them by the total number of reference sites for that class. Each class's omission error is usually calculated separately. This will allow us to assess each class's categorization accuracy and error.

The commission errors are related to the classed results. These are places that were classed as reference sites in the classified map but were left out (or omitted) from the correct class. Commission errors are calculated by looking for incorrect classifications on classified sites. This is accomplished by summing up the wrong classifications for each class and dividing them by the total number of categorized sites for that class.

When all parts of the error matrix are considered, the Kappa coefficient (κ) is a measure of agreement between two maps. A statistical test is used to calculate the Kappa Coefficient, which is used to assess the correctness of a categorization. Kappa essentially assesses how well the classification performed in comparison to simply assigning values at random, i.e. if the classification outperformed chance. The Kappa Coefficient can be all between -1 and 1. The classification is no better than a random classification if the value is 0. The classification is much worse than random if the number is negative. A close approximation to 1 suggests that the classification is much better than random. Referring to the previous study, the computed Kappa coefficient, the performance of an image classification method can be grouped into five categories (Richards et al. 2013): poor ($\kappa \leq 0.4$), moderate ($0.4 < \kappa \leq 0.6$), good ($0.6 < \kappa \leq 0.75$), excellent ($0.75 < \kappa \leq 0.8$), and most perfect ($\kappa > 0.8$). According to these criteria, we target accuracy Kappa coefficients of higher than 0.76 as the best accuracy. In terms of an error matrix, it is defined as follows:

$$\kappa = (\mathbf{Obs} - \mathbf{Exp}) / (1 - \mathbf{Exp})$$

For the target accuracy, since we used multi satellite, setting a target based on input data is inappropriate. Hence, we select the target standard accuracy based on the Otsu method. Referring to the previous study, the computed Kappa coefficient, the performance of an image classification method can be grouped into five categories (Richards et al. 2013): poor ($\kappa \leq 0.4$),

moderate ($0.4 < \kappa \leq 0.6$), good ($0.6 < \kappa \leq 0.75$), excellent ($0.75 < \kappa \leq 0.8$), and almost perfect ($\kappa > 0.8$). According to these criteria, we target accuracy Kappa coefficients of higher than 0.76 as the best accuracy.

6.4 Results and Discussion

6.4.1 Region a)

Figure 6.4 shows comparison of ALOS-2 HH+HV and GSW classified water and non-water image for region a). White color indicates water region while, dark color indicates non-water area. The spatial distribution affected by surface water flooding across Region a) from the GSW indicates that the highest proportion of areas under flooding are located at the paddy field lower region area, whilst the upper area no flooding detected even existence of paddy field in that region as well. Upper region where the paddy field located is actually closed to recreational forest area. This varied distribution suggests that the levels of risk of surface water flooding are determined by factors associated with topography and land use. However, because the range of backscatter intensity in flooded agricultural fields was nearly the same as that of non-flooded agricultural fields, HH+HV from ALOS-2 incorrectly identified paddy field as flood region. As a result, a histogram-based classification approach is likely to generate inaccurate results [9].

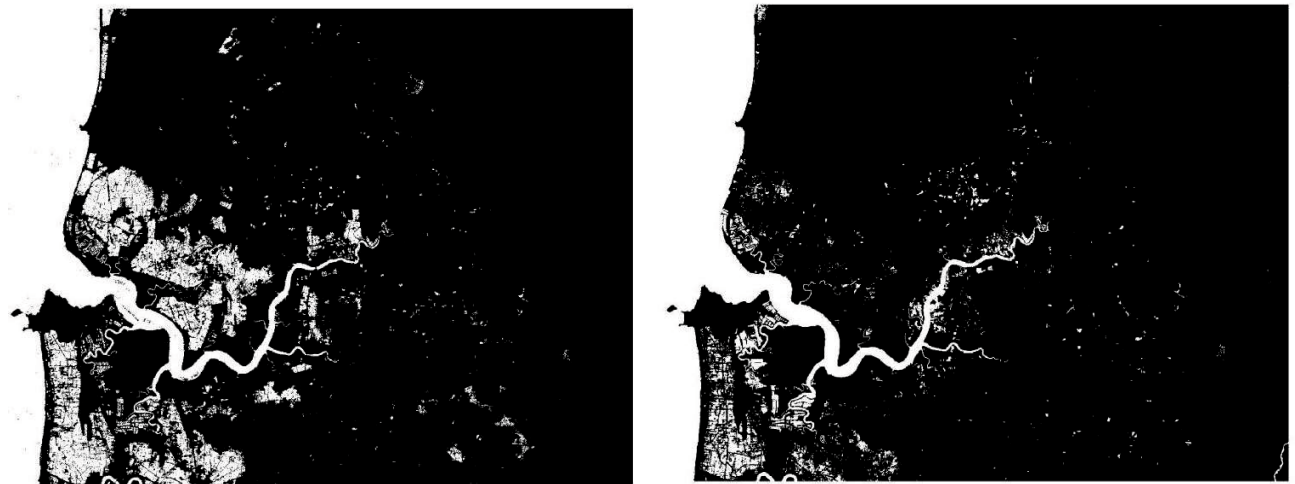


Figure 6.4: a) ALOS-2 HH+HV b) GWS binary image. White color indicates water region while, dark color indicates non-water area.

The results of Errors of Omission, Errors of commission and Kappa coefficient for Region a) was shown in Table 6.1. The temporary water class is generally delineated with satisfactory accuracies. The results agreed well with the GSW data; the accuracy was Kappa = 0.780. The floods detected outside the validation area were likely the overland floods of rice-paddy fields.

Table 6.1: Result of Errors of Omission, Errors of commission and Kappa coefficient for Region a)

	% Commission	% Omission	Estimated Kappa
Water	1.645	6.665	0.903
Non water	26.009	7.608	0.687

Kappa = 0.780 Kappa Variance = 0.000001 Observe Correct = 1012371 Total Observe = 1086530 % Observed Correct = 93.175

6.4.2 Region b)

For region b), see Figure 6.5 for a comparison of ALOS-2 HH+HV and GSW categorized water and non-water images. The color dark denotes a water area, while the color white denotes a non-water area. The same with Region a) spatial distribution affected by surface water flooding across Region b) from the GWS indicates that the highest proportion of areas under flooding are located at the open paddy field. Overall, other area other than paddy field was affected with flood. However, HH+HV from ALOS-2 identified more apparent paddy field as flood region compared to GSW. HH+HV from ALOS-2 more sensitive to water presence in mixed land cover. The overestimation of paddy fields was primarily due to the higher spatial resolution (6.25 m) of ALOS-2 compared with that of GSW data (30 m).

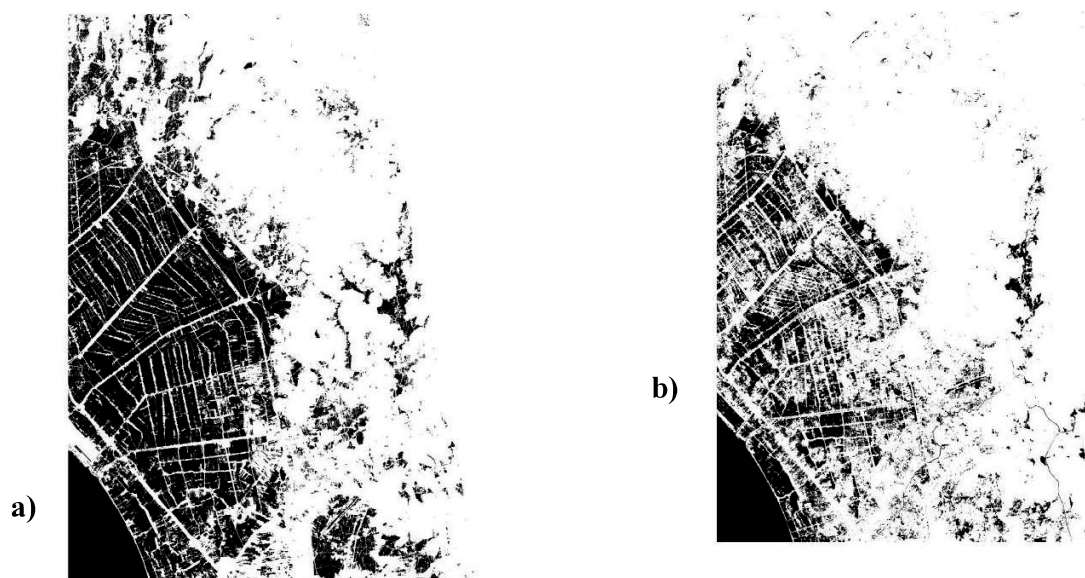


Figure 6.5: a) ALOS-2 HH+HV b) GSW binary image. Dark color indicates water region while, white color indicates non-water area.

The results of Errors of Omission, Errors of commission and Kappa coefficient for Region b) was shown in Table 6.2. The temporary water class is generally delineated with

moderate accuracies. The results agreed with the GSW data with the accuracy was $Kappa = 0.520$. The overestimation of paddy fields was mostly due to ALOS-2's higher spatial resolution (6.25 m) when compared to GSW data (30 m). Uncertainties may also arise from a change in land use during the period when the ALOS-2 image was acquired in 02 October and the monthly September GSW data.

Table 6.2: Result of Errors of Omission, Errors of commission and Kappa coefficient for Region b)

	% Commission	% Omission	Estimated Kappa
Water	49.512	14.411	0.389
Non water	4.022	19.622	0.787

Kappa = 0.520
 Kappa Variance = 0.000
 Observe Correct = 1259753
 Total Observe = 1548276
 % Observed Correct = 81.364

6.4.3 Region c)

Figure 6.6 shows a comparison of ALOS-2 HH+HV and GSW categorized water and non-water photos for region c). A water area is indicated by the color dark, while a non-water area is shown by the color white. The land use of Region c) is more heterogenous compared to Region a) and b). HH+HV ALOS-2 misclassified airport runway as water area while GSW did not classified them as flood. The same case with paddy field, airport runway range of backscatter intensity in water and flat areas is the same.

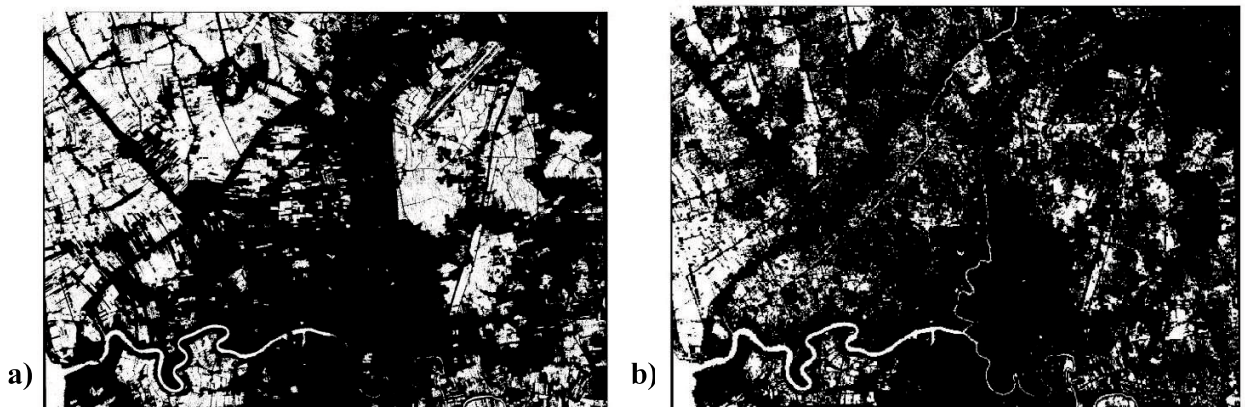


Figure 6.6: a) ALOS-2 HH+HV b) GSW binary image. Dark color indicates water region while, white color indicates non-water area.

Table 6.3 shows the findings of Errors of Omission, Errors of Commission, and Kappa coefficient for Region c). The temporary water class is generally delineated with poor accuracies.

The results poorly agreed with the GSW data with the accuracy was $Kappa = 0.312$. The overestimation of flat manmade and overestimation of paddy field might be due to high difference of spatial resolution of both satellites data.

Table 6.3: Result of Errors of Omission, Errors of commission and Kappa coefficient for Region c)

	% Commission	% Omission	Estimated Kappa
Water	9.374209	23.523928	0.478027
Non water	62.723948	36.136961	0.235454

$Kappa = 0.312$
 $Kappa \text{ Variance} = 0.000002$
 $Observe \text{ Correct} = 220246$
 $Total \text{ Observe} = 296784$
 $\% \text{ Observed Correct} = 74.210874$

6.5 Discussion for Chapter V

A surface water map can show where floods from surface water could occur as a result of local rainfall. Based on a suggested approach to extract flood extent using ALOS-2, this section explores the spatial distribution of surface water flooding and land use that affect people's exposure to flooding in the Kedah area. To corroborate our findings, we used global surface water (GSW), to be as referral map to associate flood occurrence with the surface water flood.

Overall accuracy of the estimated area of water in area a) is almost perfect agreement with value 0.903. This is because there is an open water area and ALOS-2 HH+HV can classified open water area well. Upper region where the paddy field located is actually closed to recreational forest area. This varied distribution suggests that the levels of risk of surface water flooding are determined by factors associated with topography and land use. However, because the range of backscatter intensity in flooded agricultural fields was nearly the same as that of non-flooded agricultural fields, HH+HV from ALOS-2 incorrectly identified paddy field as flood region. As a result, a histogram-based classification approach is likely to generate inaccurate results for area b), the major area is a paddy field area and there is some area misclassified as water thus, the accuracy drops to 0.780. Overall, other area other than paddy field was affected with flood. However, HH+HV from ALOS-2 identified more apparent paddy field as flood region compared to GSW. HH+HV from ALOS-2 more sensitive to water presence in mixed land cover. The overestimation of paddy fields was primarily due to the higher spatial resolution (6.25 m) of ALOS-2 compared with that of GSW data (30 m). For area c), there is a built-up zone where, urban area and airport located in this area. ALOS-2 HH+HV misclassified an airport as water area then reducing the accuracy to 0.315. HH+HV ALOS-2 misclassified airport runway as water area while GSW did not classified them as flood.

The same case with paddy field, airport runway range of backscatter intensity in water and flat areas is the same.

References

- [1] Ninla Elmawati Falabiba, Wisnu Anggaran, Affiifi. Mayssara A. Abo Hassanin Supervised, B.B Wiyono, Ninla Elmawati Falabiba, Yong Jun Zhang, Yong Li, and Xu Chen. “濟無 No Title No Title No Title.” Paper Knowledge. Toward a Media History of Documents 5, no. 2 (2014): 40–51.
- [2] Hafsat Saleh Dutsenwai, Baharin Bin Ahmad, Abubakar Mijinyawa, and Adamu. I. Tanko. “37 Fusion of SAR Images for Flood Extent Mapping in Northern Peninsula Malaysia.” International Journal of ADVANCED AND APPLIED SCIENCES 3, no. 12 (2016): 37–48. <https://doi.org/10.21833/ijaas.2016.12.006>.
- [3] Department of Irrigation and Drainage (Jabatan Pengairandan Saliran), Malaysia, 2018. Laporan Banjir Tahunan 2017/2018. Available online: <https://www.water.gov.my/> (accessed on 02nd February 2019).
- [4] Environment Agency. “What Is the Risk of Flooding from Surface Water Map? (UPDATED VERSION),” no. April (2013).
- [5] Liu, Jie, Zhiwei Xu, Fulong Chen, Fang Chen, and Lu Zhang. “Flood Hazard Mapping and Assessment on the Angkor World Heritage Site, Cambodia.” Remote Sensing 11, no. 1 (2019): 1–19. <https://doi.org/10.3390/rs11010098>.
- [6] Department of Irrigation and Drainage (Jabatan Pengairandan Saliran), Malaysia, 2014. Laporan Banjir Tahunan 2013/2014. Available online: <http://jpsweb.terengganu.gov.my/index.php/en/laporan-banjir> (accessed on 10 February, 2014).
- [7] Lindersson, Sara, Luigia Brandimarte, Johanna Mård, and Giuliano Di Baldassarre. “A Review of Freely Accessible Global Datasets for the Study of Floods, Droughts and Their Interactions with Human Societies.” WIREs Water 7, no. 3 (2020): 1–20. <https://doi.org/10.1002/wat2.1424>.
- [8] Pekel, Jean François, Andrew Cottam, Noel Gorelick, and Alan S. Belward. “High-Resolution Mapping of Global Surface Water and Its Long-Term Changes.” Nature 540, no. 7633 (2016): 418–22. <https://doi.org/10.1038/nature20584>.
- [9] Moya, Luis, Yukio Endo, Genki Okada, Shunichi Koshimura, and Erick Mas. “Drawback in the Change Detection Approach: False Detection during the 2018 Western Japan Floods.” Remote Sensing 11, no. 19 (2019): 1–19. <https://doi.org/10.3390/rs11192320>.

Chapter VII Conclusion

7.1 Dissertation Conclusions

This dissertation has dealt with the use of multi-temporal flood detection. However, it is limited. Otsu method has been widely accepted and its capability quantitatively and qualitatively proved. Many techniques are used to extract water using optical satellite images and SAR images, from image processing techniques to image classification techniques: change detection, visual interpretation using RGB composition, supervised classification, image texture algorithms, and active contour models. Meanwhile, for optical images, some examples are spectral indices, RGB False Color Composites (FCCs), supervised classification, unsupervised ISODATA classification, and dynamic thresholding. Presently, even though an abundance of near real-time data is available, the data and information still appear to be underutilized by decision-makers in the disaster response phase due to a few constraints, such as the limitation on (1) time and capacity of a person to process new geospatial datasets, (2) accessibility of near real-time data, (3) compatibility of user platforms and geospatial data formats, (4) knowledge of the data availability and inadequacy of data latency, and (5) understanding of the end-user demand on the product and timing needs. Despite various notable efforts, to this end, existing satellite technology or any single data product could not solve the current challenge from the decision-making standpoint. Hence, to fulfill decision-makers' needs, there is a demand to develop a process for improving the end product for an effective disaster response rather than developing a new system to improve flood operation. This study focuses on the standardization of multi-source remote sensing data by using the same processing platform. In this context, standardization in the meaning to reduce the inconsistency of dataset.

In mid-September 2017, Typhoon Doksuri caused heavy rain and high winds in northern Peninsular Malaysia, which resulted in a flood event in Kangar and Arau districts. In this study, Landsat 7, Landsat 8, MODIS, ALOS-2, and Sentinel-1 images were integrated to estimate changes in the floodwater extent during the September 2017 event. A simple and operational flood monitoring methodology is presented. The integration of data from different sensors enables a higher observation frequency, which is favorable for flood response and increases the reliability of the estimation of flood duration. The degree of economic loss and structural damage is highly dependent on the flood duration.

Based on our findings in Chapter IV, it shows that MNDWI shows the best and the most stable water index for the automatic thresholding of a multispectral image and total backscatter of HH+HV ALOS-2 and VV+VH for Sentinel 1 is preferable for automatic thresholding. After the integration, due to differing spatial resolution. MODIS was also heavily impacted by cloud, with the majority of pixels being deleted, resulting in no water area being recorded on September 14th and 15th. On the 14th and 8th of October, MODIS was also heavily impacted by cloud, with the majority of pixels being deleted and no water area being recorded. To achieve low bias and low variance datasets, statistical procedures such as polynomial regression can be used to overcome bias and variance.

Chapter V, the analysis proceeds by applying 2D wavelet transform to fuse two images with the same observation period using grid based. Finally, the of temporal distribution of water classification was compared with the ground water level data. Overall the two data did not have a

good relationship, but there is an assumption that there is a delay of flooding event associated with the increase of river water level. Therefore, the results are acceptable for further analysis.

In Chapter VI, we discussed about function of surface water map to indicate where surface water flooding could occur as a result of local rainfall. This chapter explores the spatial distribution of surface water flooding and land use that affect people's exposure to flooding in Kedah area based on proposed method to extract flood extent using ALOS-2. We suggest that varied distribution levels of risk of surface water flooding in the study area are determined by factors associated with topography and land use.

To summarize the main conclusion:

1. For automatic classification of multisatellite data, Otsu thresholding worked robustly. In this study, we have found multiSAR classification gives better result with dual polarization and optical images gives improve result with MNDWI as an input data when using for automatic surface water mapping.
2. Use of multisatellite data increases the frequency of observation which help in capturing the short-span flood as well as provides the possibility to track the changes of water extent during the long extent event.
3. Automatic classification accuracy of multisatellite data improves when using multiSAR data along with the multispectral data. In our study, the classification results from multispectral data improves from $r=0.498$ to $r=0.64$ when integrate with multiSAR.

7.2 Future Work

We would like to focused on developing image datasets with combination of both a high temporal resolution and medium spatial resolution by harmonizing the time-series of SAR and multispectral images. We would like to proposed a methodology that highly accurate, with high temporal resolution and medium spatial resolution and it was possible to harmonize that can help resolve issues arising from inconsistency of satellite imaging datasets.

To conclude the limitation and future works, we summarized them below:

1. Lack of calibration for data harmonisation among various sensors.
2. Varied topographical and climatical places can be selected for comparison of performance of the recommended method.
3. Accuracy assessment is still a challenging factor that restrict the high-level application of multi-source remote sensing satellites data. Collaboration for the data sharing with the local government should be promoted to investigate the accuracy and improved the consistency of multisource satellite data.