Doctoral Dissertation
Academic Year 2018

A Semantic Deforestation Interpretation System
with the Closest Semantic-Ellipsoid Algorithm
and L-Band Synthetic Aperture Radar Satellite Images

Graduate School of Media and Governance
Keio University

Irene Erlyn Wina Rachmawan
ABSTRACT

The detection of deforestation by remote sensing technologies has been one of the most important research issues in forest monitoring over the last decades. However, only identifying the area of change is usually not sufficient to understand how critical the effects are on the environment including increased CO$_2$ emissions, loss of biodiversity, and soil degradation. To interpret the causes of the detected forest loss and the full impacts upon an ecosystem, additional expert knowledge is required. A deforestation monitoring system is one kind of a cyber-physical system requiring a sensing, processing, and actuation (SPA) procedure. A fundamental SPA for environmental monitoring devotes for a single computational goal, which can be applied for deforestation area detection, estimation of the area size, and visualization of the area on the map. The need for experts’ interpretation, however, causes the redundancy of a simple SPA procedure. Thus, the first objective of this study is to develop a new deforestation analysis system that can provide semantic interpretations of the deforestation area. The second objective is to reduce the redundancy in the SPA procedure for deforestation monitoring. For the first objective, this study realizes two computational modules in sensing and processing procedure to achieve a deforestation interpretation. The first module detects the deforestation area using ALOS-2/PALSAR-2 L-band synthetic aperture radar (SAR) images as the primary dataset, which can be seen as a sensing procedure. This study proposes a new automatic deforestation detection algorithm that groups together the points with a specific temporal pattern of gamma nought values in the tropical rainforest. In the second module, the meaning is assigned to the detected deforestation area. The novel algorithm, named closest semantic-ellipsoid algorithm, is proposed to address the critical challenges to determine the ecological semantic meaning from the uncertainties condition in the assessment of the impacts of deforestation, which can be seen as a processing procedure. The algorithm calculates the semantic meaning by computing the distance between an observed deforestation area and semantic ellipsoids representing meanings of forest-situations in an ellipsoid form. The system realizes the interpretation of five contexts of deforestation: i) causes, ii) land cover change, iii) soil degradation, iv) soil erosion risk, and v) CO$_2$ emission. For the second objective, as a new concept of the computation, an integrated semantic space for both sensing and processing is created to avoid the repetition of the SPA procedure. It functions as a dimensional pooling for all related parameters of deforestation analysis. After the creation of an integrated semantic space, the SPA procedure is applied to the space. The experimental results of this study indicate that by using the proposed system, not only the deforestation detection has higher accuracy performance compared with standard clustering methods, but the semantic interpretation calculation by the closest semantic-ellipsoid has indicated more information-rich retrieval results.

Keywords: Deforestation, Semantic computing, ALOS-2/PALSAR-2, Automatic Detection, Deforestation.
ACKNOWLEDGEMENT

All praises and thanks are due to God Almighty for his compassion, His mercy, and His relief so that I can complete my study and finish this dissertation.

My immense gratitude and appreciation addresses to my academic supervisor, beloved Professor Yasushi Kiyoki, Faculty of Environment and Information Studies, Keio University, who has the attitude and the substance of a genius, for his supervision. I am deeply grateful to him for his support and advices. He always gives me broader insights for making good research and continually encourages me in regard to my research activity throughout my academic life at his research laboratory. I am deeply thankful to him and proud of being his student. Without his guidance and persistent help this dissertation would not have been possible.

I wish to thank my co-supervisor, Professor Emeritus Hideyuki Tokuda and Professor Wangling Yang of Graduate School of Media and Governance, Keio university. I consider it an honor to work with Professor Tokuda. I am really inspired by his deep experiences, spirit, and works. I also experienced nice discussion regarding my research with Professor Yang. I am grateful for his invaluable intellectual simulation and support. I also would like to appreciate my academic committee, Professor Yoshiyasu Takefuji and Professor Jin Mitsugi, for their important and encouraging comments.

I am indebted to my international co-supervisor, Ali Ridho Barakbah from Electronic Engineering Polytechnic Institute of Surabaya, for his guidance, compassion and support. My doctoral study would not have been possible without his support and encouragement. I also wish to thank Dr. Shiori Sasaki, of Graduate School of Media and Governance, Keio university, who are very supportive and motivation toward my research progress. I have learned a lot from her generous support and valuable advices.

I wish to thank Earth Observatory Research Center (EORC) JAXA for providing insightful internship experience at JAXA. A thank you to my respectful internship supervisor, Dr. Takeo Tadono, for providing the important opportunity in remote sensing field. I also thank all EORC ALOS-group members for their kind support toward my research and study. I am deeply gratitude for the opportunity to learn and have inspiring discussion about remote sensing.

My special thanks to my beloved Multimedia Database Laboratory (MDBL) member for their supports and helps, Chalisa Vessomnai, Jinmika Widjidechakul, Piyaporn Nurarak, Yuka Toyoshima, Hanako Fujioka, my wonderful friend, Ikue Yawata, and all MDBL members who always supports me throughout my study. It gives me great pleasure in acknowledging the support to one and all, which directly or indirectly, have lent their helping support for my research and study.
For those ultimately, I never forgot their supports, my deepest gratitude also goes to my beloved parent brother, and partner for all great support, Rachmad Imawan, Windarti, Bimo Laksono and Christian Koyama, for all of their encouragement.

IRENE ERLYN WINA RACHMAWAN
TABLE OF CONTENT

1. INTRODUCTION ......................................................................................................... 2
   1.1. Tropical Deforestation ............................................................................................. 4
   1.2. Observations of Tropical Deforestation and Unsolved Problems ................................ 5
   1.3. Research Question .................................................................................................... 5
   1.4. Research Objective ................................................................................................... 8
   1.5. Outline of Study ....................................................................................................... 10
   1.6. Contributions ........................................................................................................... 11

2. STATE OF THE ART SEMANTIC COMPUTING, REMOTE SENSING AND TROPICAL
   DEFORESTATION ....................................................................................................... 13
   2.1. The Semantic Computing ........................................................................................ 13
   2.1.2. Principle of Semantic Computing .......................................................................... 15
   2.2. Remote Sensing ....................................................................................................... 17
   2.2.1. Synthetic Aperture Radar Imaging ........................................................................ 18
   2.2.2. Optical-Satellite Imaging using LANDSAT 8 ........................................................ 25
   2.3. Tropical Deforestation ............................................................................................. 28
   2.3.1. Proximate Causes of Tropical Deforestation .......................................................... 29
   2.3.2. Deforestation Spatial Patterns ............................................................................. 29
   2.3.3. Soil and Vegetation Effects of Tropical Deforestation ............................................ 30
   2.3.4. Carbon Emission Effects of Tropical Deforestation ............................................... 33
   2.4. Detecting and Monitoring the Deforestation ............................................................. 34

3. THE SEMANTIC DIMENSIONAL CONTROL: THE EXTENDED VERSION OF THE
   MATHEMATICAL MODEL OF MEANING (MMM) ...................................................... 36
   3.1. Semantic Dimensional Control for Deforestation Interpretation ............................... 43
   3.1.1. The Construction of Semantic Space for Applying Semantic Dimensional Control to Interpreting
         Deforestation Activity ................................................................................................. 43
LIST OF FIGURES

Figure 1.1. The proposal of this study ................................................................. 8
Figure 1.2. The system architecture of the proposed semantic deforestation search system.......9
Figure 1.3. Schematic modularity of the proposed semantic deforestation system. ..............10
Figure 2.1. Technical coverage of semantic computing. (redrawn after [21])......................17
Figure 2.2. ALOS-2 PALSAR in orbit (left) and in the laboratory (right)...........................22
Figure 2.3. Illustration of conventional ALOS-2/PALSAR-2 polarization modes..................24
Figure 2.4. Deforestation phenomenon with one cycle of deforestation..........................29
Figure 2.5. Typology of the forest and non-forest spatial patterns ..................................32
Figure 2.6. Carbon emission from different land use.. .............................................33
Figure 3.1. The classical mathematical model of meaning originally...............................36
Figure 3.2. Semantic dimensional control: merging physical space and meaning space. ....42
Figure 3.3. Deforestation semantic space. ....................................................................45
Figure 3.4. Sensing process in semantic deforestation interpretation space. .......................47
Figure 3.5. ALOS-2 specifications; http://www.eorc.jaxa.jp/ALOS-2/en/about/palsar2.htm.....49
Figure 3.6. Global 25m resolution PALSAR-2/PALSAR mosaic & forest/non-forest map ......50
Figure 3.7. Global Land Analysis and Discovery (GLAD). .............................................51
Figure 3.8. Peru as study area.....................................................................................52
Figure 3.9. Indonesia as study area..............................................................................53
Figure 3.10. Sensitivity of ALOS-2/PALSAR-2 data. ....................................................54
Figure 3.11. L-Band backscattering value characteristic based on tropical rainforest types.....55
Figure 3.12. Seasonal L-Band backscattering value characteristic on tropical forest .............. 56
Figure 3.13. Deforestation stage characteristic ................................................................. 57
Figure 3.14. Bar plot of backscattering value of three forest types in tropical countries .......... 58
Figure 3.15. Diagram view of the proposed Density L-Band Algorithm (DELSAR) ............ 59
Figure 3.16. Peak point derived from probability ............................................................. 61
Figure 3.17. Dimensional control from sensing to processing in single semantic space ....... 62
Figure 3.18. An artificial test image; (a) original image (b) image with a fabricated speckle.... 63
Figure 3.19. SAR filter comparison of ALOS-2/PALSAR-2 data for rainforest region ......... 64
Figure 3.20. Five cycles of ALOS-2/PALSAR-2 data ....................................................... 65
Figure 3.21. Application of DELSAR on Peru ................................................................. 66
Figure 3.22. Shapefile result derived from our proposed idea .......................................... 66
Figure 3.23. Result of detection of deforestation area research activity ............................ 67
Figure 3.24. The result of proposed method ................................................................. 68
Figure 3.25. measurement of agreement between user and experimental result ............... 70
Figure 3.26. Projecting representative value to semantic space ....................................... 71
Figure 4.1. The comparison of semantic projection .......................................................... 81
Figure 4.2. Property of ellipse semantic projection .......................................................... 82
Figure 4.3. Illustration of acquiring meaning in semantic space ........................................ 83
Figure 4.4. Problems of assigning meaning in ellipsoid distribution to a point .................. 84
Figure 4.5. Property of ellipse ......................................................................................... 85
Figure 4.6. The problem definition to find closest point on the ellipse ................................. 86
Figure 4.7. Calculates distance from point to closest ellipse ........................................... 87
Figure 4.8. The problem definition to find outside point on the ellipse .............................. 88
Figure 4.9. Plots of closest distance in Euclidean space .................................................. 90
Figure 4.10. Interpretation of deforestation areas in the context of causes ........................................ 91
Figure 4.11. Interpretation of deforestation areas in the context of Soil Degradation ......................... 92
Figure 4.12. Interpretation of deforestation areas in the context of LULC ........................................ 93
Figure 4.13. Interpretation of deforestation areas in the context of carbon emission .................... 94
Figure 4.14. Set Operation of basic Boolean Logic............................................................................ 97
Figure 4.15. Set Operation of basic Multi-Valued Logic ............................................................... 97
Figure 4.16. Set Operation of the Closest Semantic-Ellipsoid Algorithm ...................................... 98
Figure 4.17. Interpretation Representative Function .......................................................................... 100
Figure 5.1. The general view of procedure for deforestation interpretation .................................. 104
Figure 5.2. Data Training and Data test derived from automatic deforestation detection................. 105
Figure 5.3. Interpretation of soil condition in deforestation area..................................................... 106
Figure 5.4. Normalized Burning Ratio value on (a) forest fire area (b) logging.............................. 107
Figure 5.5. Normalized Burning Ratio value on (a) forest fire area (b) logging.............................. 108
Figure 5.6. NBR value of deforestation because of forest fire and illegal logging ......................... 109
Figure 5.7. NBR value of deforestation because of forest fire and illegal logging ......................... 109
Figure 5.8. Deforestation caused-based effect on temperature of soil ............................................ 110
Figure 5.9. Salinity value and temp value of cause-based deforestation........................................ 110
Figure 5.10. Deforestation cause-based effect on temperature of soil .......................................... 111
Figure 5.11. Salinity value and temp value of cause-based deforestation....................................... 111
Figure 5.12. Semantic interpretation in context of Plantation....................................................... 112
Figure 5.13. Semantic retrieval result for Plant-growth semantic space....................................... 113
LIST OF TABLES

Table 2.1. PALSAR-2 instrument parameters................................................................. 23
Table 2.2. Summary of Spectral Characteristic of Landsat 8......................................... 27
Table 2.3. Landsat-8 Observatory Capabilities ............................................................. 27
Table 2.4. Variables list of proximate causes................................................................. 31
Table 3.1. Summary of Deforestation Semantic Space ................................................... 46
Table 3.2. Comparison of results of selected filtration of the artificial speckled image .... 64
Table 3.3. Confusion Matrix of proposed DELSAR ......................................................... 69
Table 3.4. Standard evaluation of Soil Erosion Risk......................................................... 72
Table 3.5. Standard evaluation of Soil Erosion Risk......................................................... 74
Table 4.1. Interpretation of deforestation areas in the context of causes.......................... 91
Table 4.2. Interpretation of deforestation areas in the Soil Degradation for plantation context. 93
Table 4.3. Interpretation of deforestation areas in the context of LULC......................... 94
Table 4.4. The basic flow of basic Boolean logic............................................................ 99
Table 4.5. The basic flow of fuzzy logic........................................................................... 99
Table 4.6. The basic flow of the Closest Semantic-Ellipsoid Algorithm............................ 99
CHAPTER 1

INTRODUCTION

This Chapter:

- Introduces problems domain
- Describes research scopes, objectives and contributions
- Underlines uniqueness and proposed algorithm
- Outlines the structure of this study
1. Introduction

Technology is growing by giant leaps and bounds in the era of big data. A large volume of data has become an integral part almost every sector. Provided that, the environmental sector can also no longer afford to remain insulated from the technology connect. Today, the rapid progress of the environmental monitoring technology, including satellites that are capable of seeing the earth surface, is enabling the acceleration of data gains related to impact on various environmental fronts. The use of satellite images is promising to provide unseen opportunities to observe environmental problems.

Nevertheless, with the development of the satellite-based Earth Observation (EO), significant challenges to monitoring environmental phenomena have increased [1]–[3]. Near real-time monitoring and analysis of land-cover change, desertification, and forest-cover loss becomes crucial. In the current trend of data processing, satellite data can be regarded as big data. The satellite-based EO market is emerging today and is expected to exponentially increase over the next decades, due to breakthroughs regarding data collection and storage. Various sources of imagery are known for their differences in spectral, spatial, radioactive and temporal characteristics. The growing advancement in the remote sensing technology and data proliferation led to its expansion in land assessment techniques for earth change study and discover valuable knowledge and information. Accordingly, any kinds of big data need to be analysed to predict or understand certain patterns in order to gain meaningful knowledge. Commonly, EO data were analysed by the traditional and physics-based concepts with image and signal processing techniques to mapping the state of the earth surface. Remote sensing research uses primary two approaches: the first is the traditional visual approach, in which the actual observed EO image is analysed subjectively by human interpretation. The second approach is commonly used in computer-implemented pattern recognition techniques applied to an EO image to classify natural-object patterns into different classes. The majority of remote sensing work has been focused on the natural environment, such as deforestation monitoring [4], [5] However, the typical results remotely-sensed data analysis that has been introduced by remote sensing researchers are natural resource mapping and sources of input data for environmental processes modelling [6], [7]. The
various information and patterns that carry rich structures and semantics meaning are not to be discussed well in remote sensing domain.

The challenge then becomes how to analyse “big data” to increase knowledge gain. Large data-repository with lack of analysis techniques has been described as data rich but information poor situation. The design of database application needs to move away from traditional Extract, Transform, Load (ETL) paradigm that requires static pre-modelling efforts toward more discovery-like analysis, interactive and more context-sensitive with the overall goal of producing an analytical result to what is the most probable context that is semantically relevant with user’s preference. To address these needs, various research approaches on multimedia data mining and knowledge discovery have been proposed. Among them, semantic computing has attracted numerous research interests showing its great potential to extract meaning from multimedia databases. Semantic computing is a field that addresses the computational process to expressed user intentions to retrieve, manipulate, or even create the content, where content may be any type of data including text, numeric, images, video, remotely-sensed images, and so on. It addresses all types of information resources including data, process, and people. Major advances in semantic computing facilitate activities that enable users to construct, manipulate or retrieve computational content based on semantics ("context", "meaning", "intention" or "sentiment"), where the content of multimedia database may be video, audio, sensor data, or satellite data.

The remotely-sensed image consists of multiple parameters, complexity and variability information. Thus, in order to provide a semantic analytical result, data can be aggregated or viewed as multidimensional data. By providing multidimensional data views, the construction of remote sensing analytical system can provide a system to detect natural environmental change semantically. A variety of remote sensing techniques have been developed to measure various parameters on the land surface. In the field of deforestation monitoring, these techniques encompass gamma ray, near to far infrared, and thermal infrared spectrometers, passive microwave radiometers, as well as active radar sensors. Among all these satellite technologies, imaging radar, due to its sensitivity towards the dielectric surface properties, its weather independent day and night operation capabilities and its potential to acquire also subsurface information, is the most suitable approach to monitor tropical deforestation. Efforts have been made to develop knowledge-based technologies and expert-systems, which commonly rely on user’s domain knowledge. Several researchers addressed earth observation issues for analysing the environmental problem.
The environmental monitoring system commonly constructed for specific problem whereas the primary information driven by the expert’s knowledge. However, it has not connected yet directly the queries of the multiple contexts to a specific environmental problem. This chapter seeks to boost remote sensing researches for deforestation monitoring by giving an overview of essential challenges, the connection between intentions and content, a brief introduction on tropical deforestation.

1.1. Tropical Deforestation

Loss of forest cover affects climate. One of the significant processes in forest change linked with human activity is deforestation, an act of forest clearing to convert forest to another land use. Humid tropical forests need special attention due to the continuity of considerable pressure on forest cover and condition in this region. Deforestation is defined as the conversion from the forest (open, close or fragmented) to non-forest uses such as such as agriculture, urban, and other non-vegetated areas [8][9]. Deforestation itself is a concept that easily understood, but the costs of deforestation activity vary with the context on deforestation cycle. Forest area itself is a concept that easily understood, but the formal definition of forest varies with the context: there are differences in forest definitions from legal or ecology point of view. To understand the cost of deforestation, it becomes extremely important to use exactly the same unambiguous definitions of the forest at both points in time – otherwise, the stated deforestation cannot be properly understood.

Global and local forest cover maps have been produced by researchers as an act of forest conservation with the utilization of satellite imageries [4], [10], [11]. Although forest map classically used to represent forest conditions, it is likely difficult for non-expert society to interpreting the meaning of forest cover map. The most critical issue is not only on the assessment of deforestation from satellite data but more on how to represent the information that society can understand the effects of deforestation on global environmental change[12]–[14]. To provide further understanding about deforestation impact, diverse forms of information are necessary to give a deep understanding. Environmental phenomenon with high temporal changes such deforestation activity contains a considerable amount of information, including temporal changing, spatial changing, and transformation of environmental conditions.
1.2. Observations of Tropical Deforestation and Unsolved Problems

Generally, main data sources for environmental monitoring derived from satellite remote sensing benefit the observation and disclose forest surface change caused by deforestation activity which the result often represented in a map. Satellites are the only tool that produces data and view of vast forest area in our earth regularly. Observation Satellites are roughly divided into two types based on their sensor type: optical sensor and a microwave sensor [15]. While Optical sensors observe reflected visible lights and infrared rays to form imagery data with varies of information derived from multiband sensor acquired in daylight, yet particularly in tropical regions, frequent cloud cover and haze make it challenging to collect cloud-free temporal image data [2]. One of most potential imagery data for deforest detection is synthetic aperture radar (SAR) data derived from the satellite with microwave sensors. The imageries data from SAR are acquired either by day or by night. Since the availability of data from the different satellite is easy to be acquired, various deforestation analysis has mainly been proposed. Detection of forest and non-forest area using SAR L-Band is very likely to have high advantages because radar L-Band can be reflected directly by trunk, and stem of the trees, moreover, the cloud coverage problem is not appear in the L-Band data [16]–[19]. Other satellite images have been primarily used are Landsat data, well-known to provide free satellite imageries and regular observation acquisition, the challenge is in the cloud coverage which very likely to be captured in its images. These heterogeneous satellite images can provide very different information that might be complement with each other to bring more knowledge about deforestation. Along with the diversity of data, the most notable challenge is to produce in-depth deforestation information beyond the total area lost in each year. In the field of forest study, it is pointed out that two of the main causes of the severe problem caused by deforestation are soil degradation and carbon sequestration. The driver of deforestation itself will make a massive difference to the condition after forest clearance has been done.

1.3. Research Question

Deforestation is defined as the conversion from forest (open, close or fragmented) to non-forest uses such as such as agriculture, urban, and other non-vegetated areas. Deforestation activities in tropical regions has been a major problem for decades with most of the activity being illegal. Illegal action to clear the forest certainly causes harm to local and global environmental conditions. Therefore, immediate deforestation detection is important to prevent further destruction to these
forest environments. This use of forest resources in a non-sustainable manner can affect the structure and function of a wide range of environmental factors. Although approaches for early deforestation detection has been largely proposed in recent research efforts, such kind of forest monitoring first requires informative multi-temporal datasets. For all the above reasons, this study aims to answer the following questions to improve the understanding of deforestation activities and their impacts on local and global environmental factors using advanced technology solutions.

1) What kind of remote sensing data is suitable to develop robust tropical deforestation detection algorithm and how the data can be used to aid the forestry communities in prevention of illegal deforestation activities.

2) What are the most significant impacts of deforestation activities upon different ecosystems on local and global scale and how this knowledge be portrayed as assimilative knowledge for reaching true intelligence system to better understand deforestation effects.

3) What is the optimum natural language processing (NLP) procedure to provide interpretable information about environmental degradation due to deforestation activities.

This study intended to answer the listed research questions. Currently, information regarding deforestation effects is insufficient, most current research focuses only on detection of a deforested area. This study is tailored to model new services for deforestation analysis that can produce not only a detection result but also a further explanation about the effects of observed deforestation area. In order to develop this model, two technical modules are taken. This study combines two analytical steps from two domain knowledge, remote sensing, and knowledge modelling to build the new model. Because deforestation area is the point of interest in this study, the first analytical step is obviously created to detect deforestation area from the dataset. The primary dataset that can be used for recognizing the change in tropical rainforest area is a satellite image. There are two options of satellite data stored in image format can be utilized: active sensor satellite and passive sensor satellite. It is very critical to select which is the most appropriate dataset beforehand because the technical approach to analyse those datasets are completely different. In this study, active sensor L-Band Synthetic Aperture Radar (SAR) is used. Detail explanation in regard to the active and passive sensor and how L-Band SAR is more appropriate than the other type of dataset is discussed exclusively in Chapter 2. The step on how the dataset is being processed by remote
sensing method is discussed in Chapter 3 and the discussion is written to answer the first listed research question.

The scope of the study covers the remote sensing methodology that is employed in order to produce the detection result deforestation area. The new algorithm is called DELSAR (Deforestation Detection using L-Band SAR) is proposed as a new approach to detecting the change in forest region only using L-Band SAR using the distribution of change in given temporal time. The result of this step is in polygon format indicating the place of deforestation area. There is no further knowledge other than location and the total area of deforestation area is registered in this first step. To achieve this study goal to develop a system that also can provide further information such as its significant effects, the technique from knowledge modelling domain is being employed as the second step to portray the knowledge to the database. In knowledge modelling, the cause-based knowledge is called associative knowledge. Unfortunately, the common method in associative knowledge normally can only be devoted only for a single context. Meanwhile, physically, deforestation can cause effects more in one context such as its effect on the soil, air, and surrounding plantation. Third listed research question interrogates the most possible optimum solution to describe multi-context knowledge to the database. One of the most promising methods to cover multiple context recognition is semantic computing. In this study, a new algorithm to acquire its effect meaning from multiple contexts is proposed and discussed in Chapter 4. In addition, during the completion of this study, a new research question had been founded. Generally, all computational procedure in the environmental analysis which contains sensing, processing and actuation procedure, always be done separately. What separate process is mean that all sensing data, processing procedure, actuation or visualization always be done sequentially in order to produce one computational result. If there are two computational results is needed to be done, then all sensing data, processing procedure, actuation or visualization has to be done twice, and so on. Based on this statement, writer of this study found a new question on how to formulate the best solution to avoid the redundancy of processing procedure if new computational result needs to be done in the future. Combination of sensing, processing space is the key idea, so-called single semantic space. The detailed answer lays on the second part of Chapter 3. Figure 1.1. explains the original idea of single semantic space as an expansion concept of the Mathematical Model of Meaning (MMM) [20]. Single semantic space functions as a polling space for all dimension can be used for any computational result in the future. The necessary
dimensions are selected based on the particular interest of computation result. In this study, the discussions of all technical proposals are applied to this single semantic space concept.

**Figure 1.1.** The proposal of this study: new expansion of MMM [20], to focused on sensing and processing. Creating a single semantic space concept to process SPA in same space.

### 1.4. Research Objective

The objective of this research is to propose a deforestation analysis system that can provide not only detection result but also a further explanation about the effects of observed deforestation area. Current research on deforestation monitoring is mainly focused on the development of a detection algorithm, but this research focuses not only on the detection but also the interpretation of deforestation area. The final result of this study was made to counter the research questions. The system has a retrieval function to find the deforestation area with most similar condition with the semantic query given by users. To realize this system in this study, a new algorithm to query the semantic meaning of deforestation area, so-called the Closest Semantic-Ellipsoid Algorithm
Algorithms is proposed. The Closest Semantic-Ellipsoid Algorithm represents the meaning of environmental parameters inspired by the movement of ellipsoid from the motions of water droplet. The algorithm determines the closest meaning by the assumption that the ellipsoid in a circular motion will hit the closest object first. This algorithm is robust for representing interval parameters to identify the meaning of deforestation. In this study, a distinctive segmentation technique to automatically detect deforestation area from L-Band SAR data is also introduced. The results of this process are areas of deforestation in polygon format and saved in shapefile format. This study also introduces a representative point selector to generate representative points from each deforestation polygon for processing the meaning of each deforestation area.

Finally, to achieve the meaning for deforestation area, this study implements an interval based similarity measurement in order to tie the interval environmental standard with the deforestation point in the process of similarity calculation. The Closest Semantic-Ellipsoid Algorithm is proposed for the semantic meaning acquisition. Figure 1.2. shows the system architecture of the proposed semantic deforestation system. The schematic modularity of the proposed semantic deforestation system is shown in Figure 1.3.

Figure 1.2. The system architecture of the proposed semantic deforestation search system.
1.5. Outline of Study

This study is composed of six chapters organized in a hierarchal manner, with each chapter building upon the previous one. Looking at the study in some more details, following this introduction, Chapter 2 will then give the state of the art of the underlying semantic computing and L-Band SAR concept to understand the characteristic of deforestation activity. This chapter summarizes related literature review, portray where the gaps in the literature are and show how this research substitute the gaps. Chapter 3 the detailed explanation of the proposed algorithm, Semantic Computing will be given. This chapter will consist of the concept, flowchart, an initial experiment. Chapter 4 discusses the new proposal of semantic dimensional control techniques of single space for sensing and processing to interpret deforestation. In order to produce meaning from SAR data, the new automatic deforestation detection using L-Band SAR and its performance comparison also will be introduced. Quantitative data analysis also will be explained in this chapter. The essential semantic computing application and the utilization of deforestation monitoring are then discussed in Chapter 5. Chapter 6 will end this study by presenting issues that remain open for future research, and by providing some reflection upon the conducted research.
1.6. Contributions

The main contribution of this research is to realize the context-based analysis of the remotely-sensed images in real environment problem. It plays an important role in the use of advanced and automatic methods and techniques to provide better ways and approaches for remote sensing analysis, searching, and retrieval system. The contributions of this research encompass three fields: semantic computing for environmental study, automatic deforestation detection, and context-based deforestation search system.

1. Dimensional-Control for Sensing and Processing
   In this research, the new computational mechanism as called semantic dimensional control is proposed. The concept of semantic dimensional control is to combine all the process, starts from sensing, processing, and actuation into one semantic space. Hence, A semantic space is created as space can handle any machine learning algorithm for data processing before representing various contexts which correspond to its subspaces. This proposal creates a new calculation process. The semantic interpretation combines the physical dimension and meaning dimension, so Sensing, Processing, and Actuation performed in the same space.

2. The Closest Semantic-Ellipsoid Algorithm
   This proposed algorithm is created to give meaning to the environmental phenomenon or also referred to as an ecological phenomenon. The guidelines for measuring the phenomenon are called Standard. The environmental standard is given in the domain range. The level of measurement refers to the relationship among the values that are assigned to the attributes for a variable. It determines the closest meaning by calculating the radius passed the meaning from the original data.

3. Automatic Deforestation Detection Algorithm using L-Band SAR (DELSAR)
   The detection of deforestation activity using L-Band SAR remains challenging. In this study, a new algorithm namely DELSAR is proposed to improve the precision of deforestation detection in the tropical rainforest using L-Band SAR.

4. Semantic Deforestation Interpretation System. This study presents a new system mechanism to context-based deforestation analysis system with deforestation context to detected deforestation area.
STATE OF THE ART:
Semantic Computing, Remote Sensing and Tropical Deforestation

This Chapter:

- Describes state of the art of semantic computing, remote sensing and tropical deforestation.
- Discusses related works on semantic computing, remote sensing for forest monitoring
2. State of the art Semantic Computing, Remote Sensing and Tropical Deforestation

2.1. The Semantic Computing

As science progresses, there is a major evolution in scientific understanding on remotely sensed data. However, existing remote sensing analysis are not extensible, provides only a fixed and limited range of analysis options and do not have the capability to search for objects with specific meaning. The current data processing requires mechanisms with the ability to make meaningful information from the data drawn from diverse analysis result and sources. This calls for new technology that can provide much more advanced integration capabilities, particularly the meaning of the data, to be shared. As a synthesis of knowledge after over 10 years of development, Natural Language Processing (NLP) has been gaining attention from data scientist. By definition, the field of computing to capture expressed user intentions to help understand the meaning, retrieve, manage, manipulate, or even create the content, where “content” refers to anything including media database such video, audio, and text is proposed as in [21]. It combines elements of natural language analysis and data mining. However, the language analysis cannot compute the meaning, cannot easily interpret the media that normally easy to be interpreted by human. Therefore, the semantic computing as proposed by [20] indicates that the semantic computing is necessary to connecting the human intentions with computational content. This connection can go both ways: computing existing content according to the user’s interest or author’s intentions. The main scope of this chapter is to provide a brief overview of the basic concepts of the mechanism of semantic computing.


The Semantic Computing technology as proposed by [20] is a novel computing to decode meaning from content. This enables a computer system to have a human-like understanding and reasoning. Over the last decades, active research of semantic computing using a mathematical model of meaning has been done. One of the applications of the MMM model for multimedia image data is proposed by [22] indicates the application by introducing a "Kansei" and semantic associative search method based on MMM. Proposed a metric of visual impression, constructed from 180-
dimensional vector space called as colour-impression space; each dimension corresponds to a specific adjective representing humans' colour perception and aiming to create a context-dependent query processing mechanism to generate a ranking by considering the temporal transition of each video's visual impressions on viewers' emotion. Besides the application of MMM model for multimedia image data, another interesting application has been found in the different type of content such as music, as proposed by [23] to extract semantic metadata for culture-based ethnic music rather than only focusing on Western classical music. By using typical music samples coinciding with the creation of a cultural music filter to generate a musical element impression transformation. The study on how to implement a context in the image retrieval system by recognising the most important features in the image search by connecting the user impression to the query has been introduced by [24]. Application of MMM in video content has been done by [25] with extracting dynamic adaptability of a multimedia search system to the contexts of its users and proposes a multi context-adaptive indexing and search system for video data.

Currently, the research of semantic computing by utilising MMM grows from understanding the context of a multimedia database to a variety of challenging diverse dataset and problem. [26] Describing the possibility of MMM utilisation for extraction of gut-microbes information to discover potentially existing bacteria-combinations for classifying nationalities in human attribute data. A dengue location-construction risk calculation method for analysing disease-spread has been proposed by [27]. A new dimension to measure the vulnerability of land use concerning with pattern of human moving. Additionally, a cross-cultural computing system that for multilingual analysis was introduced by [28] by utilising phonological-semantic metadata of multiple languages which were extracted based on two main aspects of language: form and meaning. Efforts on mining the information using semantic computing have also been tried in the environmental study to analyse environmental problems. A work by [29] presenting the result of the utilisation of semantic computing to detect the critical contaminates points and identifies the effect class of multiple-water-quality-parameters. Additionally, [30], [31] describes extending usage of MMM for environmental. [30] proposed a multi-dimensional computing model, MMM and a multi-dimensional space filtering method with, adaptive axis adjustment mechanism to implement differential computing method environment change in various aspects and contexts, to discover what is happening in nature. All these approaches were showing the potential of semantic
computing to analyse multimedia data and other data from common scientific areas such as life science, earth science, and social science.

However, the research area in how to apply semantic computing in remotely-sensed and measures of environmental processes data, such as temperature, radiance, infrared, and pressure does not yet commonly discuss. It would be desirable to describe the semantics of remotely-sensed data using semantic metadata to increase interoperability as well as provide contextual information.

2.1.2. Principle of Semantic Computing

In principle, to realize semantic computing to create the connection between content and user can be made via (1) semantic analysis, which analyses content with the goal to convert it to a semantic description; (2) semantic integration, which integrates content and semantic from multiple knowledge domain; (3) semantic services, which utilize content and semantics to solve defined problems; and (4) service integration, which integrates different kinds of services to provide more powerful services; and (5) semantic interface, which attempts to interpret naturally expressed user opinion or preference [21]. However, in this study, only semantic analysis, semantic integration and semantic interface are used and thus semantic service and service integration will not be considered in the following.

a. Semantic Analysis

Semantic analysis is the foundation of semantic computing: it provides the information resource for semantic integration and services. Semantic analysis may be the most developed part among the five layers of semantic computing. Works and researches on semantic analysis have been discussed in sub-chapter of development of semantic computing. The output of the semantic analysis is a semantic description of the content.

b. Semantic Integration

Semantic integration considers the semantic description derived from the semantic analysis process that is presented in different formats and integrates such information before it can be presented to the user. The application of semantic integration includes, but not limited to, database schema integration, data exchange, ontology integrations, ontology mapping, and semantic conflict resolution to ensure semantic interoperability among heterogeneous information.
c. Semantic Services

A major advantage of semantic computing is providing more powerful computing services for all kinds of users. Example of semantic services can be found in following applications: (1) web search, including automatic question answering and information retrieval, (2) multimedia database processing, with the focus on content-based retrieval and (3) domain-specific applications, which designed to support the interoperability of machine to machine interaction.

d. Service Integration

The mapping between ontologies is a major area in service integration for semantic computing. Automated and scalable solutions are also sought due to the vast number of services. For realising practical scenarios, service integration aims to provide methods between different service and filling the significant gap between different service.

e. Semantic Interface

To achieve the goal of providing more powerful computing services to all kinds of users, a portable and friendly user interface is required. This is especially important when cell phones become more capable. Standard graphical user interface (GUI) techniques such as browsing, menu trees, and online help may be far less appealing for the next generation applications.

Techniques to realized semantic computing have appeared as independent pieces in various computers science fields such as natural language processing, artificial intelligence, multimedia database processing, computational linguistics, and software engineering. An illustration of the technical coverage of semantic computing is shown in Figure 2.1. To addresses semantic synergetic interaction, the semantic computing combines and glues these pieces together into an integrated theme. While current semantic computing researches have focused on specific fields, considering semantic computing basic idea as an integrated discipline, it is important to address the integration of different types of content with expert knowledge domain and their corresponding methods to address complex requests from the user.
2.2. Remote Sensing

Remote sensing is defined as the acquisition or measurement of specific properties of phenomena, objects or materials with recording device not physical contact. The potential advantage from remote sensing, especially if done by satellites, is the regular revisiting time in an area of interest, makes observation as well as regular monitoring possible. In the domain of remote sensing and the instrument is mounted on an aircraft or satellite to capture the data of the objects or areas in the ground. The remote sensing data is divided into several different categories using various criteria. The most important criteria are listed as follow: First, the wavelength range: Principally, the remote sensing data can be acquired by different sensor technology. The remote sensing technology refers to a mechanism to capture the energy between the target object on the ground and the sensor on the remote platform. Practically, the only feasible alternative is by using electromagnetic radiation of specific wavelengths such as microwave, optical, optical, or thermal. Depending on the wavelengths type, the electromagnetic radiation may be available from a natural source or it must be provided and transmitted by the instrument. The instrument to capture the electromagnetic radiation from a natural source such sun is called a passive sensor, which is found in the optical satellite. If the instruments have the capability to transmit radiation and measure the amount of energy returning from the target, then it is called an active sensor.
The use of different wavelength is useful to analyse object on the earth surface. The reason for this is the characteristic of both data can be an excellent complementary for providing complete information. The useful wavelengths for forest monitoring are in the visible, infrared, and microwave region. Second, The pixel size: There are many varieties of sensors to produce the size on the ground of the picture elements(pixels). The finest pixels in current optical instruments are about 60 cm whereas the coarsest pixel size is available up to tens of kilometres. The ground resolution of an instrument is more difficult to be defined than the pixel size. For the standard remotely sensed image, the resolution is close to the pixel size.

The most appropriate pixel size for application of forest monitoring depends on the detail of information needs. It is very usual that the image size is larger when the pixel size is made finer. The data produced by a remote instrument typically documents the number of Electromagnetic energy relations with the observed target. The electromagnetic radiation or EMR is defined as all energy that travels with the speed of light in the harmonic wave pattern. Commonly the wavelength in electromagnetic spectrum can be captured by the remote instrument. The visible light is one of the categories of EMR; other types include the infrared, microwave, and gamma rays. There are several interactions that are possible to have happened to an object on the ground: absorption, reflection, scattering, or emission. An important class of such active imaging systems is radar operating in the microwave region of the electromagnetic spectrum. The main discussion on this sub-chapter is to provides a brief overview of the basic concept of remotely sensed satellite images. More detailed information can be found in the dedicated literature [32]–[36].

2.2.1. Synthetic Aperture Radar Imaging

Synthetic aperture radar (SAR) imaging system utilizes the motion of the platform that carries it to simulate a large antenna, or aperture. In short, as the platform moves, the system transmits multiple, successive radio pulses toward targeted objects or areas. The returned signals are then synthesized into a single image with higher spatial resolution than could be captured with a smaller, physical antenna. SAR employs signal processing to synthesize a two-dimensional image of Earth’s surface from all received signal. Due to this active operation mode, SAR sensors are independent to solar illumination and thus capable of night and day time acquisitions. In additions, the length of radio wave avoids the effects of clouds, fog, smoke, rain, etc. Nevertheless, generally
speaking, imaging SAR systems allow an almost all-weather continuous global scale Earth monitoring.

a. Development of SAR
The origin of the synthetic aperture concept appears from the work of Carly Wiley of the Goodyear Aircraft Corp. in the early 1950s. As described in a paper by [34], it is concluded that higher along-track resolutions could be enabled by the frequency analysis of the reflected signal than that permitted by physical beam along-track width. First imaging radar established in 1978 by SEASAT satellite. The capability in general terrain discrimination and the detection of a target also has been explained [37]. The SEASAT SAR operated at L-Band with the 23.5 cm in wavelength and a single polarization channel was employing horizontal transmit and horizontal receive (HH). [38] indicates that the SEASAT SAR observed Earth for only 105 days due to a massive electric system failure, but it demonstrated the potential of imaging radar and motivated the launching of several follow-on space-borne SAR mission in the 1980s and 1990s. The European Space Agency (ESA) launched its first Earth-observation satellite ERS-1 in 1991, followed by ERS-2 in 1995, Envisat-1 in 2002, and SENTINEL-1 continues the C-band SAR Earth Observation heritage Envisat-1 was launch in 2014. Japan Aerospace and Exploration Agency launched its first SAR satellite JERS in 1992, followed by ALOS-1 in 2006 and ALOS-2 in 2014. Today, many space-borne and airborne SAR systems are available. They are competitive with and complementary to multi- and hyperspectral radiometers as the primary remote sensing instruments. At the time of this study writing, the state of the art civil SAR satellites in orbit are namely the Canadian RADARSAT Constellation and the German TanDEM-X (DLR).

b. Principle of SAR
Radar imaging provides a two-dimensional image of radar reflectivity by illuminating and receiving the backscattered field of an area with microwave pulses. In SAR imaging, a natural scene is characterised in terms of its three-dimensional reflectivity function describing the density distribution of scattering targets in the scene. In this sense, the SAR imaging process can be regarded as the projection of this three-dimensional scene reflectivity function onto the two-dimensional range-azimuth image space. Consequently, the physical information content of the SAR image is nothing more than the band-limited projection of the scene reflectivity into the SAR image geometry [32]. The reflectivity function of the scene depends mainly on the frequency, the polarisation, and the imaging geometry. As stated at the beginning of this chapter, radar imaging
systems operate within the microwave region at frequencies from 3 MHz up to 300 GHz with corresponding wavelengths from 1 mm to 100 m. Most commonly, civil radars nowadays operate at X-, C-, S-, L-, and P-band. In regards to polarisation, SAR systems employ linearly polarised antennas (vertically and/or horizontally) in a single-, dual-, or fully polarimetric mode. In single polarisation mode, the pulse is transmitted in a single polarisation defined by the antenna, and the backscattered signal is received in the same polarisation. The most typical mode of the dual-polarisation system is to transmit in a single polarisation and to receive at two orthogonal polarisations. The example case is the dual-polarisation mode of ALOS-2/PALSAR. The complete polarimetric information in the form of the scattering matrix is measured by a fully polarimetric system. These systems also referred to as quad-polarised, are capable of simultaneous transmission and reception in two orthogonal polarisations, completely retaining the relative phase information.

c. RADAR backscattering coefficient
SAR shall be considered as a measurement instrument. It is important to note that the quantitative use of SAR data, as opposed to the qualitative, requires calibrated images. The detailed explanation of RADAR backscattering coefficient can be found in [39]. In other words, to comparing data from different sensors and/or modes, geophysical parameters extraction by using models, multitemporal analysis, etc., can only be applied using well calibrated SAR data. The mechanism to establish the relation between the pixel values of radar image and the physical observable is called radiometric calibration. It can be considered as a two-step process: i) relative calibration accounting for the relative relationship within the image, and ii) absolute calibration to establish absolute observables comparable between different SAR images with different imaging geometries (Freeman, 1992). For distributed targets, the intensity information of the SAR image is expressed in terms of the radar brightness and the radar backscattering coefficient. The radar brightness $\beta^0$ corresponds to the average radar cross section (RCS) per unit image area, i.e. the pixel or resolution cell, in $dB$ and is the standard radiometric product for uncalibrated radar images. It is a direct result of the amplitude of the received signal expressed in terms of the digital number $DN$ as shown in Eq. 2.1

$$\beta^0 = 10 \times \log_{10} \left( \frac{DN^2}{K} \right)$$ (2.1)

where $K$ is the so-called absolute calibration constant, which is derived - in the ASAR and PALSAR-2 case - from measurements over precision transponders during the Cal/Val periods. For
the ASAR sensor aboard ENVISAT K is 55 dB (Rosich & Meadows, 2004) and for the PALSAR sensor aboard the ALOS K is -83 dB (Shimada et al., 2009). The radar backscattering coefficient $\sigma^o$ is defined as the average RCS per unit ground area in $dB$. Hence, $\sigma^o$ can be obtained by normalizing $\beta^o$ to the ground patch corresponding to the projection of each pixel onto the ground with the Eq. 2.2

$$\sigma^o = \beta^o \times \sin \theta$$  

(2.2)

The angle $\theta$ is the local incidence angle (LIA), also known as angle of incidence (AOI), defined as the angle between the incident radar beam to the surface normal (Figure. 2.6.). The radiometric resolution describes the ability of a SAR sensor to discriminate differences in 30, and thus indicates its quality as a measurement instrument. It can be seen that the incidence angle is important in order to obtain the normalized intensity observable. As mentioned above, the values of $\sigma^o$ are defined by the physical and electrical properties of the target, by the wavelength and polarization, as well as by the radar look angle.

d. ALOS-2/PALSAR

Advanced Land Observing Satellite-2 or ALOS-2 is the next L-Band SAR satellite mission after ALOS (Daichi) which launched by JAXA on May 24, 2014. The mission of ALOS-2 is to acquire continuous data for cartography, regional disaster monitoring and observation. ALOS-2 which shown in Figure 2.2. is equipped with L-Band SAR, named PALSAR-2. ALOS revolves around the earth in the sun-synchronous orbit of 628 km and 97.9degree inclination. ALOS-2 has 207 orbits during a 14-day revisit cycle. PALSAR operates at L-band with a centre frequency of 1.24 GHz and a corresponding wavelength of 22.9 cm. The antenna is 2.9 m x 9.9 m in size. Besides the beam steering capability to adjust the off-nadir angles, the main advantage of this design is the fact that the system can operate in quad-polarised configuration. To cover broad observation area, ALOS-2/PALSAR-2 has a wide incidence angle (8° - 70°) electronic beam steering with dual-side looking observations from the satellite ground track. ALOS-2/PALSAR-2 has wide swath coverage for polarimetric observation and utilises a type of polarimetry as single, dual and quad (full-pol.) as a standard mode, and compact (or hybrid) as an experimental mode. Full-pol. Mode on ALOS-2 is a system which realises transmitted polarisation by replacing horizontal/vertical by
turns with an interval of PRI (Pulse Repetition Interval). Detailed technical information can be found in JAXA ALOS-2/PALSAR-2 official handbook [40].

![ALOS-2 PALSAR in orbit (left) and in the laboratory (right).](image)

(Images courtesy of JAXA)

Figure 2.2. ALOS-2 PALSAR in orbit (left) and in the laboratory (right).

The specification of Table 2.1 is defined for an incidence angle of 37° above the equator. The polarisation acronyms are as follows: 1) SP: Single Polarization, 2) DP: Dual Polarization, 3) FP: Full Polarization, and 4) Compact Polarization (experimental mode). Figure 2.3. indicates the selection of single polarisation (HH/HV/VV), dual polarisation (HH + HV, VV + VH), and quad polarisation (HH + HV + VV + VH) and the compact polarimetry is enabled by the TRMs. In L-Band SAR, the ionospheric effects like Faraday rotation and phase delay have to be considered and carefully corrected. The quat polarimetry mode uses the pulses of H and V which increase the PRF and result in a narrow swath.

The ALOS-2/PALSAR-2 has three modes: 1) Spotlight mode: the most detailed observation mode with 1 by 3 meters resolution (observation width of 25 km), 2) Strip Map mode: A high-resolution mode with the choice of 3, 6 or 10 meters resolution (observation width of 50 or 70km), and 3) ScanSAR mode: A broad area observation mode with observation width of 350km or 490km, and resolution of 100m or 60m respectively. In ScanSAR observation mode, ALOS-/PALSAR-2 ScanSAR mode is that the burst overlap ratio should be larger than 90% in 1 sigma for the five subswath modes for the interferometric analysis [41]. Since the ScanSAR mode...
provides large-area coverage, it is suitable to monitor large forest areas. In this study, ScanSAR data from ALOS-2 are being used to monitor the tropical rainforest deforestation.

Table 2.1. PALSAR-2 instrument parameters

<table>
<thead>
<tr>
<th>Launch Date</th>
<th>May 24, 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Launch Vehicle</td>
<td>H-IIA launch vehicle from TNSC</td>
</tr>
<tr>
<td>Band, wavelength</td>
<td>L-band, 22.9 cm</td>
</tr>
<tr>
<td>Polarization</td>
<td>Single / dual / full / compact (compact polarization is an experimental mode)</td>
</tr>
<tr>
<td>Look direction</td>
<td>Right or left</td>
</tr>
<tr>
<td>Beam steering range</td>
<td>Elevation: ±30°; Azimuth: ±3.5°</td>
</tr>
<tr>
<td>Mode</td>
<td>Spotlight Stripmap ScanSAR</td>
</tr>
<tr>
<td></td>
<td>Ultra-fine High-sensitive Fine</td>
</tr>
<tr>
<td>Polynomial</td>
<td>SP/DP SP/DP/FP/CP SP/DP/FP/CP SP/DP</td>
</tr>
<tr>
<td>Incidence angle</td>
<td>8° to 70° range</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>84 MHz 84 MHz 42 MHz 28 MHz 14 MHz</td>
</tr>
<tr>
<td>Ground resolution</td>
<td>3 m (rg) x 1 m (az) 3 m 6 m 10 m 100 m</td>
</tr>
<tr>
<td>Swath</td>
<td>25 km (rg) x 25 km (az) 50 km 50 km (FP:30 km) 70 km (FP:30 km) 350 km 5 looks</td>
</tr>
<tr>
<td>S/A: azimuth</td>
<td>20 dB 25 dB 20 dB 23 dB 20 dB</td>
</tr>
</tbody>
</table>
e. RADAR backscattering coefficient in forest regions

First cloud-free L-band observations of forests from space were provided by Seasat and Shuttle Imaging Radar in 1978 and 1994 [42]. The JERS-1 SAR from 1992 to 1998 provided the first L-band SAR for forest characterization, mapping and monitoring. JERS-1 L-band SAR was able to produce spatially and temporally contiguous (SAR) data sets over the tropical belt on the Earth, through the generation of semi-continental 100 m resolution, image mosaics [43], [44]. Deforestation observation using L-band SAR typically exposed a lower L-band backscatter value for non-forest areas compared to forests [45]. At the time when this study is written, (PALSAR-2), aboard on the ALOS-2 was the only commercial L-band SAR satellite in orbit with the only low frequency in orbit, has a unique potential to observe deforestation activities. Research and observations of forest region by successor satellite, ALOS/PALSAR, however highlight that the increased backscatter (Particularly at HH polarization) above typical values of forest usually followed by the initial clearance of forest area. This backscattering behaviour arises because of enhanced scattering from woody debris, cut stumps, or parts of a fallen tree [46]. In case of forest clearance activity such deforestation, declining backscattering value can be distinctly found after woody debris removal procedure. This is because of the transition from double bounce to surface

![Diagram of ALOS-2/PALSAR-2 polarization modes](image)

Figure 2.3. Illustration of conventional ALOS-2/PALSAR-2 polarization modes (same as implemented on ALOS-2/PALSAR), image credit: JAXA
scattering is highly associated with surface roughness [3]. In the case of ALOS PALSAR on global observations, [17], [45] has shown the applicability of L-Band SAR for forest monitoring. Details on standard backscattering value sensed by ALOS-2/PALSAR-2 data at tropical rainforest regions can be found in the literature [3]. In regards to literature, results derived by observed global tropical rainforest from 2007 to 2010, shows an overall decrease in γ₀ was observed at both HH and HV polarization, with this being (at the global level) respectively −0.040 and −0.028 dB yr⁻¹ for forest and non-forest combined, −0.106 and −0.031 for forest, and −0.032 and −0.016 for non-forest areas. The decreases in HH and HV γ₀ for the forest areas in Southeast Asia, Amazonia and Central Africa were −0.004 to −0.045 and −0.007 to −0.038 dB yr⁻¹. With this well-defined base knowledge, it is derived the most commonly straightforward approach for detecting the change in the forest is by applying a threshold value [3], [17], [46]–[48]. The advantages of using thresholding method are that it reduces the complexity of data processing and simplifies the process for discriminating forest regions [6], [17], [49]–[51]. However, it is necessary to evaluate the capability and stability of threshold value over space and time. The backscattering value in forest regions is not only affected by forest clearance activity such as deforestation, but also by other environmental conditions such as seasonal change, and ground surface moisture. While the usage of multi-temporal L-Band SAR has been studied by many researchers [18], [49], [52], [53], still, only a few research has been done on the automatic deforestation detection using PALSAR data that able to show its accuracy in practical use.

2.2.2. Optical-Satellite Imaging using LANDSAT 8

Since the early of the 1970s, Landsat satellites have continuously provided multispectral images of the Earth. The first Landsat was launched in 1972 with two imagers - four-band Multispectral Scanner with an 80-meter resolution and a Return Beam Vidicon. The second and third version of Landsat, so-called Landsat 2 and Landsat 3, launched in 1975 and 1978, were having the similar configuration. The fourth version of Landsat was launched with the MSS, and a new type of instrument called the Thematic Mapper (TM). The instrument upgrades include the improvisation of the ground resolution up to 30 metres and three new wavelength bands. Landsat 5, a copy version of Landsat 4, launched in 1984 and acquired data for 28 years - 23 years beyond its 5-year design life. The next version, Landsat 6, equipped with an additional 15-meter panchromatic band, was lost immediately after launch in 1993. The following version, Landsat 7 was launched in 1999. The SLC or Scan Line Corrector was reported to be failed in May 2003. The Landsat 7 still orbiting
the earth and gathering the information although the SLC failure makes the analytical step of this
data challenges. The newest version of Landsat is Landsat 8, launched on February 2013. It
consists of two new sensors—the Operational Land Imager (OLI) and the Thermal Infrared Sensor
(TIRS). These two sensors provide seasonal coverage of the global landmass at a spatial resolution
of 30 meters (visible, NIR, SWIR); 100 meters (thermal); and 15 meters (panchromatic). Landsat
8 was developed as a collaboration between the U.S. Geological Survey (USGS) and NASA. The
spectral coverage and radiometric performance (accuracy, dynamic range, and precision) are
designed to detect and characterise the multi-decadal land cover change in concert with historical
Landsat data. Landsat 8 measures different ranges of frequencies along the electromagnetic
spectrum – a colour, although not necessarily a colour visible to the human eye. Each range is
called a band, and Landsat 8 has 11 bands as shown in Table 2.2. Since the latest Landsat version
that operates is Landsat 8, this study using this data for the process the meaning and the
specification about the observatory capability is listed in Table 2.3.

Landsat 8 OLI band data is served in Digital Number Value, which can be converted to
Reflectance value to describe nature object property. Reflectance is measurement number to
represent the object which been lighted by the sun. Its value describes the number, which further
uses to a different plant, water, soil, and any other object in earth surface. To convert DN data to
reflectance conversion, the standard equation for rescaling coefficients listed in the metadata file
used and inserted in Eq. 2.3 called Top of Atmosphere reflectance (TOA) is applied to this study.

\[ \rho_\lambda = M_\rho * Q_{cal} * A_\rho / \sin \theta_{se} \]  \hspace{1cm} (2.3)

Where:
\( \rho_\lambda \) = TOA reflectance, ranges from 0 to 1.
\( M_\rho \) = Multiplicative rescaling factor from the metadata.
\( A_\rho \) = Band-specific additive rescaling factor from the metadata.
\( Q_{cal} \) = Quantized and calibrated standard product pixel values (DN).
\( \theta_{se} \) = Local sun elevation angle
Table 2.2. Summary of Spectral Characteristic of Landsat 8

<table>
<thead>
<tr>
<th>Waveband Name</th>
<th>Band Number</th>
<th>Wavelengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal</td>
<td>1</td>
<td>0.433–0.453</td>
</tr>
<tr>
<td>Blue</td>
<td>2</td>
<td>0.450–0.515</td>
</tr>
<tr>
<td>Green</td>
<td>3</td>
<td>0.525–0.600</td>
</tr>
<tr>
<td>Red</td>
<td>4</td>
<td>0.630–0.680</td>
</tr>
<tr>
<td>NIR</td>
<td>5</td>
<td>0.845–0.885</td>
</tr>
<tr>
<td>SWIR</td>
<td>6</td>
<td>1.560–1.660</td>
</tr>
<tr>
<td>SWIR</td>
<td>7</td>
<td>2.100–2.300</td>
</tr>
<tr>
<td>Panchromatic</td>
<td>8</td>
<td>0.500–0.680</td>
</tr>
<tr>
<td>Cirrus</td>
<td>9</td>
<td>1.360–1.390</td>
</tr>
<tr>
<td>Long Wavelength IR</td>
<td>10</td>
<td>10.6–11.2</td>
</tr>
<tr>
<td>Long Wavelength IR</td>
<td>11</td>
<td>11.5–12.5</td>
</tr>
</tbody>
</table>

Table 2.3. Landsat-8 Observatory Capabilities

<table>
<thead>
<tr>
<th>Scenes/Day</th>
<th>~650</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSR Size</td>
<td>3.14 Terafit, file-based</td>
</tr>
<tr>
<td>Sensor Type</td>
<td>Pushbroom (both OLI and TIRS)</td>
</tr>
<tr>
<td>Compression</td>
<td>~2.1 Variable Rice Compression</td>
</tr>
<tr>
<td>Image D/L</td>
<td>X-Band Earth Coverage</td>
</tr>
<tr>
<td>Data Rate</td>
<td>384 M bits/sec, CCSDS Virtual Channels</td>
</tr>
<tr>
<td>Encoding</td>
<td>CCSDS, LDPC FEC</td>
</tr>
<tr>
<td>Ranging</td>
<td>GPS</td>
</tr>
<tr>
<td>Orbit</td>
<td>705 km Sun-Sync 98.2° inclination (WRS-2)</td>
</tr>
<tr>
<td>Crossing Time</td>
<td>~10:11 AM</td>
</tr>
</tbody>
</table>
2.3. Tropical Deforestation

Research on understanding the dynamics of land-use and land-cover change and its effects on climate change has attracted many researchers in the last decade [14], [54]–[57]. One of very common environmental transformation is the activity to convert forest to non-forest usage. Deforestation is the process of clearing and changing the main function of the forest. Forest has never stayed in one condition. Globally, the extent of the world’s forest continues to decline as human population growth, makes the demand for food and space increase. Deforestation activities proven to comes in many ways, including forest fire, logging, and mining. these causes of deforestation are relatively bringing different effect on nature [9], [58]–[60]. According to [61] the area of forest in the tropical climatic area decreased rapidly compared to subtropical area. Often, in tropical countries, the majority of deforestation activities are the conversion of forest for raising livestock such agriculture, cattle farm, etc. [60], [62], [63] with the most common method to clear the tree is by using fire. The causes and impacts of deforestation are illustrated in Figure 2.4. In small farmer perspective, fire clears the forest in a fast and economical way. However, a fire significantly affects soil properties because organic matter located near the soil surface is rapidly destroyed. Global and local forest cover maps have been produced by researchers as an act of forest conservation with the utilization of satellite imageries. Although forest map classically used to represent forest conditions, it is likely difficult for non-expert society to interpreting the meaning of forest cover map. The most important issue is not only on the assessment of deforestation from satellite data but more on how to represent the information that society can understand the effects of deforestation on global environmental change. To provide sufficient information to analysing deforestation, diverse forms of information are necessary to forgive deep understanding. Environmental phenomenon with high temporal changes such deforestation activity contains a considerable amount of information, including temporal changing, spatial changing and transformation of environmental conditions.
2.3.1. Proximate Causes of Tropical Deforestation

In deforestation literature [13], [58], [64], [65], typically proximate causes of deforestation are grouped into three broad area: 1) agricultural expansion, 2) wood extraction, and 3) expansion of infrastructure). Factors and rationales involved in tropical deforestation are broke down here by a limited number of three deforestation causes. The proximate causes of deforestation refer to the clearance activity by human activities that have direct effects on the environment [63]. The list of variables of proximate causes shown in Table 2.4. However, several causes have drawn by study case in situ remain broad and non-aggregate to three broad categories [65].

2.3.2. Deforestation Spatial Patterns

Even though, deforestation activities come in vast variations of the cause. Few tropical deforestation footprints can be described with several common spatial patterns. [66] specify six possible patterns that happened across the forest in the tropical belt: geometric, corridor, fishbone, diffuse, patchy, and island. The geometric pattern can be recognised from large-scale deforestation for recent sector activities. Corridor pattern indicates the clearance of forest region/s for road
infrastructure. The deforestation for planned settlement schemes can be commonly associated with fishbone pattern. Transmigration, resettlement, and colonisation are part of the settlement scheme. Diffuse pattern relates to smallholder agriculture activity, where the area is relatively small in size but spread over specific wide area. The high population density that can cause deforestation can be identifying in an irregular pattern. This type of pattern is related to processes of permanent cultivation for food, predominantly for cash crop production. The last commonly recognisable pattern, island, is highly associated with the development of peri-urban areas. It is also caused by the expansion of agriculture and extension of infrastructure — the illustration of deforestation given in Table 2.4.

2.3.3. Soil and Vegetation Effects of Tropical Deforestation

The study from [67] founds that there is a significant effect of severe climate events caused by deforestation and this affects the agricultural systems. For ten years, starts from 1990, flood frequency had increased, and floods lasted longer as natural forest cover declined in tropical countries. Only a small number of modelling studies have begun to link tropical deforestation to water resources and crop productivity. Modelling low to high forest loss outside of protected areas in the Amazon [68] over half of the Amazon, soy yields drop by 25%; in a third of the area, they decline by 60%. Several studies indicate local impacts on agriculture that large-scale forest conversion in one region could lead to remote climatic impacts that limit food production in other regions. The environmental damages caused by deforestation practice vary from canopy removal with resultant changes in both above-ground and below-ground microclimate; soil compaction or loss of soil structure. The topsoil removal with resultant changes in soil physical and chemical properties is also a part of damages caused by deforestation. Not only that, the volatilisation of plant nutrients by fire, followed by the return of nutrients in deposited ash; damage to soil surface root mat, soil microbial population and seed bank. Dedicated studies for soil effect and vegetation effects can be found in the literature [69]–[73]. The broader view of soil and vegetation effects of tropical deforestation include the effects of different deforestation practices on soil and vegetation, deforestation effects on microclimate, soil physical properties, soil erosion, chemical properties, soil biology, the change in soil conditions with time after deforestation and effects on vegetation itself. In this study, we focus on the deforestation effects only on microclimate, soil physical properties, soil erosion and effects on vegetation which can be detected only by remote sensing methods and not by laboratory measurement.
**Table 2.4.** Variables list of proximate causes.

Rewrite from (Helmut J. Geist & Lambin, 2001)

<table>
<thead>
<tr>
<th>Agriculture Expansion</th>
<th>Shifting Cultivation</th>
<th>Traditional Shifting Cultivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Colonist Shifting Cultivation</td>
</tr>
<tr>
<td></td>
<td>Permanent Cultivation</td>
<td>Subsistence Agriculture (food, small holder)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commercial Agriculture (small holder, large-scale)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agricultural Development Project ()</td>
</tr>
<tr>
<td></td>
<td>Cattle Ranching</td>
<td>Small holder Cattle Ranching</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Large-scale Cattle Ranching</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unspecified</td>
</tr>
<tr>
<td></td>
<td>Colonization, transmigration, resettlement</td>
<td>Spontaneous transmigration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local transmigration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Military transmigration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Estate Settlement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industrial forestry plantation settlement</td>
</tr>
<tr>
<td>Wood Extraction</td>
<td>Commercial wood extraction (clearcutting, selective harvesting) Permanent Cultivation</td>
<td>State-run logging (selective, clear-cutting)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Private company logging (selective, clear-cutting)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;Growth coalition&quot;-led logging</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Illegal (illicit, undeclared) logging</td>
</tr>
<tr>
<td></td>
<td>Fuel Wood Extraction</td>
<td>Domestic uses (rural, urban)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industrial uses (rural, urban)</td>
</tr>
<tr>
<td></td>
<td>Pole wood Extraction</td>
<td>Domestic uses (rural, urban)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industrial uses (rural, urban)</td>
</tr>
<tr>
<td></td>
<td>Charcoal Extraction</td>
<td>Domestic uses (rural, urban)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industrial uses (rural, urban)</td>
</tr>
<tr>
<td>Infrastructure extension (INFRA)</td>
<td>Transport Infrastructure</td>
<td>Roads (public, military, logging, mining, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Railroads</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rivers &amp; tributaries</td>
</tr>
<tr>
<td></td>
<td>Market Infrastructure</td>
<td>Public infrastructure (food markets, storage, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Private infrastructure (sawmills, food markets, etc.)</td>
</tr>
<tr>
<td></td>
<td>Public Services</td>
<td>Water &amp; sanitation facilities, electrical grids, etc.</td>
</tr>
<tr>
<td></td>
<td>Settlement expansion</td>
<td>(Semi-)urban settlements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rural settlements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Military defense villages</td>
</tr>
<tr>
<td></td>
<td>Private enterprise Infrastructure</td>
<td>Hydrology Development</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Oil Exploration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mining (gold, coal, tin ore, etc.)</td>
</tr>
</tbody>
</table>
In the context of microclimate, light, temperature and moisture regime are important triggers for seeds. Surface albedo (reflectivity) of forests is low in the visible wavelengths as much of the [74] indicate that deforestation and regrowth after clearance activity could increase the albedo by 1.5 fold. The removal of forest canopy in the tropical forest is likely to cause a greater increase on albedo than elsewhere [75]. Consistently higher soil temperature have been reported by [76] after deforestation activity compared to an undisturbed forest, and the differences appear to depend both on the season and method of forest clearance. The researcher has actively studied the effect of deforestation in physical properties. Deforestation and subsequently tillage practices resulted in almost a 20% increase in bulk density, 50% decrease in organic matter and total nitrogen, a 10 to 15% decrease in soluble ions comparing to the undisturbed forest soil [77]. The main effect of these changes in soil physical properties on plant growth is through loss of topsoil macroporosity. Deforestation also encourages erosion, and the most rapid soil loss occurred early during initial land clearance. The most study indicates that high erosion rates are due to the lack of protection of the soil surface [76], [78]–[82]. Soil nutrient level also very variable after deforestation, especially after burning, due to the ash distribution and composition of ground [83]–[86]. The detailed study on soil effects of deforestation can be found at deforestation book literature [87]. The typology of the change from forest to non-forest in spatial patterns can be seen from Figure 2.5.

**Figure 2.5.** Typology of the forest and non-forest spatial patterns [10].
2.3.4. Carbon Emission Effects of Tropical Deforestation

Forest plays an integral role in mitigating climate change. Not only are forests one of the most crucial carbon storage, but forest also uses the carbon for photosynthesis. With that being said, the forest can use carbon to produce clean oxygen. However, rapid land use conversion in support of human development have led to the removal of natural forest cover with forest carbon loss. The conversion of forest to agriculture or urban areas produces larger emissions as shown in Figure 2.6. Calculating carbon loss from deforestation activity is more complicated than it is for other sectors, like energy or transportation. That is because the carbon emission from transportation and energy sector is only in a single-way flow, however, the flow of carbon between forests and the atmosphere happens in two directions. In order to measure carbon emission from deforestation, two kinds of measurement are needed to estimate GHG emissions from deforestation: the rate of change in the forest cover (or “deforestation rate”) and the amount of the stored carbon in the forest (or “carbon stock”). The standard methodology described in [88]–[90] and the IPCC guidelines [91] in calculating gross carbon emissions from deforestation assuming immediate carbon release at forest clearing.

![Figure 2.6. Carbon emission from different land use. The blue arrows indicate absorption and orange arrow indicates emission.](image-url)
2.4. Detecting and Monitoring the Deforestation

It is understandable that deforestation brings several negative impacts on the environment. This makes it vital to regularly monitor tropical rainforest in order to report deforestation and forest degradation. Researches on detecting and monitoring forest region have been a very active area in earth observation. However, the forest has a large area, and it is challenging to monitor deforestation activity by visiting all the observed forest area. Therefore, remote sensing technology is beneficial to detect and monitor deforestation. The most commonly found technique to detect deforestation is by using optical satellite imageries. The literature on detecting deforestation using Landsat can be found in works of literature [92]–[97]. The analysis of Southeast Asia tropical deforestation and Amazonian tropical deforestation exhibits different pattern and trend over the last decades in these two large tropical rainforest areas. The quantitative result from indicates the total forest degradation and loss of primary forest in Sumatra in the past two decades from 1990 is up to 7.54 Mha. More than 90 percent was the degraded state was cleared [4]. [98] Present the result of deforestation monitoring by using Landsat TM/ETM+ images from 1984 to 2004. From this study, it was found that deforestation rate for the period 1984-2004 was 2.47% yr(-1) in the 7295 km(2) study area, but decreased to 1.99% and 2.15% in 2000-2002 and 2002-2004, respectively. In other hands, efforts to detecting deforestation in tropical rainforest using optical satellite imageries remain challenging since in humid climate area. SAR images can penetrate the clouds and can provide clear images whenever there is a satellite pass, hence provide a potentially valuable tool for monitoring forest changes. [99] using the JAXA ALOS PALSAR L-band radar ScanSAR HH polarisation with repeat images every 46. Days, over Sumatra. Thus, the data providing much more frequent explicit data more than other satellite images by using the Fine-Beam Dual (FBD) image pairs with HH and HV polarisations. According to the study, the temporal analysis result shows that deforestation in Riau can be identified by large values of the temporal standard deviation, but high detection rates are associated with high false alarm rates, particularly in swamp forest. That is, after reflecting the current existing researches, there are several types of research aspect of deforestation research is left untouched. At this point, the background knowledge related to this study has been given. It has become obvious that it is still uncommon to find the research on how to interpret analysis result from remote sensing data to a semantic domain. Besides, in the domain of semantic computing, the method to compute Remote Sensing data content also never been discussed.
THE SEMANTIC COMPUTING FOR DEFORESTATION INTERPRETATION
With Semantic Dimensional Control

This Chapter:

• Shows the ability of semantic dimensional control to merge the SPA process into one semantic space.

• Describes the mechanism to control the value from one dimension to another dimension using Reinforcement Learning.

• Implementing the Processing function in one dimension to automatically detect deforestation.
3. The Semantic Dimensional Control: the Extended Version of The Mathematical Model of Meaning (MMM)

The mathematical model of meaning (MMM) [20] research has been actively applied to multimedia data. The mechanism of MMM applies as a fundamental framework for realising the semantic associative search for extracting information by giving its context words. The mechanism in MMM determines the correlation between content and context for realising dynamic context recognition. Dynamic context recognition is created using the mathematical foundation for obtaining similar content by giving the user’s impression and the contents of the information as a context. Thus, the noted point is that in MMM the mechanism of any computational procedure before context calculation has never been discussed. A semantic space in MMM is created as a space for representing various contexts which correspond to its subspaces. The computation recognises a context and then selecting a subspace that corresponding to it. The illustration of MMM can be found in Figure 3.1. As one of the promising applications of MMM to the nature, this study can be declared as a new proposal of the concept of deforestation interpretation with meanings.

![Figure 3.1](image.png)

**Figure 3.1.** The classical mathematical model of meaning originally proposed by Kiyoki, Y., *et al* [20].
The main objective of this study is to give the meaning to nature. To achieve the main objective, this study expands the MMM basic idea of dimensional control into a new framework that capable to interpret environmental conditions automatically. The framework is designed to assign the meaning into a specific natural object captured by the satellite images in regards to the deforestation phenomena. The main feature of this approach is that the semantic associative search is performed dynamically in the orthogonal semantic space in order to perform environmental analysis based on human context. The first step of the framework is to create the semantic space for computing the similarity between the meaning of environmental condition and measurement values derived by the satellite images which has the potential to capture the information about deforestation conditions. The ecological effects of major global environmental phenomena such as deforestation on natural resources vary from soil to atmospheric effects. That is, to delve into the consequences of deforestation phenomena, the advance system should cover the analysis of deforestation impacts on multiple natural resources and have the mechanism of context recognition. In this study, the advance system uses the expansion of the MMM which realises the dynamic context recognition for deforestation phenomenon. The deforestation interpretation result is presented based on a captured context, which represents the impact of a deforestation area.

The primary MMM mechanism of dimensional control is suitable to capture the context which very essential for environmental information retrieval. The relation between content values (the satellite image) and the contexts (deforestation impacts) determine the meaning of deforestation. The computation for realising dynamic context recognition is fundamentally important for multimedia information acquisition.

The advantages and original points of the MMM are as follows:

(1) The semantic associative media search based on a semantic computation for words is realized by a mathematical approach. This media search method surpasses the search methods which use pattern matching for associative search. Users can use their own words for representing impression and data contents for media retrieval and do not need to know how the metadata of media data of retrieval candidates are characterized in databases.

(2) Dynamic context recognition is realized using a mathematical foundation. The context recognition can be used for obtaining multimedia information by giving the user's impression and the contents of the information as a context. A semantic space is created as a space for representing
various contexts which correspond to its subspaces. A context is recognized by the computation for selecting a subspace.

The MMM procedures can be summarized as follows according to the original paper [20]:

1. Creation of the semantic space:
   To provide the function of semantic associative search, basic information on m data items ("data-items for space creation") is given in the form of a matrix. Each data item is provided as fragmentary metadata which are independently represented one another. No relationship between data items is needed to be described. The information of each data item is represented by n features as n-dimensional vector. The m basic data items are given in the form of an m by n matrix $\mathbf{M}$. For given m basic data items, each data item is characterized by n features. Then, each column of the matrix is normalized by the 2-norm in order to create the matrix $\mathbf{M}$. The eigenvalue decomposition of $\mathbf{M}^T\mathbf{M}$ is computed as indicate in Eq. 3.1. By using this matrix $\mathbf{M}$, the orthogonal space is computed as the metadata space MDS.

$$
\mathbf{M}^T\mathbf{M} = Q \begin{bmatrix}
\lambda_1 & & \\
& \ddots & \\
& & \lambda_v \\
0 & & \\
& \ddots & \\
& & 0
\end{bmatrix} Q^T, 0 \leq v \leq n \tag{3.1}
$$

The orthogonal matrix $Q$ is defined by

$$Q = (q_1, q_2, \ldots, q_n)^T$$

We call the eigenvectors "semantic elements." Here, all the eigenvalues are real and all the eigenvectors of $\{q_1, q_2, \ldots, q_v\}$ are mutually orthogonal because the matrix $\mathbf{M}^T\mathbf{M}$ is symmetric. The orthogonal semantic space MDS is created as a linear space generated by linear combinations of $\{q_1, q_2, \ldots, q_v\}$. Where that $\{q_1, q_2, \ldots, q_v\}$ is an orthogonal basis of MDS.

2. Representation of information resources and contexts in n-dimensional vectors:
   Each of the information resources is represented in the n-dimensional vector whose elements correspond to n features used in Eq. 3.1. These vectors are used as "metadata for information resources". The information resources are the candidates for semantic associate search in this
model. Furthermore, each of context words, which are used to represent the user's impression and data contents in semantic associative search, is also represented in the $n$-dimensional vector. These vectors are used as "metadata for contexts."

3. Mapping information resources into the metadata space $MDS$:

Metadata items (data-items for space creation, metadata for information resources and metadata for context words) which are represented in $n$-dimensional vectors are mapped into the orthogonal metadata space.

4. Semantic associative search:

When a sequence of context words which determine the user's impression and data contents is given, the most related information resources to the given context are extracted from information resources in the metadata space by selecting and applying one of the metrics defined in the semantic space. (The most correlated information resources to the given context are extracted in the selected subspace.)

Several information retrieval methods, which use the orthogonal space created by mathematical procedures like Latent Semantic Indexing (LSI) [100], have been proposed. The MMM is essentially different from those methods using the Singular Value Decomposition (SVD) (e.g. the Latent Semantic Indexing (LSI) [100], and Principal Component Analysis (PCA) [101]) method. The essential difference is that the MMM model provides the important function for semantic projections which realizes the dynamic recognition of the context. The concept the MMM is reflecting this dynamic subspace selection where all dimension is equally important before the context is given. The weighting will be calculated and the most related dimension will have more priority after the context is given and this realizes the dynamic recognition of the context. That is, the context-dependent interpretation is dynamically performed by computing the distance between different media data, information resources and words. While the concept of the weighting in the statistical method like PCA is to measure how important a particular dimension is in the beginning, so the high variance axes are treated as principal components and low variance axes are treated as noise, in MMM all dimensions treated as potential information in the different dimension after subspace selection.
The context-dependency in MMM works as a weighting procedure for dynamically selecting a subspace from the entire orthogonal semantic space, according to a context. However, the original MMM works for processing procedure, when all initiated dimension is used for processing procedure. In this study, the original version of MMM is extended to tackle the redundancy process in the current computational method where all processing procedure, such as sensing and processing are done separately. Sensing or data acquisition, processing the datasets to retrieve the meaning, and visualise the result is done sequentially and design for achieving one specific result. With this kind of computation method, sensing, processing, and actuation (SPA) have completely to be done in order to achieve one desired research result. In this study, an extended function of MMM called semantic dimensional control is proposed. The concept of semantic dimensional control is to combine all the process, starts from sensing, processing, and actuation into one semantic space. The concept of this model can be found in Figure 3.2.

Hence, A semantic space is created as space can handle any machine learning algorithm for data processing before representing various contexts which correspond to its subspaces. The advantages and original points of the proposed methodology are as follows:

1. Allows the Sensing, Processing, and Actuating in the same space based on the concept of dimension control. With this mechanism, processing the data with any machine learning algorithm can be easily done in one subspace. The learning results/contents such as clustering or segmentation can be projected in any different dimension for giving words or meaning realized by a mathematical approach to those results/contents. Since the semantic space is dynamic, this new extension enables users to add their own words for representing learning results/contents for retrieval, and do not need to know how the contents retrieval candidates are characterized are characterized in the system.

2. Dynamic context recognition is created using a mathematical foundation. The context recognition can be used for obtaining information by giving the user’s context and the contents of the information as a context. A semantic space is created as a space for representing various contexts which correspond to its subspaces. A context is recognized by the computation for selecting a subspace.
The extended Version of MMM Consists of:

1) A set of features is given which includes the feature of media content, such as x and y axis in remotely sensed image. Each feature is normalized. Features in dataset is referred as dimension in Semantic Space. That is, an \( n \) dimension of semantic space is created.

2) Dimensional controls for utilizing machine learning or expert knowledge is applied to process the remotely sensed image. In term of image segmentation, a subspace of the orthogonal semantic space corresponding to remotely sensed image is selected according to the feature user need to produce segmented/clustered data.

3) The representative point of segmented/clustered data is computed. This representative point is used as the data projection to another semantic space in order to give dynamic meaning to each segment/cluster.

4) Performs the standard procedures of MMM:
   a. A set of \( m \) words is given, and each word is characterized by \( n \) features. That is, an \( m \) by \( n \) matrix is given as the data matrix.
   b. The correlation matrix with respect to the \( n \) features is constructed. Then, the eigenvalue decomposition of the correlation matrix is computed and the eigenvectors are normalized. The orthogonal semantic space is created as the span of the eigenvectors which correspond to nonzero eigenvalues.
   c. Images and context words are characterized by using the specific features(words) and representing them as vectors.
   d. The images and context words are mapped into the orthogonal semantic space by computing the Fourier expansion for the vectors.

5) A set of all the projections for interprets environmental condition from the orthogonal semantic space to the invariant subspaces (eigenspaces) is defined in the ellipsoidal form. Each subspace represents a phase of meaning, and it corresponds to a context or situation.

6) A subspace of the orthogonal semantic space is selected according to the user’s context on the environmental phenomena’s context, which are given as a context represented by a sequence of words.

7) Apply the Closest Semantic-Ellipsoid Algorithm to calculate the most correlated environmental phenomena to the given context is extracted in the selected subspace.
The concept of Semantic Dimensional Control for Environmental Interpretation

![Diagram of Semantic Dimensional Control](image)

**Figure 3.2.** Semantic dimensional control: merging physical space and meaning space into one semantic space.
3.1. Semantic Dimensional Control for Deforestation Interpretation

The idea of semantic dimensional control is to enable the sensing, processing and actuating (SPA) media contents in one single semantic space, illustration given in Figure 3.2. Therefore, the semantic space has to contain all the information in advance to handle all of these processes. Since there is no limitation to the number of space that a semantic space can hold, it is possible to register as many information as possible. The application of semantic dimensional control in this study is focused on interpreting deforestation activity. Deforestation is the process of clearing and changing the main function of forest and one of major global change happens in several decades.

For interpreting deforestation condition, first, we have to define the sensing data. In this regard, remotely sensed image, L-Band SAR and Landsat is used. The processing procedure requires the segmentation computation to found deforestation area since in original remotely sensed image has no distinct information about deforestation area. Then, after deforestation has been found, the dynamic meaning of each deforestation is calculated using the Closest Semantic-Ellipsoid Algorithm in semantic space. Lastly, the actuation also can be performed in semantic space. In semantic deforestation system, the actuation is a facility where the user can interact with the system to retrieve the deforestation area with a given context.

3.1.1. The Construction of Semantic Space for Applying Semantic Dimensional Control to Interpreting Deforestation Activity

The Semantic space constructed from the essential parameters derived from remotely-sensed data, L-Band and Optical image respectively. A data matrix $M$ is created in regards to semantic space for deforestation activity. The big semantic space combining the content space from remotely sensed data: spatial and temporal space with meaning space or basic words that are used to explain all the vocabulary of deforestation condition. In meaning space, when $m$ terms are given as the vocabulary entries in the deforestation analysis and $n$ features are utilized to explain all these terms, an $m$ by $n$ matrix $M$ is constructed for the space creation. Each term in the matrix $M$ is characterized by the $n$ features $(f_1, f_2, \cdots, f_n)$. For given $d_i = (1, 2, \cdots, m)$, the data matrix $M$ is defined as the $m \times n$ matrix whose $i$-th row is $d_i$. 
Therefore, the dimension of deforestation semantic space describes as follows:

1. *Gamma-nought* ($\sigma^0$) from HH Polarization; value for measuring backscattering coefficient from SAR data derived from microwave satellite. HH - for horizontal transmit and horizontal receive.

2. *Gamma-nought* ($\sigma^0$) from HV Polarization; value for measuring backscattering coefficient from SAR data derived from microwave satellite. HV - or horizontal transmit and vertical receive.

3. *Band 1;* 45-0.515µm (Visible) wavelength for aerosol effect detection

4. *Blue filtered* spectral; 45-0.515µm (Visible) wavelength. This axis of the spectrum is used to analyse the characteristics of the water, the land-use, the soil, and the plants. This part is easy to be affected by air diffusion.

5. *Green filtered* spectral; This axis is used to models how the brain interprets the context based on how dominant the colour is, for example, has wavelength 0.52-0.60µm (Visible): this axis of the spectrum reflects the chlorophyll.

6. *Red filtered* spectral; wavelength 0.63-0.6100 meter resolution, Improved thermal mapping and estimated soil moisture (Visible): this red filter reflects the chlorophyll of the plants with green leaves, makes the dominant of received red scattering is important to get data.

7. *infrared filtered* spectral is part of invisible wavelength. This axis provides the information that barely seen by human eyes. Has wavelength 0.76- 0.90µm makes this filter important for ecology

8. *Short-wave Infrared (SWIR) 1;* To discriminate the content of soil and biomass

9. *Short-wave Infrared (SWIR) 2;* To improve the soil moisture detection

10. *Panchromatic Band;* Improve the resolution of optical satellite image

11. *Cirrus Band;* Eliminate the cloud effect on tropical forest

12. *Band 10 – TIRS 1;* 10.60- 11.19µm and has 100 meter resolution, thermal mapping and estimated soil moisture

13. *Band 11 – TIRS 2;* 11.50 - 12.51µm and has 100 meter resolution, thermal mapping and estimated soil moisture

14. *Biomass cover;* Total vegetation cover remaining in deforested area
15. *Soil Salinity*; Indicates the soil chemical property such as Na, K, Ca, Mg.

16. *Soil temperature*; surface temperature of the deforested area also needs to be measured. Standard Landsat 8 product acquired not only by the Operational Land Imager but also by Thermal Infrared Sensor.

17. *Velocity of deforestation*; The speed of change from forest to non-forest area.

18. *Shape of deforestation*; Deforestation driver can be identified by its speed of change. The primary driver of deforestation is either by a forest fire or illegal logging. As both activities have extremely different behaviour to forest clearance, the observation of speed of change would be one of the essential parameters to observe the degree of deforestation change. The formal definition of deforestation speed of change is the rate at which deforestation area spreads.

    The speed of wildfire in the forest can move as fast as 22 kilometres per hours.

19. *Density*; The total density of point change from forest to non-forest.

20. *Elevation*; A digital elevation model is a bare-earth raster grid referenced to a vertical datum

The list of intervals of the dimension explained above is represented as Table 3.1 and illustrated in Figure 3.3.

*Figure 3.3. Deforestation semantic space.*
Table 3.1. Summary of Deforestation Semantic Space

<table>
<thead>
<tr>
<th>Axis</th>
<th>Interval</th>
<th>Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 1 – Coastal Aerosol</td>
<td>0.435 - 0.451</td>
<td>Coastal and aerosol studies</td>
</tr>
<tr>
<td>Blue</td>
<td>125 - 210</td>
<td>distinguishing deciduous from coniferous vegetation</td>
</tr>
<tr>
<td>Green</td>
<td>110 - 225</td>
<td>Emphasizes peak vegetation</td>
</tr>
<tr>
<td>Red</td>
<td>130 - 220</td>
<td>Discriminates vegetation slopes</td>
</tr>
<tr>
<td>Near Infrared (NIR)</td>
<td>120 - 225</td>
<td>Emphasizes biomass content and shorelines</td>
</tr>
<tr>
<td>Short-wave Infrared (SWIR) 1</td>
<td>0 - 255</td>
<td>Discriminates moisture content of soil and vegetation; penetrates thin clouds</td>
</tr>
<tr>
<td>Short-wave Infrared (SWIR) 2</td>
<td>0 - 255</td>
<td>Improved moisture content of soil and vegetation and thin cloud penetration</td>
</tr>
<tr>
<td>Band 8 - Panchromatic</td>
<td>0.503 - 0.676</td>
<td>15 meter resolution, sharper image definition</td>
</tr>
<tr>
<td>Band 9 – Cirrus</td>
<td>1.363 - 1.384</td>
<td>Improved detection of cirrus cloud contamination</td>
</tr>
<tr>
<td>Band 10 – TIRS 1</td>
<td>10.60 – 11.19</td>
<td>100 meter resolution, thermal mapping and estimated soil moisture</td>
</tr>
<tr>
<td>Band 11 – TIRS 2</td>
<td>11.50 - 12.51</td>
<td></td>
</tr>
<tr>
<td>HH Polarization</td>
<td>(-14) - (-10)</td>
<td>Discriminates vegetation biomass</td>
</tr>
<tr>
<td>HV Polarization</td>
<td>(-13.5) - (-9.5)</td>
<td>Emphasizes forest region</td>
</tr>
<tr>
<td>Biomass</td>
<td></td>
<td>Total vegetation cover remaining in deforested area</td>
</tr>
<tr>
<td>Salinity</td>
<td></td>
<td>Indicates the soil chemical property such as Na, K, Ca, Mg.</td>
</tr>
<tr>
<td>Soil Temperature</td>
<td></td>
<td>surface temperature of the deforested area also needs to be measured.</td>
</tr>
<tr>
<td>Velocity</td>
<td>0 to &gt;30 (km/day)</td>
<td>Retrieve speed of deforestation change</td>
</tr>
<tr>
<td>Shape</td>
<td>[line, geometric, fishbone]</td>
<td>Discriminate structure and pattern deforestation area</td>
</tr>
<tr>
<td>Density</td>
<td>[dense, sparse]</td>
<td>Discriminate type of change</td>
</tr>
<tr>
<td>Elevation</td>
<td>0 to &gt;20</td>
<td>Topography mapping</td>
</tr>
</tbody>
</table>
3.2. Sensing Dimension for Deforestation Analysis

After defining the space in single semantic space, it is then possible to perform the sensing procedure which applied in sensing dimension. The sensing procedure is corresponding to the processing of L-Band SAR from ALOS-2/PALSAR-2 satellite image to detect deforestation area. In order to provide accurate detection of deforestation area, it is necessary to consider the scientific details of remotely sensed image behaviour and how it responds to the change of natural object before and after deforestation. All the procedures of sensing including pre-processing, normalisation, and geometric correction are performed in the sensing dimension of single semantic space as explained in the previous sub-chapter. This sub-chapter will explain briefly about the mechanism and procedure of sensing in dimensional control. Figure 3.4. illustrates the implementation of dimensional control in sensing dimension. The dataset maps into sensing process contain sensor data, ancillary data, and validation data.

The concept of Semantic Dimensional Control for Environmental Interpretation

![Diagram of Sensing process in semantic deforestation interpretation space.](image)

**Figure 3.4.** Sensing process in semantic deforestation interpretation space.
Temporal dataset from ALOS-2/PALSAR-2 is mandatory to produce deforestation detection. Deforestation is not an area that can be detected only by a single temporal image. In this discussion, the L-Band SAR from ALOS-2/PALSAR-2 referred to as the dataset. The conventional method to detect deforestation is to detect the area of change from forest object to non-forest object in dataset $t_1$ (temporal dataset before deforestation) and dataset $t_2$ (temporal dataset after deforestation). However, the simple differential computation between $t_1$ and $t_2$ is not enough. This sub-chapter indicates how other parameters such as seasonal change, type of canopy, or even the growth period of plants can affect the deforestation detection. The further discussion in this sub-chapter covers the proposal of a new algorithm called DELSAR, the abbreviation for deforestation detection using L-Band SAR, which detecting deforestation with the consideration of these important parameters in the differential calculation.

a. Data Acquisition

This study utilizes L-Band SAR from ALOS-2/PALSAR-2 as the primary dataset to detect deforestation area. The ancillary dataset is used as complementary to produce a comprehensive result. In the following discussion, the process to utilise primary data set and how to put mask a classification raster as the secondary dataset is explained. Masking process is important to mark the forest regions as the area of interest.

*Primary dataset: ALOS-2/PALSAR-2 dataset*

As defined in the overview of this chapter, ALOS-2/PALSAR-2 dataset provided in this study are Dual Polarization consist of HH and HV polarisation, and HV polarisation is used mainly for identifying the forest characteristic. The automatic deforestation algorithm solely relies on backscattering value from HV. It is because HV dataset tends to show more sensitivity toward the change in the forest region. Therefore it is significantly useful to identify the deforestation. In addition to the polarisation, several aspects should be taken into account to determine the most appropriate characteristic of the dataset. One of the aspects is the observations mode ALOS-2/PALSAR-2. The three major observation modes can be seen from Figure 3.5, in which can be brake down into spotlight, stripmap and ScanSAR. The other observation modes have higher resolution but lower temporal coverage which makes data acquisition become significantly difficult. However, despite the advantage, this type of dataset remains challenging for making a high precision deforestation detection result with simple differential computing because of the
changes that can be captured occurs in the small number of pixels. Variability in vegetation structure has such a significant impact on long-wavelength radar backscatter intensity that affects the interpretation of temporal change detection.

<table>
<thead>
<tr>
<th>Observation mode</th>
<th>Spotlight</th>
<th>Ultrafine [3m]</th>
<th>Stripmap</th>
<th>High sensitive [6m]</th>
<th>Fine [10m]</th>
<th>ScanSAR</th>
<th>Normal</th>
<th>Wide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth (MHz)</td>
<td>84</td>
<td>84</td>
<td>42</td>
<td>28</td>
<td>14</td>
<td>28</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Resolution (m)</td>
<td>3 x 1 (Rg x Az)</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>100</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incidence angle (deg.)</td>
<td>8 - 70</td>
<td>8 - 70</td>
<td>8 - 70</td>
<td>20 - 40</td>
<td>8 - 70</td>
<td>23.7</td>
<td>8 - 70</td>
<td>8 - 70</td>
</tr>
<tr>
<td>Swath (km)</td>
<td>25 x 25 (Rg x Az)</td>
<td>50</td>
<td>50</td>
<td>40</td>
<td>70</td>
<td>30</td>
<td>350 (5 scans)</td>
<td>490 (7 scans)</td>
</tr>
<tr>
<td>Polarization*</td>
<td>SP</td>
<td>SP/DP</td>
<td>SP/DP/CP</td>
<td>FP</td>
<td>SP/DP/CP</td>
<td>FP</td>
<td>SP/DP</td>
<td>SP/DP</td>
</tr>
<tr>
<td>NESZ (dB)</td>
<td>-24</td>
<td>-24</td>
<td>-28</td>
<td>-25</td>
<td>-26</td>
<td>-23</td>
<td>-26</td>
<td>-26</td>
</tr>
<tr>
<td>S/A (dB)</td>
<td>Rg</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>25</td>
<td>20</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Az</td>
<td>20</td>
<td>25</td>
<td>20</td>
<td>20</td>
<td>23</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

* SP: HH or HV or VV, DP: HH+HV or VH+VH, FP: HH+HV+VH+VV, CP: compact pol. (experimental)

Figure 3.5. ALOS-2 specifications; http://www.eorc.jaxa.jp/ALOS-2/en/about/palsar2.htm

The straightforward method for detecting deforestation using ALOS-2/PALSAR-2 data is to detect significant changes in the time series difference of the backscattering coefficient by using a threshold value, which referred as simple differential computing in the previous discussion. Inasmuch as the backscattering value in forest relatively has a stable range value, this value range is still highly correlated to the forest location, type of forest, and above ground biomass (AGB). Standard threshold for global forest region has been studied and shows precise knowledge on backscattering range. The highest forest region backscattering variability of L-Band SAR was found in Amazon, with the range between -11.24 to -16.75. This is related to richer variability of forest species. The small change of AGB or other environmental condition, in the Amazon forest region, might visible in ALOS-2/PALSAR-2 within this range. Lead to the possibility of deforestation detection miss-detection result if thresholding value is uniformly applied. The further observation indicates the behaviour of the backscattering coefficient in the forested and deforested area also contributed by ground surface moisture and residual woody debris left after
felling. If physical forest characteristics determine backscatter value, the aspect of how to define the characteristics of forest changes are vital to increasing our understanding of deforestation situation. Considering not only the temporal change of backscattering value but also the characteristic of backscattering change could improve our knowledge on deforestation activity.

**Ancillary dataset**

Three following ancillary datasets are used as the complementary dataset to analyse the change detected in ALOS-2/PALSAR-2 dataset.

1) *The global forest/non-forest map (FNF)*

The FNF is a mask map for forest region produced by JAXA. The classification algorithm produces the FNF map by performed to 25m resolution ALOS-2/PALSAR-2 mosaic (as shown in Figure 3.6.). The high and low backscatter in HV is classified as "forest" region and "non-forest" region. The term forest used in the FNF map refers to FAO definition in which describe that forest is the natural forest with the area larger than 0.5ha and forest cover over 90%. Since the radar backscatter from the forest depends on the region (climate zone), the classification of Forest and non-forest regions is conducted by using the region dependent threshold of backscatter.

![Figure 3.6](http://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/fnf_index.htm)
2) *ALOS DSM*

The 30-meter resolution digital surface model (DSM) is the global-scale elevation data derived by Japanese Satellite, ALOS by JAXA (Japanese Aerospace Agency). The excellent thing about ALOS DSM is that currently, by the time this study is written, is the most precise global-scale elevation data based on the stereo mapping from PRISM of ALOS.

3) *The Shuttle Radar Topography Mission (SRTM)*

SRTM is topographical information that is fundamental to many geospatial and/or Earth observation applications and analysis which is initiated by NASA (National Aeronautics and Space Administration) and NGA (National Geospatial-Intelligence Agency)

*Validation Dataset*

*The Global Land Analysis and Discovery* (GLAD) is the dataset produced by the Department of Geographical Sciences laboratory at the University of Maryland. The dataset is dedicated to investigating methods, causes, and impacts of the Land-use change in global scope. The primary dataset used to produce the GLAD dataset come from the satellite images. The GLAD project has many expert partners as their collaborator such as the World Resources Institute to create global operational data on forest extent and change, as a part of the Global Forest Watch project that can be checked on http://www.globalforestwatch.org/. The GLAD dataset is commonly used for validation for forest algorithm developer due to the number of human resources that frequently checked the dataset. Although the dataset itself contains detection error of deforested places, however, there is no other better evaluation dataset are available right now. The sample of GLAD data availability can be seen in Figure 3.7.

![Figure 3.7. Global Land Analysis and Discovery (GLAD) result used as ground truth data measurement.](image-url)
b. Study Area

Deforestation happens mostly in tropical countries. To assess the performance and stability of the proposed idea, this study applies the proposed system and methodology to three different rainforests in two states, Peru and Indonesia. The actual broader target application of this study is the global-scale tropical rainforest, the feature for analysing and learning the spatial and temporal difference for detecting deforestation should be robust for different tropical forest. However, due to the limitation of research boundaries, this study observes only three different places. The utilised data is ALOS-2/PALSAR-2 between 2016 and 2017 with the 46-day interval of each data. The data is analysed retrospectively by the proposed procedures and done from scratch using python. Each data is pre-processed beforehand by using Improved Lee Filter using 7x7 window size.

Peru, Peruvian Amazonian

Located in Western South America, The Peruvian Amazon is highly diverse both ecologically and ethnically, home to approximately 65 different Indigenous groups. Over the last decades, illegal logging has become a severe problem in the Peruvian Amazon. In 2012 the World Bank estimated that 80% of Peru’s timber exports are illegally harvested. Peru as a study area, to simplify the explanation of the experimental study. Half of Peru territory is covered by dense forest. Deforestation rate is likely to grow from past decade which makes Peru as one of the good study areas to observe the deforestation phase using temporal difference change. The location and area of Peru can be seen from Figure 3.8.

![Study Case: Peru](image)

**Figure 3.8.** Peru as study area.
Indonesia, Kalimantan Tengah

Indonesia is a developing country located in Southeast Asia. Indonesian government policies used to favour forest-based industrialisation. The total forest area lost from long term deforestation in Indonesian since 1980 is up to 72 per cent of total forest cover. Large multinational pulp companies have cleared large areas of forest in Indonesia. The burning of forests and logging have been the most popular way to change the land use from forest to non-forest application makes Indonesia the world's third largest emitter of CO₂. The location and area of Indonesia can be seen from Figure 3.9.

![Study Case: Indonesia](image)

**Figure 3.9.** Indonesia as study area.

3.2.1. Learning the Pattern of Deforestation Area

a. Analysis and observation: characteristic captured by ALOS-2/PALSAR-2

The primary way to understand the changes over earth surface using ALOS-2/PALSAR-2 data is by observing the backscattering value in forest region over different temporal data. The basic interpretation of SAR data is that the backscattering intensity value is corresponding to the roughness surface. The rougher the monitored surface, the stronger the signal can be captured by the satellite sensor. In forest region, the backscattering value varies depending on forest type, canopy, wind, and growth rate, although it is likely to be in between -11 to -15 dB (please refer to Figure 3.10 for the actual dataset comparison). The sensitivity of ALOS-2/PALSAR-2 data over the roughness of observed objects can be a challenge to understanding the changes on the earth
surface and sometimes leads to miss-detection of deforestation area. The usual miss detection of deforestation occurs if the algorithm detects a slight change of backscattering value even it might be because of sensor fault or some small changes due to seasonal changes.

Figure 3.10. Sensitivity of ALOS-2/PALSAR-2 data.

b. Tropical rainforest characteristic

The tropical rainforest is located at tropical and subtropical latitudes. Though these forests occur in climates that are warm year-round and may receive several hundred centimetres of rain per year. The tropical rainforest is characterised by an enclosed canopy and high species diversity, the definition of the forest itself is varies depending on the country. In Indonesia, there are three types of forest commonly seen. Wet tropical forest, dry tropical rainforest, and commercial agriculture which has a dense canopy, such as rubber and acacia. Interestingly, when we look at the
backscattering variance over these forest type regions, we could find slightly different backscattering value as shown in Figure 3.11. This study assesses the variability of backscattering of different forest type in Indonesia. This study found that the variability of the primary wet forest is much lower than the backscattering variability in the primary dry forest. It is because generally wetter conditions the backscattering value is reduced in L-Band HV, primarily through double bounce interactions between the trunks and ground and direct ground returns. This finding profoundly has a similar result with the dedicated research on the soil moisture impact on L-Band SAR as found in [17].

<table>
<thead>
<tr>
<th>Forest Type</th>
<th>PALSAR-2</th>
<th>Optical Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Dry Forest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma naught(dB)</td>
<td>-13.76~ -11.02</td>
<td>© Landsat-7</td>
</tr>
<tr>
<td>Primary Wet Forest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma naught(dB)</td>
<td>-14.50~ -12.47</td>
<td>© Landsat-7</td>
</tr>
<tr>
<td>Commercial Agriculture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma naught(dB)</td>
<td>-15.8~ -14.21</td>
<td>© Landsat-7</td>
</tr>
</tbody>
</table>

**Figure 3.11.** L-Band backscattering value characteristic based on tropical rainforest types.

c. Seasonal change behaviour on tropical rainforest

The tropical rainforests normally are located in a climate where there is no dry season throughout the year. All the time the rainforest have average of 60 mm precipitation at least. Therefore tropical
Tropical Forest Seasonal Change

Research Challenge

<table>
<thead>
<tr>
<th>Seasonal Period</th>
<th>PALSAR-2</th>
<th>Optical Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry season</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-12.46~ -11.89</td>
<td></td>
</tr>
<tr>
<td>Wet Season</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-13.77~ -11.91</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.12.** Seasonal L-Band backscattering value characteristic on tropical forest.

d. **Deforestation stage**

The deforestation stage means the different stage in forest transition. It is started from the undisturbed forest as stage 0 in deforestation stage. The first stage indicates the preparation of agricultural industrial where all vegetation is destructed or burned, but a large amount of woods and unburned biomass remains on the ground. In the second stage, the clearing process of
remaining biomass is done to complete the land use change to some agricultural industrial or other purposes. The second process also can be considered as an expansion period. The third or final stage is the condition of the fully prepared deforested area. By this stage, the land is completely clean from the forest biomass, and carbon that once locked in the forest biomass has been released to the atmosphere. The stage of deforestation can be detected using ALOS-2/PALSAR-2 data HV polarisation because the sensitivity of HV to surface roughness. However using HV polarisation only, ALOS-2/PALSAR-2 requires more than two months to identify deforestation. This is because of the backscattering effect of woody debris on the ground after log or slash burn. Some small land change might happen in a short time, and make the algorithm under-detect the activity. Comparison of deforestation stage capture by L-Band SAR is indicated in Figure 3.13. In this study, the stage of deforestation captured by HV polarisation is studied. The value of backscattering fluctuation happens in different deforestation area is recorded to provide better deforestation detection as well as to give meaning to the deforested area based on the stages.

**Figure 3.13.** Deforestation stage characteristic.
To bring a clear comparison toward tropical rainforest characteristic and behaviour depending on the spatial and temporal correlations, Figure 3.14., indicates the value difference between tropical rainforest in Indonesia and Peru as well as the seasonal effects on backscattering value of tropical rainforest as a bar chart. It is shown that each forest has different backscattering characteristic. After looking at the behaviour of forest and its characteristic in the spatial and temporal domain, it is important to consider on these parameters before proceeding ALOS-2/PALSAR-2 data to detect the change in forest region and identified that changes as deforestation activity.

![Backscattering Value](image)

Figure 3.14. Bar plot of backscattering value of three forest types in tropical countries.

3.2.2. Deforestation Detection with L-Band SAR (DELSAR) Algorithm

To keep the discussion focused on the dimensional control mechanism of expansion of MMM, this study process deforestation detection in the sensing dimension. The processing in the subspace of L-Band SAR is utilised to detect deforestation area and produce the shapefile of deforestation area.
The focused on this sub-chapter is to detect deforestation area automatically. In regards to the previous discussion about satellite images, it is shown in the literature in Chapter 2 that radar image is more robust to measure objects on earth especially tropics because of its penetration characteristic. Therefore, to detect deforestation, ALOS-2/PALSAR-2 ScanSAR dual-polarisations data were used. In this subchapter, a new algorithm [102] specifically made to process L-band SAR will be discussed. The outline of the algorithm includes three modules of processing. The temporal difference operator takes two ALOS-2/PALSAR-2 imageries which are $X_1 = \{X_1(i,j), 1 \leq i \leq H, 1 \leq j \leq W\}$ and $X_2 = \{X_2(i,j), 1 \leq i \leq H, 1 \leq j \leq W\}$. Where H and W are the height and width of the ALOS-2/PALSAR-2 imageries, respectively, that are acquired over the same area at different times, indicates as cycle $c_1$ and $c_2$. As output, a new vector difference contains temporal difference is produced and saved in local memory. The diagram view of the system can be seen in Figure 3.15.

![Diagram view of the proposed Density L-Band Algorithm (DELSAR).](image)

The procedure for deforestation detection using ALOS-2/PALSAR-2 data is as follow:
1. Find N representative magnitude of temporal change, hereafter will be referred as “peak points” of temporal change with the calculation from Eq. 3.2.

\[
P_{ij} = \mathbb{1}_{ij} \in \left\{ \sum_{i=1}^{N} p \left( \mathbb{1}_{ij} - \mu \right)^2 * \alpha \right\}, \text{where } \sum_{i=1}^{N} p_{ij} x_i \tag{3.2}
\]

where \( P_{ij} \) is the selected peak point in which temporal difference \( \mathbb{1} \) should be lower than standard deviation multiply by constant variable of ALOS-2/PALSAR-2 forest region threshold

2. For each point in peak point, assign \( r \) radius around the peak point

3. If value of gamma naught observed temporal (\( Q \)) (refer to temporal 2(c2)) lays between given threshold probability of ALOS-2/PALSAR-2 amplitude, then grow a member point.

\[ \forall P, Q: \text{if } Q \in r \text{ is radius density-connected member from } P \text{ wrt. } r \text{ and } \text{min}_\text{changes, then } Q \in \text{Pneigbors} \]

4. count number of point within the neighbors in \( r \) radius

5. Check the number of points within the \( r \) radius

- If number of member points are more than \( \text{min}_\text{changes} \), recursively expand the area by applying the same criteria to the points inside the \( r \) radius. Then mark peak point and its member as deforestation area

- If number of member points are less than \( \text{min}_\text{changes} \), but consist of two or more points, then mark the peak point and its member as natural change

- If peak point has no member of points then mark the peak point as noise.

6. If two peak points are connected by a member point, group to peak points and its member into one group. Append peak points to array, multiple peak points should be memorized.

7. Calculate total area of membership union of “peak point” and its member if the mark is deforestation area.

8. If a stopping criterion (e.g., maximum length of detected peak points or iteration function) is not yet fulfilled, go to step 2
With the rule within \( r \) radius (procedure 5), tracking of “density structure” is achieved by characteristic of peak point as indicated by Figure 3.16. Obtaining representative magnitude of change by applying threshold uniformly to all ALOS-2/PALSAR-2 gamma naught value remains challenging. Although filtering has been applied prior temporal difference process, certain noise caused by nature object response is still remained and possibly lead to miss-detection of deforestation area. The definition of density structure here is the degree of compactness of a temporal difference (magnitude of change) between two temporal ALOS-2/PALSAR-2 imageries for identifying deforestation.

![Figure 3.16. Peak point derived from probability threshold (a) and core temporal point (P) in the middle of density expansion (b).](image)

To obtain reliable range, ALOS-2/PALSAR-2 imageries calibration and filtering is necessary to be applied prior to the differential process. The change magnitude of image difference is defined as

\[
\Delta d(i,j) = \| x_{ij}^{c1} - x_{ij}^{c2} \|. 
\]

Where \( x_{ij}^{c1} \) is a gamma naught value obtain in \( c_1 \) and \( c_2 \) respectively, and correspond to column \( i \) and row \( j \). Thus, since the gamma naught value lays within negative value, a higher value in pixel obtained in vector difference normally corresponds to a high probability of change.

The algorithm of deforestation monitoring runs inside of the same single deforestation semantic interpretation space as shown in Figure 3.17. This is indicating that after the sensing process, the processing process also been done in the same single space. The later discussion in this sub-chapter focused only on subspace \( x \), \( y \) and \( t \) in respect to physical space which aims to detect deforestation. All of the mentioned datasets proceeds in the space of sensing in single semantic space.
b. Implementation of DELSAR

The pre-processing phase includes the radiometric correction of remote sensing data. The atmospheric correction in the visible range (the case of Landsat) and speckle reduction of SAR data were carried out.

1. Pre-processing: ALOS-2/PALSAR-2 de-speckling process

The main drawback of SAR images is the presence of speckle, a signal dependent granular noise, inherent of all active coherent imaging systems, that visually degrade the appearance of images. Speckle noise is a variance of microwave backscattered signal due to the interaction of different terrain geometry, moisture, wavelength, polarisation, and view angle that interferes image interpretation for various applications. Speckle may severely diminish the performances of automated scene analysis and information extraction techniques, as well as it may be harmful in applications requiring multiple SAR observations, like automatic multi-temporal change detection. That is, the noise creates a lot of difficulty in interpreting the image. For these reasons, preliminary processing of real-valued detected SAR images aimed at speckle reduction is of crucial importance for many applications. Such a pre-processing, however, should be carefully designed to avoid spoiling useful information, such as the local mean of backscatter, point targets, linear features, and textures. This study compared four most well-known SAR filters: Lee Filter, Frost filter, enhanced Lee Filter and Mean Filter.
The Mean Filter: works by averaging all values within the filter window in SAR image. The concept of mean filter is to average the n-window speckle noise into one value. Therefore the concept it does not remove the noise from the image. Theoretically, in the term of SAR image, the bright and dark speckle pixels inside of the filter window should decrease the noise. However, the result of this filter often makes the image result appears blur, most of the time they loss the information details and reduce the spatial resolution.

The Frost filter: reduce the noises by using the concept of the minimum mean square error algorithm. The Frost filter kernel adjusts to the local statistical contents of the image. The goal is to suppress the speckle while at the same time maintaining the information on the edges.

Lee filter: The Lee filter is reportedly the first model-based filter devoted to reducing the speckle noise. Works by utilising the minimum mean square error algorithm.

Enhanced Lee filter or Lee-sigma filter: This filter is based on the Gaussian distribution model. It averages only the pixels within a certain standard deviation range. By the performances of each filter in an artificial homogenous test (see Figure 3.18.), the Enhanced Lee Filter is the optimal filter for speckle noise suppression (Table 3.2.). The filter is then applied to ALOS-2/PALSAR-2 data for forest region. The result of the filter comparison can be seen in Figure 3.19.

![Figure 3.18. An artificial test image; (a) original image (b) image with a fabricated speckle](image-url)
Table 3.2. Comparison of results of selected filtration of the artificial speckled image

<table>
<thead>
<tr>
<th>Filter</th>
<th>Window Size</th>
<th>Mean dif.</th>
<th>dif. std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee filter</td>
<td>4</td>
<td>5.6</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4.8</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>3.7</td>
<td>3.2</td>
</tr>
<tr>
<td>Frost filter</td>
<td>4</td>
<td>5.9</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4.7</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>4.0</td>
<td>3.6</td>
</tr>
<tr>
<td>Enhanced Lee filter</td>
<td>4</td>
<td>4.9</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4.1</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>2.9</td>
<td>3.8</td>
</tr>
<tr>
<td>Mean filter</td>
<td>4</td>
<td>5.3</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4.4</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>3.3</td>
<td>2.9</td>
</tr>
</tbody>
</table>

**Filter used for comparison**
- LEE Filter
- FROST Filter
- Enhanced LEE Filter
- Mean Filter

**Example Dataset**
PALSAR-2 Cycle 39 S08W075

Figure 3.19. SAR filter comparison of ALOS-2/PALSAR-2 data specified for rainforest region.
2. Segmentation and deforestation detection of ALOS-2/PALSAR-2 dataset

**Figure 3.20.** Five cycles of ALOS-2/PALSAR-2 data. This method detects the orange blocks indicating the captured deforestation activity.

Dataset was acquired from January 7th, 2016 to January 24th, 2017 with a 46-day interval of each data as shown in Figure 3.20. Implementation of our proposed procedure is done from scratch using python. Each image was filtered using Improved Lee Filter using 7x7 window size. Following application, results are given to demonstrate our model performance for detecting deforestation. The first initial step is to determine peak point to represent the significant change between c1 and c2. By using the rule 1 in Section 3.A, the result of peak point (white dots) as shown in Figure. 3.21. (c). After the stopping rule in Section 3, a chained member will be considered as one deforestation area as shown in Figure. 3.22. (a) red polygons indicate the remaining highest density which considered as deforestation area, and white points are peak points that were indicated as noise. Figure 3.22. (b) shows the speed of change, where the highest speed is referred to the forest fire. Purple polygons have the speed of change in about 2.9 km/day. Figure 3.22. (c) indicates the road has been expanded through forest region as indicated in the orange-coloured circle.
Figure 3.21. Application of DELSAR on Peru whereas (a) cycle 59 (c1); before deforestation (b) cycle 63 (c2); after deforestation (c) temporal difference with peak point of c1 and c2.

Figure 3.22. Shapefile result derived from our proposed idea (a) structure of density (b) speed of change; purple indicate fast speed change (c) expansion pattern indicate new road is establish.
3. Polygonise and producing the deforestation area detection

After detecting the deforestation area in the previous session, the next activity is to extract the information. Here is the candidate of deforestation will be injected by any detailed spatial information such as location, longitude latitude, the total cover of deforestation area, the id and priority even the accuracy of detected deforestation area. Then, polygonise the detected deforestation area to serve the data into the system as shown in Figure 3.23.

![Figure 3.23. Result of detection of deforestation area research activity.](image)

When observing deforestation area in the tropical country, the use of ALOS-2/PALSAR-2 imagery is advantageous since it can penetrate cloud and observe the object in day and night time. Therefore, this study aims to construct new method temporal difference learning of SAR data for identifying deforestation stage. This study outlines two further processing steps which capable to automatically detects deforestation and learns the temporal contrast of deforestation area. With the combination of this two process, the excellent accuracy deforestation detection using ALOS-2/PALSAR-2 data could be achieved. The result of the system is represented in Figure 3.24.
Figure 3.24. The result of proposed method.
4. The implementation performance of DELSAR and the performance measurement with other deforestation detection

The DELSAR has been implemented to detect the deforestation on Peru and Indonesia. In this study, we analyse the performance using Kappa analysis. Kappa is an index that considers observed agreement with respect to a baseline agreement. The formula of kappa is in Eq. 3.3

$$\kappa \equiv \frac{P_o - P_e}{1 - P_e} = 1 - \frac{1 - P_o}{1 - P_e}$$

(3.3)

where $p_o$ is the relative observed agreement among user and the analysis result (identical to accuracy), and $p_e$ is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly seeing each category. The physical interpretation of this kappa evaluation would be that this is how close, the class label result drawn by our proposed neural network gets to the actual GLAD data as the ground truth data as shown in Confusion Matrix of Result and GLAD data in Table 3.3 and the performance of DELSAR can be seen in Figure 3.25.

<table>
<thead>
<tr>
<th></th>
<th>Deforestation</th>
<th>Non-Deforestation</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deforestation</strong></td>
<td>1995</td>
<td>588</td>
<td>2583</td>
</tr>
<tr>
<td><strong>Non-Deforestation</strong></td>
<td>1761</td>
<td>236656</td>
<td>238417</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>3756</td>
<td>237244</td>
<td>241000</td>
</tr>
</tbody>
</table>

Table 3.3. Confusion Matrix of proposed DELSAR
3.2.3. Representative Point of Deforestation Area

After feature extraction, each deforestation areas are created and stored in the data repository in shapefile format as shown in Figure 3.25. These deforestation areas are used for the matching process to the metadata of a deforestation query based on context. Semantic computing in this study is used as an effort to give meaning to content. Semantic computing utilises semantic operator to acquire the closest meaning to the content. Meanings of a domain knowledge projected into semantic space are called semantic projections. There are several existing variations of semantic operator determined by the type of problems. Points represent average semantic projection makes the utilisation of general distance calculation between points is commonly utilised.

**Figure 3.25.** measurement of agreement between user and experimental result
Figure 3.26. Projecting representative value to semantic space.

3.3. Dimensional Control: Semantic Meaning for Interpreting Deforestation Activity

The dynamic meaning interpretation of given deforestation area according to the given deforestation context is realised through the selection of semantic subspace from the entire semantic space that consists of dimensions derived from remote sensing sensors. That is, the vectors of the deforestation area in the semantic subspace have norms that set according to given context. The semantic or meaning interpretation is performed dynamically as projections in semantic space in regards to the selected subspaces. The most highly correlated meanings from given knowledge table are projected to the Semantic Deforestation Interpretation System. The projection calculates the relationships between the representative context and deforested area in semantic space. The most significant value which has the highest values of the projection are obtained then used for selecting the deforestation features among n dimension. In this study, four contexts of deforestation are discussed: Deforestation cause, deforestation impact on soil salinity, deforestation impact on soil erosion risk, and land use change after deforestation activity. Deforestation interpretations are based on up to date research and literature for tropical deforestation.
3.3.1. Projects the Deforestation Area to Semantic Subspace: Soil Condition Effect on Soil Erosion

Soil erosion is considered to be a major environmental problem since it seriously threatens natural resources, agriculture, and the environment. It is one of the main processes that reduce soil productivity by removing fertile topsoil layers. Moreover, soil erosion often causes adverse downstream effects, such as the sedimentation of soil material in reservoirs and lakes or damage to infrastructural facilities. Disastrous environmental consequences caused by soil erosion can be in the form of overflow from rivers or reservoirs after soil erosion, pollution of natural waters, or adverse effects on secondary soil salinization. Overall, severe soil erosion has been and is degrading soil and water resources. It leads to the degeneration of climatic and ecological environments and hinders the development of society and economy. The methodology used to assess erosion risk requires three parameters as inputs: slope gradient, land-use, and vegetation fraction cover. The slope gradient is calculated using DEM. The land-use map and vegetation fraction cover are extracted from Landsat TM respectively. In this study, the soil erosion risk was divided into six grades based on the aforementioned factors: slight, light, moderate, severe, more severe, and extremely severe. Trends in the overall erosion risk and erosion grade dynamic changes in time and space can be revealed by analysing the erosion grade in the same region with different phases. In qualitative methods, the qualitative factors integration firstly classifies some significant factors based on a specific standard, and the classified factors are then integrated according to the specific formula to create an erosion risk map. The parameter to interpret the soil erosion risk can be seen from Table 3.4.

<table>
<thead>
<tr>
<th>Ground Cover</th>
<th>Vegetation Cover</th>
<th>Slope Gradient (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5–8</td>
<td>8–15</td>
</tr>
<tr>
<td>Non-cultivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;75%</td>
<td>slight</td>
<td>slight</td>
</tr>
<tr>
<td>60%–75%</td>
<td>slight</td>
<td>light</td>
</tr>
<tr>
<td>45%–60%</td>
<td>slight</td>
<td>light</td>
</tr>
<tr>
<td>30%–45%</td>
<td>slight</td>
<td>light</td>
</tr>
<tr>
<td>&lt;30%</td>
<td>slight</td>
<td>moderate</td>
</tr>
<tr>
<td>Slope-cultivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>slight</td>
<td>light</td>
<td>moderate</td>
</tr>
</tbody>
</table>

Table 3.4. Standard evaluation of Soil Erosion Risk
3.3.2. Projects the Deforestation Area to Semantic Subspace: Soil Condition Effect on Plant Growth

Salinity reduces water availability for plant use. High salt levels hinder water absorption, inducing physiological drought in the plant. The soil may contain adequate water, but plant roots are unable to absorb the water due to unfavourable osmotic pressure. This is referred to as the osmotic or water-deficit effect of salinity. Plants are generally most sensitive to salinity during germination and early growth. The second effect of salinity is shown when excessive amounts of salt enter the plant in the transpiration stream and injure leaf cells, which further reduces growth. This is called the salt-specific or ion-excess effect of salinity (Greenway and Munns, 1980). Symptoms may include restricted root growth, marginal or leaf tip burning/browning, inhibited flowering, reduced vigor, and reduced crop yields. EC testing is a reliable way to assess how salts are affecting plant growth. The EC of soil or water is influenced by the concentration and composition of dissolved salts. Salts increase the ability of a solution to conduct an electrical current, so a high EC value indicates a high salinity level. Generally, an EC (1:5) water extract <0.15 will not affect plant growth. In discussing remote sensing of saline soils, one must distinguish between salinity at the soil surface (sometimes visible as salt crusts) and salinity in the soil root zone (i.e., the soil volume down to a depth of about 3 to 5 feet). Soil-root zone salinity affects plant growth, and it is the salinity indicator of most significant interest in plant growth assessments [103]. When a plant is stressed by root zone salinity, an increase in crop reflectance occurs in the blue (B), green (G) and red (R) ranges of the electromagnetic spectrum (e.g., leaves turn from green to hues of yellow and/or red), and a decrease occurs in the near-infrared (NIR) range. Specifically, vegetation indices such as the Normalized Difference Vegetation Index (NDVI) or the Canopy Response Salinity Index (CRSI), calculated from satellite multispectral reflectance data, can be used to infer root zone salinity within a satellite image pixel. [104] developed a prediction model for WSJV root zone salinity using CRSI as a predictor variable. The CRSI is defined as in Eq. 3.4.

\[
CRSI = \frac{(NIR \times R) - (G \times B)}{(NIR \times R) + (G \times B)}
\]  \quad (3.4)

Salinity prediction is calculated from CRSI values using Landsat 7 ETM+ canopy reflectance data with a resolution (pixel size) of 32.8 × 32.8 yards (900 square meters).
The regression for soil salinity can be seen in Eq. 3.5

\[ EC_e = 26.3 + EC_j + 0.02 \times RAIN_j + 3.35 \times TEMP_j \]  

(3.5)

where: \( EC_e \) is soil salinity (deciSiemens per meter, \( dS/m \), see Box 1), the subscript \( j \) indicates the year of the maximum CRSI value, RAIN (mm) is the total rainfall for the year and TEMP (°C) is the average daily minimum temperature for the year. Soil salinity is quantified as the electrical conductivity of a saturated soil paste extract \( (EC_e, dS/m) \). Soil salinity is quantified as the electrical conductivity of a saturated soil paste extract \( (EC_e, dS/m) \). The parameter of soil salinity interpretation is described in Table 3.5.

<table>
<thead>
<tr>
<th>Value</th>
<th>Rating</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 0.15</td>
<td>Very Low</td>
<td>Plants may be starved of nutrients.</td>
</tr>
<tr>
<td>0.15 - 0.50</td>
<td>Low</td>
<td>If soil lacks organic matter. Satisfactory if soil is high in organic matter.</td>
</tr>
<tr>
<td>0.51 - 1.25</td>
<td>Medium</td>
<td>Okay range for established plants.</td>
</tr>
<tr>
<td>1.26 - 1.75</td>
<td>High</td>
<td>Okay for most established plants. Too high for seedlings or cuttings.</td>
</tr>
<tr>
<td>1.76 - 2.00</td>
<td>Very High</td>
<td>Plants usually stunted or chlorotic.</td>
</tr>
<tr>
<td>&gt;2.00</td>
<td>Excessively High</td>
<td>Plants severely dwarfed; seedlings and rooted cuttings frequently killed.</td>
</tr>
</tbody>
</table>

**Table 3.5. Standard evaluation of Soil Erosion Risk**

c. **Projects the Deforestation Area To Semantic Subspace: Carbon Condition**

Deforestation and forest degradation are the main source of CO₂ emissions in the ‘forestry and other land use’ sector that accounts for 12% of anthropogenic CO₂ emissions. Emission can be calculated by the total area loss with the carbon stock in a tropical rainforest. explicit changes in carbon stocks will be far easier to implement going forward. First, carbon emissions from deforestation, forest degradation and other losses from non-forested ecosystems can be monitored against this high-resolution carbon map over time. This can be done at low cost using free Landsat imagery.
d. Projects the Deforestation Area to Semantic Subspace: Land Use Land Change

Few tropical deforestation footprint can be described with several common spatial patterns. [66] specify six possible patterns that happened across forest in tropical belt: geometric, corridor, fishbone, diffuse, patchy, and island. The geometric pattern can be recognized from large-scale deforestation for modern sector activities. Corridor pattern indicate the clearance of forest region/s for road infrastructure. The deforestation for planned settlement schemes can be commonly associated with fishbone pattern. Transmigration, resettlement, and colonization are part of the settlement scheme. Diffuse pattern relates to smallholder agriculture activity, where the area is relatively small in size but spread over specific wide area. The high population density that can cause deforestation can be identify in patchy pattern. This type of pattern is related to processes of permanent cultivation for food, predominantly for cash crop production. The last commonly recognizable pattern, island, is highly associated with the development of peri-urban areas. It is also caused by the expansion of agriculture and extension of infrastructure.
THE CLOSEST SEMANTIC-ELLIPSOID ALGORITHM
New Function for The Semantic Interpretation of Environmental Phenomena

This Chapter:
- Underlines the problem of classical Boolean logic for interpreting environmental phenomena
- Introduces new meaning assignment techniques, called as the Closest Semantic-Ellipsoid Algorithm, with a fast-computational time for assigning meaning to deforested area
4. The Closest Semantic-Ellipsoid Algorithm

This research proposes an algorithm to acquire meaning from environmental phenomena as described in the previous chapter. In this study, a new algorithm to obtain the meaning, called as the Closest Semantic-Ellipsoid Algorithm is proposed to assign the interpretation from an environmental phenomenon which projected as a range value in semantic space. The Closest Semantic-Ellipsoid Algorithm is inspired by the loop movement of the water droplet from a point when the water drop to approach the closest object. The proposed algorithm determines the closest meaning by calculating the loop passed the meaning from the original data point. This algorithm is proposed to locate the nearest context that has the closest meaning and most substantial volumes to be close to the data context. The forthcoming subsections explains the detail of the Closest Semantic-Ellipsoid Algorithm.

4.1. Uncertainty Problem in Interpreting the Environmental Phenomena

Human expert reasoning for interpretation to environmental phenomena involves not only intelligent but also rules of the environmental standard that constrain decision of interpretation making. Traditionally the environmental standard classifies the measurement value, as called parameter value, from the environmental sensor into several condition categories to presenting meaningful quantitative measures of environmental results and establishing whether or not the problem of environmental exists. The categories in the environmental standard represent the level of the condition such as the health or quality level which has been fed by the experts. The conventional approach for classifying environmental condition is fundamentally based on the idea that the calculation of selected parameters value that is within a specified range defined of a category in the formal environmental standard. This study is mainly focused on deforestation phenomena, in which the interpretation depends on the formal environmental standards relates to deforestation impacts. There are several traditional calculations to measure the interpretation of environmental phenomena such as numerical approach as represented, e.g., by pattern matching that is supported by classical Boolean logic rule. However, in the Boolean logic rule, the truth interpretation values of parameters may only be the truth values, true and false in a category. The problem of classical Boolean logic then falls in the inability to handle the interpretation of imprecise or uncertain information that often occurs in environmental monitoring modelling.
research. The uncertainty of information in environmental monitoring or ecological research inherited from:

1) The presence of inaccurate data cause by measurement error
2) The existence of missing value that fulfilled by the approximate estimations
3) Incomplete, vague expert knowledge or even subjective rules to determine the analysis

This study demonstrates the type of logical approach that has huge potential to assign the interpretation of environmental phenomena in where the truth value may fall in the range between completely true and completely false. There are several options to process the uncertainty problems using statistic, e.g. belief interval, probabilistic such as Bayes theorem [105], or fuzzy logic [106]. However, in the following discussion, the advantages and challenges of these methods will be compared with the proposed the Closest Semantic-Ellipsoid Algorithm.

4.2. Semantic and Multi-Valued Logic for Environmental Phenomena Interpretation

The fundamental of classical logic permits only the conclusions of a condition which are only true or false. But in contrast, in the domain of environmental problem, there are propositions with variable answers, in such as one might be found when interpreting a soil temperature from the degree of Celsius into categories such as "high", "medium" or "low". This underlines the problem of environmental phenomena analysis in where the knowledge of the expert is not always well defined and crisp. In such instances, the truth of semantic interpretation from environmental phenomena appears as the result of reasoning from inexact values which the interpretation lays on the spectrum between true and false. The environmental interpretation can be approximate the meaning of the linguistic variable that not brittle and exact. The type of logical approach that can provide the interpretation based on the degree of truth is called multi-valued logic whereas can describe better semantic in imprecise categories (such as "high", "medium" or "low") using some conceptual scale or semantic hierarchy. In this study, a new proposal based on the concept of multi-valued logic is proposed to fulfil the goal of this study for giving the interpretation to environmental phenomena, especially in deforestation phenomena.

To provide a semantic interpretations system for knowledge representation of the environmental phenomenon, it should be understood that there is a strong relationship between knowledge and language. In fact, the main objectives of knowledge representation used by the human being are languages. In the environmental phenomena monitoring, there is a vagueness
problem that can be found in the process of acquiring the meaning of the observed phenomena. The vagueness refers to an important problem in semantics according to philosophy where the interpretation is fall in the range between *completely true* and *completely false*. The value of environmental phenomenon sometimes said to has vague information if the boundaries of interpretation criteria are indeterminate or whether a borderline case does, or does not, satisfy a given vague concept. The condition of a phenomenon “extremely bad” is vague because there seems to be no crisp parameter value distinction at which condition from bad become extremely bad. With that being said, the particular difficult problems in the semantic of languages in the environmental problem are to deal with uncertainty and ambiguity. In logic, there is a propositional calculus theory that accepts the concept of more than truth values so-called the multi-valued logic.

This study initiates the research of semantic for environmental phenomena with the ideas to attach the vague semantics to the concept of linguistic variables. The objective of the discussion in this chapter is to develop a specific method that is capable of capturing and representing the vagueness and ambiguity as well as the uncertainty inherent in environmental phenomena monitoring. This is done through formal specification algorithm called the Closest Semantic-Ellipsoid Algorithm. The second and third questions of this study underline the description on how to find effective methods for the representation and retrieval of human knowledge. To provide interpretable information about environmental degradation due to deforestation activities, that has more than one context in which has $n$ uncertain interpretation in each context, the discussion in this section addresses the method automatic knowledge representation system. The proposed the Closest Semantic-Ellipsoid Algorithm is equipped with the means of resolving the uncertainty of environmental phenomena, especially for deforestation phenomena, by adopting the concept of multi-valued logic for semantic interpretation of a phenomenon.

### 4.3 The Ellipsoid Projection for Multi-Valued Logic: A New Approach for Interpreting the Uncertainty Problems in Environmental Phenomena

The main problem in interpreting the environmental phenomena is to give a meaning to the environmental object. Conventional Boolean logic ignores the continuous characteristic of ecological parameters and uncertainty of data. The multi-valued logic approach interprets the meaning according to classes that do not have sharply defined boundaries. In this study the ellipsoid is selected as a representative shape to express the meaning of environmental phenomena.
The guidelines for measuring some ecological conditions of ecological phenomena are called environmental standard. The environmental standard is given in the domain range. The level of measurement of the environmental condition to produce meaning refers to the relation between the value of a parameter that is captured by a sensor. The measurement in the environmental study is presented as interval measurement. In interval measurement, the distance between attributes has a different meaning. For example, the measurement temperature in the unit of Fahrenheit, the gap between 30 degrees to 40 degrees will be the same from the distance from 70 degrees to 80 degrees. However, the interval value interpretation is different for these two cases. Therefore, it is not suitable to compute the simple average value of an interval value to give the meaning to the environmental variable. Because of the problem with this ordinal scale, the standard point representation for acquiring meaning in semantic computing is not appropriate.

Figure 4.1 illustrates the ellipsoid shape to representing the non-crisp boundaries in comparing to the basic Boolean logic, implemented as if-else conditional computing, which can be represented as a cube. The ellipsoid shape can represent the environmental standard better than cube shape. While the common calculation for interpreting the environment measurement value uses conditional computing, the degree of uncertainty to give meaning in environmental phenomena remains challenging. In this study, we proposed the new information interpretation shown as an ellipse to represent the set of interpretation value which derived from the expert knowledge consists of multiple environmental parameters in Euclidean space. The characteristic of the ellipse is useful for describing environmental standard intervals for the meaning of single new observations (prediction intervals). The property of standard ellipse in the conic section is adopted in this study. The ellipse area that is representing the class range of environmental standard is determined by the by the length of its interval value. In the conic section, the area of ellipse is indicated with the horizontal and vertical rim, in two dimensions, as the X-axis and Y-axis, in which is equal to the mean (Range * I) where the mean and range refer to the X or Y variable, and I is the current value of the coefficient field. This definition of ellipsoidal parameter also applied to the ellipsoid in the Closest Ellipsoid-Semantic Algorithm where the rims indicates the interval of the interpretation classes in the environmental phenomena. An ellipse as or a part of circle
family in two dimensions as called as called conic sections, is formed by the intersection of a plane with a right circular cone. The general equation of an ellipse centred at (h,k) is shown in Eq. 4.1:

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$$  \hspace{1cm} (4.1)

The ellipse describes in this study is following the conic sections of the ellipse. The property of ellipse as the semantic projection is shown in Figure 4.2, where the ellipse is drawn in two dimensions and the ellipse has centre C at the origin, the major axis locates at (+a,0) and (-a,0), and in which the minor axis vertices at (+b,0) and (-b,0). The standard equation of ellipse indicates how the sum of the distances from two foci to any point on the ellipse is a constant is at the length of the major axis. To define the ellipse in Euclidean space, the parametric equation of ellipse are varying on a different dimension with, \(x = f(t), \ y = g(t)\ \ t \in D.\)

Figure 4.1. The comparison of semantic projection between proposed ellipse range with conditional computation.

Why not Cube?

Because it can represent degree of uncertainty.
Property of Ellipse Semantic Projection

Standard Parametric Representation in 2d using the sine and cosine functions

\[(x, y) = (a \cos t, b \sin t), \quad 0 \leq t \leq 2\pi\]

Where
- \(a\) is range domain of parameter of 1\(^{\text{st}}\) dimension
- \(b\) is range domain of parameter of 2\(^{\text{nd}}\) dimension

For 3D:

\[(x, y, z) = (a \cos t \sin u, b \sin t \sin u, c \cos u), \quad 0 \leq t \leq 2\pi, \quad 0 \leq u \leq \pi\]

Figure 4.2. Property of ellipse semantic projection.

4.4. The Closest Semantic-Ellipsoid Algorithm

The proposed Closed Semantic-Ellipsoid Algorithm is categorised as a higher-level procedure to interpret the environmental phenomena, that is created to make a better meaning acquisition. This algorithm calculates the semantic meaning by computing the distance between an observed deforestation area and semantic ellipsoids representing meanings of forest-situations in an ellipsoid form. The Closest Semantic-Ellipsoid Algorithm has the features calculate the closest meaning from environmental problems to semantic projection or the environmental standard. To achieve the result, the Closest Semantic-Ellipsoid Algorithm has to approximate formulas for the Point-to-ellipse and point-to-ellipsoid distance problem. The illustration of the point to ellipse data is presented in Figure 4.3. The feature of spatial distance calculation to determine the meaning of deforestation phenomena is the main proposal of this algorithm. The novelty of this proposal is that by using the Closest Semantic-Ellipsoid Algorithm, it is possible to designate many logical values at a time. That possibility, which seems quite natural, was ignored by other originators of many-valued logics.

4.4.1. The Closest Ellipsoid-Semantic Algorithm: The Basic Concept

In the previous section, we have defined semantic information about deforestation meaning as shown in Table 3.4. and Table 3.5. If the deforestation area is represented as a point, the semantic meanings are represented as the ellipsoids, and the goal is to find the closest semantic ellipsoid
that can represent the deforestation point, then the question is, how to find the closest ellipsoid to
the given point. There are two computational questions that have to be answered to answer the
question. 1) How to find the point on an ellipse that is closest to the point of $X$ outside of the ellipse.
After we find the closest point on an ellipse, since semantic means the entire element of an ellipse
is essential, then the question on 2) How to find the ellipse that has most of the area that is closest
to the point of $X$ outside of the ellipse. While the distance between the point of $X$ outside of the
ellipse to the nearest point on an ellipse can be solved by the distance equations, for any ellipses
the critical problem is how to find the closest point. The Closest Ellipsoid-Semantic Algorithm has
two main calculation: 1) To find the closest distance between the closest point on a semantic
ellipsoid to a deforestation point and 2) How to determine the closest semantic-ellipse that has the
most semantic area that is closest to deforestation point as indicated in Figure 4.3. and 4.4.

Figure 4.3. Illustration of acquiring meaning in semantic space.
Figure 4.4. Problems of assigning meaning in ellipsoid distribution to a point.
4.4.2. The Closest Ellipsoid-Semantic Algorithm: The Calculations

The closest distance is a key parameter of data to their interpretation/meaning in environmental phenomena. The discussion in this sub-chapter derives an analytic expression for the distance of closest approach of the centres of \( n \) arbitrary hard ellipses which indicates the deforestation semantic-meaning as function of their orientation relative to the line joining their centres. We describe our method for solving this problem, give the solution, illustrate our results, and discuss its usefulness in modelling and simulating systems of deforestation interpretation. In the general ellipsoid shape in 2 dimensions, the ellipse points from the ellipse with the centre point \( C \) located at the origin in parametric equation are written in Eq. 4.2.

\[
\begin{align*}
  x &= \alpha \cos \varphi \\
  y &= \alpha \sin \varphi
\end{align*}
\] (4.2)

The formal definition of ellipse shape will be applied to representing the semantic meaning. In this dissertation the proposed calculation is written in two steps; 1) To determine the closest point on the ellipsoid to a deforestation point and 2) determining the closest semantic meaning. Figure 4.5. indicates the property of ellipse and how the calculation refers the standard ellipse property.

![Property of ellipse](image)

**Figure 4.5.** Property of ellipse
a. Determining Closest Ellipse to Point

Supposed that X point is deforestation area and the ellipse is meaning as shown in Figure 4.6, analytically finding the smallest distance between a point and an ellipse can be solved by the iterative method. To illustrate the solution, Figure 4.6 shows the bounding circle around the given point P(x, y), which passes through the nearest point on the ellipse. From the figure, it is clear that the closest point is such that a line drawn from it to the given point must be perpendicular to the shared tangent of the ellipse. Any other points would be outside the circle and so must be further away from the given point.

![Diagram showing the problem definition to find closest point on the ellipse with the point X outside of ellipse.](image)

**Figure 4.6.** The problem definition to find closest point on the ellipse with the point X outside of ellipse

The calculation to find the closest point here is by pick eight points in the radians of the ellipse. Then calculate the distance between point to ellipse points. The point with the shortest distance will be the starting point for the next iterative process by derivative of curvature as shown in Figure 4.7. To demonstrate the applicability of the method, a necessary condition for the point on the ellipse to be the closest point to the deforestation X point is perpendicular to the tangent vector in the ellipse E with the definition of \((\vec{X} - \vec{E}(\theta)) \cdot \vec{E}(\theta) = 0\).
Distance from point to an ellipse

1) Pick Starting Point
   a) Pick 8 points every phi/4 radians on ellipse
   b) Calculate distance from data to ellipse points
   c) Ellipse point with shortest distance is starting point

2) Iterate to Pick Next Closest Point, by derivative of curvature ellipse

3) Iterate to Pick Next Closest Point, by derivative of curvature ellipse

4) Iterate to Pick Next Closest Point, by derivative of curvature ellipse

Figure 4.7. Calculates distance from point to closest ellipse.

The condition always converges after a couple of iterations using the derivative curvature of Eq. 4.3. The desired distance is measured by Euclidean distance $|\vec{X} - \vec{E}(\theta)|$.

$$f(\theta) := (a^2 - b^2) \cos \theta \sin \theta - xa \sin \theta + y b \cos \theta \quad (4.3)$$

b. Determining Closest Ellipse-Semantic Meaning

The outer bound point calculates the closest meaning in this dissertation on the ellipse. An ellipse defines meaning, and to assign the meaning, we have to consider the total area that is closer to point X to the ellipses, the illustration of the whole area that is closer to the point X is indicated by Figure 4.4. This study proposes the calculation to find the outer bound of the ellipse to realise the estimate of the closest total area of the ellipse. It is defined by the normal line of the ellipse, in which perpendicular to the deforestation point X. To find the point, in this study, we should find the equation of a line from point to inner bound. This subsection contains a stand-alone description of an efficient outbound of an ellipse search algorithm. It is intended to be sufficiently detailed to
allow a straightforward implementation. For efficiency, the recursive operations discussed in the previous subsection have been restructured into a looping procedure. We consider the normal line of the ellipse from acquired closest point in an ellipse as can be seen in Figure 4.8.

![Diagram of ellipse and deforestation point](image)

**Figure 4.8.** The problem definition to find outside point on the ellipse with the point X outside of ellipse

In this dissertation, to find the outer boundaries of the meaning, the normal line of an ellipse is used to imitate the radius expansion to reach ellipsoidal meaning

1. Find f(x), it is determined from the inner bound point of the semantic ellipse and the point of deforestation by using the linear equation as shown in Eq. 4.4

   
   \[
   f(x) = \left( \frac{px - e_i x}{py - e_i y} \right) x + b
   \]

   
   (4.4)

2. Calculate the outside point that intersects to the given line function of Eq. 4.5 considering the ellipse function in Eq. 4.1 as follow

   \[
   x = \pm \frac{ab}{\sqrt{a^2 py^2 + b^2 px^2}} px ; \quad y = \pm \frac{ab}{\sqrt{a^2 py^2 + b^2 px^2}} py
   \]

   (4.5)

3. Calculate the total distance between deforestation point X to the outer bound. The desired distance is the closest measured distance by Euclidean distance \( |\vec{x} - \vec{E}(\theta)| \).
4.4.3. The Closest Ellipsoid-Semantic Algorithm: The procedures

1. Set \( SP \in R \). Where \( SP \) is semantic projection of a context in \( N \) dimensional space
   
   - **In two dimension:**
     
     \[
     x = a \cdot \cos(t) \quad y = b \cdot \sin(t) ;
     \]
     
     - \( a \) is radius of dimension 1
     - \( b \) is radius of dimension 2

2. Calculate the closest inner bound point from ellipse to point \( \text{in\_boundE}_i \)
   
   1. Choose a point \( a \) on the ellipse
   2. Compute the distance \( r \) between \( a \) and \( p \)
      - \( \text{ev}_x = (a^2 - b^2)/a \cdot \cos^3(t) \)
      - \( \text{ev}_y = (b^2 - a^2)/b \cdot \sin^3(t) \)
   3. On the circle centered at \( p \) with radius \( r \), find the other intersection \( b \)
   4. Set \( a' \) to the midpoint of \( a \) and \( b \) on the ellipse
   5. Return to step 2

3. Calculate \( d \leftarrow \text{dist}(P, \text{in\_boundE}_i) \)

4. Calculate \( \text{out\_boundE}_i \) by \( \text{secant}(\text{in\_boundE}_i) \)

5. Calculate \( d \leftarrow d + \text{dist}(\text{in\_boundE}_i, \text{out\_boundE}_i) \)

6. Assign Meaning by \( d_{min} \)
4.4.4. The Closest Ellipsoid-Semantic Algorithm: The Initial Experimental Result

To demonstrate the performance of the Closest Semantic-Ellipsoid Algorithm, in this sub-chapter the performance of the initial experiment to acquire the area closest to the given ellipse is applied. The plots show good convergence after three iterations over a wide range of eccentricities as shown in Figure 4.9.

![Figure 4.9. Plots of closest distance in Euclidean space.](image)

4.4.5. The Closest Ellipsoid-Semantic Algorithm: The Implementation on the Deforestation Interpretation

This section shows the demonstration result of the Closest Semantic-Ellipsoid Algorithm for deforestation phenomena interpretation. The deforestation area is the area resulted from the DELSAR algorithm explained in Chapter 3. The deforestation areas are then stored in the database and the meaning is calculated by the Closest Semantic-Ellipsoid Algorithm with the expert knowledge shown in Table 3.4 and 3.5.
In Figure 4.10, we demonstrate four deforestation area interpretation in the term of the causes. The meaning is calculated by the summary of the distance between closest point to the outside point of a semantic ellipse. The deforestation point $X$ have different interpretation in regards to the HH and HV value derived from ALOS-2/PALSAR-2 data. The interpretation of each example deforestation area can be seen in Table 4.1.

**Table 4.1.** Interpretation of deforestation areas in the context of causes

<table>
<thead>
<tr>
<th>Area</th>
<th>Country</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peru North</td>
<td>Peru</td>
<td>Forest fire</td>
</tr>
<tr>
<td>Peru South</td>
<td>Peru</td>
<td>Cultivation</td>
</tr>
<tr>
<td>Peru Ucayali</td>
<td>Peru</td>
<td>Local Agriculture</td>
</tr>
<tr>
<td>Kalimantan Timur</td>
<td>Indonesia</td>
<td>Cultivation</td>
</tr>
</tbody>
</table>
The experimental study not only applies to the cause of deforestation, the effect of deforestation on soil degradation also calculated using the Closest Semantic-Ellipsoid Algorithm as shown in Figure 4.11. The detail results of interpretation of deforestation in term of soil degradation can be seen in Table 4.2.

![Figure 4.11. Interpretation of deforestation areas in the context of Soil Degradation for plantation](image)

The experimental study for interpreting the conversion area or LULC using the Closest Semantic-Ellipsoid Algorithm can be seen in the Figure 4.12. The detail results of interpretation of deforestation in term of soil degradation can be seen in Table 4.3
Table 4.2. Interpretation of deforestation areas in the context of Soil Degradation for plantation

<table>
<thead>
<tr>
<th>Area</th>
<th>Country</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peru North</td>
<td>Peru</td>
<td>Slight Soil Degradation</td>
</tr>
<tr>
<td>Peru South</td>
<td>Peru</td>
<td>Slight Soil Degradation</td>
</tr>
<tr>
<td>Peru Ucayali</td>
<td>Peru</td>
<td>Extremely Severe</td>
</tr>
<tr>
<td>Kalimantan Timur</td>
<td>Indonesia</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

The last interpretation is the deforestation impact on carbon emission that can be seen in Figure 4.13. All the semantic ellipses are reflecting the expert knowledge on each context, and the Closest Semantic-Ellipsoid Algorithm realizes the new calculation to interpret the nature by the semantic ellipse distance calculation.

Figure 4.12. Interpretation of deforestation areas in the context of LULC
Table 4.3. Interpretation of deforestation areas in the context of LULC

<table>
<thead>
<tr>
<th>Area</th>
<th>Country</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peru North</td>
<td>Peru</td>
<td>Modern Sector</td>
</tr>
<tr>
<td>Peru South</td>
<td>Peru</td>
<td>Resettlement</td>
</tr>
<tr>
<td>Peru Ucayali</td>
<td>Peru</td>
<td>Traditional Agriculture</td>
</tr>
<tr>
<td>Kalimantan Timur</td>
<td>Indonesia</td>
<td>Resettlement</td>
</tr>
</tbody>
</table>

Figure 4.13. Interpretation of deforestation areas in the context of carbon emission
4.5. The comparison of the Closest Semantic-Ellipsoid Algorithm with Many-Valued Logics

Some knowledge in the human brain has the concept of many-valued or uncertain sets rather than basic Boolean logic. Especially in the domain of environmental or ecological research, many uncertain inferences contain many meanings. This theory encourages the proposal of the Closest Semantic-Ellipsoid Algorithm in this study. The previous sub-chapter has described the procedures to obtain the rule of uncertain sets for environmental or ecological phenomena interpretation. The mathematical detail about the uncertain set itself can be found in the literature [107]. The concept of many-valued logic has been actively introduced by many researchers in the fuzzy sets or random sets research domain. However, the Closest Semantic-Ellipsoid Algorithm is enabled to make better use of imprecise ecological data interpretation based on the vague and limited expert knowledge. It is because the features of the Closest Semantic-Ellipsoid Algorithm can describe as follow:

1) The representation of the interpretation in ellipsoid shape handles the uncertainty degree

2) The ellipsoid interpretation is defined in uncertainty set.

3) The wave motion-like let all the member in the domain of analysis to have the interpretation

The discussion on the theory of meaning assignment using the Closest Semantic-Ellipsoid Algorithm is described in the mathematical set theory explanation and the computational performance explanation.

4.5.1. Comparison in Mathematical Set Theory: The concept of membership function

The basic Boolean logic is interpreted as the crisp system that unable to handle the uncertainty value that commonly found in ecological research. The basic Boolean logic normally presented in the theory of set and the concept is certainly different from the concept of many-valued logic. The discussion in theory of set surrounds the fact of when an object is being inside of the set or not. An object has true meaning if it is inside of a true set. For the fuzzy or random set, the theory serves that an object can be in between two sets, moreover with no strict true or false border value. In the fuzzy world, the partnership expresses the condition value in which two fuzzy sets involved. This will determine the meaning of an object based on its value of membership. From the definition, it is clear that fuzzy-set can handle the uncertainty value of an object, in which the traditional set theory not. However, it is also should be noted that in fuzzy-set, an object should be a member of
a set or two sets to have a meaning. The characteristic of set and fuzzy-set underlines the novelty of this new proposed multi-valued logics algorithm called the Closest Semantic-Ellipsoid Algorithm. The Closest Semantic-Ellipsoid Algorithm profoundly describe that an object can have a meaning, even it does not fall in a set. An object exists in the uncertainty set and will have the same meaning to the closest total set. By definition, a meaning or a semantic is a gathering together into a whole of definite, distinct objects of our perception or thought. Therefore, the theory of meaning set and how an object belongs to a set is defined differently in this study. To have the visual understanding between basic Boolean logic represented by set theory, fuzzy-set or random-set and semantic-set, please refer to Figure 4.14., 4.15., and 4.16. The set theory is closely linked with the basic Boolean logic operation in which the correlation between object and set is called implication of the statement in where if the first expression is true then B is also true. The part of the domain that is not in a set can be written as \( A^c \) or \( \sim A \). The set operation that is presenting the basic Boolean logic indicates that if there is an object that not in the set, then the object does not belong to any meaning, or in another word, the object is meaningless as shown in Figure 4.14. (b) and (d). It is because by definition the outer area of the sets is the complement element. Therefore, object red in Figure 4.14. (b) and object green in Figure 4.14. (d) are meaningless.

The same concept is also can be seen in Figure 4.15. which describes the fuzzy-set. The uncertain value provides large varieties of interpretation of set to interact. This can be observed from the border of the fuzzy set. The border tolerates the uncertainty as seen in section (d) of Figure 4.15. (it is 40% belong to set A, and 60% reside to set B). Therefore, purple object has B interpretation wherein the term of fuzzy logic this process is essential and is called as a membership function. The concept of fuzzy logic is claimed to be close to realising the human brain mechanism to interpret a phenomenon. [108], [109]. However, in the fuzzy set, the object that is not in a set or two sets is also defined as a meaningless object. This is never a case of the human brain; there always be an interpretation that human can attach to each phenomenon. Therefore, in this study, the underline problem on interpreting the environmental phenomena is addressed and this achieved by the concept of the Closest Semantic-Ellipsoid Algorithm, all objects (which in this study is corresponding to ecological phenomena) will always have a meaning. The assignment of the meaning of each object is calculated by water droplet-like motion as shown in Figure 4.16.: (a) set A in a domain; (b) two objects lay on a domain; the blue object is in the set, and red object is not in the set. (c) set A and set B in a domain, and (d) two objects lay on a domain; the purple object
is in the set A, and the green object is not in the set A or B, however with the mechanism of the Closest Semantic-Ellipsoid Algorithm in (e), object green has B interpretation.

Figure 4.14. Set Operation of basic Boolean Logic: (a) set A in a domain; (b) two objects lay on a domain; blue object is in the set and red object is not in the set. (c) set A and set B in a domain, and (d) two objects lay on a domain; purple object is in the set A and green object is not in the set.

Figure 4.15. Set Operation of basic Multi-Valued Logic (usually found in concept of fuzzy logic): (a) set A in a domain; (b) two objects lay on a domain; blue object is in the set and red object is not in the set. (c) set A and set B in a domain, and (d) two objects lay on a domain;
Figure 4.16. Set Operation of basic Multi-Valued Logic using the Closest Semantic-Ellipsoid Algorithm. (a) set A in a domain; (b) two objects lay on a domain; blue object is in the set and red object is not in the set. (c) set A and set B in a domain, (d) two objects lay on a domain, and (e) the Closest Semantic-Ellipsoid Algorithm

4.5.2. Comparison in Computation of Basic Boolean Logic, Multi-Valued Logical, and the Closest Semantic-Ellipsoid Algorithm of the Semantic Relation of Environmental Phenomena.

Interpreting environmental or ecological phenomena that have uncertainty value is challenging. The Closest Semantic-Ellipsoid Algorithm attempts to provide the interpretation of environmental phenomena that resembles the human brain to handle imprecise and vague information. The Closest Semantic Ellipsoid Algorithm approach is specifically designed for ecological phenomena. However, the research for meaning acquisition is a very active field. In mechanical control, the
fuzzy logic algorithm has been actively developed to support the mechanism of meaning acquisition before executing an output action. However, the significant differences between the proposed system, the Closest Semantic-Ellipsoid Algorithm another logical approach in term of conceptual and fundamental background has been discussed in the previous sub-chapter. The details on computation performance, such as pseudocode and rules will be explained in this chapter including the membership function discussion that is very important for meaning acquisition. The comparison of procedures are as follow:

Table 4.4. The basic flow of basic Boolean logic

<table>
<thead>
<tr>
<th>The flow of the Basic Boolean Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialization: Define the semantic variables, meaning and terms</td>
</tr>
<tr>
<td>2. Construct the rule base</td>
</tr>
<tr>
<td>3. Calculate the object status to the most suitable rule base</td>
</tr>
<tr>
<td>4. Produce the meaning for the object only in the set</td>
</tr>
</tbody>
</table>

Table 4.5. The basic flow of fuzzy logic

<table>
<thead>
<tr>
<th>The flow of the Multi-Valued Logic in fuzzy logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialization: Define the semantic variables, meaning and terms</td>
</tr>
<tr>
<td>2. Define the membership functions</td>
</tr>
<tr>
<td>3. Define the rule base</td>
</tr>
<tr>
<td>4. Convert crisp data to fuzzy values</td>
</tr>
<tr>
<td>5. Calculate the object status to the most suitable rule base by degree of membership</td>
</tr>
<tr>
<td>6. Produce the meaning by defuzzification to convert the output data to crisp output for the object only in the set</td>
</tr>
</tbody>
</table>

Table 4.6. The basic flow of the Closest Semantic-Ellipsoid Algorithm

<table>
<thead>
<tr>
<th>The flow of the Closest Semantic-Ellipsoid Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialization: Define the semantic variables, meaning and terms</td>
</tr>
<tr>
<td>2. set the ellipsoid as representative of meaning function</td>
</tr>
<tr>
<td>3. Calculate the object status to the most suitable rule base on wave-like distance</td>
</tr>
<tr>
<td>4. Produce the semantic meaning based on the result of wave-like distance</td>
</tr>
</tbody>
</table>
According to the basic Boolean logic, the membership or interpretation of an object is crisp, between true or 1 and false or 0. In the fuzzy-set, the grade of membership for all its members thus is normally a real number between 0 and 1. The Closest Semantic-Ellipsoid Algorithm perform the distance calculation, in which the membership of an object is determined by the distance between an object to the entire element set of a context. To bring brief procedural comparison, Table 4.4, 4.5, and 4.6 is describing the flow and rules of basic Boolean logic, multi-valued logic presented by fuzzy logic, and the Closest Semantic-Ellipsoid Algorithm. The membership function for those three methods can be seen in Figure 4.17.

Figure 4.17. Interpretation Representative Function for Basic Boolean logic, the imprecision multi-valued logic in fuzzy-set and the Closest Semantic-Ellipsoid Algorithm

Apart from the ability to give the meaning of an object that is not in the set, the distinction between basic Boolean logic, typical many-valued logic such as fuzzy logic, and the Closest Semantic-Ellipsoid Algorithm for interpreting ecological phenomena, the different of each method can be divided as follows:

a) Interpretation Representative Function

The function to project the interpretation representation in the many-valued logic is defined as interpretation representative function. In the basic Boolean logic, the interpretation representative function works as a step function, define the crisp distinction between true or false. In the domain
of fuzzy-set, it is generally known as membership function which presents the concept of degree of membership. The fuzzy membership function sets the value of the object between 0 to 1, inclusively. The membership function in fuzzy typically is a triangular or trapezoidal function. However, only one parameter can be treated at a time in the processing of the membership function. The proposed the Closest Semantic-Ellipsoid Algorithm defined the ellipsoidal shape to define the interpretation representative function as shown in Figure 4.2. The n-dimensional parameters can be treated in one time. The relation between “multi-logic sets theory” and “statistics and probability theory” is that in the probability theory, realization of events is based on the classical crisp logic as 0–1 logic, i.e., an event occurs or does not occur. When the boundaries of classes that reflect the events are precise, such logic is valid. For example in the probabilistic, when a dice has been rolled, the event of “coming up 1 or 2” is a precise event and it has a precise probability. But the event of “coming up a little number” is an imprecise event since its boundaries cannot be stated; consequently, its probability cannot be designated. In which the multi-logic algorithm has more admissible results.

b) Rule to acquire the meaning and the effect on computational performance

The many-valued logic such as fuzzy logic, to determine the meaning of a parameter, it has to establish the rule beforehand. It utilises the fuzzy associate memory (FAM) to map the information of the fuzzy input and its output. The current fuzzy rule is determined by the result calculation between object and membership function. The designer has to define the rule of meaning interpretation in fuzzy logic, and this makes the fuzzy logic static. However, due to the conditional computing if-else procedure that described as the rule after the calculation of membership function makes the computational cost of fuzzy logic cheap. For this reason, therefore fuzzy logic is prevalent to be found in some mechanical hardware. Although it is the computational cost is very cheap, some cases of the fuzzy logic application might have more than hundreds of rules. The construction of the rule should be done carefully because the change of the rule might be resulting to change the entire system rule. The proposed system, the Closest Semantic-Ellipsoid Algorithm support the dynamic interpretation. The number of initial knowledge does not require a complete set; it is because the meaning calculated not depending on the set characteristic but rather to the wave-like distance mechanism. As a trade-off from the dynamic mechanism, the Closest Semantic-Ellipsoid Algorithm has more expensive computational cost. It is because the way to acquire the meaning needs to check the distance in Euclidean space for each object.
SEMANTIC DEFORESTATION INTERPRETATION and RETRIEVAL SYSTEM

This Chapter:

- Shows the ability of semantic dimensional control to merge the SPA process into one semantic space.

- Describes the mechanism to control the value from one dimension to another dimension using representative point learning.
5. Semantic Deforestation Analysis and Retrieval System

After automatic deforestation detection algorithm to produce deforested regions derived from L-Band SAR satellite images, interpreting the situation of its damage is the next essential process to understand the meanings (semantics) of the phenomena fully. The goal of semantic computing application to deforestation monitoring is to plug the semantic gap and enrich the meaning and information value of deforestation analysis. On the basis of this study, it is expected that semantic computing address two core problems: First, to understand the meanings (semantics) of deforestation situation on the various value of remotely sensed data. Second, to map semantics interpretation of experts with the media database content for content and context retrieval. The semantic computing works by selecting the best or most informative dimension as information/feature extraction in high-dimensional data to address the challenges for analysing and giving meaning to deforestation area. This dimension selector is called sub-space selection and applied to remotely sensed data. The subspace selection in this study is based on theory, and it requires a model and/or knowledge of the deforestation analysis in the attempt to achieve optimal performance to interpret the environmental problem. The determinations of the semantic projection for all contexts in the semantic domain were carried out from the environmental standard. In the following discussion, the detail description of the application is given. The second interest here is on how to derived suitable information retrieval for environmental phenomena. The basic principle of a search engine is based on Natural Language Processing (NLP) techniques, for instance, string matching. Semantic computing shows its potential to lead the future direction of the search engine, aiming at context indexing by understanding the user intention. In this chapter, the detail discussion about the proposed idea to apply semantic computing for interpreting deforestation phenomena is provided followed by the detailed description of the dimensions created from remotely sensed data, which are used later on for the development of a new deforestation retrieval system. The computational steps for the deforestation interpretation can be seen in Figure 5.1. The image analysis for detecting deforestation using Phased Array Type L-band SAR-2 (ALOS - 2 / PALSAR - 2) as actual data is done by DELSAR algorithm. The follow up steps are realizing deforestation interpretation system through the deforestation meaning space for interpreting the meaning of Deforestation, we set image data mapping function and language information mapping function for meaning interpretation on top measure and realize them.
Figure 5.1. The general view of procedure for deforestation interpretation
5.1. Semantic Deforestation Analysis System

To give semantic meaning to each deforestation area, the system first obtains the necessary dimension to portray information from the environmental standard with the principle of subspace selection, refer to Chapter 3. The mechanism to give the meaning to each deforestation area is illustrated in Figure 5.1. The system retrieved deforestation area from automatic detection algorithm. There is 8 deforestation detected within 45. The deforestation polygons produced from the algorithm in Chapter 3 is divided into two part; one part for learning and constructing semantic space, and other is for testing the retrieval space. After retrieving 30 deforested pixel points as Def1, Def2, …, Def30 (Figure 5.2.) as training sample to create semantic space. After semantic space has been created, then, all deforested test pixels will be mapped onto semantic space.

As the experiment result of semantic analysis, we show 6 closet point of deforestation as the result. By applying semantic analysis to determine the effect of deforestation in soil condition, we could interpret reflected “substances (material)” of deforestation area in spectrum domain into human language in different context. The results of effect classification parameter as effect classification of deforestation effect found 4 classes. Shows the different point of deforestation activity and which word has the closest meaning to interpret each point. Our proposed idea is to interpret effect of soil condition in deforestation activity in language interpretation by using semantic analysis by using soil moisture and texture, soil temperature, and soil salinity. After converting multispectral
images from DN to Reflectance, and process the language interpretation and semantic analysis, the semantic matrices for deforestation is produced.

As the experiment result of semantic analysis, we show 6 closet point of deforestation as the result. By applying semantic analysis to determine the effect of deforestation in soil condition, we could interpret reflected “substances (material)” of deforestation area in spectrum domain into human language in different context. The results of effect classification parameter as effect classification of deforestation effect found 4 classes. Shows the different point of deforestation activity and which word has the closest meaning to interpret each point.

<table>
<thead>
<tr>
<th>long</th>
<th>lat</th>
<th>Word1</th>
<th>Word2</th>
<th>Word3</th>
<th>Word4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3047.00</td>
<td>2110.00</td>
<td>0.057</td>
<td>0.197</td>
<td>0.247</td>
<td>0.377</td>
</tr>
<tr>
<td>2237.00</td>
<td>3047.00</td>
<td>0.061</td>
<td>0.153</td>
<td>0.203</td>
<td>0.333</td>
</tr>
<tr>
<td>2007.00</td>
<td>3045.00</td>
<td>0.098</td>
<td>0.042</td>
<td>0.092</td>
<td>0.222</td>
</tr>
<tr>
<td>2002.00</td>
<td>3030.00</td>
<td>0.035</td>
<td>0.105</td>
<td>0.155</td>
<td>0.285</td>
</tr>
<tr>
<td>2045.00</td>
<td>2998.00</td>
<td>0.121</td>
<td>0.019</td>
<td>0.069</td>
<td>0.199</td>
</tr>
<tr>
<td>1769.00</td>
<td>3272.00</td>
<td>0.174</td>
<td>0.034</td>
<td>0.056</td>
<td>0.146</td>
</tr>
<tr>
<td>1925.00</td>
<td>2956.00</td>
<td>0.131</td>
<td>0.045</td>
<td>0.095</td>
<td>0.189</td>
</tr>
<tr>
<td>1929.00</td>
<td>2989.00</td>
<td>0.121</td>
<td>0.053</td>
<td>0.103</td>
<td>0.199</td>
</tr>
<tr>
<td>2066.00</td>
<td>2173.00</td>
<td>0.071</td>
<td>0.101</td>
<td>0.151</td>
<td>0.249</td>
</tr>
<tr>
<td>1892.00</td>
<td>2946.00</td>
<td>0.156</td>
<td>0.026</td>
<td>0.076</td>
<td>0.164</td>
</tr>
<tr>
<td>2101.00</td>
<td>2952.00</td>
<td>0.118</td>
<td>0.258</td>
<td>0.308</td>
<td>0.438</td>
</tr>
<tr>
<td>2125.00</td>
<td>3033.00</td>
<td>0.143</td>
<td>0.006</td>
<td>0.056</td>
<td>0.177</td>
</tr>
<tr>
<td>2244.00</td>
<td>3188.00</td>
<td>0.193</td>
<td>0.253</td>
<td>0.303</td>
<td>0.412</td>
</tr>
<tr>
<td>2232.00</td>
<td>3304.00</td>
<td>0.135</td>
<td>0.029</td>
<td>0.079</td>
<td>0.185</td>
</tr>
<tr>
<td>1787.00</td>
<td>3184.00</td>
<td>0.236</td>
<td>0.096</td>
<td>0.046</td>
<td>0.084</td>
</tr>
<tr>
<td>2062.00</td>
<td>2898.00</td>
<td>0.141</td>
<td>0.005</td>
<td>0.055</td>
<td>0.179</td>
</tr>
<tr>
<td>1265.00</td>
<td>6679.00</td>
<td>0.428</td>
<td>0.288</td>
<td>0.238</td>
<td>0.114</td>
</tr>
</tbody>
</table>

Figure 5.3. Interpretation of soil condition in deforestation area.

The system will detect the input value and use range effect for classification, which design and provide in Table 3.3. from knowledgebase discovery in deforestation knowledge area as a knowledge database for analysis process of semantic Analysis. Deforestation occurs in two different ways, forest fire or logging, not many information about cause of deforestation are served. our main proposed idea is to detect the cause of deforestation automatically, In this experimental result, we apply NBR to measure the burning scars of forest. Lower value (map as darker colour in Figure 5.4.) identified the burn scar of the deforestation place. The hypothesis could be given that cause of deforestation could be determine by NBR values. Figure 5.4. indicates the knowledge about the degree of severity of fire.
Figure 5.4. Normalized Burning Ratio value on (a) forest fire area (b) logging.
It is shown that the value of indices of land with forest fire, is higher(a) compare to the land with deforestation caused by illegal logging. The comparison data shown in Figure 5.5.
5.1.1. Soil Moisture

Soil moisture and texture are analyzed by using Red and NIR Reflectance from study site in Indonesia. The closer data with the H soil line, the more moisture the soil is, this method particularly useful and sensitive for monitoring changes in soil water content over time. The smaller value of the red reflectance with higher value of NIR, the dryer the soil condition will be. Figure 5.7. indicating the correlation between red and NIR to determine the soil moisture. To mining interesting knowledge in deforestation area, the system select 5 unique point from study area, where every point should be representing soil moisture after deforestation activity. The droughts and moisture of the soil could be identified by the value of NIR and Red Reflectance, with high NIR Reflectance value and small value of Red.

5.1.2. Soil Temperature

Soil temperature was analysed by temperature measurement from Landsat and represented in Celsius Degree. The surface or soil temperature at study area in 2017 between 21° C ~ 38° C. While normal soil temperature for tree to grow is in between 20° C ~ 26° C. Soil temperature was highly dependent on the cause of deforestation. Deforestation because of forest fire has a higher temperature compared to logging activity as shown in Figure 5.8.
Figure 5.8. Deforestation caused-based effect on temperature of soil

5.1.3. Soil Salinity

Major constraint to identify salinity is correlation between spatial and temporal variability in soil profile. Spectral data acquisition in nowadays research does not allow high precision information to be extracted from the entire soil profile, since only the Earth surface is observed. But, it is still reflecting the change of salinity of soil after deforestation effect. Soil salinity index is present in range between 0 and 1. The higher value of salinity index showing the higher soil salinity in those area the salinity of soil in Study Area have a brief correlation with deforestation activity occur in March 2014, where the salinity in Figure 5.9. shows high number of saline index which reach around 0.5.

Figure 5.9. Salinity value and temp value of cause-based deforestation.

Reflecting on the value of Figure 2.9, we can analyse that the smaller amount of the red reflectance with a higher value of NIR, the drier soil condition will be. We are indicating the correlation between red and NIR to determine the soil moisture. To interesting mining knowledge in the deforestation area, the system selects five unique points from the study area, where every point should representing soil moisture after deforestation activity. The droughts and moisture of the soil could be identified by the value of NIR and Red Reflectance, with high NIR Reflectance value and small value of Red.
Major constraint to identify salinity is correlation between spatial and temporal variability in soil profile. Spectral data acquisition in nowadays research does not allow high precision information to be extracted from the entire soil profile, since only the Earth surface is observed. But, it is still reflecting the change of salinity of soil after deforestation effect. Soil salinity index is present in range between 0 and 1. The higher value of salinity index showing the higher soil salinity in those area, and as seen in Figure 5.11., the salinity of soil in Study Area have a brief correlation with deforestation activity occur in March 2014, where the salinity in Figure 5.10. shows high number of saline index which reach around 0.5.

<table>
<thead>
<tr>
<th>long</th>
<th>lat</th>
<th>soil salinity fire</th>
<th>long</th>
<th>lat</th>
<th>soil salinity logging</th>
</tr>
</thead>
<tbody>
<tr>
<td>3442.00</td>
<td>4136.00</td>
<td>0.36</td>
<td>3442.00</td>
<td>4136.00</td>
<td>0.47</td>
</tr>
<tr>
<td>3447.00</td>
<td>4142.00</td>
<td>0.36</td>
<td>3447.00</td>
<td>4142.00</td>
<td>0.41</td>
</tr>
<tr>
<td>2462.00</td>
<td>4127.00</td>
<td>0.47</td>
<td>2462.00</td>
<td>4127.00</td>
<td>0.58</td>
</tr>
<tr>
<td>3422.00</td>
<td>4151.00</td>
<td>0.29</td>
<td>3422.00</td>
<td>4151.00</td>
<td>0.55</td>
</tr>
<tr>
<td>3443.00</td>
<td>4159.00</td>
<td>0.34</td>
<td>3443.00</td>
<td>4159.00</td>
<td>0.54</td>
</tr>
</tbody>
</table>

**Figure 5.11.** Salinity value and temp value of cause-based deforestation.

### 5.2. Semantic Retrieval of Deforestation Detection

Context dependent for interpreting nature condition is applied to determine the effect of deforestation in this research. The users have to specify the category of context in which most appropriate to represent their intention. Context of plant growth and soil degradation were selected as the semantic words analysis. The combination from three semantic spaces as proposed to determine the meaning of deforestation area stated in Table 3.3, Table 3.4 and spatial Pattern in Figure 3.26. In order to put an analysis in context of plant Growth, soil moisture, temperature and salinity were used to give the meaning. Semantic meanings are interpreted by measuring the distance
between the semantic interpretations in Table 3.2, 3.4, and spatial Pattern in Figure 3.26. with the close observance point. As the experiment result of semantic analysis, we retrieve 6 closet point of deforestation as the result. By applying semantic analysis to determine the effect of deforestation in soil condition, we could interpret reflected “substances (material)” of deforestation area in spectrum domain into human language in different context. The results of effect classification parameter as effect classification of deforestation effect found 4 classes. Figure 5.12. Shows the different point of deforestation activity and which word has the closest meaning to interpret each point. The system will detect the input value and use range effect for classification, which design and provide in Table 3.3, Table 3.4 and spatial Pattern in Figure 3.26. from knowledgebase discovery in deforestation knowledge area as a knowledge database for analysis process of semantic Analysis.

<table>
<thead>
<tr>
<th>Long</th>
<th>Lat</th>
<th>Word1</th>
<th>Word2</th>
<th>Word3</th>
<th>Word4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3047.00</td>
<td>2110.00</td>
<td>0.057</td>
<td>0.197</td>
<td>0.247</td>
<td>0.377</td>
</tr>
<tr>
<td>2237.00</td>
<td>3047.00</td>
<td>0.061</td>
<td>0.153</td>
<td>0.203</td>
<td>0.333</td>
</tr>
<tr>
<td>2007.00</td>
<td>3045.00</td>
<td>0.098</td>
<td>0.042</td>
<td>0.092</td>
<td>0.222</td>
</tr>
<tr>
<td>2002.00</td>
<td>3030.00</td>
<td>0.035</td>
<td>0.105</td>
<td>0.155</td>
<td>0.285</td>
</tr>
<tr>
<td>2045.00</td>
<td>2998.00</td>
<td>0.121</td>
<td>0.019</td>
<td>0.069</td>
<td>0.199</td>
</tr>
<tr>
<td>1789.00</td>
<td>3272.00</td>
<td>0.174</td>
<td>0.034</td>
<td>0.056</td>
<td>0.146</td>
</tr>
<tr>
<td>1925.00</td>
<td>2956.00</td>
<td>0.131</td>
<td>0.045</td>
<td>0.095</td>
<td>0.189</td>
</tr>
<tr>
<td>1929.00</td>
<td>2989.00</td>
<td>0.121</td>
<td>0.053</td>
<td>0.103</td>
<td>0.199</td>
</tr>
<tr>
<td>2006.00</td>
<td>2173.00</td>
<td>0.071</td>
<td>0.101</td>
<td>0.151</td>
<td>0.249</td>
</tr>
<tr>
<td>1892.00</td>
<td>2946.00</td>
<td>0.156</td>
<td>0.026</td>
<td>0.076</td>
<td>0.164</td>
</tr>
<tr>
<td>2101.00</td>
<td>2952.00</td>
<td>0.118</td>
<td>0.258</td>
<td>0.308</td>
<td>0.438</td>
</tr>
<tr>
<td>2125.00</td>
<td>3033.00</td>
<td>0.143</td>
<td>0.006</td>
<td>0.056</td>
<td>0.177</td>
</tr>
<tr>
<td>2244.00</td>
<td>3118.00</td>
<td>0.193</td>
<td>0.253</td>
<td>0.303</td>
<td>0.412</td>
</tr>
<tr>
<td>2232.00</td>
<td>3304.00</td>
<td>0.135</td>
<td>0.029</td>
<td>0.079</td>
<td>0.185</td>
</tr>
<tr>
<td>1787.00</td>
<td>3184.00</td>
<td>0.236</td>
<td>0.096</td>
<td>0.046</td>
<td>0.084</td>
</tr>
<tr>
<td>2062.00</td>
<td>2898.00</td>
<td>0.141</td>
<td>0.005</td>
<td>0.055</td>
<td>0.179</td>
</tr>
<tr>
<td>1265.00</td>
<td>6679.00</td>
<td>0.428</td>
<td>0.288</td>
<td>0.238</td>
<td>0.114</td>
</tr>
</tbody>
</table>

**Figure 5.12.** Semantic interpretation in context of Plantation

The system will detect the input value and use range effect for classification, which design and provide in Table 3.3 from knowledgebase discovery in deforestation knowledge area as a knowledge database for analysis process of semantic Analysis.
Semantic retrieval results also been tested to another semantic space of deforestation we provide in Table 2.4. In Figure 5.13, we retrieve the location that has the semantic meaning of: (a) Critical Soil Condition, lead a high risk of soil desertification. Due to the low water contained in observed soil.. and (b) Normal soil Condition, moisture index indicate that soil is hold proper amount of water and soil temperature is not to high, and will not lead any desertification in short period of time. It clearly shows that different point of deforestation activity and which word has the closest meaning to interpret each point can be retrieved to fulfil the semantic context inputted by user.
CONCLUSION AND PERSPECTIVES

This Chapter:

- Concludes this study.

- Indicates the perspective and future challenges.
6. Conclusion and Perspectives

The result of the study on the variety of satellite images led to the new research direction, what is the most appropriate dataset for deforestation detection. The Phased Array Type L-band SAR-2 (ALOS-2/PALSAR-2) satellite images are used to detecting deforestation due to its capability to penetrate cloud which is very powerful to observe the tropical area.

In this study, there are two technical proposals. One is the deforestation detection for L-Band SAR which applied to the sensing dimension, so-called DELSAR or deforestation detection for L-Band SAR. In this study, semantic dimensional control is proposed. The second one is the Closest Semantic-Ellipsoid Algorithm, a new algorithm to acquire semantic meaning. In the processing domain, there is a distinct difference in the value of gamma-naught of L-Band data in the wet and dry season. In addition, the canopy cover is also contributing to the different value of gamma-naught. During the deforestation activity process, the timber is covering the exposed soil and makes the strong backscattering value. These factors that affect the gamma-naught value from The Phased Array Type L-band SAR-2 (ALOS-2/PALSAR-2) data are founded at this study and it is shown that DELSAR can deal with this obstacle. The experimental study also found that the gamma-naught in the wet area has higher variance in about ± 2 db compare with the dry season. With DELSAR all these variance and changes have been taken into account to determine deforestation position. By considering the density of the change instead of only leaning on the change of backscattering value.

In the processing procedure, it has been developed that the interpretation from environmental data usually represented by the interval value of a parameter. This extends the challenges to interpret the environmental problem that full of uncertainties. The technical question on how to representing the knowledge to the database also has been answered by the mechanism to present and acquire the meaning of environmental problem as described in the new proposal, the Closest Semantic-Ellipsoid Algorithm. the Closest Semantic-Ellipsoid Algorithm works specifically to tackle environmental meaning acquirement, where the value of semantic projection is not presented as a point but as an ellipsoid. The interval value of each parameter in the semantic dimension is treated as an ellipse. The technique to calculate the closest distance is by point to point calculation but by closest area of an ellipsoid. In the result of this result, it is founded that
the meaning assignment using the Closest Semantic-Ellipsoid Algorithm is more appropriate for deforestation monitoring in comparing with the other method using the point as semantic projection.

Reflecting on the result of the experimental study, semantic computing for interpreting environmental phenomena, such as deforestation, allows an automatic meaning computation based on context rather than relying on imperative methods operating on meaningless values. By linking specific natural language and meaning related deforestation phenomena, i.e. the cause of deforestation, the effects to soil, carbon emission, and land use change after deforestation, instead of just using calculation of meaningless values, the system with semantic computing can make inferences about deforestation that might not be hard-coded into the system from the start. Notably, its potential for the interpretation of deforestation phenomena using constitutes an essential contribution to the environmental study, as well as to the traditional database analysis for producing meaningful information. Taking into account, the temporal data is prerequisite for any learning algorithm for understanding the deforestation phenomena, therefore remote sensing data were employed for this study. The sensing procedure used in this study aimed to detect deforestation area in tropical rainforest automatically. In Chapter 3, new automatic deforestation detection specifically designed for L-Band SAR, DELSAR, was introduced. The performance of the algorithm was tested against commonly used forest ground truth data, GLAD. The comparison of the accuracy performance between DELSAR and widely used classification algorithms such as classical K-means, DBScan, and Meanshift algorithm were discussed. The state of environmental phenomena can be acquired by environmental quality standard, in which standard is widely used to specify the desired state of deforestation phenomena.

The major impediment for the interpretation of deforestation phenomena from L-Band SAR data using semantic computing lies in the computational of the closest meaning to the deforestation area. the Closest Semantic-Ellipsoid Algorithm is the new key to overcome this problem because it allows a computational of closest meaning in the interval domain which commonly found in the environmental study. In result, the meaning can be acquired by the formula derived from the trigonometry-based semantic algorithm. The fundamental principles of the Closest Semantic-Ellipsoid Algorithm have been introduced in Chapter 4. The proposed essential semantic interval-dimensions derived from heterogonous satellite images, the combination of L-Band SAR and optical, namely: HV gamma-naught, red, green, blue, NIR, SWIR channel; and combinational
features: soil temperature, soil moisture, temporal change density, temporal change velocity, shape, and texture. Afterward, semantic computing is employed as an analysis model to explain significant knowledge of deforestation activity. The results show that integrated independent dimensions from both the optical and SAR domains have the potential for presenting for aspect-based deforestation assessment and to enable the design of robust deforestation interpretation system. The experiment result confirms the ability of the proposed approach to semantic computing for ALOS-2/PALSAR-2 data to provide context-dependent forest monitoring and mapping. Chapter 5 contains the complete discussion on retrieval performance of deforestation interpretation system. This study demonstrates the combination of radar/active satellite imageries to complementing the optical/passive sensor satellite imageries particularly in the tropics where optical sensors are often constrained by the presence of haze, smoke, or clouds. In this study, independent dimensions of Landsat Thematic Mapper imagery (Landsat-8) and The Phased Array Type L-band SAR-2 (ALOS-2/PALSAR-2) were combined to create an integrated interpretation of environmental condition for a study area in the deforestation zone of global tropical rainforest.
REFERENCE


2004.


