# DETECTING SLOW, GRADUAL CHANGES WITH REMOTE SENSING: FUZZIFICATION OF RANDOM FOREST CLASSIFICATION FOR COASTAL MANAGEMENT

BAKUL A. PATIL
August, 2021

SUPERVISORS:
*Dr. ir. Wietske Bijker*

Dr. *Mariana Belgiu*

# DETECTING SLOW, GRADUAL CHANGES WITH REMOTE SENSING: FUZZIFICATION OF RANDOM FOREST CLASSIFICATION FOR COASTAL

BAKUL PATIL

August, 2021
Enschede, The Netherlands

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Earth Observation Science (EOS)

SUPERVISORS:
Assistant Professor, Dr. ir. W. (Wietske) Bijker
Assistant Professor, Dr. M. (Mariana) Belgiu

THESIS ASSESSMENT BOARD:
Professor, Dr. ir. A. (Alfred) Stien (Chair)
Dr. Ratna S. Dewi  (External Examiner, Badan Informasi Geospasial, Indonesia)

# ABSTRACT

The city of Semarang faces the risk of frequent coastal and tidal flooding in combination with coastal erosion and land subsidence. Therefore, monitoring shorelines has a great significance in providing information on its dynamic nature, resource management, and evaluation of potential risks for sustainable coastal development and management. Changes in the shoreline are gradual and slow, and often, they are overlooked for a long time until it becomes a big problem. However, shorelines do not have clear boundaries as it is a transition zone between land and water. Therefore, it is a vague object. Hence, the soft classification will provide better information for the fuzzy shoreline boundary rather than hard classification. Being dynamic in nature, shorelines are challenging to identify, and therefore, their position contains a degree of uncertainty. This study focuses on understanding the problem at hand and aims at developing an approach to study the slow and gradual change in the shoreline using remote sensing and machine learning techniques. The probabilistic results of random forest give membership degrees for classes in the context of fuzzy logic. Moreover, these fuzzified results are used to determine three measures of uncertainty, namely, Confusion index, Ambiguity index, and Fuzziness for the study area. The implementation of the results of this study is further explored in contributing to the indicators of Integrated Coastal Zone Management. The uncertainty maps can also show the vulnerable areas to while disaster risk management and the areas that can be possibly reclaimed. The uncertainty in the classification shoreline over five years will result in uncertainty maps that coastal planners can use for sustainable spatial planning or disaster risk management. This study applies an approach that can be used as a tool for Coastal monitoring and management and disaster risk management.

# ACKNOWLEDGEMENTS

I would like to thank my family and friends back home who were more enthusiastic than me, for my dream to study at ITC and the strength they gave me in these dark times of the pandemic. I would not be motivated to keep doing better if not for my parent's loving words and my friend's support; no matter what time it was, they were always there for me.

I would like to thank ITC for the Foundation scholarship because of which I could fulfill my dream to study abroad.

I express my gratitude to my research supervisors, Wietske and Mariana, who guided me patiently and answered my silly questions. I thank them that they never doubted that I could do it. Thank you for believing in me and giving me the freedom to work my way. I thank you both for insightful discussions and to always letting me know my strengths and weakness, I learned a lot from you.

A special mention to Luc Boerboom for giving me courage and support in the time when I was struggling a lot with my thesis and was feeling very lost. You went out of your way to support me when I hit rock bottom. I am thankful to you for your kind words. They gave me the strength to start again. Thank you for always checking up on my progress and well-being.

I am very grateful to my family here in the Netherlands, especially Thom, who helped me during my health crisis, motivated me when I didn't even have the motivation to get out of bed, took care of me when I would even forget to eat or drink water, and who was always very patient and kind to me even at my worst. I thank you, Marja and Dick van Harten, for your loving and inspiring words and for always making me feel so much at home.

I am thankful to my friends at ITC for the insightful discussions and beautiful memories.

Lastly, I would like to thank all the people who were directly or indirectly involved in my thesis process.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

## 1.1. Background and Motivation

The consequences of global climate change are overseen until they become life-threatening for humankind and their interests. The rapid changes in the landscapes we observe nowadays result from overlooking the slow and gradual changes in the landscapes for a significant period. The major consequence of global climate change is a sea-level rise that has significantly changed the coastlines worldwide and engulfed the land. Climate change is resulting in a high risk of coastal flooding (Harwitasari, 2009). The coastal systems possess much complexity as it is a transition between land and water. It is directly affected by both terrestrial and marine processes.

The coastal areas are dynamic regions with constant changes every day, making them strategically challenging locations in managerial aspects. (Dewi, Bijker, Stein, & Marfai, 2016). This makes monitoring these areas of great significance. Monitoring coastlines will provide helpful information, such as the annual rate of erosion of coasts, which can be used to predict future problems or hazards. Furthermore, shoreline mapping and monitoring allow us to know its spatial distribution and trend analysis (Dewi et al., 2016). In order to obtain this information, we need good techniques for image classification.

With the help of remote sensing techniques, it has become easier to monitor the landscapes and study how and why they are changing. Remote sensing imagery can be considered as a Euclidean space, where the number of bands and pixels represents a point in the space dimensions (Alonso, Malpica, & De Agirre, 2011). Every pixel can be considered a feature space where specific spectral band values represent one or more classes. Image classification has made it possible to see the land cover that exists and its use by humans. Image Classification has always been tricky (Dewi et al., 2016); it is impossible to achieve the perfect classification that depicts exact land cover/land use. In Supervised classification, every pixel of the satellite image is classified to depict a particular ground type or class for the better representation of the feature space (Alonso et al., 2011). It is easy to classify the pixels that cover 100% single class attributes. However, the classification gets difficult in two cases; a) the mixed pixels (Gebbinck, 1998; Loosvelt et al., 2012), where the pixel represents two classes (say, land and water), in which case the dilemma is which class to be assigned as two classes share the feature space. b) the spectral resemblance of the classes (Gebbinck, 1998; Hong et al., 2019; Verbeiren, Eerens, Piccard, Bauwens, & Van Orshoven, 2008), where two or more classes overlap in the feature space of the pixel. This results in a lot of resemblance as the classes overlapped, making it very challenging to separate the classes or assign the pixel a class (e.g., settlement and barren land). The classification becomes more challenging when a pixel represents the boundary of two classes, or it represents more than one class. In the ideal case where every class has unique spectral values would result in perfect classification, whereas in practice, spectral signatures for few classes tend to overlap or are very similar (Lee, Warner, & Virginia, 1996). Image classification has been enhanced by Machine learning techniques. With Machine learning, a methodology can be developed to extensively study the single class pixels as well as mixed pixels in the satellite imagery (Cai, Zhang, Yan, Zhang, & Banu, 2021; Tuda & Luna-Maldonado, 2020; Wang, Fan, & Wang, 2021).

Studying coastline comes with its challenges with mixed pixels and overlap of more than one class. These types of pixels are assigned the value of the dominant class in hard or crisp classification. This type of classification does not show gradual changes in the land cover classification in space and time at pixel level. In studying slow and gradual changes, we need to consider soft or sub-pixel classification (Verbeiren et al., 2008). To address this, an approach combining GIS and Machine learning techniques can be used.

Random forest classifier, a supervised machine learning classifier, has achieved much appreciation in the last two decades for its excellent outputs of classification and high processing speed (Breiman, 2001). Random forest is a popular machine-learning technique nowadays due to its good classification capability with relatively reduced training samples and user-defined parameters (Belgiu & Drăgut, 2016; Breiman, 2001; Robnik-Šikonja, 2004). Random Forest can solve overfitting in decision tree models to a certain extent (Breiman, 2001). Moreover, it is also famous as it addresses the Hughes phenomenon (Alonso et al., 2011) by introducing randomness in two stages of the classification: first, random sample selection for constructing trees, and second, the random selection of a user-defined number of variables (Alonso et al., 2011; Belgiu, 2018). Random forest is a robust classifier that can efficiently handle high-dimensional and imperfect data and gives promising results (Breiman, 2001). However, this method proves effective when the dataset has crisp values (Bayram et al., 2017). The crisp data has very definite boundaries, unlike fuzzy sets. Therefore, a feature like a coastline with no definitive boundary can hold a degree of uncertainty is represented by the hard or crisp classification. De Matteis et al. (2015) concluded in their study that although the results of RF are very accurate and promising, with fuzzy logic, it can better manage the noise in the datasets. De Matteis et al. (2015) also mentioned the advantage of fuzzy logic applications and the lack of their application. Therefore, it would be interesting to use the fuzzy logic for RF results and use it in a real-life application.

This concept is the primary motivation of the study, to apply fuzzy logic with the probabilities of classes obtained by Random forest to see the class proportion of a specific area which in turn can depict the degree of uncertainty in the class membership. The application of fuzzy logic and GIS for the decision-making process has been proved very useful for site suitability of potential urban spaces and for arguing the selection of a specific area (Arefiev, Terleev, & Badenko, 2015).

The approach of random forest with fuzzy logic is referred to in this study as "Fuzzification of random forest results (FoRF)." The use of the results of the FoRF approach is explored for coastal management in this study and how it can contribute to looking for vulnerable areas with high uncertainty. Also, how it can be integrated with the decision-making process for possible scenarios for mitigation and the degree of uncertainty, it holds in its impact.

## 1.2. Integrated Coastal Zone Management (ICZM)

Integrated coastal zone management is a multidisciplinary process that can be iterated to achieve sustainable management plans for coastal zones, approved by European Parliament and the Council concerning European coastal threat. ICZM is an indicator-based approach to assess sustainable coastal development (European Environment Agency, 2000a; Marti, Katrien, Borg, & Valls, 2007). This ICZM involves the whole process of data acquisition, planning and management on a broader scale, decision-making, and monitoring of implementation plans (European Environment Agency, 2000a). ICZM integrates participation and involvement of government, local authorities, citizens, and all other stakeholders to evaluate the societal goals in a particular area along the coasts and work towards meeting these objectives through assessing with appropriate indicators (European Environment Agency, 2000a). ICZM aims to work on long-term solutions to balance objectives in every aspect, namely, environmental, economic, social, cultural, and recreational, that fall within the boundary of the natural ecosystem of a given area (European Environment Agency, 2000a).

Moreover, the integration in the name also means integrating objectives, resources, and efforts from all policy areas and administration levels to meet ICZM objectives for Sustainable coastal management (European Environment Agency, 2000a). These objectives are achieved when the assessment by the predefined indicators is met. The ICZM approach is necessary as currently, the visions of different administrative levels are often fragmented. Therefore, it seemed necessary to look for implementation of FoFR results in the existing information system on the same objective data for an integrated vision of the different visions and interests in the coastal areas.

## 1.3. Management of Semarang coast, a *"wicked"* problem

According to Maplecroft's (Nurhidayah, 2019; Verisk Maplecroft, 2021) climate change vulnerability index, 1,500 of Indonesia's islands are predicted to be underwater by the year of 2050 due to sea-level rise (Nurhidayah, 2019). Experts have predicted through their models, 25 cm to 50 cm sea-level rise in 2050 and 2100, inundating many coastal cities of Indonesia. The impact of sea-level rise is also accelerated by the land subsidence issue in these areas (Nurhidayah, 2019). IPCC has inferred that "*without adaptation, hundred* *million people worldwide will be displaced due to land loss by the year 2100*" (Nurhidayah, 2019; Oppenheimer & Glavovic, 2019).

Being a coastal city, Semarang faces many challenges such as tidal floods, coastal abrasion, subsidence of land, and sea intrusion (Hadi, 2017). All these hazards significantly impact the citizens, communities, and industries that reside here, posing a severe threat to citizens' health, socio-economic development, economic stress, and delineation of the property value (Hadi, 2017). The causes of these phenomena are the cutting of mangrove trees along the coast, land-use change, and other human activities such as withdrawing water, stress on the land for providing shelter to the overgrowing population, and massive industrial growth (Hadi, 2017; Harwitasari, 2009; Nurhidayah, 2019; Oppenheimer & Glavovic, 2019).

Sea level rise (SLR) poses a significant threat to the settlements, cultural heritage, infrastructures, and habitat of several species sensitive to these changes. The people with low economic status who depend upon natural resources to earn their living are considered highly vulnerable (Nurhidayah, 2019). The people living here are highly prone to natural disasters as they do not have enough money to relocate. People keep living there despite all the discomfort.

Semarang is amongst the 100 resilience city programs funded by Rockefeller Foundation (Hadi, 2017). The city is expected to come up with innovative strategies to deal with these problems. Various government and local initiatives are being carried out to temporarily deal with tidal floods and floods (Hadi, 2017). However, the more integrated the initiative, the more long-term sustainable coastal management can be achieved. In this section, it was seen how the coast of Semarang is affected by natural as well as human activities, and the effect of this combination makes it a very wicked problem involving various factors affecting each other, forming a wicked cycle that is very difficult to break. For instance, the extraction of water by citizens and industries causes land subsidence, which aggravates the already increasing problem of sea-level rise, causing flooding and tidal floods, which compel people to add levels to their homes, resulting in land subsidence as the load-bearing capacity of the soil decreases by more extraction of water (Suripin & Helmi, 2015).

In the wickedness framework by Spatial Engineering (see Annex A), the problem is known to some extent. However, the stakeholders are fragmented, and their consensus is relatively less on the issue, therefore making the problems of Semarang coast a "wicked" problem. The studies are going on to understand the situation and the dynamic nature of the coastline better. Furthermore, the conflict of the coastline adds to the wickedness of lack of information as the indicative boundary might not be accepted by the people. On the other hand, according to Nurhidayah (2019), the existing government policies have fragmented vision and stakeholder conflicts due to the lack of involvement of citizens in decision-making. From which we can conclude that the stakeholder consensus is yet to be achieved. Since the problem is very complex to comprehend, it needs to be solved bit by bit. Therefore, this research will focus on the dynamic areas that have a lot of uncertainty and how the results of this research can contribute to ICZM indicator assessment for sustainable coastal management.

### 1.4. Problem Statement

The coastlines are vague objects as they do not have a crisp boundary. The coastlines are transition zones where the land gradually meets the water. Therefore, the pixels representing this feature space have mixed pixels representing one or more classes, and there are pixels with a high resemblance between two or more classes. These challenges need to be addressed. This research will look into the mixed pixels and study the membership degree of the classes it represents. The transition zone of the coastline with a continuous gradient can be seen ending abruptly when represented by crisp classification. The indicative boundary represented by the crisp classification can hold a degree of uncertainty. Stakeholders' can argue the indicative boundary of the coastline classified by the crisp classification, which can be subjective as people use different thresholds to separate land and water. Thus, there is a need to understand the coastline with its fuzzy nature to make better management decisions. The uncertainty of the classification and overlap of more than one class which can create confusion while assigning classes, needs to be considered to understand the nature of the coastline and the problems near it.

The limitation of existing GIS methods is that they are based on Boolean logic, but the wicked problems often cannot be answered by Boolean logic as there can be more than one answer. For this reason, decision-making with the conventional Boolean logic or crisp logic is not efficient (Karabegovic, Avdagic, & Ponjavic, 2006). Therefore, there is broad scope to explore fuzzy logic in the decision-making process. The combination of the fuzzy logic with the random forest can be seen as an exciting extension. It can be seen as the modified version of the existing crisp method of random forest that can assess the uncertainty of the classification, which can contribute to effective coastal management. By knowing the ambiguous areas, the coastal planners can identify the vulnerable zones and prioritize areas while making risk management plans. An informed decision of where the coast needs to be protected because of the changes occurring, be it slow and gradual, and whether they are at high risk. Identifying the slow and gradual changes will also help predict the effects of coastal flooding and other disasters, providing an interdisciplinary approach to the wicked problem as it combines the knowledge of two fields, spatial information systems, and spatial planning and governance.

## 1.5. Research objectives and questions

**Main Objective**: To apply the fuzzy logic with the random forest method to determine the uncertainty of the gradual changes in areas near shorelines that can be further used in the decision-making process for sustainable coastal management.

**Sub- Objectives:**

a) To produce fuzzified probabilistic results by the random forest algorithm.
b) Determine the uncertainty of the predicted classes and visualize the uncertainty maps.
c) To explore how the results of FoRF can be linked to indicators of the ICZM

**Research Questions**

a) How can the results of random forest be fuzzified?
b) How can the uncertainty be measured?
c) Which ICZM indicators can be improved by the FoRF results and how?

## 1.6. Conceptual Framework

This MSc thesis will address the research mentioned above with an approach using an ensemble of existing machine learning and remote sensing methods (FoRF) to assess the uncertainty in the classification of the coastline to detect and monitor slow and gradual changes in the coastline in space and time over five years. In addition, we will also investigate how this knowledge can be further used to meet the ICZM indicators for decision-making processes.

When it comes to sustainable urban planning, several heterogeneous data needs to be integrated before making any decision, and many factors have to be considered (Sideris, Bardis, Voulodimos, Miaoulis, & Ghazanfarpour, 2019). In a recent study by Sideris et al. (2019), the random forest has been used to predict different urban spaces' suitability for a particular use. Analysis can be challenging when the boundaries are fuzzy, let alone spatial planning in this situation. As the boundaries of coasts are dynamic, it carries a certain degree of uncertainty that decision-makers must take into consideration for managing coastal areas sustainably. Like Sideris et al.'s (2019) study, the results of FoRF can provide a degree of uncertainty in the classification maps used for ICZM plans. This can be useful for taking appropriate measures to plan the coastal cities for safety during hazards and assessing the vulnerability index for ICZM indicators. The Uncertainty maps can also show vulnerable zones where the measure for conservations is needed. For instance, the maps would help in the identification of highly uncertain zone, and how many people are vulnerable (as they reside in that area). by contributing to the ICZM indicators, the degree of uncertainty for making any plan for the coastal management can help in preventing the future damage or hindrance in the action plans.

Figure 1. depicts the concepts used for this study; it is a comprehensive view of the study. The concepts of supervised machine learning are used in remote sensing to assess the uncertainty in the classification and show ambiguous and vulnerable areas. This will help study the changes in the vulnerable zones of shoreline and act as a monitoring tool.

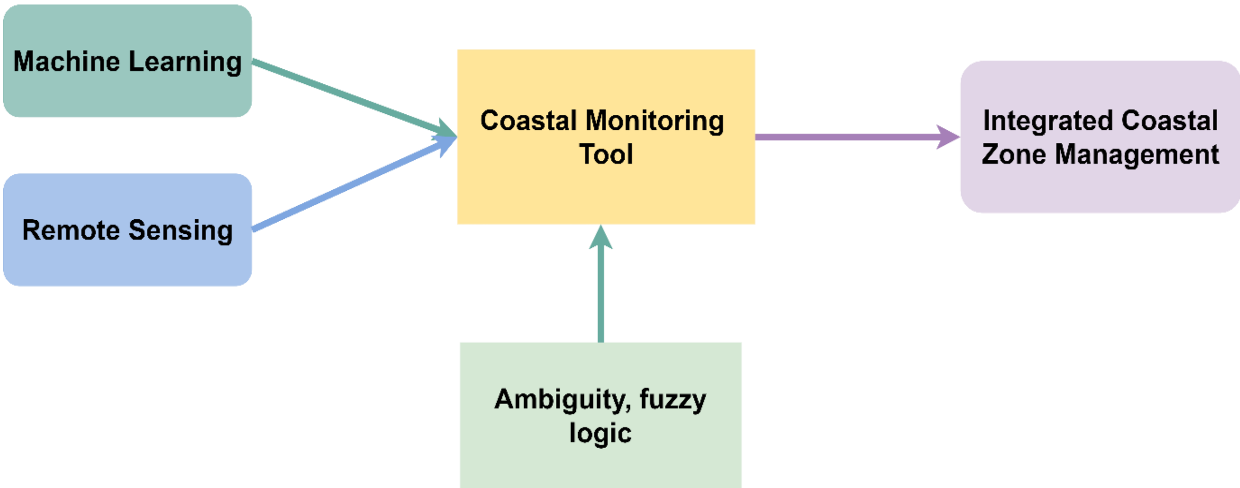



Figure 1 Conceptual diagram for the study

Moreover, the results obtained from this study can be helpful in the determination of the uncertain areas which are vulnerable. The results of this study can be interpreted by ICZM experts for the evaluation of indicators and scenario development and their uncertainty of implementation for the decision-making process. Figure 2 depicts the study's approach, the broader view of the workflow of the study than the detailed methodology. Data acquisition is made in the first phase, then in the processing phase, the images and other indices are defined as input variables for the algorithm, and the ambiguity maps are obtained.. Further, the use of the results is explored for ICZM indicators.

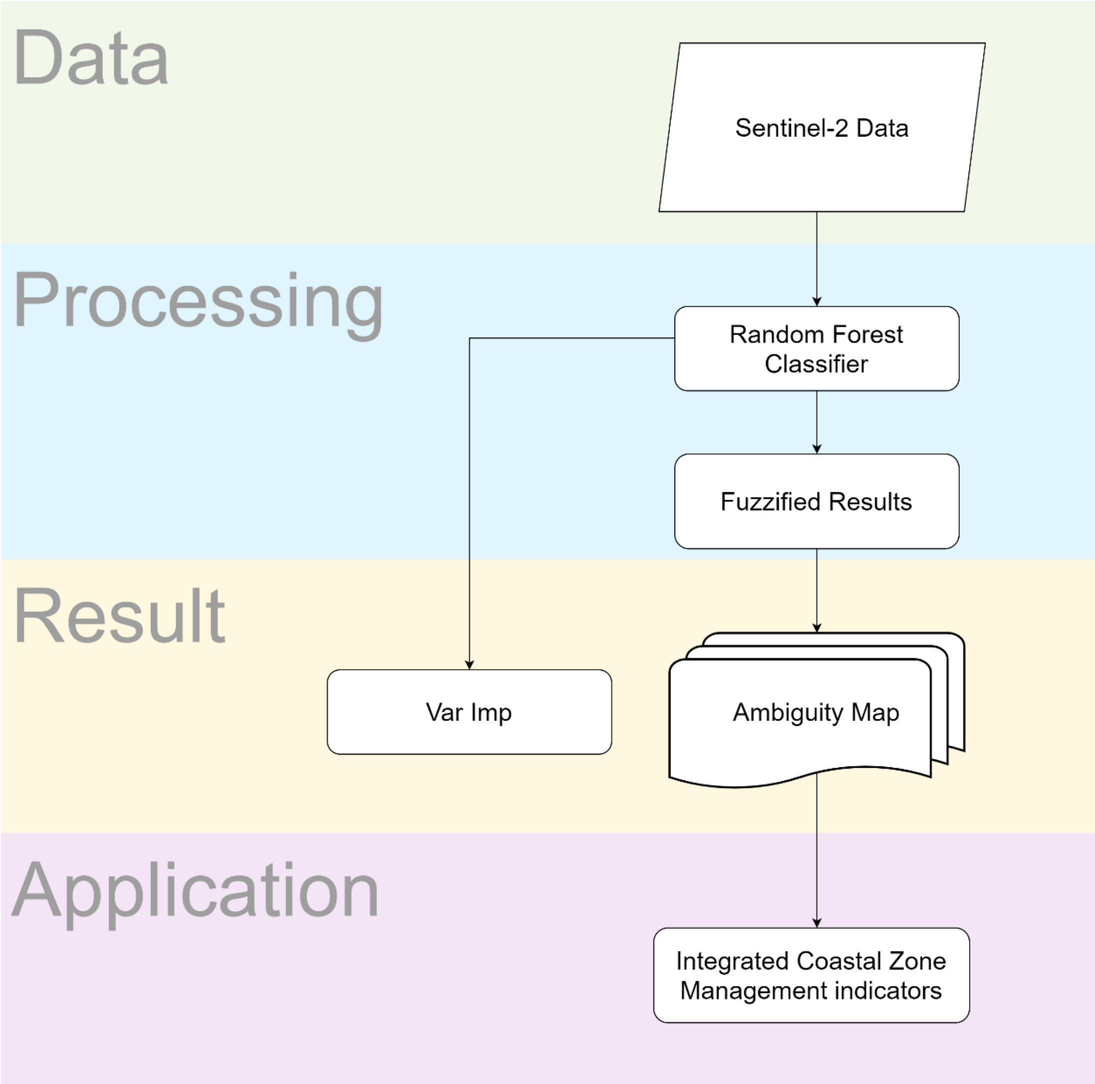



Figure 2 Approach of the study

## 1.7. Outline of the document

The outline of the document is mentioned for the reference of the reader. The first chapter contains a brief introduction of the topic and its background information. It also contained the motivation behind the MSc research topic. An attempt has been made to sufficiently introduce all the aspects of the research. The section further discusses how the coastal management of Semarang has become a wicked problem and the problem statement this study will look into, along with the research objectives and questions for this study. Furthermore, the first chapter also contains the conceptual approach to give a broader idea to the reader about what this thesis will contain.

The second chapter, Literature review, involves all the studies and research that have already been done and is the basis of this study. Chapter two will present the relevant work and mention all the existing problems and research gaps in detail. This chapter will consist of the relevant literature that will lead to the formulation of the methods for this study.

The third chapter, Materials, and methodology, present the data and other materials and software used for this study. It will mention all the processing and methods done and their justification. This chapter will argue why the methods are appropriate and the workflow in detail.

The fourth chapter, Results, will consist of the results of this study and their significance and contribution to ICZM indicators. Followed by this chapter is the fifth chapter, discussions of the results and their interpretations are represented. Discussions will present the personal views of the researcher and the limitations of the study.

The sixth chapter, Fuzzy random forest- a tool for ICZM, discusses the implementation of the work done in this study in the process of ICZM. Lastly, the seventh chapter presents the conclusions drawn from the entire study.

# 2. LITERATURE REVIEW

## 2.1. Remote Sensing for Coastline

The application of remote sensing can be seen in many fields, such as urban monitoring, detection of phenomena like floods and fire from remotely sensed data (Camps-Valls, 2009). Valuable information can be gained using remote sensing data. For example, we can monitor natural landscapes and manmade settlements. This information proves useful for policy and decision-makers, even tourism (Camps-Valls, 2009; Camps-Valls, Tuia, Gómez-Chova, Jiménez, & Malo, 2012). The recent rapid development of remote sensing technologies and an increasing need for precise coastline monitoring leads to the need for a renewed look at remote sensing capabilities for coastal monitoring (Camps-Valls et al., 2012; Toure, Diop, Kpalma, & Amadou, 2019; Tran & Tran, 2009).

In the past, the extraction of the information about coastline was done with photogrammetry (Appeaning Addo, Walkden, & Mills, 2008; Gens, 2010; Pardo-Pascual et al., 2018), however recently with the launch of satellites with sensors for the optical and microwave section of the electromagnetic spectrum paved the new way (Gens, 2010). The most frequent method for detecting shorelines is still a visual interpretation, which has some drawbacks and is very subjective in every way (Gens, 2010; Zhang, Yang, Hu, & Su, 2013). The shoreline can be identified more objectively using tidal datum markers or unsupervised classification systems (Zhang et al., 2013). However, as the dimensionality of the pixels is high in the satellite imagery, noise, and high spectral resolution, it became difficult to extract the coastline with the traditional automated mapping approaches, namely, (1) edge detection (Mason & Davenport, 1996; Zhang et al., 2013); (2) band thresholding methods, wherein selected user-defined threshold value (Liu & Jezek, 2004; S.-J. Tang, 2009); (3)the classification method that identifies land and water (El-Deen Taha & Elbeih, 2010; Khatami, Mountrakis, & Stehman, 2017; Muslim, Ismail, Khalil, Razman, & Zain, 2011; Tuda & Luna-Maldonado, 2020); (4) studying two or more data sources (e.g., multispectral imagery and airborne lidar data) (Lee & Shan, 2003; Zhang et al., 2013). Therefore, the application of machine learning in remote sensing has given better results for this problem of classification of the features (Camps-Valls, 2009; F. Tang & Ishwaran, 2017; Wang et al., 2021).

## 2.2. Coastline classification with indices

Landsat and other remote sensing satellites have been providing digital images in infrared bands with well-defined land-water interfaces since 1972 (Alesheikh, Ghorbanali, & Nouri, 2007). As a result, remote sensing imagery and image processing techniques may offer a solution to some of the challenges associated with creating and updating coastal maps (Alesheikh et al., 2007). The most common application of index techniques for water bodies distinguishes the water from the background using a threshold value (Acharya, Subedi, & Lee, 2019).

To distinguish water from the background, McFeeters (1995) recommended a zero threshold for NDWI to distinguish water from the non-water background. As the value of NDWI is positive for water features, they are enhanced from their background (McFeeters, 1995; Xu, 2006). On the other hand, Xu (2006) concludes that the zero threshold cannot distinguish water bodies from the built-up area, and to achieve that, he used shortwave infrared (SWIR) band instead of near-infrared (NIR) in McFeeters' NDWI, which he named as Modified NDWI (MNDWI).

Automated Water Extraction Index (AWEI) is relatively more accurate in classifying edge pixels than MNDWI (Feyisa, Meilby, Fensholt, & Proud, 2014). Feyisa et al. (2014) developed the method AWEI to identify water bodies with noise or urban background. This could be very useful if one wants to see coastal or tidal flooding. The interesting functions of AWEI are: AWEIsh removes the shadow pixels, and $AWEI_{nsh}$ distinguishes urban areas (Feyisa et al., 2014).

Furthermore, it is recommended to use NDWI, NDVI, and $AWEI_{nsh,}$ and MNDWI to identify the mixed water pixels in shallow water or rivers (narrow) as an automated method for extraction of water bodies (Acharya et al., 2019; Jiang et al., 2018; Wicaksono, Wicaksono, Khakhim, Farda, & Marfai, 2019). Therefore, it would be interesting to use these indices as input parameters for the approach of this research.

Amongst all the indices $AWEI_{nsh}$ gave the best results in the distinction of best land cover class (Jiang et al., 2018; Wicaksono et al., 2019); however, a new index, especially for sentinel-2 images, was proposed, the Sentinel Water Mask (SWM) which gave results equally good (Milczarek, Robak, & Gadawska, 2017). The visual assessment and statistical results of SWM were very good with 96%accuracy in water detection with more contrast between water and non-water than other indices (Milczarek et al., 2017). Therefore it would be a useful input parameter for this research.

The table 1 presents all the indices that will be used in this research as input parameters with their formulas (for Sentinel-2) and sources. The indices will help in distinguishing between water and non-water areas better based on the above discussed literature. The literature provides a context on how the indices were formulated and their capability of shoreline extraction

.

Table 1 Indices used in the study with their formulas and sources

| INDICES | FORMULA | SOURCE |
|---|---|---|
| **NDWI** | (Green-NIR)/(Green+NIR) | (McFeeters, 1995; Wicaksono et al., 2019) |
| **MNDWI** | (Green-SWIR1)/(Green+ SWIR1) | (Xu, 2006) |
| **Sentinel Water Mask** | (Blue+Green)/(NIR+SWIR1) | (Milczarek et al., 2017) |
| **AWEI (nsh)** | 4*(Green-SWIR1) - (0.25*NIR+2.75*SWIR2) | (Feyisa et al., 2014; Wicaksono et al., 2019) |

## 2.3. Machine learning in remote sensing

The supervised classification methods became more popular recently due to their flexibility in handling diversity in the appearance of an object in question for an image scene (Belgiu & Drăgut, 2016; Niemeyer, Rottensteiner, & Soergel, 2014). Decision trees are a powerful tool amongst other machine learning tools for regression and classification, where each tree node is split based on attributes with the help of information gain or Gini impurity (Rokach & Maimon, 2006). The tree ends with the leaves in the decision tree, consisting of pixels of one class only, i.e., pure classes. The advantage of a decision tree classifier is the accuracy in its classification and robustness. However, the disadvantage of overfitting cannot be overseen in complex level classifications (Zhou, Zhang, Zhou, Guo, & Ma, 2021). The input data highly influence decision trees, and there are chances that a part of the decision tree, i.e., the subtree, is repeated in the tree resulting in overfitting (Zhou et al., 2021).

It is very easy to use as it has very few user-defined parameters; moreover, the results can be interpreted with ease. However, the decision trees tend to overfit, which could lead to misclassification of few pixels. This could be solved by the Random Forest classifier, which is technically an ensemble of several decision trees that are formed by a subset of random attributes and a random subset of the samples. (De Matteis et al., 2015).

### 2.3.1. Random forest

Among existing supervised machine learning, RF classifier has gained a lot of appreciation over the past two decades (Breiman, 2001) due to high accuracy classification and high-speed processing (Bonissone et al., 2010; De Matteis et al., 2015). Random forest is a forest of the decision trees working together as classifiers. Each tree selects a subset of the attributes randomly, after which every tree votes for a predicted class. The final class in the random forest is assigned by the majority of the votes for a particular class. Figure 3 shows the concept of the random forest (Ruiz Hernandez & Shi, 2018).

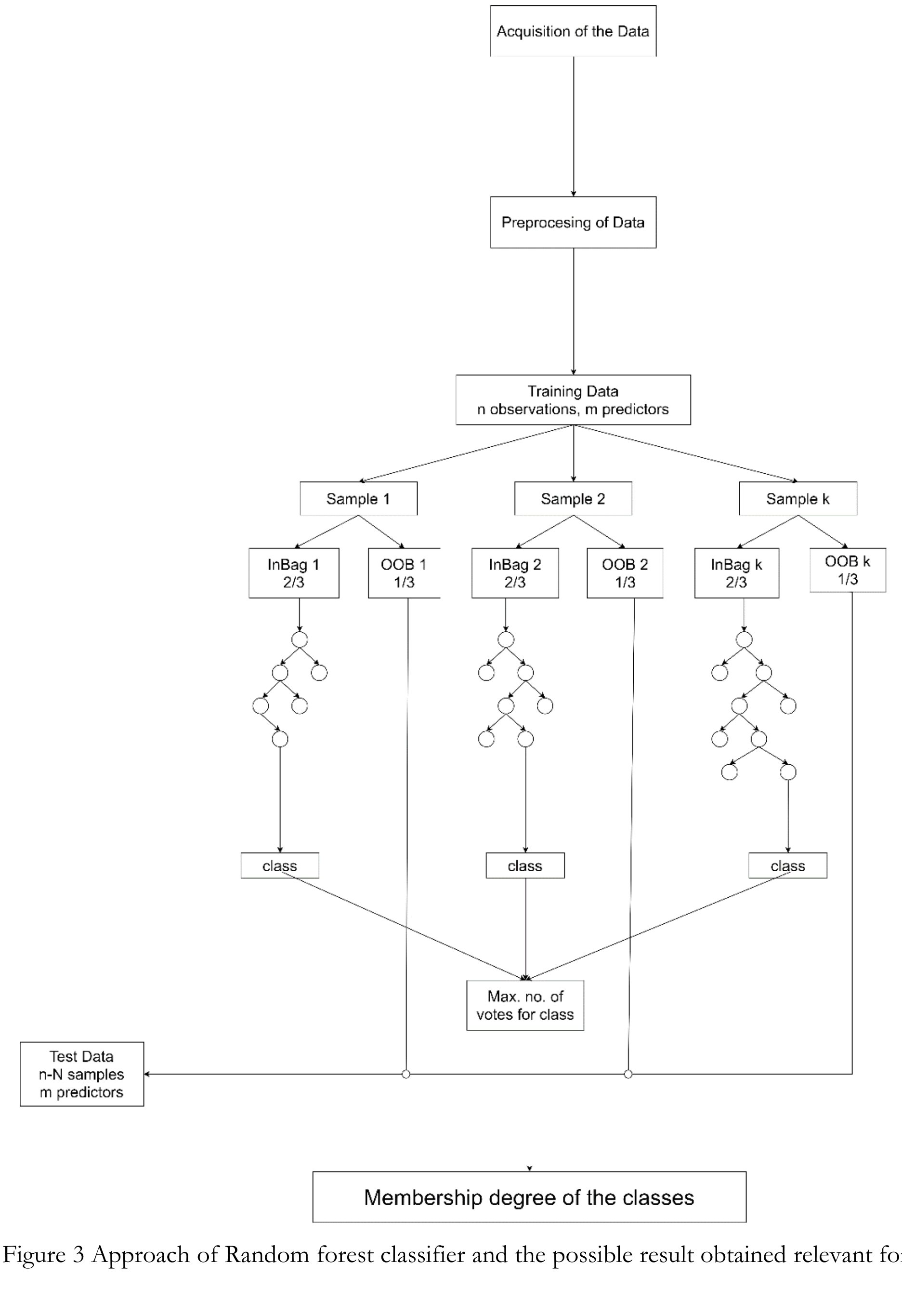



Figure 3 Approach of Random forest classifier and the possible result obtained relevant for this research

Random forest is widely used for the prediction of the classes in the case of classification. Random attributes are considered to form these decision trees, which makes the classification accurate (Breiman, 2001).
Many studies have used RF in various sectors, from developing a generic classifier to analyze hyperspectral data (Ham, Chen, Crawford, & Ghosh, 2005) to predict a compound classification based on their structure in the field of cheminformatics (Svetnik et al., 2003). In vegetation mapping, by aerial pictures, RF outperformed the maximum likelihood classifier which is traditionally used in vegetation mapping (Feng, Liu, & Gong, 2015). Unlike the other parametric techniques such as the maximum likelihood classifier, a non-parametric technique like RF is more efficient in classifying as it uses bootstrapping and random subsets in its model to form decision trees (Millard & Richardson, 2015).

Multiple classifier systems are preferred more than individual classifiers (Bonissone et al., 2010). Shoreline extraction done with the RF classifier by Bayram et al. (2017) showed that it works best in high-resolution images. However, for detecting the slow and gradual change of the dataset with fuzzy boundary, for instance, shorelines, wetlands, soft classification is needed. Although the random classification has promising results and has greatly managed imperfect datasets with crisp values, it still needs to be explored in the context of fuzzy (De Matteis et al., 2015).

## 2.4. Fuzzy logic

The fuzzy logic was first introduced by L.A. Zadeh (1965) in his article "Fuzzy Sets-Information and Control" (Zadeh, 1965), which disagreed with the assumption that the area of consideration has a defined boundary. In other words, he denied the binary logic of 0 and 1, or yes and no; or black and white, as there is always a grey area that has a "fuzzy" boundary somewhere between 0 and 1. According to Ross (1995), fuzzy set logic is used in many fields and technical applications for wicked problems (Kuncheva, 2001; Ross, 2010). Also, combining GIS techniques with Fuzzy set theory is an unconventional way to study the land cover (Kurtener & Badenko, 2003). Bonissone et al. (2010) used the "forest of fuzzy decision trees" to infer that the method has the strength of the randomness that highly diversify the trees predictors in the RF, and the flexibility of fuzzy logic for decision trees are aggregated into a method that can efficiently manage the uncertainty (i.e. pixels with more than one membership) in the datasets. Coastal monitoring can be efficiently carried out by detecting slow and gradual changes, this has been done in a study by Dewi et al. (2016) where fuzzy logic has been used to extract the shoreline using fuzzy c- means (FCM). Therefore, it would be interesting to use the probabilistic results in fuzzy logic context for shoreline detection and its uncertainty.

## 2.5. Fuzzification of RF

The RF classification is an ensemble of several trees that give the classification results and probabilistic results (Breiman, 2001). The module also integrates several other user-defined functions, such as calculating variable importance, and proximities (Loosvelt et al., 2012). Random Forest models are very confusing to interpret and the quantification of uncertainty becomes very difficult due to its "black-box" characteristics (Baake, 2018). Amongst other approaches for looking at the local level classification quality, using probabilities of classes as the degree of memberships for the quality of prediction, using tree votes in a random forest, or probabilities in neural networks are very efficient (Khatami, Mountrakis, & Stehman, 2017). To consider probabilistic results of classification as the degree of class membership is used as one of the factors to quantify uncertainty (Baake, 2018). The concept is for a particular pixel. The higher the probability of the degree of membership for a given class, the lower the uncertainty for that class. This method is prominently used especially in medical predictions (Escobar et al., 2015; Gurm et al., 2014), any forecasts, ranking, calculation of anticipated utility, and in observational studies that deal with uncertainty (Olson & Wyner, 2018).

Uncertainty evaluation tools provide uncertainty maps, which can reflect the classification accuracy. It shows locally where the mixed pixels are located with high uncertainty, in other words, where the "unreliable pixels" exist (Roodposhti, Aryal, Lucieer, & Bryan, 2019).
Conventionally the evaluation of the accuracy of claasification of the map is done at the class level or globally by looking at the confusion matrix. However, this does not give much information about the uncertainty in the classification or the spatial variation of uncertainty throughout the classification (Loosvelt et al., 2012). It is preferred to know the degree of membership of a pixel rather than just to know which class it belongs to. The soft classification gives the membership degree for each pixel and allows quantification of uncertainty at the pixel level (Peters et al., 2009). Therefore, for assessing the uncertainties in random forest's classification, fuzzification of its results is a very good methodology to locally calculate the membership for classes (Loosvelt et al., 2012). In this way, ambiguity in the class members can be quantified for the hard classification (Loosvelt et al., 2012).

## 2.6. Variable Importance

Variable Importance is defined as "how much" a variable is used to get an accurate prediction. It is intentionally designed to rank the predictors unbiasedly (objectively, not with any experts' shadow) for the variables with high correlation. This tool is provided in the random forest as well as its more robust version is now available in the package fuzzy forest, which is an extension of random forest (Conn, Ngun, Li, & Ramirez, 2019). Variable importance addresses the problem of knowing what feature is highly important as often the features are highly correlated which can be unknown at the beginning. It helps the experts know what feature highly influences the classification results. This is a very useful tool as it reduces the error of the results being biased as it is done by experts, hence it will give a transparency amongst the stakholders when assgining importance values for decision-making. Identifying the ranks of the feature of importance has been a potential topic for extensive research in the field of machine learning, decision-making, and statistics (Conn et al., 2019). The Variable importance measures (VIMs) in the random forest provide another approach for model-based feature selection, where a model is designed to do a particular task (Breiman, 2001). This tool will give quick and unbiased results, which makes it very efficient, especially in the decision-making process.

## 2.7. Measures of Uncertainty

Uncertainty in GIS results can be called information on imperfection as the information is not a hundred percent sure (Karabegovic et al., 2006). The uncertainty reflects the ambiguousness of the mixed pixels. The area that shows a specific pattern of ambiguousness could be useful to make assumptions about the area. This study shows that the evaluation of uncertainty is an asset in land use management. Furthermore, this can be easily assessed by the probabilities obtained by the random forest algorithm (Loosvelt et al., 2012).

There are three measures of uncertainty, namely, Confusion Index, Ambiguity Index, and fuzziness (Hofmann, 2016; Loosvelt et al., 2012; Roodposhti, Aryal, Lucieer, & Bryan, 2019; Siler & Buckley, 2004). Hofmann (2016) combined the operators “AND” and “OR” while assigning the classes to the pixel, rather than the conventional method of “IF” and “THEN,” wherein we can say at least one of the conditions are fulfilled. In fuzzification, as the conditions for classifying an object are increasingly fulfilled, the membership to a certain class also increases gradually (Hofmann, 2016). Hofmann (2016) and Siler and Buckley (2004) quantified the uncertainty in the classification per object.

As Hofmann (2016) presents in his paper, the Confusion Index is a measure of similarity amongst the classes for an entity, whereas ambiguity is the measure between the possible classification result and the classification result achieved. On the other hand, Siler and Buckley (2004) recommend the evaluation of fuzziness by looking at the classes in a pixel assigned with equal membership (=0.5), which is the highest fuzziness value possible. As suggested by the above literature studied, the measures of uncertainty in the classification can help look for the gradual and slow changes as it detects the objects occuring gradually with the “AND” and “OR” operators—this way, the gradual transition of the land to water can be seen more minutely.

## 2.8. Coastal Management

The interaction of people with Indonesia's coastal zone is often more complex and needs careful monitoring (Harwitasari, 2009; Marfai & King, 2008; Suripin & Helmi, 2015). This is an issue that recently caught the government’s attention too (e.g., Program Pengentasan Kemiskinan)(Sukristijono Sukardjo & Pratiwi, 2015).

In the coastal zone, various different and highly productive ecosystems such as coral reefs, mangroves, sand dunes, and seagrasses can be found. Due to globalization, anthropogenic activities have increased at an alarming rate, especially on the coasts for trading, transport of heavy goods, water resources, and food. These activities have put a lot of pressure on coasts (Marfai & King, 2008; Nayak, 2004). Therefore, it is necessary to protect these coastal ecosystems and ensure sustainable development (Nayak, 2004; Suripin & Helmi, 2015). Ensuring this necessitates constant monitoring to obtain consistent data on ecosystems, coastal processes, and dangers of natural disasters.

### 2.8.1. Semarang's problems

For several years, coastal exploitation of the Semarang coast has led to complex problems forming a cycle of processes affecting each other, which is very challenging to break. Several stakeholders have realized that anthropogenic effects on the environment have long-term detrimental effects on the ecosystem and human activities (Hadi, 2017, 2018; Suripin & Helmi, 2015). Coastal zone management issues are exacerbated by rising demand from inside provinces with diverse socio-economic structures and from coastal communities, which are experiencing fast population development (Hadi, 2018; Harwitasari, 2009; Marfai & King, 2008).

Furthermore, the alarming rate of exploitation of land, water, and other coastal resources and degradation of coastal areas, caused disruption of environmental processes due to degrading quality of environment and loss of aquatic habitat as well as terrestrial biodiversity (Marfai & King, 2008). This has severe effects on coastal ecosystems’ health, affecting food accessibility (Hadi, 2017; Nayak, 2004); along with this, it also affects the economic development of people residing in these areas, their health, and the environment they live in.

The following are the main problems and conflicts occurring in the Indonesian coastal zone:

- falling area of mangrove forest and tidal swamp, which are benefits conventional fisheries (Hadi, 2017; Sukardjo, 1999),
- The use of the coastal areas by wealthy people is inappropriate (not coastal inhabitant - for example, conversion of mangrove forest for Tambak, a fish pond) (Hadi, 2017; S. Sukardjo, 1999),
- Loss of income and livelihood for the lowest-income citizens of coastal towns (Marfai & King, 2008),

- An increase in the number of people living along the shore (Harwitasari, 2009; Marfai & King, 2008),
- Overfishing and the employment of harmful tactics (Hadi, 2018; S. Sukardjo, 1999),
- A lack of detailed ecological data on the coastal zone/resources for developers and policymakers at the regional level (Hadi, 2017, 2018; Nurhidayah, 2019),

- Agricultural policy's importance in terms of revenue and food security (S. Sukardjo, 1999),
- The extraction of groundwater resulting in a high land subsidence rate (Suripin & Helmi, 2015),
- Increase in elevation of the houses to mitigate the flooding (Harwitasari, 2009).

These anthropogenic activities on the coasts affect the coastal ecosystem and hence aggravate the degradation of the coastal regions. These problems may be socio-economic; however, they have a direct effect on the geographical and physical processes of the coasts. The result of these activities also hinders the coastal management plan and their impact if carried out. This results in coastal and tidal flooding along with a high rate of land subsidence.

### 2.8.2. Measures taken by people to solve the problem

Citizens of Mangunhardjo and Mangkang Wetan worked with seed funding of the Government of Central Java to construct a coastal belt (Hadi, 2017). So far, a stretch of 3.2 km out of the project's 3.5 km coastal belt has been built (Hadi, 2017). For this project, additional funding from the Ministry of Fisheries and Marine and other sources has been secured (Hadi, 2017; Nurhidayah, 2019). Parallel to the coastal belt, the locals have planted mangroves for further protection (Hadi, 2017; Nurhidayah, 2019).

Hadi (2017, 2018) recommends that buildings in the region are not compatible with spatial planning as they are causing land subsidence and are prone to natural disasters; hence they should be destroyed if required to make way for green open space.

In order to deal with floods, it is important to update upstream spatial planning by performing a strategic environmental assessment (SEA) (Marfai & King, 2008; Suripin & Helmi, 2015). By evaluating the 'environmental capacity,' Surpin & Helmi (2015) believes that recommendations can be given as to which areas should be utilized for green open space.

To mitigate the frequent floods, people residing in these regions simply elevate their buildings by adding another level; due to the overburden of the buildings, the land subsidence rate is higher than the sea level rise (Hadi, 2017; Harwitasari, 2009; Sukristijono Sukardjo & Pratiwi, 2015). Therefore, the disaster risk maps for the Semarang coast are in need of revision.

For this moment, the above-stated measures have been successful in reducing the impacts (Hadi, 2017; Harwitasari, 2009); however, in the long term, additional measures will be necessary if the root causes for the issues such as overfishing, over-extraction of water, increased population, overburdening of buildings and many more are not dealt with.
Overdevelopment of the coastline necessitates revisions to the spatial planning in the area, which includes the north coast of Semarang, taking into account the tidal floods in the area (Hadi, 2017; Marfai & King, 2008).

### 2.8.3. Integrated Coastal Zone Management (ICZM)

Integrated coastal management (ICM), Integrated coastal planning, or Integrated coastal zone management (ICZM) is often used to manage the coast; it is an integrated approach attempting sustainable management practices ("Integrated Coastal Zone Management (ICZM)," 2007). The methods take into account various elements of the coastal zone, such as political and geographical boundaries (Farhan & Lim, 2010; "Integrated Coastal Zone Management (ICZM)," 2007).

ICZM promotes the sustainability of coastal management by a multidisciplinary and iterative process (European Environment Agency, 2000a; "Integrated Coastal Zone Management (ICZM)," 2007). ICZM oversees the whole lifecycle of the process, and the steps are data collection, policymaking, management, and evaluation and monitoring of the measures (European Environment Agency, 2000a; Marti et al., 2007). ICZM evaluates the societal goals in a specific coastal area and takes steps to accomplish these goals with all stakeholders' informed involvement and collaboration (Akvopedia, n.d.; Farhan & Lim, 2010; Marti et al., 2007). Overseeing the whole lifecycle means integrating all applicable sectors, administrative levels, and policy areas (European Environment Agency, 2000a; Farhan & Lim, 2010; Marti et al., 2007; S. Sukardjo, 1999). Additionally, it means integrating over space and time, terrestrial and marine components (European Environment Agency, 2000b).

ICZM is a globally recognized approach multidisciplinary approach for decision-making that is flexible to give customized solutions to the diverse range of the world's as well as unique national, regional and local coastline and coastal environments and needs (Sukristijono, 2002; Thia-Eng, 1993).

In order for ICZM to be successful, it requires a comprehensive overview of the functionality, dynamic and complex level of interactions within the coastal environment (Sukristijono, 2002; Thia-Eng, 1993). However, the approach can only be used within set geographical boundaries (this can be conflicting) if all prior named factors are integrated (Marti et al., 2007; S. Sukardjo, 1999; Thia-Eng, 1993).

The goals of ICZM are below, as defined by papers regarding ICZM (Marti et al., 2007; Thia-Eng, 1993):

- To maintain the coastal systems' resilience;
- To maintain the health of the coastal environment;
- To advance in multi-level sectoral development;
- Reducing conflicts concerining the use of resources;

#### 2.8.4. Role of remote sensing in ICZM

Mapping of coastal zones and estimation of the extent and the mutations in land cover in coastal zones have been aided to a large extent by satellite imagery (Ghosh, Kumar, & Roy, 2015; Nayak, 2004; Tran & Tran, 2009). The following are the technical problems and governance hindrance with regards to coastal management derived from the literature (Appeaning Addo et al., 2008; Esmail, Mahmod, & Fath, 2019; Gens, 2010; Hadi, 2018; Marfai & King, 2008; Nayak, 2004; Nurhidayah, 2010; Toure et al., 2019; Zhang et al., 2013):

1. Coastal Ecosystems
    a. Baseline data, conservation of coastal habitats such as mangroves, coral reefs, etc.
    b. Wetland protection and reclamation for agriculture or industrial use.
    c. Use of resources sustainably

2. Coastline protection
    a. Identifying susceptible regions, such as eroded areas and development projects
    b. Coastal protection action plan and its implementation (erosion, flood protection, saltwater intrusion, etc.)
    c. Engineering constructions and dam's impact assessment on coastal erosion, and sedimentation processes.
    d. Studying the changes in the bottom topography

3. Coastal Hazards and Climate Change
    a. Sea-level rise and its impact
    b. Disaster risk management for natural disasters, including cyclones and sea-level rise, as well as human actions such as oil spills
4. Coastal development
    a. Appropriate site selection for recreational activities, landfall points, industries, aquaculture, etc.

b. Evaluation of circumstances in regulation zones, development setback–lines, megacities, and other locations.

An integrated approach has better results or impact than a non-integrated approach with respect to improving the coastal quality follows from studies (Farhan & Lim, 2010; Harwitasari, 2009; Nurhidayah, 2019). In coastal management, multi-level stakeholders should be chosen at a regional scale over a national scale and better focus on users than on uses (Farhan & Lim, 2010; Hadi, 2018; Harwitasari, 2009).
According to Aheto et al. (2016), an approach of co-management should be taken towards effective coastal management wherein the government shares some responsibilities, duties, and tasks with citizens together. Because of their proximity to the resource management regions, local government entities should be involved in rule-making and enforcement (Aheto et al., 2016).

One of the rare coastal areas in Semarang City is Mangunharjo's fully grown mangrove, which is now called a mangrove forest (Hadi, 2018). It is necessary to designate places as conservation zones through spatial planning revision, or else the land will be used by the private sector as an industrial zone (Hadi, 2018; Sukristijono, 2002).

#### 2.8.5. ICZM indicators

In order to operate ICZM, a number of indicators to monitor the quality of the implementation and the level of sustainable coastal development have been developed by a European ICZM expert group (European Environment Agency, 2000a; Marti et al., 2007). This group consists of the 20 member states and two candidate member states with a coastline (Marti et al., 2007). The recognized criteria and indicators can be split into two categories (Marti et al., 2007):

- An indicator is used to track the progress of ICZM implementation (progress indicators).
- 27 indicators (consisting of 44 measures) measuring the sustainable development along the coastline
-

The indicators are split into seven areas based on the EU ICZM Recommendation's seven aims (Marti et al., 2007), it is represented in figure 4. The indicators in each group, when combined, will aid the European Commission, the Member States, and coastal partnerships in tracking progress toward the EU Recommendation's goals for coastal sustainability (European Environment Agency, 2000a; Marti et al., 2007). This research will look into the possibility of its contribution to the assessment of ICZM indicators 25, 26, and 27.

| GOALS | INDICATORS | MEASUREMENTS |
|---|---|---|
| To control further development of the undeveloped coast as appropriate. | 1. DEMAND FOR PROPERTY ON THE COAST | 1.1. Size, density and proportion of the population living on the coast |
| | | 1.2. Value of residential property |
| | 2. AREA OF BUILT-UP LAND | 2.1. Percentage of built-up land by distance from the coastline |
| | 3. RATE OF DEVELOPMENT OF PREVIOUSLY UNDEVELOPED LAND | 3.1. Area converted from non-developed to devel oped land uses |
| | 4. DEMAND FOR ROAD TRAVEL ON THE COAST | 4.1. Volume of traffic on coastal motorways and major roads |
| | 5. PRESSURE FOR COASTAL AND MARINE RECREATION | 5.1. Number of berths and moorings for recreational boating |
| | 6. LAND TAKEN UP BY INTENSIVE AGRICULTURE | 6.1. Proportion of agricultural land farmed intensively |
| To protect, enhance and celebrate natural and cultural diversity. | 7. AMOUNT OF SEMI-NATURAL HABITAT | 7.1. Area of semi-natural habitat |
| | 8. AREA OF LAND AND SEA PROTECTED BY STATUTORY DESIGNATIONS | 8.1. Area protected for nature conservation, land scape and heritage |
| | 9. EFFECTIVE MANAGEMENT OF DESIGNATED SITES | 9.1. Rate of loss of or damage to, protected areas |
| | 10. CHANGE IN SIGNIFICANCE COASTAL AND MARINE HABITATS AND SPECIES | 10.1. Status and trend of specified habitats and species |
| | | 10.2. Number of species per habitat type |
| | | 10.3. Number of Red List coastal area species |
| To promote and sup-port a dynamic and sustainable coastal economy. | 11. LOSS OF CULTURAL DISTINCTIVENESS | 11.1. Number and value of sales of local products with regional quality labels or European PDO/PGI/TSG |
| | 12. PATTERNS OF SECTORAL EMPLOYMENT | 12.1. Full time, part time and seasonal employment per sector |
| | | 12.2. Value added per sector |
| | 13. VOLUME OF PORT TRAFFIC | 13.1. Number of incoming and outgoing passengers per port |
| | | 13.2. Total volume of goods handled per port |
| | | 13.3. Proportion of goods carried by short sea routes |
| | 14. INTENSITY OF TOURISM | 14.1. Number of overnight stays in tourist accommodation |
| | | 14.2. Occupancy rate of bed places |
| | 15. SUSTAINABLE TOURISM | 15.1. Number of tourist accommodation units holding EU Eco-label |
| | | 15.2. Ratio of overnight stays to number of residents |
| To ensure that beaches are clean and that coastal waters are unpolluted. | 16. QUALITY OF BATHING WATER | 16.1. Percentage of bathing waters compliant with the guide value of the European Bathing Water Directive |
| | 17. AMOUNT OF COASTAL, ESTUARINE AND MARINE LITTER | 17.1. Volume of litter collected per given length of shoreline |
| | 18. CONCENTRATION OF NUTRIENTS IN COASTAL WATERS | 18.1. Riverine and direct inputs of nitrogen and phosphorus in inshore waters |
| | 19. AMOUNT OF OIL POLLUTION | 19.1. Volume of accidental oil spills |
| | | 19.2. Number of observed oil slicks from aerial surveillance |
| To reduce social exclu-sion and promote social cohesion in coastal. | 20. DEGREE OF SOCIAL COHESION | 20.1. Indices of social exclusion by area |
| | 21. RELATIVE HOUSEHOLD PROSPERITY | 21.1. Average household income |
| | | 21.2. Percentage of population with a higher education qualification |
| | 22. SECOND AND HOLIDAY HOMES | 22.1. Ratio of first to second and holiday homes |
| To use natural resources wisely. | 23. FISH STOCKS AND FISH LANDINGS | 23.1. State of the main fish stocks by species and sea area |
| | | 23.2. Recruitment and spawning stock biomass by species |
| | | 23.3. Landings and fish mortality by species |
| | | 23.4. Value of landings by port and species |
| | 24. WATER CONSUMPTION | 24.1. Number of days of reduced supply |
| To recognise the threat to coastal zones posed by climate change and to ensure appropriate and ecologically responsible coastal protection. | 25. SEA LEVEL RISE AND EXTREME WEATHER CONDITIONS | 25.1. Number of 'stormy days' |
| | | 25.2. Rise in sea level relative to land |
| | | 25.3. Length of protected and defended coastline |
| | 26. COASTAL EROSION AND ACCRETION | 26.1. Length of dynamic coastline |
| | | 26.2. Area and volume of sand nourishment |
| | | 26.3. Number of people living within an 'at risk' zone |
| | 27. NATURAL, HUMAN AND ECONOMIC ASSETS AT RISK | 27.1. Area of protected sites within an 'at risk' zone |
| | | 27.2. Value of economic assets within an 'at risk' zone |

Figure 4 ICZM indicators for DEDUCE project (Marti et al., 2007)

# 3. MATERIALS AND METHODOLOGY

## 3.1. Study Area

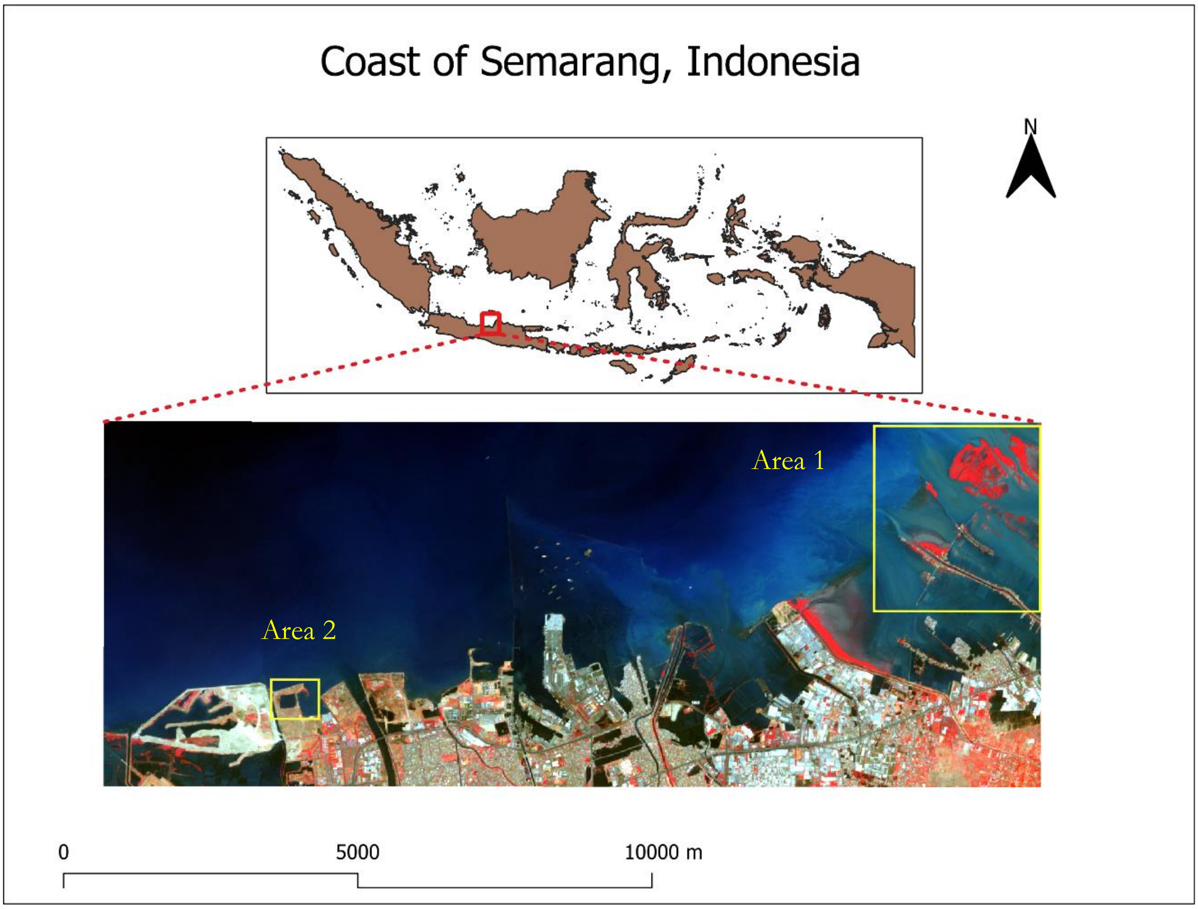



Figure 5 Semarang coast as Study area with the area to be studied for uncertainty

Semarang is one of the biggest cities of Indonesia; as summarized by Marfai (2008), a very severe impact of coastal erosion in Semarang can be seen in its land-use pattern, ecology, and the infrastructure of the low elevation areas. This study furthermore addressed the other threat of tidal flooding due to the significant subsidence of land faced by Semarang besides sea-level rise. According to (Suripin & Helmi, 2015), the sea-level data recorded from 1985 to 2008 at Tanjung Emas Harbor, Semarang, indicates that the sea level rises 3.7mm per year; furthermore, the interpretation of satellite imagery from 1991 to 2009 is that 25.6km of the total coastline of 36.6 km is eroded due to sea-level rise and land subsidence. The erosion has removed 1,764.5 ha of the coastal plain. In this context, learning some facts about the natural and human-induced calamities through the records of coastal erosion until this date of the areas of interest can be of great significance, which can help us understand the dynamic areas on the coastline and the areas of uncertainty. Therefore, figure 3 shows the study area which will be used for this study. Two subsets of this area are select to see the slow and gradual changes properly and at a local level. These areas are shown in the above figure in the yellow box.

## 3.2. Data

This study will use the Sentinel-2 data. The Sentinel-2 data for this study has been downloaded from the USGS Earth Explorer platform. Sentinel -2 is a mission by the European Space Agency (ESA), especially to monitor changes in the land surface. This mission consists of twin satellites orbiting at a phase of 180$^{o}$ in the same orbit at a mean altitude of 786 km, which gives us more frequent data with the re-visit time of 5 days at the Equator (ESA, 2015). The orbital swath of these satellites is 290 km (ESA, 2015). According to the user handbook of ESA, the twin satellites will give high temporal resolution like SPOT and spatial resolution like LANDSAT. The sentinel-2 orbit is Sun-synchronous which ensures that the angle of sunlight is the same all the time, which minimizes the shadow errors and light on the ground (ESA, 2015). This shows that the data acquired is always consistent and is very useful for the analysis of time-series data. The geographical coverage of the twin satellites extends between the latitude 56$^{o}$ South to 83$^{o}$ North (ESA, 2015).

The Sentinel-2 provides data of following resolutions (ESA, 2015):

- The temporal resolution of the satellite is the revisit frequency at a specific location. The revisit period of each satellite is ten days; however, with the combination of two satellites, the revisit period is five days.
- The spatial resolution is a satellite is a ground represented by an array of the sensor. The data is available in 10m, 20m, and 60m resolution.
- The spectral resolution is the ability of the sensors to distinguish the reflectance. The sentinel data is obtained on 13 spectral bands in visible and near-infrared (VNIR) and short-wave infrared (SWIR), as indicated in table 2.

Table 2 Spectral bands of Sentinel-2 images (ESA, 2015)

| Bands | Name of the band | Resolution (m) |
|---|---|---|
| **Band 1** | Coastal aerosol | 60 |
| **Band 2** | Blue | 10 |
| **Band 3** | Green | 10 |
| **Band 4** | Red | 10 |
| **Band 5** | Vegetation red edge (VNIR) | 20 |
| **Band 6** | Vegetation red edge | 20 |
| **Band 7** | Vegetation red edge | 20 |

| | | |
|---|---|---|
| **Band 8** | NIR | 10 |
| **Band 8A** | Narrow NIR | 20 |
| **Band 9** | Water vapour | 60 |
| **Band 10** | SWIR – Cirrus | 60 |
| **Band 11** | SWIR | 20 |
| **Band 12** | SWIR | 20 |

### 3.3. Software

The creation of training dataset and subsetting is done in QGIS software, whereas, for calculation of indices and stacking of the raster is done through Spatial modelling in ERDAS imagine. The ideas for the script are taken from Beyer (2018).

### 3.4. Tidal data

The tidal data is acquired from the Bandan Informasi Geospatial, Indonesia ("Tide prediction," 2021). The Sentinel-2 satellite passes Indonesia approximately 02:00- 3:00 h GMT. For the period of October 2015- August 2020, the tidal data was manually extracted for each of the scenes with respect to the time of acquisition. Further, this data is plotted to see the low tides and high tides. Figure 6 shows the tidal information for all 82 scenes, and it shows the cyclic pattern of the tides.

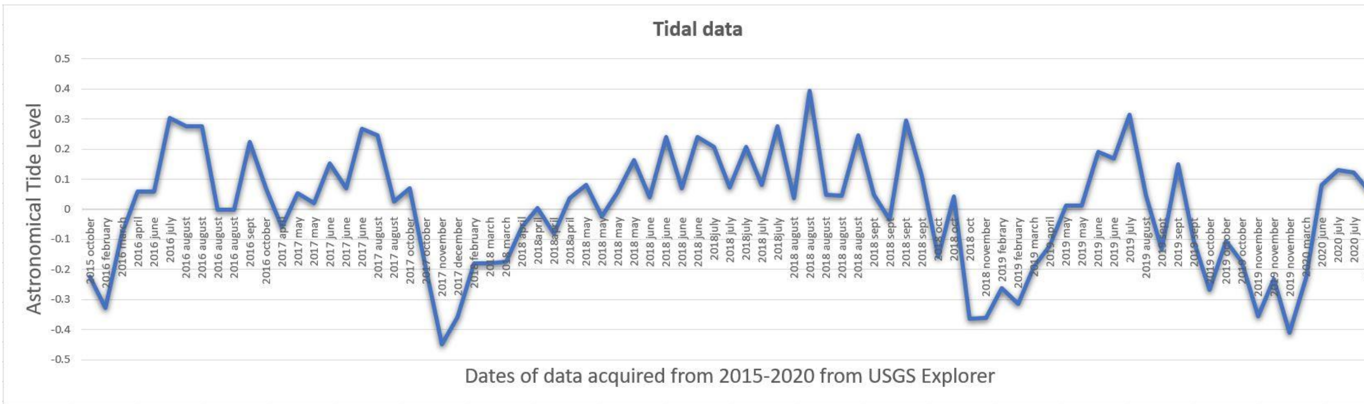



Figure 6 Graph of Astronomical tidal level (m) of Semarang coast from October 2015- August 2020 ("Tide prediction," 2021)

For the analysis, the scenarios with low tidal data are selected to see the actual changes in the coastal without the influence of the tides. The below is table 3.1. of the dates and the tide levels were chosen for further analysis.

Table 3 Tidal elevation in meters for respective acquisition date

| Acquisition Date | Astronomical Tide level (m) |
|---|---|
| **07-10-2015** | -0.223 |
| **05-01-2016** | -0.266 |
| **12-08-2016** | -0.001 |
| **09-04-2017** | -0.058 |
| **10-11-2017** | -0.448 |
| **28-02-2018** | -0.179 |
| **31-10-2018** | -0.363 |
| **23-02-2019** | -0.314 |
| **05-11-2019** | -0.356 |
| **14-03-2020** | -0.228 |
| **11-08-2020** | -0.058 |

## 3.5. Methodology

In figure 7 flowchart represents the workflow of the study, indicating the phases a particular process lies in, which makes is it easier to understand. The methodology is divided into four phases,

a. Data – in this phase, the data is acquired, such as satellite images, elevation data, and tidal data. The training dataset for every year is also prepared. This phase also consists of using the relevant spectral bands and stacking them, calculating indices. Calculating slope and aspect from the elevation data
b. Analysis – Stacking the prepared datasets and processing them through the random forest classifier comes under the analysis phase
c. Results- this phase consists of the fuzzified results of the random forest classifier and the uncertainty maps obtained from them. Variable importance can be calculated to understand what parameters influence the classification the most.
d. Application – to link the results of this study for their application for assessing indicators of ICZM

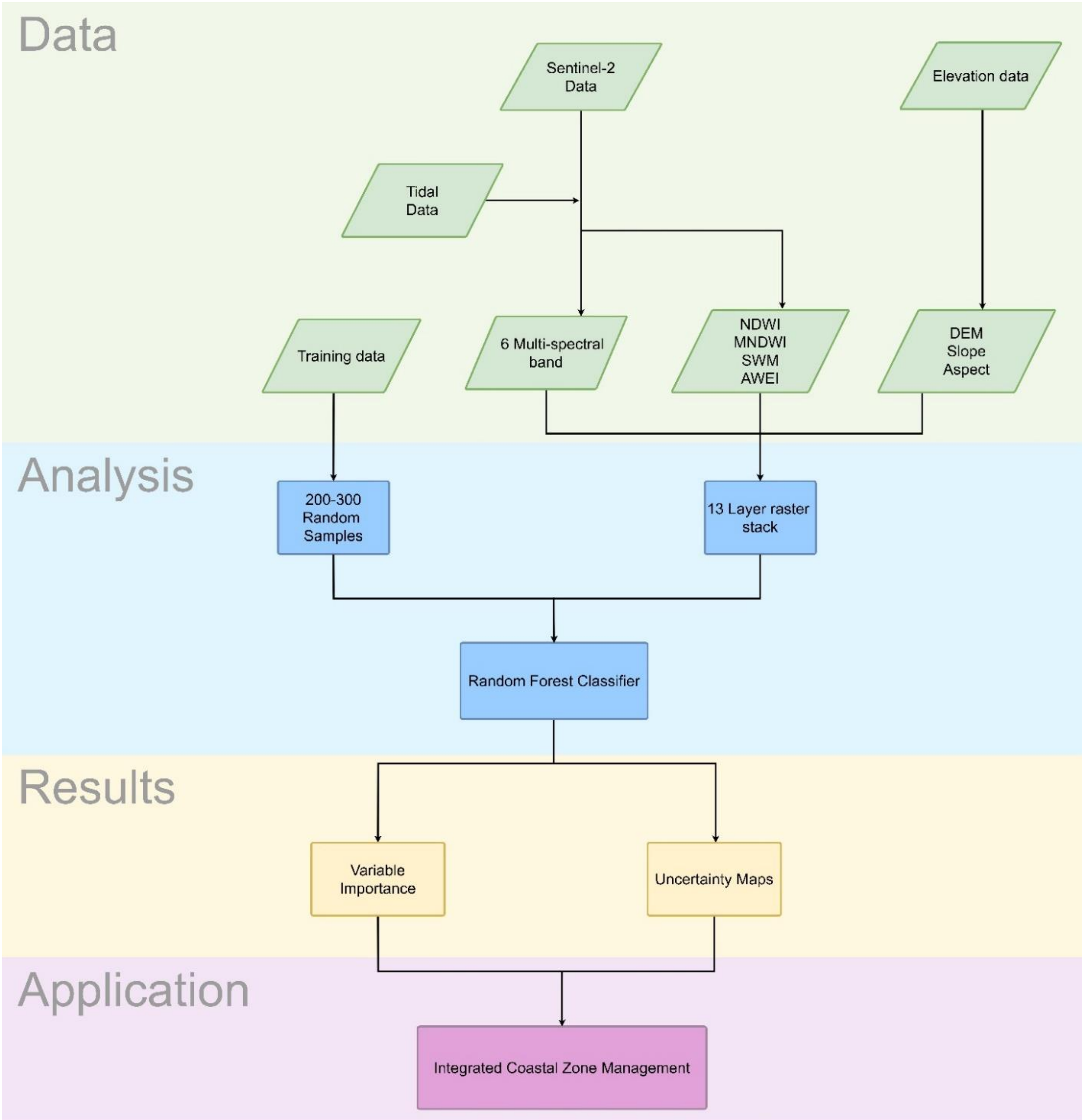



Figure 7 Methodology of the research

For the analysis, six spectral bands were used as input parameters, namely, red, blue, green, near-infrared band short wave infrared 1, and short wave infrared 2, which have different resolutions. To address this, the bands with different resolutions, SWIR1 and SWIR2, are resampled to 10m resolution with Bilinear Interpolation; as this algorithm, the value of the pixel is calculated using the distance-weighted value of the four nearest pixels. Along with this, the indices used for extracting shoreline boundary effectively are used as input parameters; namely, NDWI, MNDWI, SWM, and $AWEI_{nsh}$ were calculated in ERDAS with the Spatial Model Editor with the ratios mentioned in Chapter 2. NDWI, defined by Mcfeeters (1996), is used to observe changes in water content in water bodies. MNDWI is used to efficiently extract water areas from the settlement, vegetation, and soil noise as a prominent background which is very well the case of the Semarang coast (Xu, 2006). SWM is selected as it is very effective in the detection of water, especially in flood assessment (Milczarek et al., 2017). Lastly, $AWEI_{nsh}$ is useful in surface water mapping, which reduces the classification noise from shadows and non-water dark surfaces (Feyisa et al., 2014). This can be the case with Semarang, as the shadow of the tall buildings and roofs can cause misclassification. The model of every index was created in the spatial model editor, and the following figure 8 is the instance for the NDWI index.

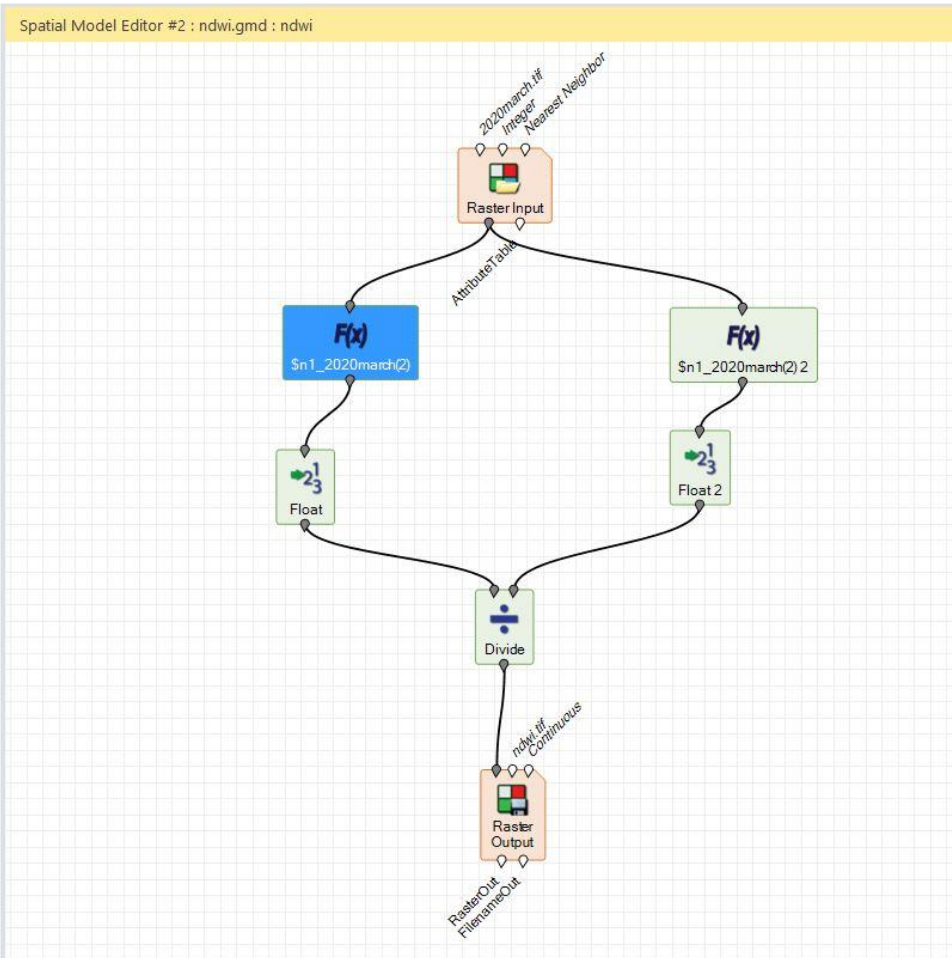



Figure 8 The NDWI index model in the spatial model editor in ERDAS

The elevation data used in this study is acquired from the national elevation data of Indonesia DEMAS (BIG Indonesia, n.d.) of 7.5m resolution and further reprojected to WGS 84 49S and resampled to 10m using Bilinear Interpolation as all the other data is in the same resolution. Slope expressed in degrees and aspect were calculated from DEM in QGIS. Finally, the raster stack was created with 13 layers consisting of six spectral bands, four indices, DEM, slope, and aspect. The point data of 300-400 samples are created as training data for every year in QGIS, classifying the points for four classes land, water, vegetation, and settlement. The script of random forest is developed in R, where the RF uses the rasters, and the training datasets are used as input data for the random forest classifier.

`

### 3.5.1. Random Forest (RF) Classifier

RF uses a bagging technique in which classifiers are constructed with a different subset of features or attributes of the original dataset; however, the bagging in the RF is done by attributes selected randomly (Breiman, 2001). For every node of the trees of random forest subset of attributes is chosen, and the best category from that subset is selected for the splitting of the tree further (Breiman, 2001). Random selection creates diversity in trees and improves overall accuracy (Belgiu & Drăgut, 2016). After bagging, about 1/3 of the samples of every tree's training dataset is left out, and they are called "out of bag (OOB)" that will act as a testing sample for the tree (Bonissone, Cadenas, Garrido, & Díaz-Valladares, 2010; Breiman, 2001; Olaru & Wehenkel, 2003). The remaining 2/3 of examples are used to build the tree, "Inbag samples" where one sample can be used more than one time in a single tree (Bonissone et al., 2010; Breiman, 2001; Olaru & Wehenkel, 2003). This entire process is known as bootstrapping (Bonissone et al., 2010; Breiman, 2001; Olaru & Wehenkel, 2003). The samples are run through the decision trees, and from all the decision tree results, the maximum voted result is assigned as the final class.

Further, to tune the RF, an $m_{try}$ with minimum OOB error is chosen for the random forest algorithm. The $m_{try}$ is the number of variables used for splitting the tree nodes. The tuning of the RF help in selecting the most optimum RF model corresponding to its mtry. Figure 9 is the graphic representation of the mtry with respect to the OOB error (explained earlier). Thus for this model mtry=3 shows the minimum OOB error, which infers that the RF model with mtry=3 will give the most optimal results.

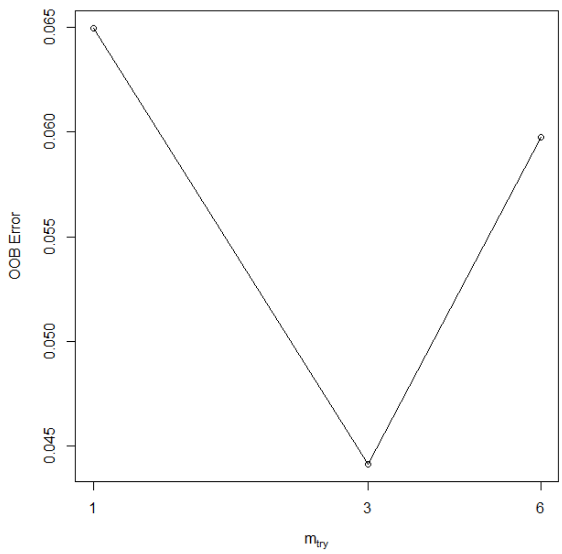



Figure 9 OOB error of the RF model for respective mtry

### 3.5.2. Variable Importance

Variable Importance measure is the function in the "randomForest" package in R which reflects which attribute was most used to get the pure classes. It generally shows the relative importance of all the attributes with respect to each other. The Variable importance feature is important to see which parameters are of higher significance for splitting classes during classification. This also helps in determining the outliers, i.e., the factors which affect the classification the least and can be removed. This also settles the dispute amongst the stakeholders which aspects will have more value in case of the decision-making process. Therefore, the variable importance can be very useful in classification to see which factors affect the most while classifying the area of interest. Following figure 10 is the example of the variable importance for one of the image analysis.

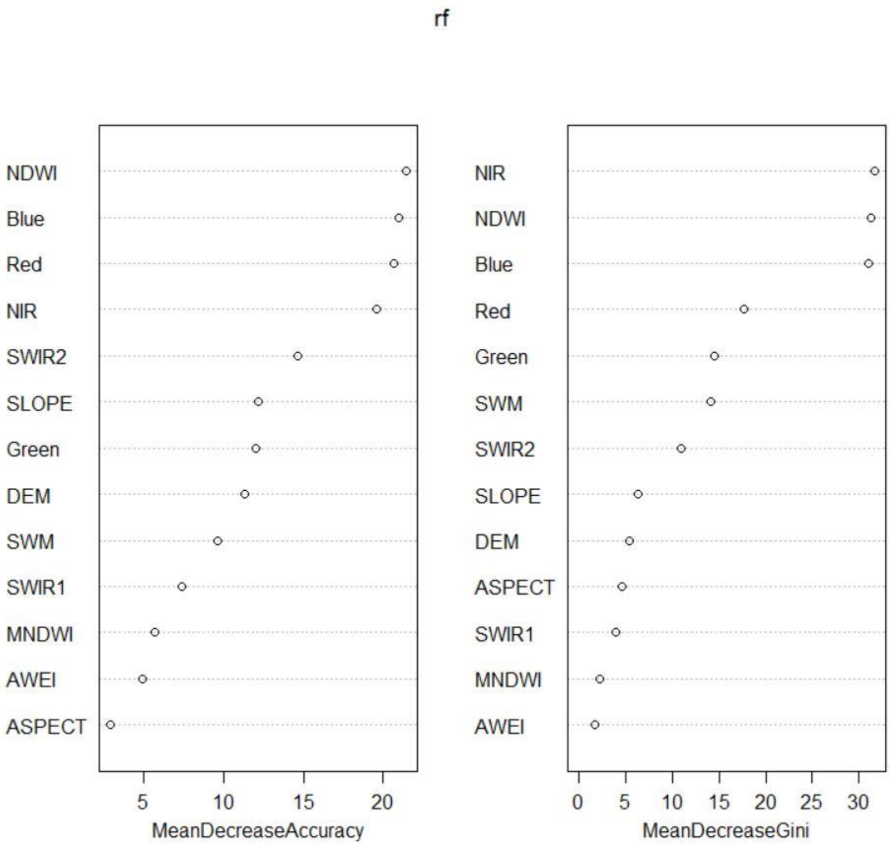



Figure 10 Variable Importance of the input parameters used in RF model

In figure 9, the variable importance is represented in two manners; the features are plotted with respect to Mean decrease accuracy and Mean decrease Gini. The mean decrease accuracy plot represents how much accuracy is lost by excluding the corresponding input variable. The more accuracy is lost, the more important the variable is considered for the successful classification. Thus, the higher the mean decrease accuracy, the higher is the importance of the variable. On the other hand, Mean decrease Gini is the mean of all the parameter's overall decrease in the impurity of the node, weighed by the number of samples reaching that node for a certain decision tree in the RF model. This measures how important a parameter contributes towards the purity (homogeneity) of nodes and leaves in the RF model. A higher MeanDecreaseGini is inferred in the higher importance of a variable.

### 3.5.3. Fuzzification of Random forest results

The two study areas are processed through the RF model, and the probabilities of classes a pixel can belong to for each pixel are determined. The mixed pixels could be represented by the degree of membership for every class. As mentioned in the literature (Baake, 2018; Khatami et al., 2017), the probabilities between 0 and 1 are obtained from random forest results can be considered as class memberships or degrees of membership. These results can be seen in a fuzzy logic concept, where the pixel rather than having one class assigned is expressed in membership degree for every class. This is another way to use the results from RF for fuzzy applications. The sum of all the class membership degrees is 1. However, for calculation purposes, only the highest and the second-highest membership degree are considered, namely, best and runner-up class, respectively. For each area, the membership function of the prominent class is examined, which shows how the class membership of the study areas is changing in space and time over the years. The best class and the runner-up classes are then extracted by ranking them. However, for calculation purposes, the best class and runner-up class membership degree values are normalized, and further, the measures of uncertainty are calculated in the context of best and runner-up classes.

### 3.5.4. Normalized fuzzy membership to the best and runner up classes

As the mixed pixels now have two class memberships, namely best class and runner-up class. The sum of the membership values should be equal to unity. Therefore, the following equation is used for normalization as the membership of best and runner-up classes do not add up to absolute 1.

$$\mu_n = \frac{\mu}{\mu_0 + \mu_i} \tag{1}$$

Where $\mu_n$ is the normalized fuzzy membership value, $\mu_0$ is the membership degree of the best class, and $\mu_i$ is the membership degree of the runner-up class.

### 3.5.5. Confusion Index

As Hofmann (2016) suggests in his paper, the confusion Index represents the possibility of the similarity between the classes (Hofmann, 2016). In other words, it represents the possibility of classes getting confused in fuzzy classification. CI value 0 is that there is no similarity between best and runner-up classes.

$$CI = 1 - (\mu_0 - \mu_I) \tag{2}$$

where $\mu_0$ is the membership degree of the best class and $\mu_i$ is the membership degree of the runner-up class.

### 3.5.6. Ambiguity Index

The ambiguity (Hofmann, 2016; Siler & Buckley, 2004) Index represents the uncertainty of the membership of the best class. In other words, it represents the ambiguity of the membership between the best and runner-up classes. The AI has the range from 0 to 1, 0 being the value of less uncertainty and one being of high uncertainty.

$$AI = 1 - \mu_0 \tag{3}$$

### 3.5.7. Fuzziness

Fuzziness (Siler&Buckley,2004) is known as to what extent the fuzzy set is not crisp. Fuzziness is higher when there is more than one class with 0.5 membership in a pixel, whereas fuzziness decreases if the memberships are closer to 0 or 1 (Siler&Buckley,2004). Therefore, if the member for a certain pixel has two similar membership of 0.5, then the fuzziness will be highest, that is 2. The fuzziness ranges from 0 to 1.

$$Fuzz = \sum_{i=0} (1-|(2\,\mu_I - 1)|) \tag{4}$$

### 3.5.8. ICZM

European Commission (European Environmental Agency, 2000) defines ICZM as an integrated iterative process that covers all the processes from data collection, spatial planning, decision making, managing as well as monitoring of the action plans. Therefore, it can be said that ICZM is the most efficient tool for sustainable coastal management and monitoring. This study attempts to connect the interdisciplinary bridge between the quantitative analysis done so far and the coastal management tool, ICZM. The results obtained by the fuzzy random forest model are then explored to determine their contribution in quantifying the measures for ICZM indicators. The qualitative study will help in determining how the results of fuzzy random forest can contribute to quantifying the measure laid by ICZM to fulfill its objectives through indicators. The qualitative analysis will further show how the results of this study will fit the ICZM approach for the sustainable coastal management of the 'wicked' Semarang coast.

# 4. RESULTS

This section presents the results of the FoRF. The results are shown corresponding to the selected two test areas showing the classification, the degree of membership of the classes, and the uncertainty considered by the model. The results are presented corresponding to the test areas; therefore, all the results are present for area 1 and then area 2 for the reader's convenience. The results are shown from 2015 to 2020, for the first half of the year at the top and the second half of the year at the bottom for the respective year. The results presented include false-color images of test areas, the membership degree map of vegetation for study areas 1 and 2, normalized fuzzy membership to best and runner up class, confusion index, ambiguity index, and fuzziness. To see more clear legends and enlarged maps see annex B.

## 4.1. False-color composition

Figure 11 and 12 shows the false-color composition (FCC) with the bands 8, band 4 and band 3 (refer Chapter 3, data) for the two study areas. The red colour in the image represents the presence of vegetation, and the water is presented by shades of blue. A decrease in the land surface covered with vegetation can be seen visually. The small patch of land seems to disappear over the period indicated by the circle.

### a) Area 1

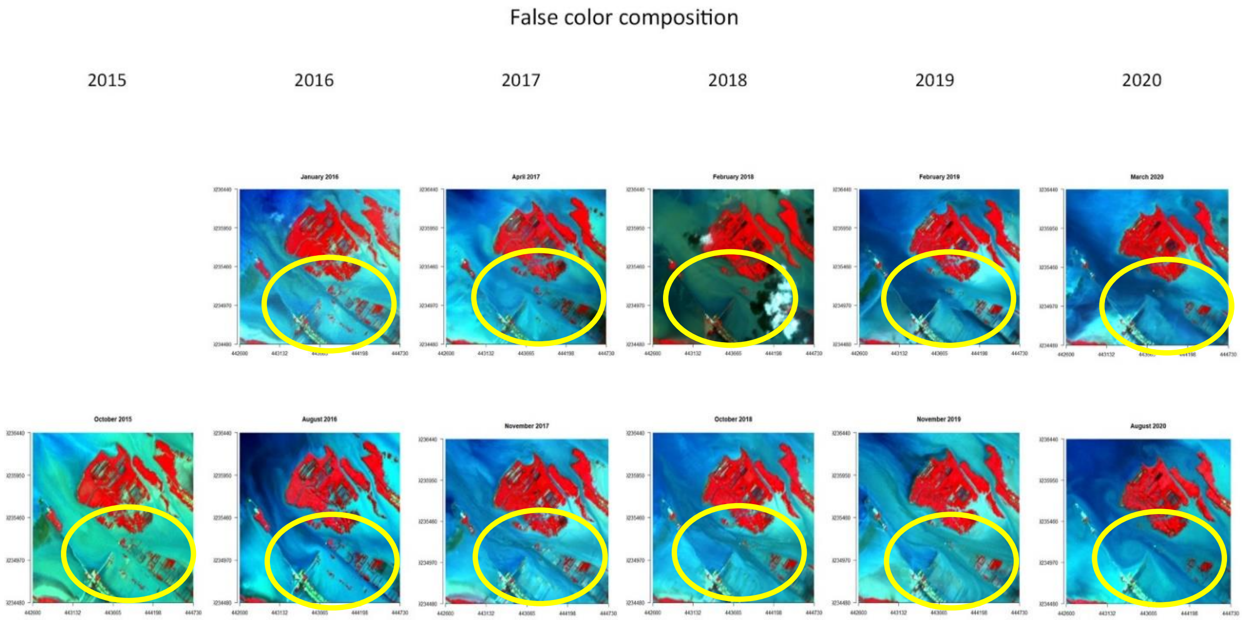



Figure 11 False colour combination of the images of study area 1 from 2015-2020 and indicated disaapearing pieces of land

b) **Area 2**

Figure 12 represents the FCC of area 2, where vegetation, water, and land (barren or settlement) can be seen. The inland water body is observed to gradually increase over the period at the coast of land. The gradual decrease of land is shown in figure 12, indicated by the circle. The vegetation is observed to increase in this area over time as well.

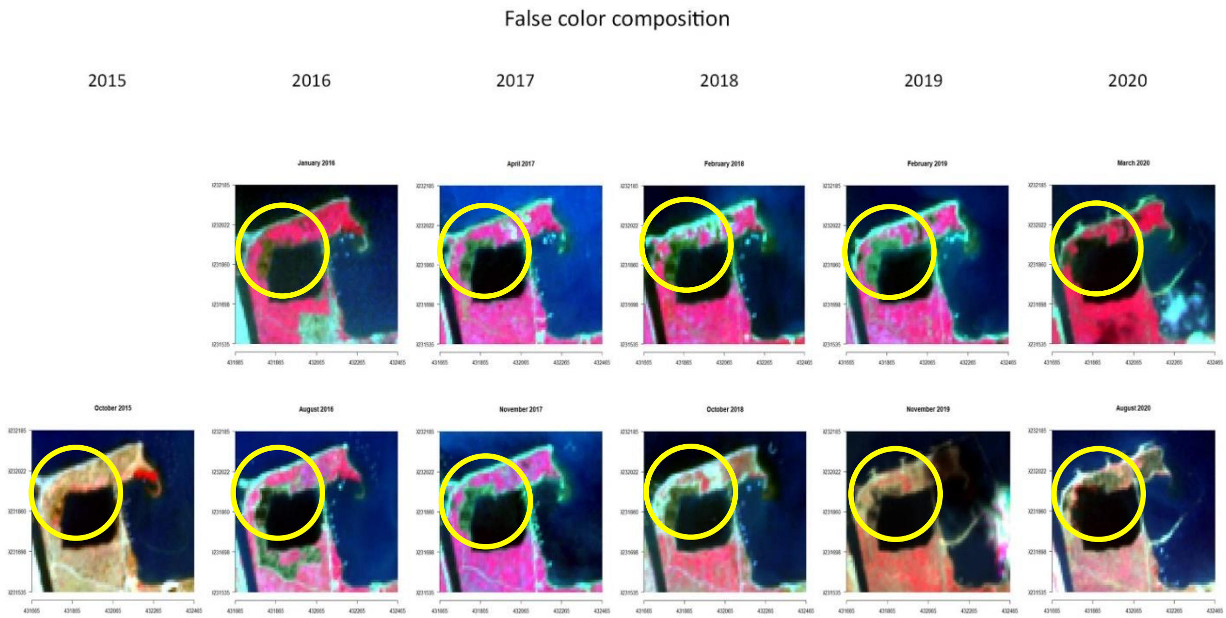



Figure 12 False colour combination of the images of study area 2 from 2015-2020 and indicated loss of land

## 4.2. Membership degree map of the dominant class

a) **Area 1**

This map represents the membership degree of every pixel with respect to the vegetation class. Based on the result of the image of 2015 (the first image studied), the dominant class appears to be vegetation. In order to study the gradual coastal changes in area 1, the same class has been chosen for the images in the following years. The membership degree of vegetation varies from 0 to 1, where one represented by the red colour indicates the full membership of the class vegetation and 0.5 represented by the yellow colour indicates 50% membership of the vegetation class. Moreover, the value 0 is represented by the colour blue, which indicates the zero membership of the vegetation class. In figure 13, the vegetation membership fluctuates in the area from 1 to 0.5 the overall surface area of the land can be seen decreasing.

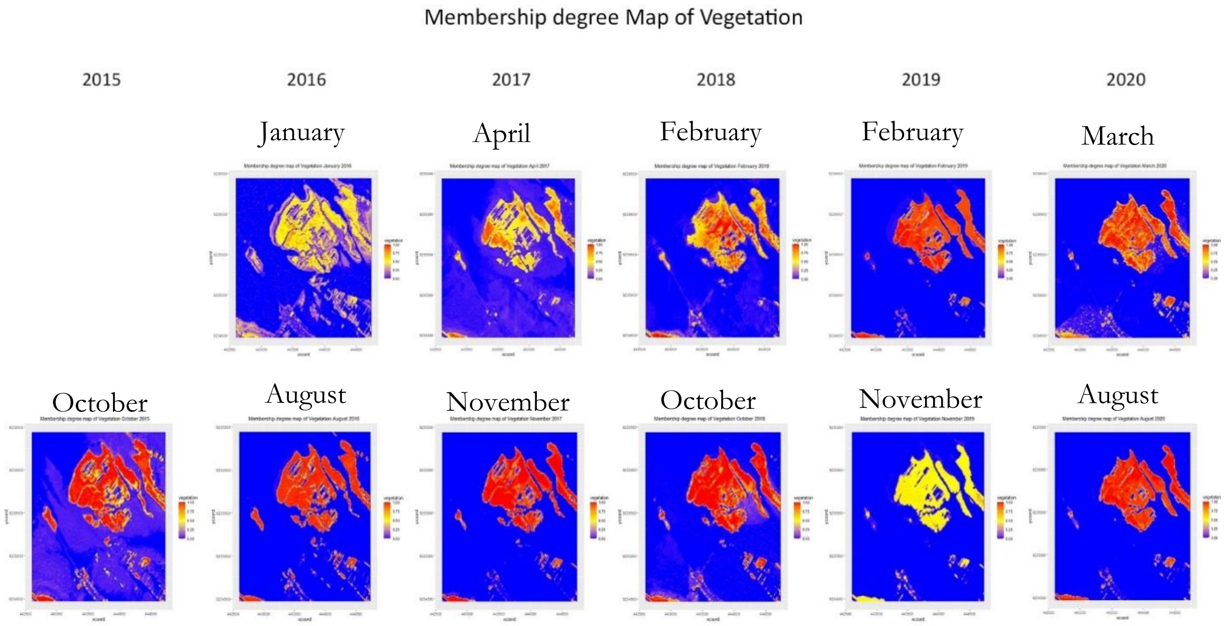



Figure 13 Map of membership degree of vegetation in area 1

### b) Area 2

Similar to area 1, the class membership studied for area 2 for land class. Therefore, in figure 14, the membership of the land varies from 0 to 1. Zero is represented by the colour yellow, which indicates no membership of the land class, whereas one represented by the colour orange represents full membership of the land class. This can be seen in figure 14, where the membership degree of land reflects the change in the land use of the area over the years. The land membership of the area drastically changes between the years 2015 and 2016. The area with high membership of land is suddenly changed into no membership of land. However, the area at the transition represents full land membership which indicates a clear boundary of the land. From late 2019 to early 2020, the membership of land is seen where there was water before.

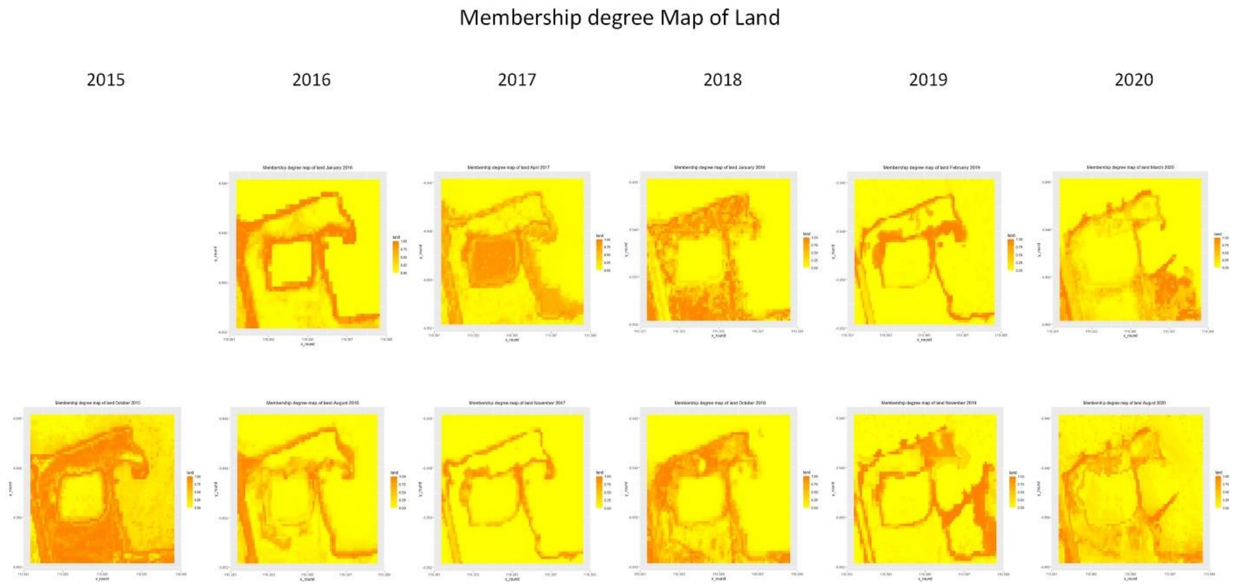



Figure 14 Map of membership degree of Land in area 2

### 4.3. Normalized fuzzy membership to the best and runner up classes

**a) Area 1**

The normalized fuzzy membership compares the relative membership degree between the best and the runner-up class, where best class means the class with the highest membership value for a pixel and the runner-up class means the second-highest class membership value. The value of the membership of best and runner-up classes is normalized, eliminating other class membership values. Therefore, each pixel is now and further in this study assessed in the context of best and runner-up class.

Fig 15 depicts the spatial distribution of the normalized fuzzy membership to the best class for area 1. The range of the normalized fuzzy membership to the best class is from 0 to 1, where the red color represents the maximum fuzzy membership degree to the best class. The yellow colour represents the 50% fuzzy membership degree to the best class. The same is represented in figure 16 for area 2. The fuzzy membership to the best class fluctuates throughout the period; however, the fuzzy membership of the pixel in the transition zone, mixed pixel, represents the 50% of fuzzy membership, which indicates the boundary of the shore.

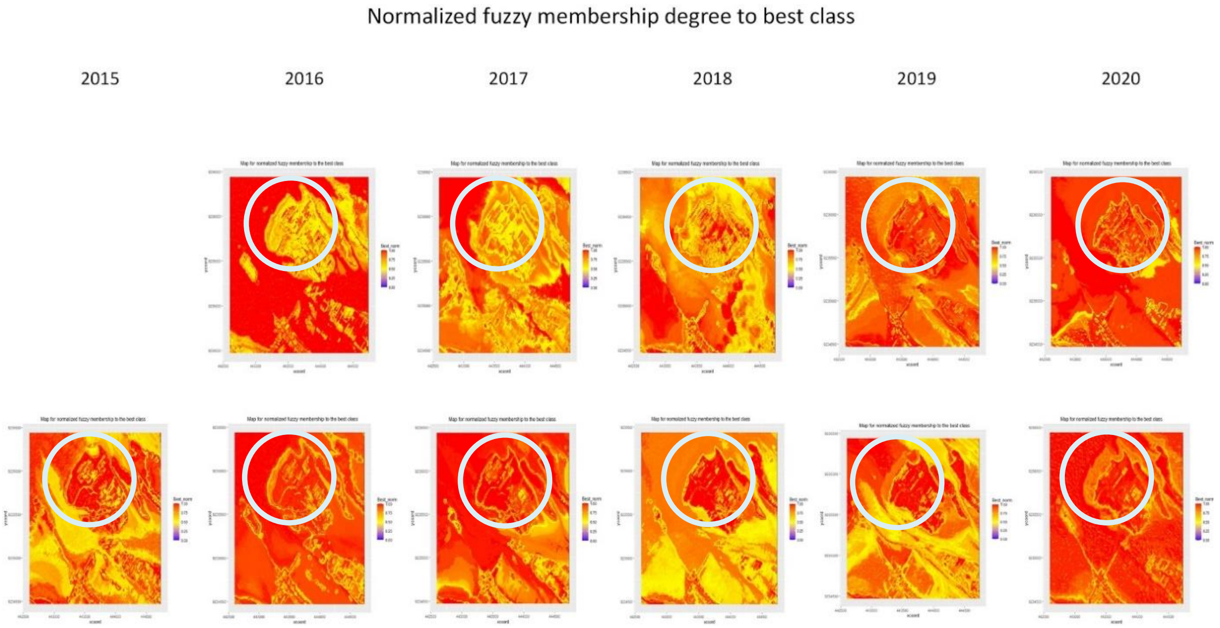



Figure 15 Map of Normalized fuzzy membership degree to best class of area 1 and indication of changing Fuzzy membership degree around the shoreline over the years

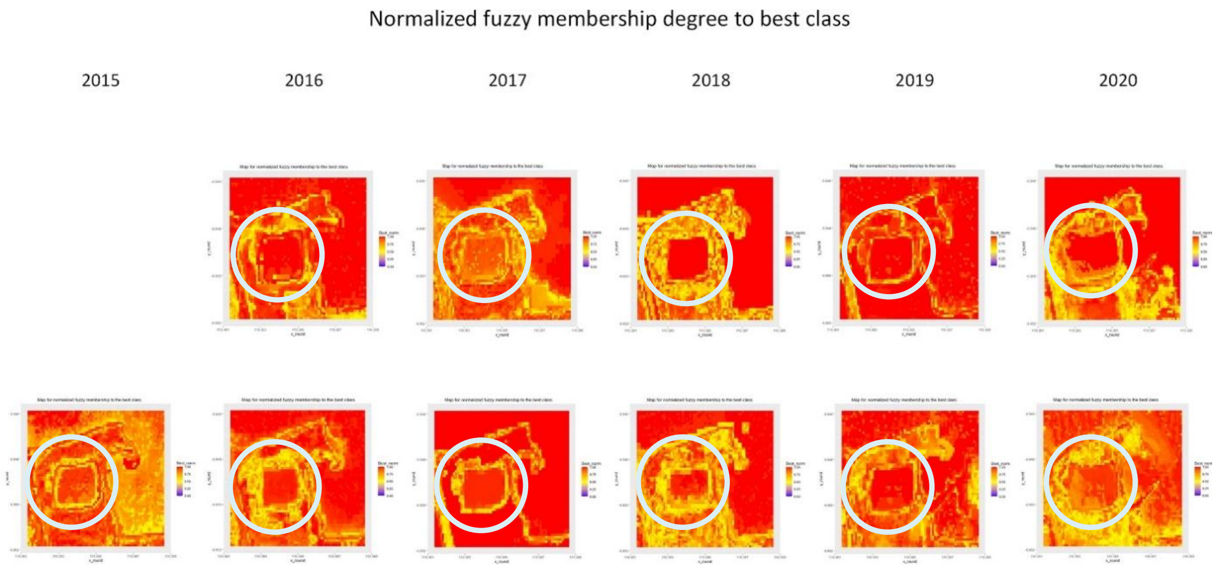



Figure 16 Map of Normalized fuzzy membership degree to best class of area 2 and indication of changing Fuzzy membership degree around the shoreline over the years

Figures 17 and 18 represent the normalized fuzzy membership degree to the runner-up class for area 1 and area 2, respectively. It reflects the inverse results as of Figure 15 and 16 for the area 1 and 2 respectively, the difference between the spatial pattern of normalized fuzzy membership degree to the best class and runner up class is that the areas recorded highest and lowest values respectively.

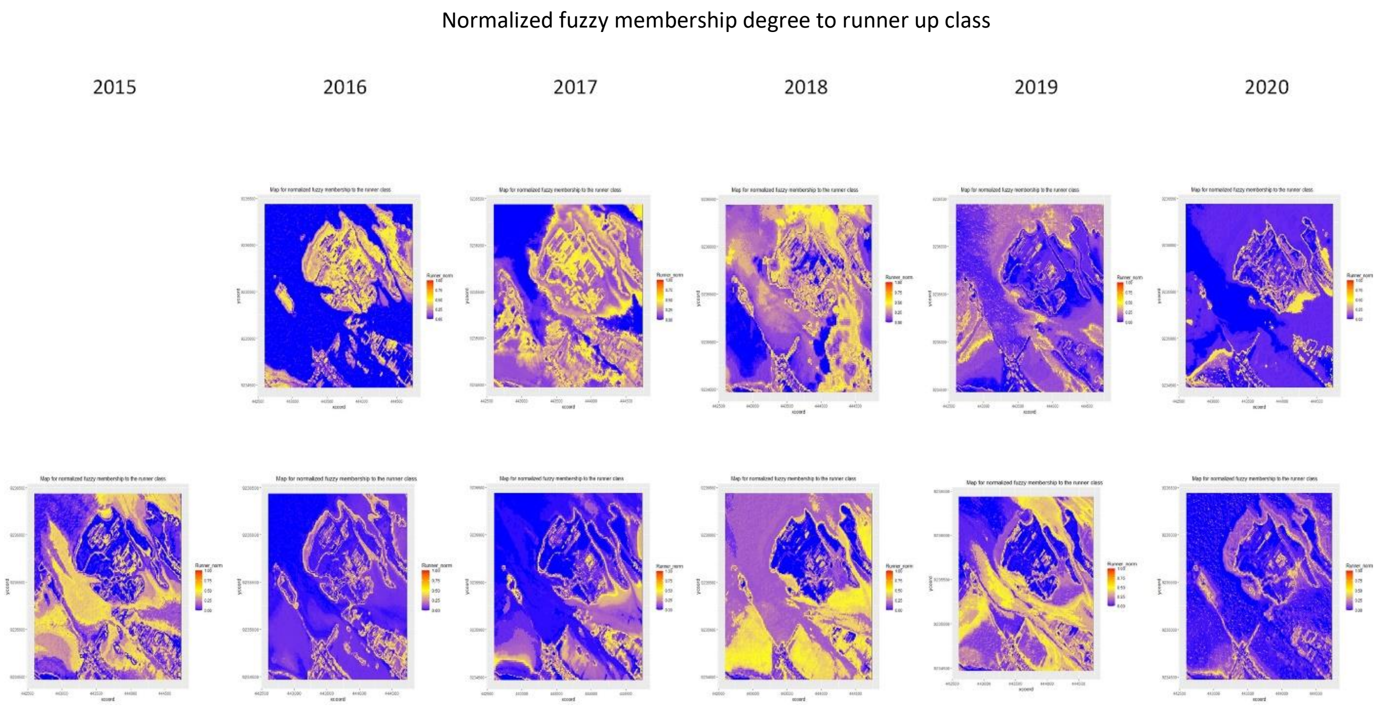



Figure 17 Map of Normalized fuzzy membership degree to the runner-up class of area 1


Figure 18 Map of Normalized fuzzy membership degree to runner up class of area 2

## 4.4. Measures of Uncertainty

Uncertainty in this study is measured by the confusion index, ambiguity index, and fuzziness. The three measures show the uncertainty in the sub-pixel classification, especially for mixed pixels.

### 4.4.1. Confusion index

The confusion index measures the similarity between the best and runner-up classes. In figures 19 and 20, the mixed pixels show the higher value of confusion index over the period for areas 1 and 2. The area with the light blue colour represents pixels with a high confusion index, whereas the area with dark blue colour represents pixels with zero confusion index. In figure 19, the value of the confusion index of the pixels around the land area fluctuates as the pixel shows the equal (0.5) membership of the best and the runner-up class. However, the mixed pixels around the transition zone of land and water (boundary) have high confusion index, and the transition zone can be clearly seen.

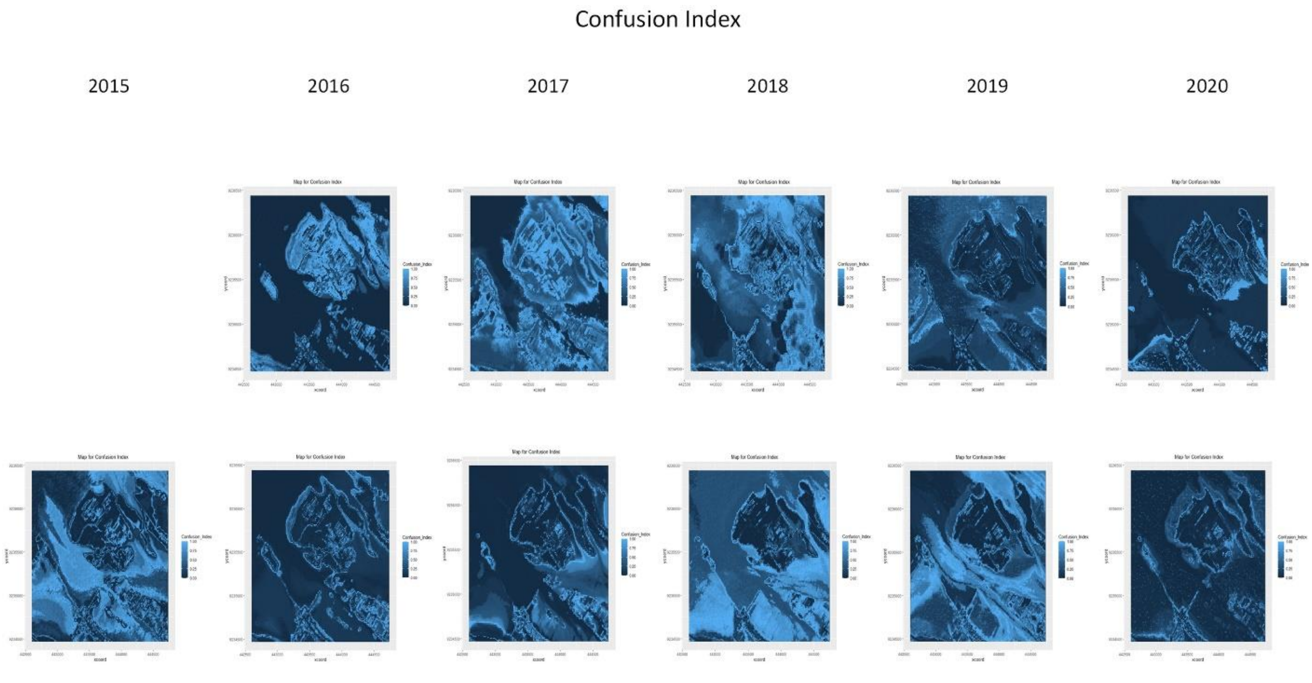



Figure 19 Map of the confusion index of the study area 1

In figure 20, the confusion index of area 2 is presented where the confusion index inside the land area seems to have high confusion index value for most of the images as the piece of land resembles best and the runner up class equally.

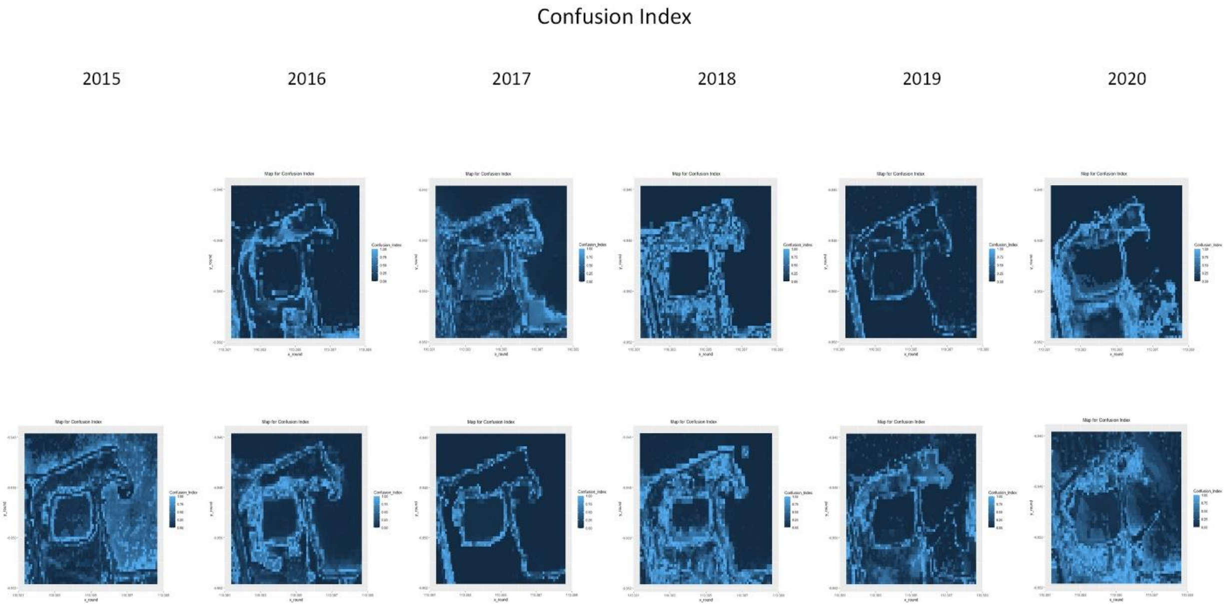



Figure 20 Map of the confusion index of the study area 2

### 4.4.2. Ambiguity index

The ambiguity index depicts the uncertainty of a pixel belonging to the best class. In figure 21, the ambiguity is high for early-years images of 2015, 2016, 2017, and 2018 on the land region of area 1. Overall the ambiguity of the mixed pixels on the transition area of the shoreline is 0.5, represented by the colour yellow. The pixels with no ambiguity are represented by blue colour. Similarly, for figure Y2, the ambiguity of the mixed pixels goes up to 0.5 high. In figure 21 and 22, the ambiguity in the mixed pixels can be seen very clearly, which identifies uncertain zones.

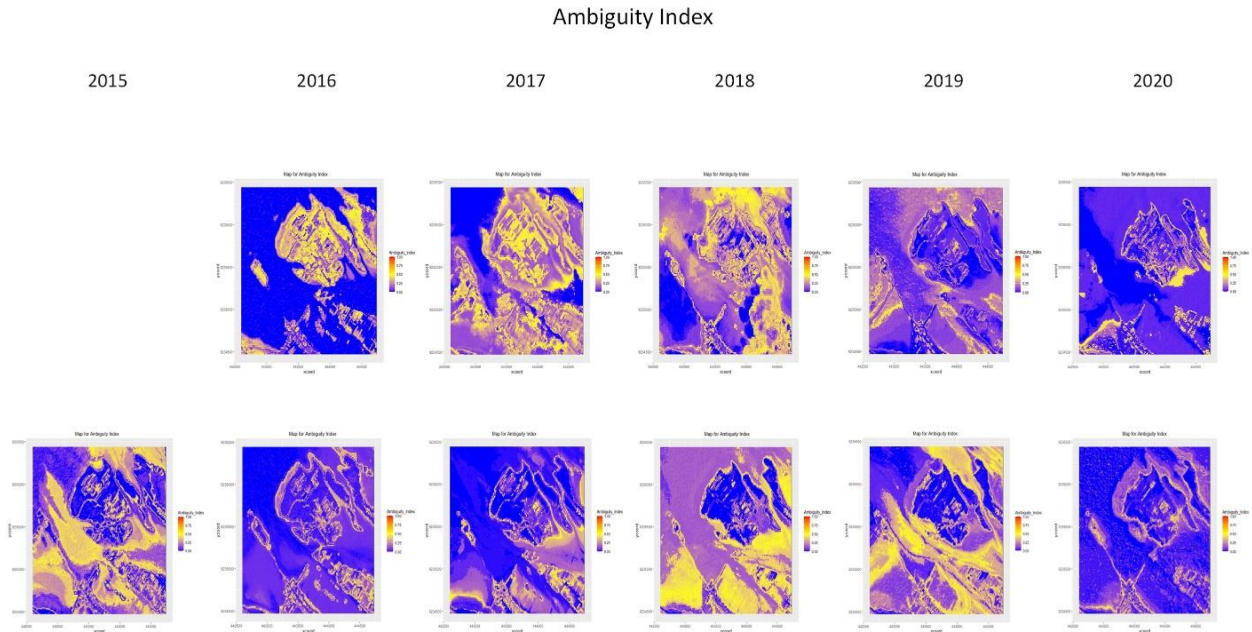



Figure 21 Map of the ambiguity index of the study area 1

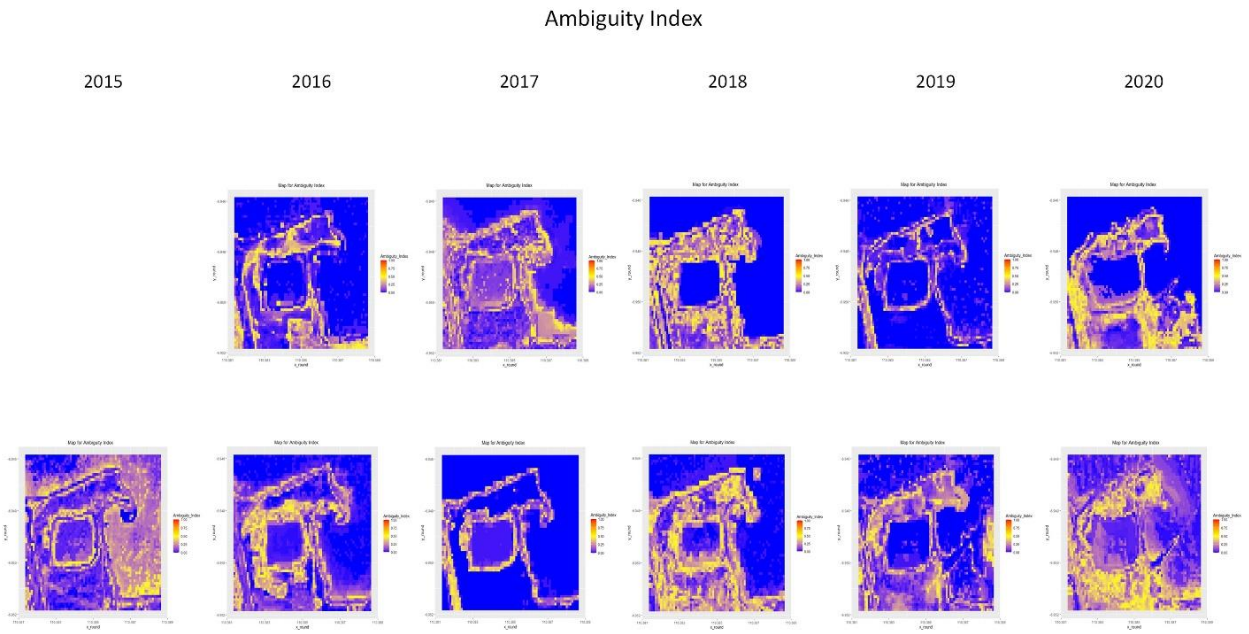



Figure 22 Map of the ambiguity index of the study area 2

### 4.4.3. Fuzziness

Fuzziness is a measure of uncertainty that indicates fuzzy areas containing a lot of uncertainty as they have a higher probability of belonging to more than one class. In Figures 23 and 24, fuzziness can be seen at the boundary of water and land, giving a clear picture of the coastline. The lighter blue color in Figures 23 and 24 represents a high value of fuzziness. All the edges in the figures have high fuzziness values.

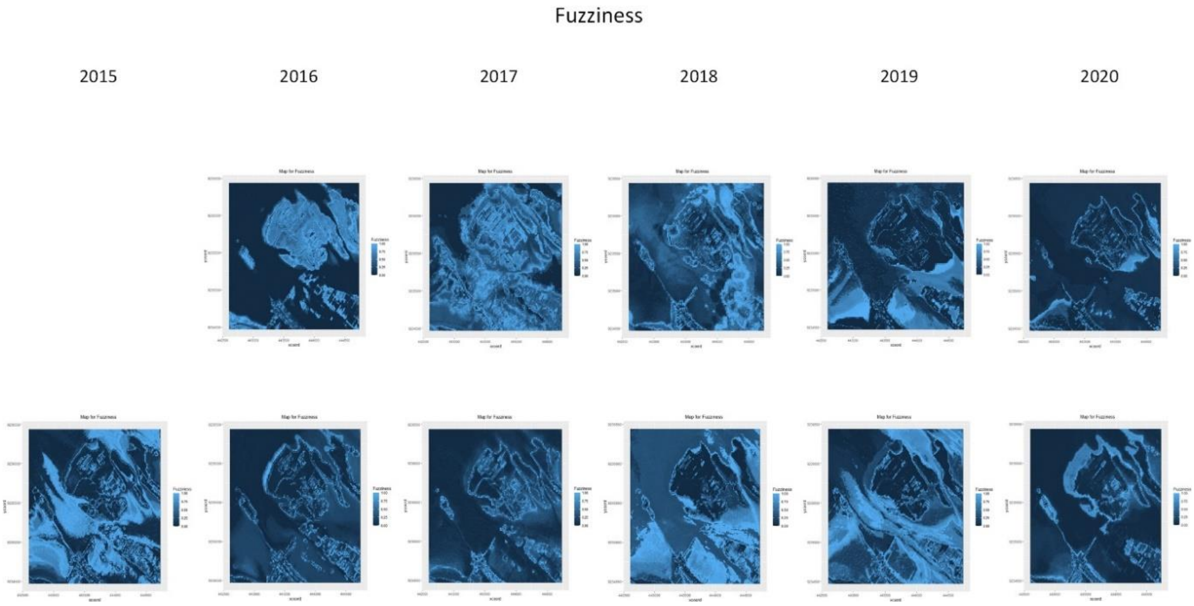



Figure 23 Map of the fuzziness of the study area 1

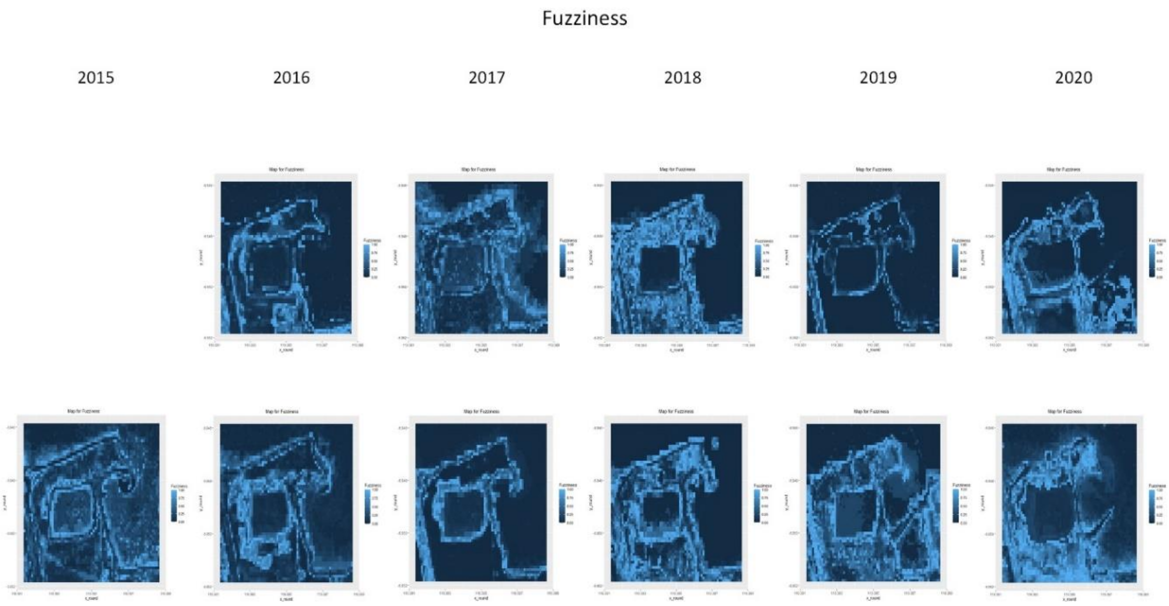



Figure 24 Map of the fuzziness of the study area 2

# 5. DISCUSSIONS

## 5.1. False colour composition

In figure 11, around the edges of the shore covered with vegetation (red), there is a gradient of blue that can be seen from light to dark blue, which can be considered a transition area from land to water. From this, we can conclude that there is an overlap of the two classes as the gradient is neither dark blue (only water) nor red or yellow (vegetation or barren land). Given the location, the gradient most likely represents the shallow water; however, this is just a hypothesis. It can also be just an instantaneous sediment flow. The patches of land can be seen disappearing visually, and the shape of the vegetation covering land can be seen decreasing; from this, it can be concluded that there is a gradual change in the coastline. However, to see the mixed pixels, an analysis has to be done.

In figure 12, the land use pattern seems to change into vegetation area as the red color represents the vegetation increases. The overall outer shape of this area seems to change over the period, which indicates gradual coastal erosion. The gradual change can be most significantly seen in the northern region of the area forming a neck. In addition, the inland waterbody seems to increase in size at the cost of land. Unlike in figure 11, in figure 12 we do not see the blue color gradient around the land, which indicates no overlap between the classes land and water.

## 5.2. Membership degree map of the dominant class in the area

a) Area 1

In figure 13, vast variations can be seen in the membership of the vegetation throughout the five years as well as within a year. The membership of vegetation varies from 0.5 to 1 between the first and the second half of the year. This can be due to the crop cycle in the dry and wet seasons. However, overall the area covered by the vegetation (range 0.5-1) appears to decrease over the years. This shows the dynamic nature of area 1. The very low vegetation membership is represented by the colour purple seen around the land area represents the land. However, the membership of the vegetation is less. This region might have the membership of water or barren land along with vegetation class. From this, we can say that there are mixed pixels and the overlap of two or more classes in the same feature space.

b) Area 2

The membership degree of the (dominant class in 2015) land seems to change drastically in 2016, possibly due to vegetation planted for the protection of the coast. Overall in the following years, the membership degree of the land fluctuates, possibly due to dry and wet seasons. A high membership of land can be seen around the outer periphery of the land, which can be considered the coast's distinctive boundary. This indicates that there is considerably less overlap of the classes. However, around the inland water, the membership of the land seems to be less distinctive over time, indicating the presence of the mixed pixels and the increase in the overlap of two or more classes (land, water vegetation). However, figure 14 also reflects how the coastline is constantly changing and that it is difficult to mark down a specific boundary of the coast for an entire year. The high membership of land is suddenly seen in late 2019 to early 2020. It could be due to the sediment flow or the land appearing due to the low tide; however, the fieldwork needs to verify this theory. Due to the changing membership of land around the periphery for the pixels, the dilemma occurs exactly what to consider as a shoreline boundary for the planning purpose.

### 5.3. Normalized fuzzy membership to the best and runner up classes

The normalized fuzzy membership map represents the spatial distribution of the highest and lowest fuzzy membership values for best and runner-up classes, respectively. In figures 14 and 15, the most interesting part of observing is the yellow colour region which indicates equal fuzzy membership of the two classes. This region represents the pixels with either a high overlap of the best and runner-up class or the approximately similar membership values for a mixed pixel.

a) Area 1

It can be seen that the fuzzy membership degree to the best class decreases, especially at the distinctive coastline. The further the pixel from the transition zone of the coast, the easier it is to classify the pixel to a certain class. In figures 15 and 16, many images have yellow or 0.5 values in the fuzzy membership map of best and runner up class, suggesting that many mixed pixels have the spectral signature of both the classes (best and runner up). This can be seen due to the shallow water around the region (refer to 5.1). This indicates that there is still land present in that region, whereas, in recent years, this seems to disappear due to the further loss in the land, which is now fully a water body. However, the mixed pixels around the transition of water and land still have a low fuzzy membership degree to the best class, making a distinct line around the coastline.

b) Area 2

Due to many mixed pixels, it is challenging to classify the study area to the best class; hence the normalized fuzzy membership to the best class is low. However, in few instances (for the years 2017 and 2019) in figures 17 and 18, the landscape might have been evident, for instance, fully grown vegetation, to get most of the pixel's normalized fuzzy membership to the best class high.

## 5.4. Measures of Uncertainty

### 5.4.1. Confusion index

Thus, the more similar value membership degrees for the pixel's best and runner-up class, the harder it gets to classify a pixel's class. As reflected in figures 19 and 20, the pixels with the higher value of confusion index also represent a pattern that indicates the coastline. With this index, we can conclude for a pixel how equal the probabilities are for the best and the runner-up class. From this, we can understand where the dynamic zones lie, the zones that more or less belong to both the best and runner-up classes.

### 5.4.2. Ambiguity Index

In other words, it tells us how uncertain the classification is, as in general, the class of a pixel is assigned by the membership of the best class of a pixel. . In figure 21, the ambiguity is high for many pixels for early-years images of 2016, 2017, and 2018. This can happen due to the early stage of vegetation growth, where the model can be ambiguous about the best class belongs to the region. For 22, the ambiguity index can tell how much uncertainty a classification contains, which has much significance while considering the uncertainty aspect for scenario building, decision making for the management plan, and impact analysis for a given action plan.

### 5.4.3. Fuzziness

Other regions in figures 23 and 24 can be seen as fuzzy due to the mixed pixels with a higher probability of them belonging to the best and runner-up class. For instance, in the most recent image from the second half of the year 2020, the transition area between land and water is indicated as fuzzy, which depicts that the fuzzy area equally belongs to water as well as land class (to say). This makes it a vulnerable zone that needs to be protected immediately or reclaimed easily with proper management action.

### 5.4.4. Discussion for measures of uncertainty

Given the three measures for uncertainty, the question of which one to use and what aspect arises. As the results for three indices, namely, confusion index, ambiguity index, and fuzziness, seem to be very similar, there are few subtle changes between them which, if compared closely, can be seen distinctively. In figure 25, the comparison of the results of the measures of uncertainty for the year 2020 was studied. The two images are from the first and the second half of the year, respectively. It was seen that the confusion and the ambiguity index maps are very similar. This happens since the region with high confusion index is bound to have high ambiguity.

On the other hand, fuzziness shows a similar pattern to some extent. The high confusion index indicates that the best and the runner-up classes are very similar, resulting in a high fuzziness value. However, we can see the difference in the confusion index and the fuzziness maps when seen minutely. The fuzziness of the mixed pixels on the coastline was very high, which gives a precise boundary of the coastline. The fuzziness maps show the crisp boundary of the coastline more precisely than the confusion or ambiguity index. The fuzziness map shows the small regions that cannot be seen in the confusion or ambiguity maps. Thus, the fuzziness map shows a more enhanced or crisp boundary of the coastline.

2020

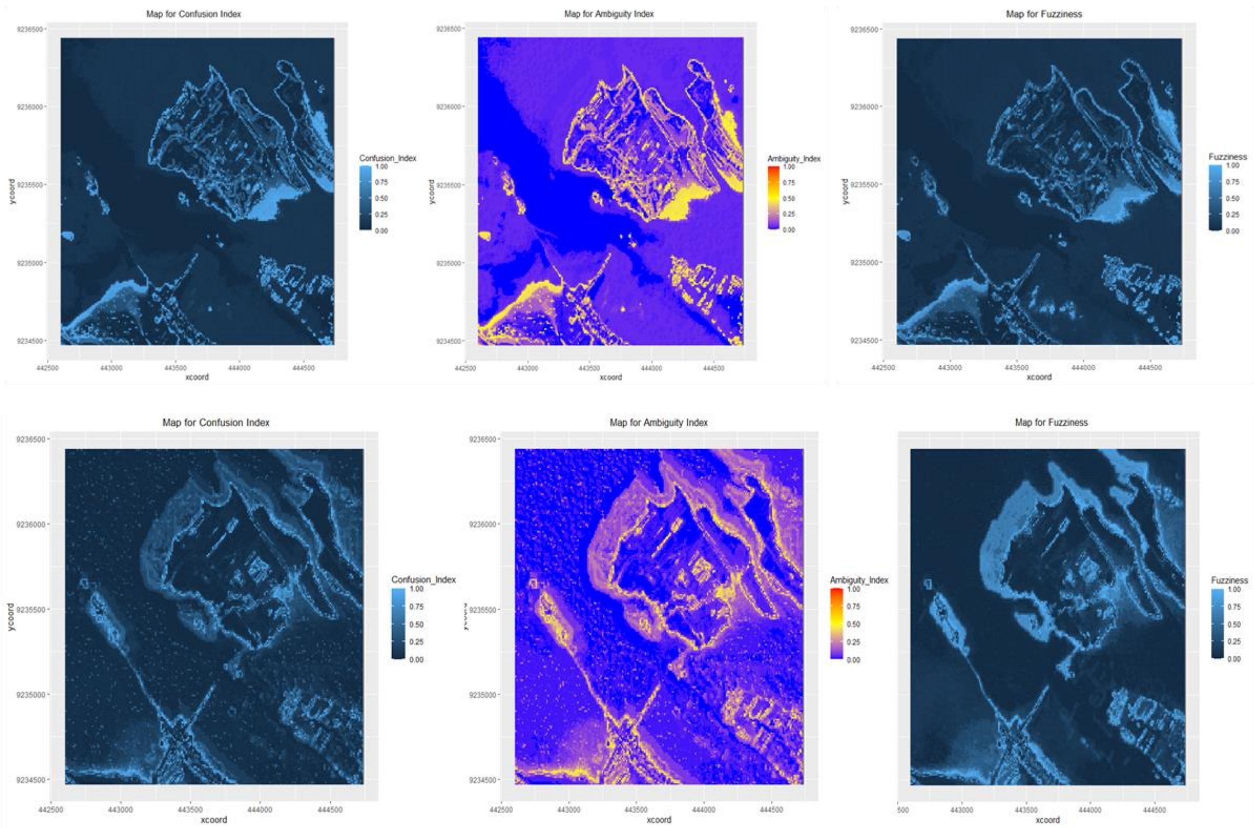



Figure 25 Comparison between the results for measures of uncertainty for the month of March and August of the year 2020

## 5.5. Overall discussions

From the results of the FoRF, we can see that it either represents the mixed pixels areas with equal fuzzy membership of best and runner-up classes or it can identify the region with overlap of two (best and runner up) classes. With these results, we can identify the uncertain area or the areas that are most difficult to study because of the mixed pixels. The changes in the degree of membership of the pixels in the time dimension also would indicate the flood areas; for example, the uncertainty of a particular area is low at first as there is a dominant class, land. However, over time, uncertainty is high, which means there is now approximately equal membership degree of two classes (land and water); the area further shows high membership of water class, which indicates floods. Therefore, these identified areas can be very helpful in mapping the flood-prone areas developed over time and the most uncertain area where we are not sure what is present without field data and experts' of the region.  The quality of the results can be improved by adding more statistical parameters like the standard deviation of the indices and their combination. The reference data from the field, along with experts' opinions, would have helped in obtaining more accurate results for this research. This research can be extended in the future by focusing on separating built-up areas and the flooded streets as the roof of the houses are wider, resulting in shadowing the roads. Another possibility of this research could be to identify the built-up areas and barren land areas by analyzing the texture as people cover their roof with tiles made of (baked) clay and spectrally resembling soil and the uncertainty that lies in the classification.

The results of this research provide a different perspective on knowing what we are not sure of or don't know is also a significant step towards addressing the issue. Knowing what we don't know is also an important step towards better solutions; at least we know where the uncertainty in the region lies the most. Although the changes can be observed visually from which a theory can be derived of what is happening, however, to understand how and why it is happening, ground truth data is required along with stakeholder engagement.

# 6. FUZZIFICATION OF RANDOM FOREST RESULTS– A TOOL FOR COASTAL MONITORING AND MANAGEMENT

The role of remote sensing in coastal monitoring is very significant. It provides periodic information that can be useful to see the shoreline movement and understand its pattern. This information can further be used for resource management on the coast. Similarly, the results of this study show the boundary of the coastline and its fuzziness which can contribute to ICZM. The fuzzy random forest model can be used as a tool for measuring indicators of ICZM.

ICZM is a multidisciplinary approach towards sustainable coastal development through an iterative process. It involves all the stakeholders to evaluate the societal goals for a particular coastal region. To assess if these goals are fulfilled for sustainable coastal development, several indicators must be measured. Thus, the fuzzy random forest model results can be used to measure indicators of the last goal of ICZM *"To recognize the threat to coastal zones posed by climate change and to ensure appropriate and ecologically responsible coastal protection."* Figure 24 shows the relevant indicator from the whole list of ICZM goals, indicators, and measurements (refer to Chapter 2, figure 4 )

| | | |
|---|---|---|
| To recognise the threat to coastal zones posed by climate change and to ensure appropriate and ecologically responsible coastal protection. | 25. SEA LEVEL RISE AND EXTREME WEATHER CONDITIONS | 25.1. Number of 'stormy days' |
| | | 25.2. Rise in sea level relative to land |
| | | 25.3. Length of protected and defended coastline |
| | 26. COASTAL EROSION AND ACCRETION | 26.1. Length of dynamic coastline |
| | | 26.2. Area and volume of sand nourishment |
| | | 26.3. Number of people living within an 'at risk' zone |
| | 27. NATURAL, HUMAN AND ECONOMIC ASSETS AT RISK | 27.1. Area of protected sites within an 'at risk' zone |
| | | 27.2. Value of economic assets within an 'at risk' zone |

Figure 26 Indicator of ICZM proposed by DEDUCE project (Marti et al., 2007)

To achieve the goal of recognizing threats to coastal zones, three indicators are used, namely, sea-level rise and weather conditions, coastal erosion and accretion, and natural, human, and economic assets at risk. As implied earlier, the fuzzy random forest tool can be used to measure the aforementioned indicators. The fuzzy random forest model can be used as a tool for the measurement 25.3., length of protected and defended coastline, by identifying the region with high vegetation or land membership degree with barely any change in the shape over the years can be considered as protected or defended coastline. Because either on that region mangrove plantation has been done, in that case, vegetation membership degree can be seen high, or a coastal belt or dyke is built, in which case the land or settlement membership degree is high.

For indicator 26.1, length of dynamic coastline, the areas with high fluctuation in the degree of membership of land and water can be seen as a dynamic coastline. On the other hand, the areas with stable high fuzziness and stable high ambiguity or confusion are not necessarily at high risk given there is no change, i.e., the fluctuation of the uncertain areas observed is very gradual or none. In this case, we can say the area is just transitioning in space dimension. However, if areas with high fuzziness/ambiguity fluctuate frequently, we can say that they are at high risk as they are transitioning in time.

Furthermore, the water areas with moderate land membership degree indicate that the water area is shallow and can be reclaimed by various measures like sediment trap by planting mangrove forests. As the model can identify the vulnerable zones, the indicator 26.3 number of people living under the 'at risk' zone can be measured for ICZM by overlaying the demographics on the vulnerable areas identified by the model to know how many people are at risk. With the identification of vulnerable zones, protected site 'at risk' areas can be known for indicator 27.1. The indicators measured by the tool can help assess if the goal of ICZM to preserve the coastline has been met or not, which can influence the decision-making and planning of the action plan.

The variable importance function in the random forest provides insights on what variable inputs are of less importance. This helps the decision-makers to identify the outliers and the most significant parameters. It saves the long debates and discussions amongst the stakeholders on what significant value should be assigned to certain input variables. This result independently provides an insight into outliers that do not affect the resulting outcome. This can be very useful during scenario development for ICZM as they will know which parameters will affect the process significantly.

Moreover, the tool helps measure the indicators, but it also directly helps soft criteria for the decision-making process. The Ambiguity map can show the degree of uncertainty in the land use map, developing safety value factors in coastal and disaster management. Also, the degree of uncertainty can be used for the impact analysis for the action plan as well as the scenario development for sustainable coastal management.

# 7. CONCLUSIONS

The coastal erosion, sea-level rise, and land subsidence have proven a threat to the environment and human security because the slow and gradual changes are overseen, and studying them has proved of great significance in identifying the vulnerable and transition areas. Semarang is severely affected by this threat, and it jeopardizes livelihood and cultural identity and may result in severe property damage. However, it is the home for people who do not know how else to live. It is going on for so long that it has become part of their lives. The legislation has several levels; hence the efforts taken by the government to address this issue are fragmented. Sustainable coastal development while protecting the assets and providing people with good quality of life is a complex issue that needs an integrated approach with all the stakeholders. ICZM provides this front to make it possible.

The probabilistic random forest results used for fuzzy logic theory to detect slow and gradual changes can significantly contribute to ICZM. The results show that the fuzzy random forest model can detect coastline detection and its uncertainty. The variable importance function in the random forest can prove a handy tool for decision-makers as it identifies the significance of every variable input. The variable importance function is a very reliable way to settle the dispute of the stakeholders on the significance of the input layers.

In this study in figure 11, the FCC shows vegetation, whereas, in figure 13, the membership degree of the vegetation was 0.5 in few cases. Therefore, it is most likely is vegetation, but there is also a good possibility that the region belongs to another class. This is the best example of the fact that not everything we see visually is entirely accurate. This validates the theory of fuzzy sets that nothing is certain there is always a degree of uncertainty between two extremes. This fuzzy logic can significantly influence the coastline and coastal management as the coastal system is complex and even more challenging to comprehend.

Moreover, comparing the measures of uncertainty maps in figure 25, we could say that the fuzziness map represents an enhanced coastline boundary without any noise (like in confusion and ambiguity maps). The areas are very well-identified and can be of great significance in ICZM. The uncertainty factors being used in ICZM can also prove the significance of fuzzy logic in the decision-making process. Making it better applicable in reality as the interpretations or decisions are never right or wrong, they always hold a degree of uncertainty, or in other words, we say, "provided if everything goes as we predicted"; instead, they are somewhere in between. This can conclude that the fuzzy theory is of great significance in disaster management. It can give a factor of uncertainty for the action plan and develop alternatives for several scenarios.

# LIST OF REFERENCES


(ESA), E. S. A. (2015). SENTINEL-2 User Handbook. *Sentinel-2 User Handbook*.

Acharya, T. D., Subedi, A., & Lee, D. H. (2019). Evaluation of machine learning algorithms for surface water extraction in a landsat 8 scene of nepal. *Sensors (Switzerland)*, *19*(12). https://doi.org/10.3390/s19122769

Aheto, D. W., Kankam, S., Okyere, I., Mensah, E., Osman, A., Jonah, F. E., & Mensah, J. C. (2016). Community-based mangrove forest management: Implications for local livelihoods and coastal resource conservation along the Volta estuary catchment area of Ghana. *Ocean and Coastal Management*, *127*, 43–54. Retrieved from https://doi.org/10.1016/j.ocecoaman.2016.04.006

Akvopedia. (n.d.). ICZM in Indonesia. Retrieved May 14, 2021, from https://akvopedia.org/wiki/ICZM_in_Indonesia

Alesheikh, A. A., Ghorbanali, A., & Nouri, N. (2007). Coastline change detection using remote sensing. *International Journal of Environmental Science and Technology*, *4*(1), 61–66. https://doi.org/10.1007/BF03325962

Alonso, M. C., Malpica, J. A., & De Agirre, A. M. (2011). Consequences of the hughes phenomenon on some classification techniques. *American Society for Photogrammetry and Remote Sensing Annual Conference 2011*, (November), 32–40.

Appeaning Addo, K., Walkden, M., & Mills, J. P. (2008). Detection, measurement and prediction of shoreline recession in Accra, Ghana. *ISPRS Journal of Photogrammetry and Remote Sensing*. https://doi.org/10.1016/j.isprsjprs.2008.04.001

Arefiev, N., Terleev, V., & Badenko, V. (2015). GIS-based fuzzy method for urban planning. *Procedia Engineering*, *117*(1), 39–44. https://doi.org/10.1016/j.proeng.2015.08.121

Baake, K. (2018). Quantifying Uncertainty of Random Forest Predictions.

Bayram, B., Erdem, F., Akpinar, B., Ince, A. K., Bozkurt, S., Catal Reis, H., & Seker, D. Z. (2017). THE EFFICIENCY of RANDOM FOREST METHOD for SHORELINE EXTRACTION from LANDSAT-8 and GOKTURK-2 IMAGERIES. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. https://doi.org/10.5194/isprs-annals-IV-4-W4-141-2017

Belgiu, M. (2018). Classification: Random Forests. Retrieved from eo4geo website: http://www.eo4geo.eu/training/classification-random-forests/

Belgiu, M., & Drăgut, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Beyer, F. (2018). Random Forest with R. Retrieved from GitHub website: https://github.com/florianbeyer/Random_Forest_with_R

BIG Indonesia. (n.d.). DEMNAS. Retrieved from http://tides.big.go.id/DEMNAS/

Bonissone, P., Cadenas, J. M., Garrido, M. C., & Díaz-Valladares, R. A. (2010). A fuzzy random forest.

*International Journal of Approximate Reasoning*, *51*(7), 729–747. https://doi.org/10.1016/j.ijar.2010.02.003

Breiman, L. (2001). Random forests. *Machine Learning*. https://doi.org/10.1023/A:1010933404324

Cai, Y., Zhang, Z., Yan, Q., Zhang, D., & Banu, M. J. (2021). Densely connected convolutional extreme learning machine for hyperspectral image classification. *Neurocomputing*. https://doi.org/10.1016/j.neucom.2020.12.064

Camps-Valls, G. (2009). Machine learning in remote sensing dataprocessing. *Machine Learning for Signal Processing XIX - Proceedings of the 2009 IEEE Signal Processing Society Workshop, MLSP 2009*. https://doi.org/10.1109/MLSP.2009.5306233

Camps-Valls, G., Tuia, D., Gómez-Chova, L., Jiménez, S., & Malo, J. (2012). Remote sensing image processing. In *Synthesis Lectures on Image, Video, and Multimedia Processing*. https://doi.org/10.2200/s00392ed1v01y201107ivm012

De Matteis, A. D., Marcelloni, F., & Segatori, A. (2015). A new approach to fuzzy random forest generation. *IEEE International Conference on Fuzzy Systems*. https://doi.org/10.1109/FUZZ-IEEE.2015.7337919

Dewi, R. S., Bijker, W., Stein, A., & Marfai, M. A. (2016). Fuzzy classification for shoreline change monitoring in a part of the Northern coastal area of Java, Indonesia. *Remote Sensing*. https://doi.org/10.3390/rs8030190

El-Deen Taha, L. G., & Elbeih, S. F. (2010). Investigation of fusion of SAR and Landsat data for shoreline super resolution mapping: The northeastern mediterranean sea coast in Egypt. *Applied Geomatics*. https://doi.org/10.1007/s12518-010-0033-x

Esmail, M., Mahmod, W. E., & Fath, H. (2019). Assessment and prediction of shoreline change using multi-temporal satellite images and statistics: Case study of Damietta coast, Egypt. *Applied Ocean Research*. https://doi.org/10.1016/j.apor.2018.11.009

European Environment Agency. (2000a). Integrated coastal zone management. Retrieved from CEC Communication 2000/547 ICZM. website: http://www.europa.eu.int/comm/environment/iczm/comm2000.htm

European Environment Agency. (2000b). Integrated Coastal Zone Management. Retrieved from https://www.eea.europa.eu/help/glossary/eea-glossary/integrated-coastal-zone-management

Farhan, A. R., & Lim, S. (2010). Integrated coastal zone management towards Indonesia global ocean observing system (INA-GOOS): Review and recommendation. *Ocean and Coastal Management*, *53*(8), 421–427. https://doi.org/10.1016/j.ocecoaman.2010.06.015

Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, *140*, 23–35. https://doi.org/10.1016/j.rse.2013.08.029

Gebbinck, M. S. . (1998). *Decomposition of Mixed Pixels in Remote Sensing Images to Improve the Area Estimation of Agricultural Fields*.

Gens, R. (2010). Remote sensing of coastlines: Detection, extraction and monitoring. *International Journal of Remote Sensing*, *31*(7), 1819–1836. https://doi.org/10.1080/01431160902926673

Ghosh, M. K., Kumar, L., & Roy, C. (2015). Monitoring the coastline change of Hatiya Island in Bangladesh using remote sensing techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*. https://doi.org/10.1016/j.isprsjprs.2014.12.009

Hadi, S. P. (2017). In Search for Sustainable Coastal Management: A Case Study of Semarang, Indonesia. *IOP Conference Series: Earth and Environmental Science*, (012054), 55. https://doi.org/10.1088/1742-6596/755/1/011001

Hadi, S. P. (2018). Integrated Community Based Coastal Management: Lesson from the Field. *IOP Conference Series: Earth and Environmental Science*, *116*(1), 8–11. https://doi.org/10.1088/1755-1315/116/1/012064

Harwitasari, D. (2009). *Adaptation Responses to Tidal Flooding in Semarang, Indonesia*. Retrieved from http://thesis.eur.nl/pub/12145/(1)33555.pdf

Hofmann, P. (2016). Defuzzification strategies for fuzzy classifications of remote sensing data. *Remote Sensing*, *8*(6). https://doi.org/10.3390/rs8060467

Hong, Z., Li, X., Han, Y., Zhang, Y., Wang, J., Zhou, R., & Hu, K. (2019). Automatic sub-pixel coastline extraction based on spectral mixture analysis using EO-1 Hyperion data. *Frontiers of Earth Science*, *13*(3), 478–494. https://doi.org/10.1007/s11707-018-0702-5

Integrated Coastal Zone Management (ICZM). (2007). Retrieved April 12, 2021, from The Coastal Wiki website: http://www.coastalwiki.org/wiki/Integrated_Coastal_Zone_Management_(ICZM)

Jiang, W., He, G., Long, T., Ni, Y., Liu, H., Peng, Y., … Wang, G. (2018). Multilayer perceptron neural network for surface water extraction in landsat 8 OLI satellite images. *Remote Sensing*, *10*(5). https://doi.org/10.3390/rs10050755

Karabegovic, A., Avdagic, Z., & Ponjavic, M. (2006). Applications of Fuzzy Logic in Geographic Information Systems for Multiple Criteria Decision Making. *Sustainable Solutions for the Information Society Proceedings of 11th International Conference on Urban Planning and Spatial Development in the Information Society ; CORP 2006, GEO MULTIMEDIA 06*.

Khatami, R., Mountrakis, G., & Stehman, S. V. (2017). Mapping per-pixel predicted accuracy of classified remote sensing images. *Remote Sensing of Environment*, *191*, 156–167. https://doi.org/10.1016/j.rse.2017.01.025

Kuncheva, L. I. (2001). Fuzzy Logic with Engineering Applications, Timothy J. Ross, (Ed.); McGraw Hill, New York, 1995, pp. 592, ISBN 0-07-053917-0. *Neurocomputing*. https://doi.org/10.1016/s0925-2312(01)00329-0

Kurtener, D., & Badenko, V. (2003). *Fuzzy Algorithms to Support Spatial Planning*. https://doi.org/10.1007/978-3-540-24795-1_14

Lee, D. S., & Shan, J. (2003). Combining lidar elevation data and IKONOS multispectral imagery for coastal classification mapping. *Marine Geodesy*, *26*(1–2), 117–127.

https://doi.org/10.1080/01490410306707

Lee, J. Y., Warner, T. A., & Virginia, W. (1996). IMAGE CLASSIFICATION WITH A REGION BASED APPROACH IN HIGH SPATIAL RESOLUTION IMAGERY. *Methods*.

Liu, H., & Jezek, K. C. (2004). A complete high-resolution coastline of antarctica extracted from orthorectified radarsat SAR imagery. *Photogrammetric Engineering and Remote Sensing*, *70*(5), 605–616. https://doi.org/10.14358/PERS.70.5.605

Loosvelt, L., Peters, J., Skriver, H., Lievens, H., Van Coillie, F. M. B., De Baets, B., & Verhoest, N. E. C. (2012). Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification. *International Journal of Applied Earth Observation and Geoinformation*. https://doi.org/10.1016/j.jag.2012.05.011

Marfai, M. A., & King, L. (2008). Coastal flood management in Semarang, Indonesia. *Environmental Geology*. https://doi.org/10.1007/s00254-007-1101-3

Marti, X., Katrien, A., Borg, M., & Valls, M. (2007). ICZM Indicators Guidelines, DEDUCE consortium. In *Department of the Environment and Housing, Government of Catalonia*. Retrieved from http://im.umg.edu.pl/images/ksiazki/2007_indicators_guidelines.pdf

Mason, D. C., & Davenport, L. J. (1996). Accurate and efficient determination of the shoreline in ERS-1 SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, *34*(5), 1243–1253. https://doi.org/10.1109/36.536540

McFeeters, S. K. (1995). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/01431169608948714#metrics-content

Milczarek, M., Robak, A., & Gadawska, A. (2017). *Sentinel Water Mask (SWM) - New Index For Water Detection On Sentinel-2 Images*. Retrieved from http://eoscience.esa.int/landtraining2017/files/posters/MILCZAREK.pdf

Muslim, A. M., Ismail, K. I., Khalil, I., Razman, N., & Zain, K. (2011). Detection of shoreline changes at Kuala Terengganu, Malaysia from multi-temporal satellite sensor imagery. *34th International Symposium on Remote Sensing of Environment - The GEOSS Era: Towards Operational Environmental Monitoring*.

Nayak, S. (2004). Role of remote sensing to integrated coastal zone management. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *35*.

Niemeyer, J., Rottensteiner, F., & Soergel, U. (2014). Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*, 152–165. https://doi.org/10.1016/j.isprsjprs.2013.11.001

Nurhidayah, L. (2010). Integrated Coastal Zone Management in Indonesia : The implementation and its challenges. *SSRN Electonic Journal*. https://doi.org/10.2139/ssrn.1666807

Nurhidayah, L. (2019). Climate Change Adaptation : Addressing Sea Level Rise Through Integrated Coastal Zone Management : Semarang and Demak as Case Study. *INTERNATIONAL NETWORK FOR GOVERNMENT SCIENCE ADVICE*. Retrieved from https://idl-bncidrc.dspacedirect.org/bitstream/handle/10625/57839/57944.pdf

Olaru, C., & Wehenkel, L. (2003). A complete fuzzy decision tree technique. *Fuzzy Sets and Systems*. https://doi.org/10.1016/S0165-0114(03)00089-7

Oppenheimer, M., & Glavovic, B. (2019). Chapter 4: Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities. IPCC SR Ocean and Cryosphere. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.- O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, M. Nicolai, A. Okem, J. Petzold, B. Rama, N. Weyer (Eds.)]. In Press.*, *Chapter 4*(Final Draft), 1–14.

Pardo-Pascual, J. E., Sánchez-García, E., Almonacid-Caballer, J., Palomar-Vázquez, J. M., de los Santos, E. P., Fernández-Sarría, A., & Balaguer-Beser, Á. (2018). Assessing the accuracy of automatically extracted shorelines on microtidal beaches from landsat 7, landsat 8 and sentinel-2 imagery. *Remote Sensing*. https://doi.org/10.3390/rs10020326

Robnik-Šikonja, M. (2004). Improving random forests. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*. https://doi.org/10.1007/978-3-540-30115-8_34

Roodposhti, M. S., Aryal, J., Lucieer, A., & Bryan, B. A. (2019). Uncertainty assessment of hyperspectral image classification: Deep learning vs. random forest. *Entropy*. https://doi.org/10.3390/e21010078

Ross, T. J. (2010). Fuzzy Logic with Engineering Applications: Third Edition. In *Fuzzy Logic with Engineering Applications: Third Edition*. https://doi.org/10.1002/9781119994374

Sideris, N., Bardis, G., Voulodimos, A., Miaoulis, G., & Ghazanfarpour, D. (2019). Using Random Forests on Real-World City Data for Urban Planning in a Visual Semantic Decision Support System. *Sensors (Basel, Switzerland)*, *19*(10). https://doi.org/10.3390/s19102266

Siler, W., & Buckley, J. J. (2004). *Fuzzy Expert Systems And Fuzzy Reasoning*. Wiley- Interscience.

Sukardjo, S. (1999). Integrated Coastal Zone Management (ICZM) in Indonesia. In W. Salomons, R. K. Turner, L. D. de Lacerda, & S. Ramachandran (Eds.), *Perspectives on Integrated Coastal Zone Management. Environmental Science*. https://doi.org/https://doi.org/10.1007/978-3-642-60103-3_13

Sukardjo, Sukristijono, & Pratiwi, R. (2015). Coastal zone space in Indonesia: Prelude to conflict? *International Journal of Development Research*, *5*(1), 2992–3012. Retrieved from https://www.researchgate.net/publication/274074770

Sukristijono, S. J. D. (2002). Integrated Coastal Zone Management (ICZM) in Indonesia: A View from a Mangrove Ecologist. *Southeast Asian Studies*, *40*(2), 200–218.

Suripin, & Helmi, M. (2015). The lost of Semarang coastal areas due to climate change and land subsidence. *International Conference on Coastal and Delta Areas*, *1*(1), 98–108.

Tang, F., & Ishwaran, H. (2017). Random forest missing data algorithms. *Statistical Analysis and Data Mining*. https://doi.org/10.1002/sam.11348

Tang, S.-J. (2009). *Investigation of coastal dynamics of the Antarctic Ice Sheet using sequential Radarsat SAR images*. (May).

Thia-Eng, C. (1993). Essential elements of integrated coastal zone management. *Ocean and Coastal*

*Management*, *21*(1–3), 81–108. https://doi.org/10.1016/0964-5691(93)90021-P

Tide prediction. (2021). Retrieved from Bandan Informasi Geospatial website: http://tides.big.go.id/pasut/index.html

Toure, S., Diop, O., Kpalma, K., & Amadou, S. M. (2019). Shoreline detection using optical remote sensing: A review. *ISPRS International Journal of Geo-Information*. https://doi.org/10.3390/ijgi8020075

Tran, T. V., & Tran, T. B. (2009). Application of remote sensing for shoreline change detection in Cuu Long estuary. In *Earth Sciences*.

Tuda, M., & Luna-Maldonado, A. I. (2020). Image-based insect species and gender classification by trained supervised machine learning algorithms. *Ecological Informatics*. https://doi.org/10.1016/j.ecoinf.2020.101135

Verbeiren, S., Eerens, H., Piccard, I., Bauwens, I., & Van Orshoven, J. (2008). Sub-pixel classification of SPOT-VEGETATION time series for the assessment of regional crop areas in Belgium. *International Journal of Applied Earth Observation and Geoinformation*. https://doi.org/10.1016/j.jag.2006.12.003

Verisk Maplecroft. (2021). Climate Change Vulnerability Index. Retrieved from Verisk Maplecroft website: https://www.maplecroft.com/risk-indices/climate-change-vulnerability-index/

Wang, P., Fan, E., & Wang, P. (2021). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognition Letters*. https://doi.org/10.1016/j.patrec.2020.07.042

Wicaksono, A., Wicaksono, P., Khakhim, N., Farda, N. M., & Marfai, M. A. (2019). *Semi-automatic shoreline extraction using water index transformation on Landsat 8 OLI imagery in Jepara Regency*. (January 2020), 52. https://doi.org/10.1117/12.2540967

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, *27*(14), 3025–3033. https://doi.org/10.1080/01431160600589179

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhang, T., Yang, X., Hu, S., & Su, F. (2013). Extraction of coastline in aquaculture coast from multispectral remote sensing images: Object-based region growing integrating edge detection. *Remote Sensing*, *5*(9), 4470–4487. https://doi.org/10.3390/rs5094470

# ANNEX

Annex A

The wickedness framework explain the structure of the problem at hand, in this case there is uncertainty in the classification which gives us information, so we can say the knowledge is uncertain. Further, the stakeholders consensus to the acknowledge the severity of the problem and usage of resource is less making the Semarang coast a “wicked problem” zone.

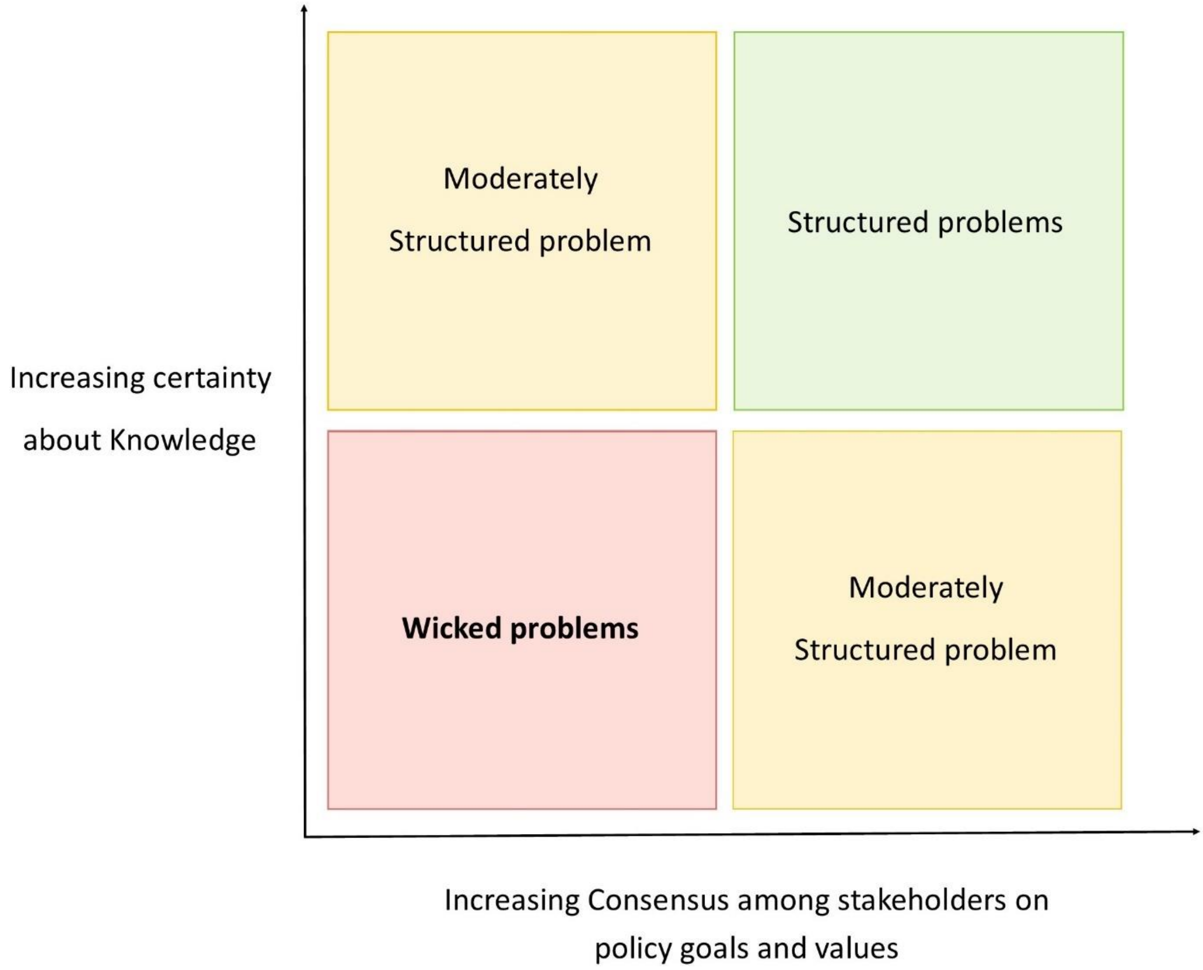
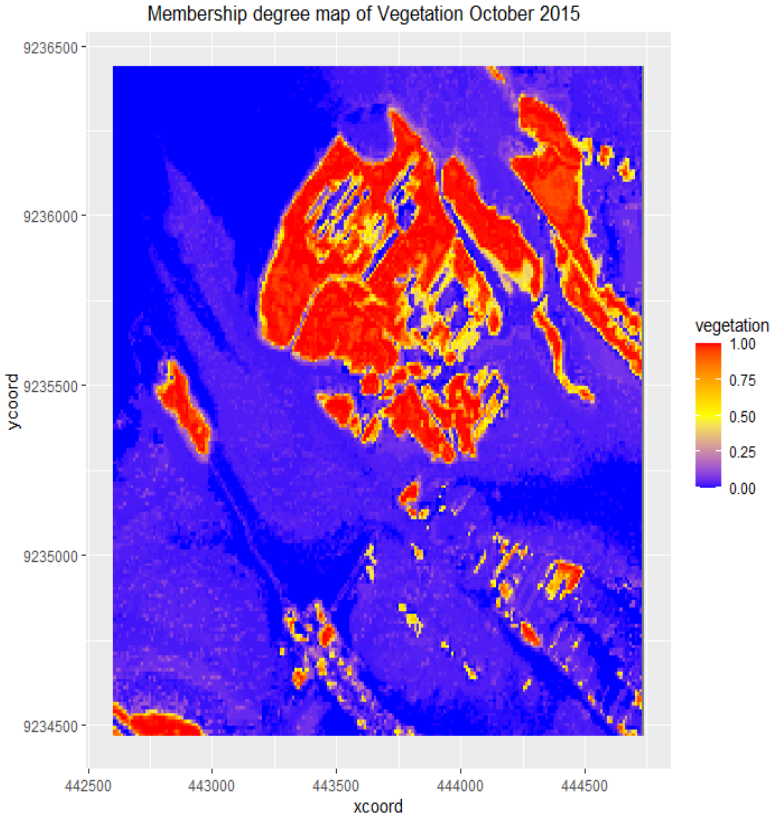


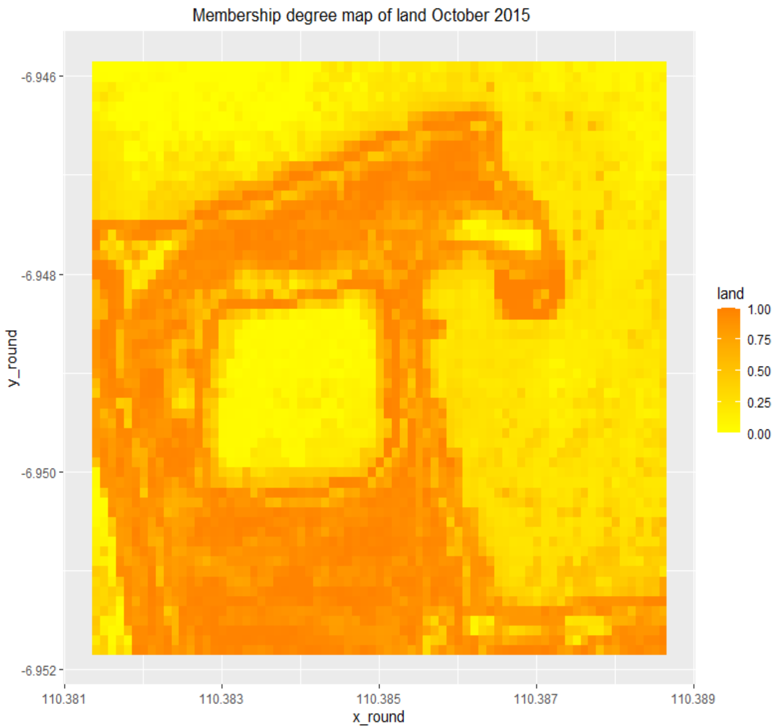

Figure: Wickedness framework

Annex B

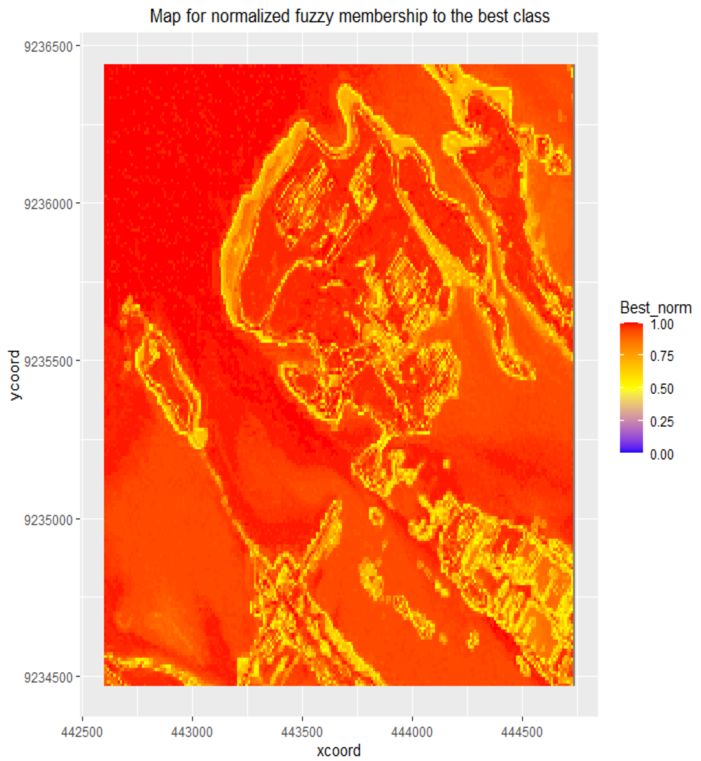
For section 4.2., the example of the map for area 1 in figure 13



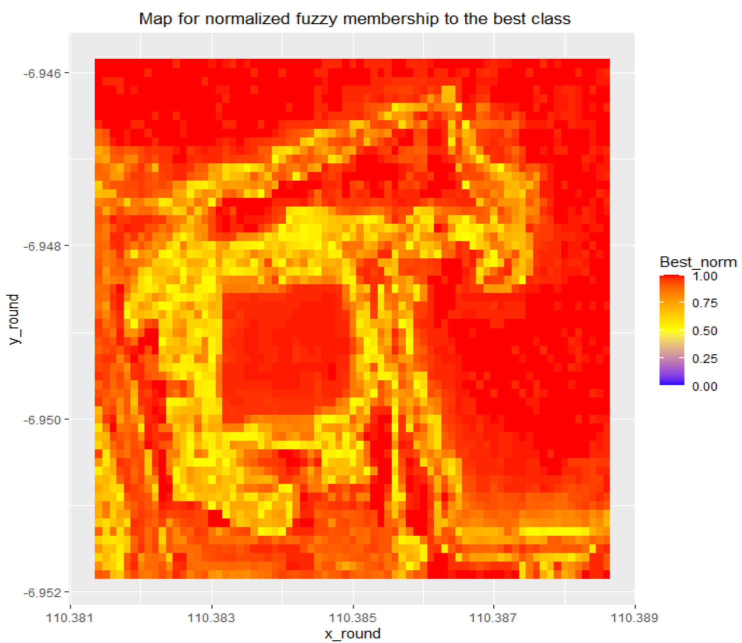

For section 4.2., the example of the map for area 2 in figure 14



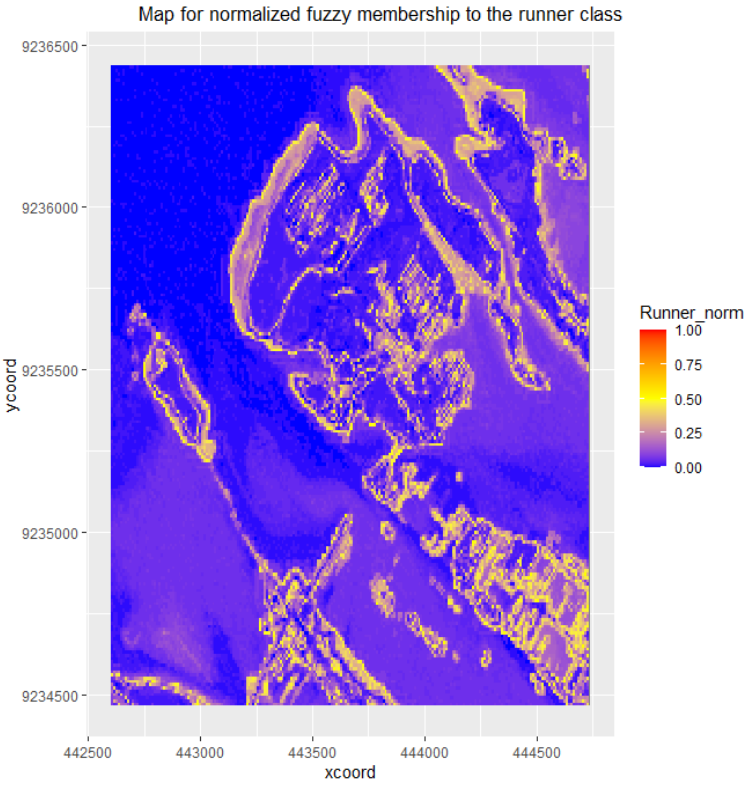
For section 4.3., the example of the map for area 1 in figure 15



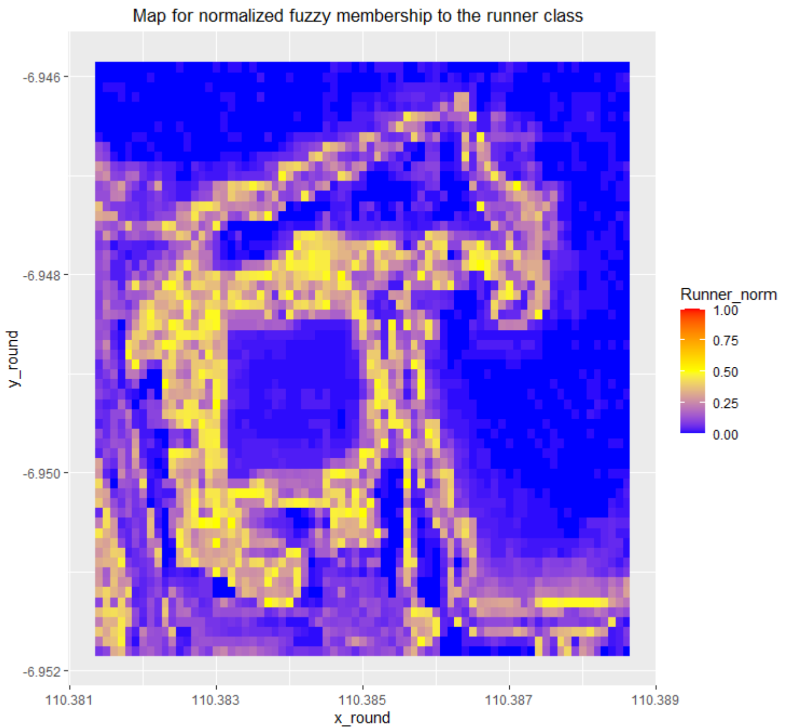

For section 4.3., the example of the map for area 2 in figure 16



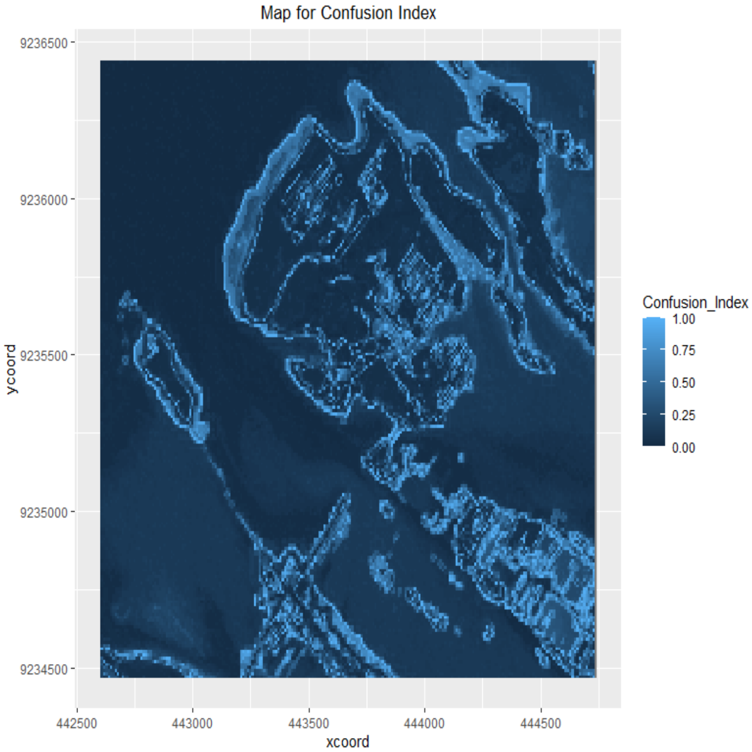
For section 4.3., the example of the map for area 1 in figure 17



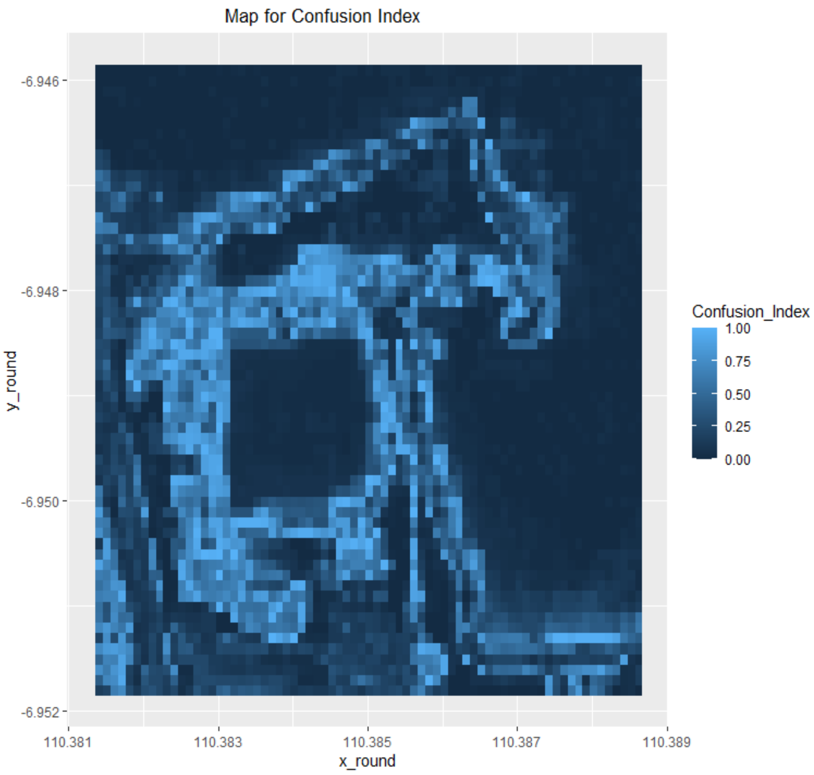

For section 4.3., the example of the map for area 2 in figure 18



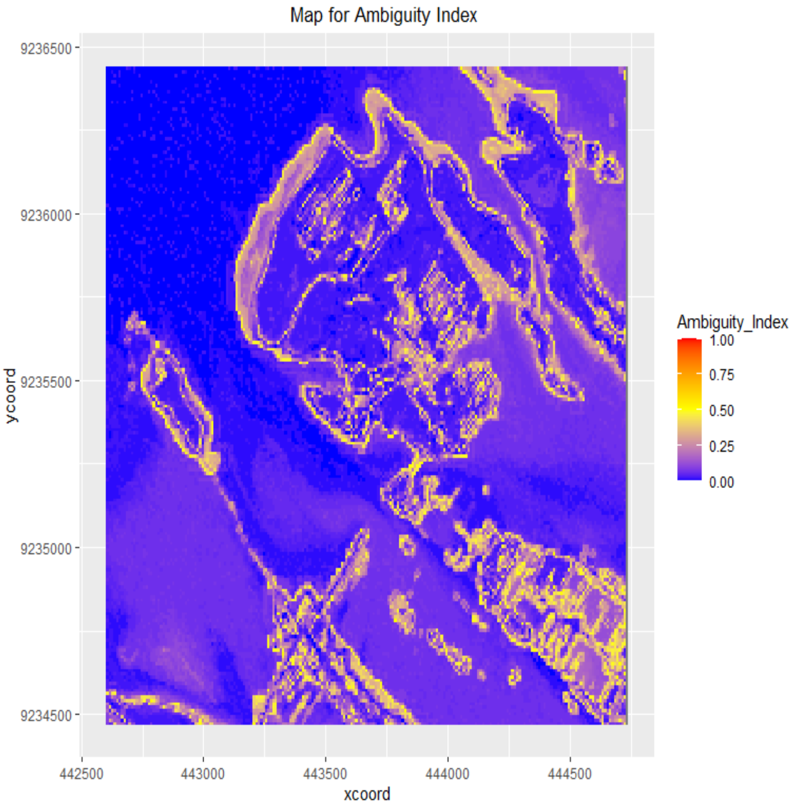
For section 4.4.1., the example of the map for area 1 in figure 19



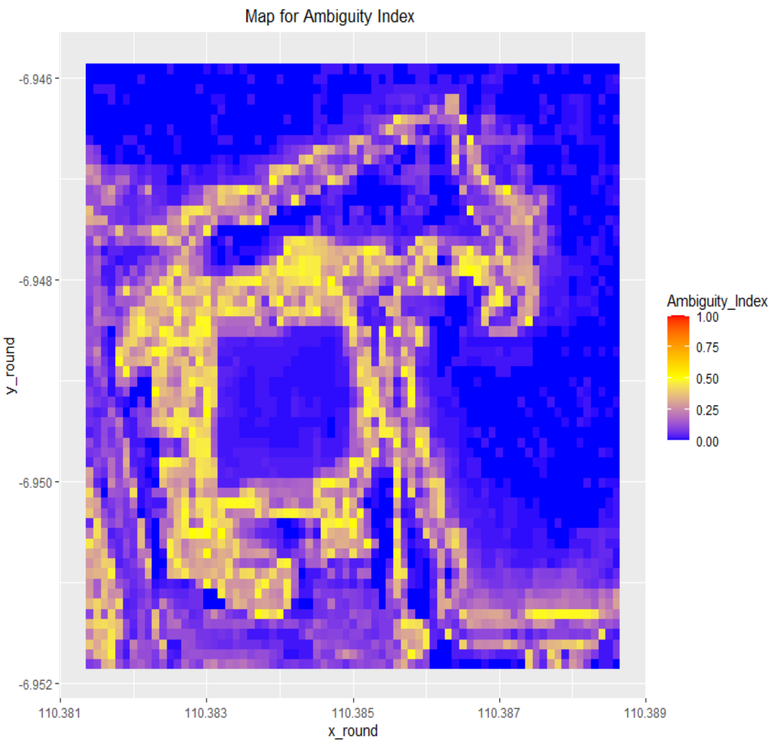

For section 4.4.1., the example of the map for area 2 in figure 20



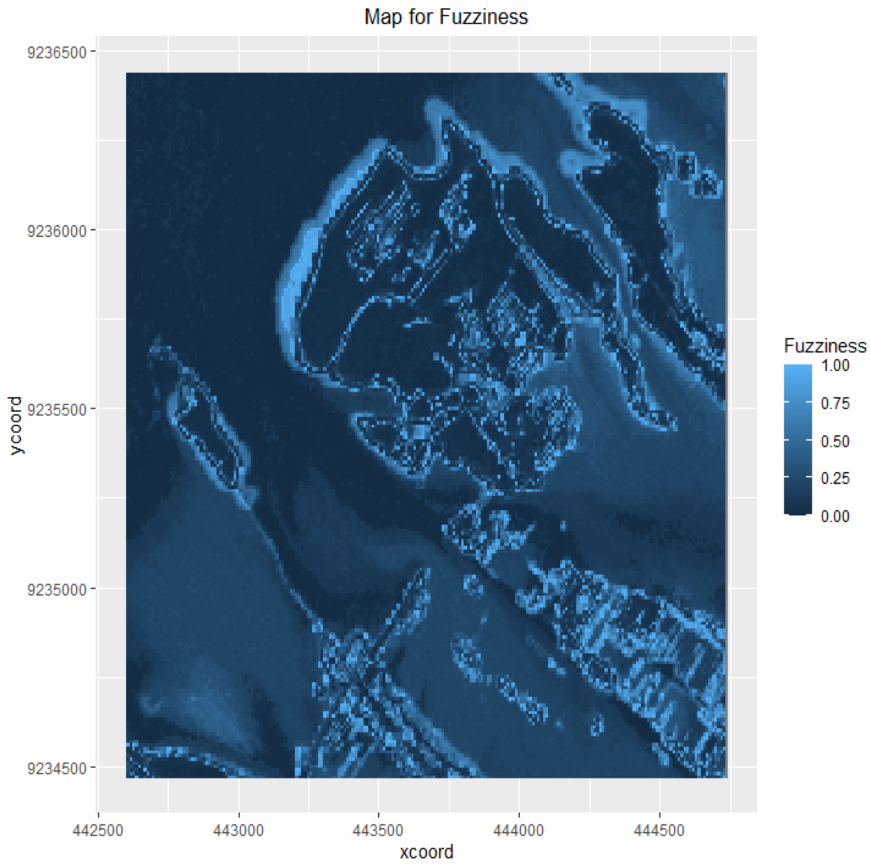
For section 4.4.2., the example of the map for area 1 in figure 21




For section 4.4.2., the example of the map for area 2 in figure 22



For section 4.4.3., the example of the map for area 1 in figure 23




For section 4.4.3., the example of the map for area 2 in figure 24