

LANDSLIDE SUSCEPTIBILITY MAPPING OF ALBAY PROVINCE,
PHILIPPINES USING DEEP LEARNING ALGORITHMS



KARL MALCOLM N. CORDOVA

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Geoinformatics
Suranaree University of Technology
Academic Year 2023

การทำแผนที่ความเสี่ยงดินถล่มด้วยวิธีการเรียนรู้เชิงลึก จังหวัดอัลบาย
ประเทศฟิลิปปินส์

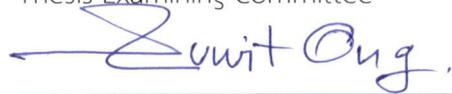


วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาภูมิสารสนเทศ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2566

LANDSLIDE SUSCEPTIBILITY MAPPING OF ALBAY PROVINCE, PHILIPPINES
USING DEEP LEARNING ALGORITHMS

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master' Degree.

Thesis Examining Committee



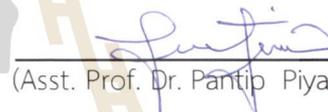
(Assoc. Prof. Dr. Suwit Ongsomwang)

Chairperson



(Dr. Tanakorn Sritarapipat)

Member (Thesis Advisor)



(Asst. Prof. Dr. Pantip Piyatadsananon)

Member (Thesis Co-Advisor)



(Dr. Siripon Kamontum)

Member



(Assoc. Prof. Dr. Songkot Dasananda)

Member



(Assoc. Prof. Dr. Yupaporn Ruksakulpiwat)

Vice Rector for Academic Affairs
and Quality Assurance



(Prof. Dr. Santi Maensiri)

Dean of Institute of Science

คาร์ล มัลคอล์ม คอร์โดวา : การทำแผนที่ความเสี่ยงดินถล่มด้วยวิธีการเรียนรู้เชิงลึก
จังหวัดอัลบาย ประเทศฟิลิปปินส์ อาจารย์ที่ปรึกษา : อาจารย์ ดร.ธรรณกร ศรีธราพิพัฒน์
อาจารย์ที่ปรึกษาร่วม : ผู้ช่วยศาสตราจารย์ ดร.พนทิพย์ ปิยะทัศนานนท์, 145 หน้า.

คำสำคัญ: ความเสี่ยงดินถล่ม, การทำแผนที่สินค้าคงคลังดินถล่ม, การเรียนรู้เชิงลึก, CNN-2D, MLP
จังหวัดอัลบาย (Albay) ได้ประสบกับภัยดินถล่มเป็นประจำ เนื่องจากมีที่ตั้งทางภูมิศาสตร์
ทางตะวันออกเฉียงใต้ของเกาะลูซอน (the island of Luzon) ทำให้จังหวัดนี้เสี่ยงต่อภัยพิบัติทาง
ธรรมชาติ เนื่องจากจังหวัดนี้ตั้งอยู่บนชายฝั่งทะเลตะวันออกของประเทศ จึงมีโอกาสถูกพายุไต้ฝุ่นซึ่ง
เข้าประเทศฟิลิปปินส์เป็นประจำ อีกทั้งจังหวัดอัลบายยังตั้งอยู่ในแนววงแหวนแปซิฟิกแห่งไฟ (the
Pacific Ring of Fire) ด้วย ดังนั้นจังหวัดนี้จึงได้รับผลกระทบจากการเคลื่อนของแผ่นเปลือกโลกด้วย
เนื่องจากมีรอยเลื่อนที่ยังคุกรุ่นอยู่ผ่านในพื้นที่ และเป็นที่ตั้งของภูเขาไฟมายอน (Mayon Volcano)
ซึ่งเป็นภูเขาไฟที่มีการปะทุมากที่สุดในฟิลิปปินส์

จากสถานการณ์เหล่านี้ การทำแผนที่ความเสี่ยงแผ่นดินถล่มจึงมีความจำเป็นอย่างยิ่ง
อย่างไรก็ตามยังขาดพื้นที่ดังกล่าว ยังขาดการศึกษาที่ครอบคลุมในสาขานี้ เนื่องจากความยากลำบาก
ในการเตรียมข้อมูลซึ่งมีขนาดใหญ่ ดังนั้นเพื่อให้สามารถสร้างแผนที่ดังกล่าวซึ่งครอบคลุมพื้นที่ขนาดใหญ่
ใหญ่ การศึกษานี้จึงใช้เทคนิคการสำรวจระยะไกลด้วยวิธีสองเทคนิคหลัก ได้แก่ การวิเคราะห์
ฮอตสปอตข้อมูล PSI (persistent scatterer interferometry) (PSI-based hotspot analysis)
และการจัดกลุ่มดัชนีสเปกตรัม (clustering of spectral indices)

โดยเทคนิคแรก เป็นการวิเคราะห์ฮอตสปอตข้อมูล PSI ใช้เพื่อระบุตำแหน่งที่มีโอกาสเกิด
การเกิดดินถล่มสูง โดยอิงตามเกณฑ์ความเสถียรของข้อมูล PSI ทั้งนี้ข้อมูล PSI สามารถคำนวณได้
จาก ภาพถ่าย SAR (Synthetic Aperture Radar) จากดาวเทียม Sentinel-1 ซึ่งประกอบไปด้วย
ภาพแนวขึ้น 62 ภาพ และภาพแนวลง 65 ภาพ ตั้งแต่วันที่ 30 กรกฎาคม ค.ศ. 2017 ถึง 31 ธันวาคม
ค.ศ. 2020 โดยสกัดข้อมูล PSI ออกมาได้ทั้งหมด 277,525 จุด ซึ่งวัดค่า V_{LOS} (line-of-sight
velocity) โดยมีเกณฑ์ความเสถียรที่ 11 มม./ปี โดยการวัดค่า V_{LOS} ถูกฉายไปตามทิศทางของความ
ลาดชันที่ชันที่สุดเพื่อวัดความเร็วการเสียรูปของความลาดชัน (slope deformation velocity)
(V_{SLOPE}) และได้ผลลัพธ์ของการวิเคราะห์ฮอตสปอตของ Getis-Ord G_i^* ดังนั้น จาก 871 จุด มีระดับ
ความมั่นใจอยู่ที่ 99% และมี 356 จุดที่มี $V_{SLOPE} > 33$ มม./ปี และจัดตั้งกล่าวจะถูกกำหนดให้เป็นจุด
ดินถล่ม

อีกหนึ่งเทคนิคที่ใช้สำหรับการทำแผนที่ความเสี่ยงดินถล่ม คือการตั้งคุณลักษณะของพื้นที่ดินถล่ม ด้วยการจัดกลุ่มดัชนีสเปกตรัมแบบ Gaussian (the Gaussian clustering of spectral indices) ในการศึกษาที่ใช้ดัชนีสเปกตรัมสองตัว ได้แก่ NDVI (Normalized difference vegetation index) และ BSI (Bare Soil Index) และข้อมูลเสริมจากภาพของความชัน (slope raster) ซึ่งข้อมูลทั้งหมดถูกนำมาใช้เป็นข้อมูลอินพุตในอัลกอริธึมการจัดกลุ่มแบบ Gaussian เพื่อหาพื้นที่ดินถล่มแบบกึ่งอัตโนมัติ โดยได้รับการยืนยันจากกลุ่มกระจุกตัวของจุดเกิดดินถล่ม โดยการเปรียบเทียบภาพถ่ายดาวเทียมก่อนเกิดดินถล่มและหลังเกิดดินถล่มจากสามเหตุการณ์ที่แตกต่างกัน ในช่วงหลังพายุไต้ฝุ่น โดยสามารถหาพื้นที่ดินถล่มได้ทั้งหมด 200 พื้นที่ จากกระบวนการประมวลผลด้วยตนเอง

ข้อมูลดินถล่มถูกแบ่งออกเป็นสองส่วน ได้แก่ ข้อมูลการฝึก 70% ซึ่งถูกใช้เป็นข้อมูลป้อนเข้าในอัลกอริธึมการเรียนรู้เชิงลึก (Deep Learning) สำหรับการหาแผนที่ความเสี่ยงดินถล่ม และข้อมูลการทดสอบ 30% ซึ่งถูกใช้เป็นข้อมูลสำหรับการประเมินความถูกต้องของอัลกอริธึม โดยได้ผลลัพธ์ดังนี้ คือ อัลกอริธึมแบบการเรียนรู้เชิงลึก ในรูปแบบ CNN (Convolutional Neural Network) และ MLP (Multilayer perceptron) ให้ความถูกต้องที่สูงถึง 92% และ 89% ตามลำดับ เมื่อเปรียบเทียบกับอัลกอริธึมทั่วไปให้ความถูกต้องที่น้อยลงมา พบว่า RF (Random Forest) ให้ความถูกต้องที่อยู่ที่ 86% ตามมาด้วย SVM (Support Vector Machine) อยู่ที่ 85% และ LR (Logistic Regression) ให้ความถูกต้องน้อยที่สุดคือ 79%

จากการประเมินพื้นที่ความเสี่ยงของดินถล่มในแต่ละจังหวัด พบว่า จากพื้นที่ทั้งหมด 2,500 ตร.กม. มีพื้นที่เสี่ยงแผ่นดินถล่มที่สูงมาก อยู่ที่ 990.92 ตร.กม. และมีพื้นที่เสี่ยงแผ่นดินถล่มที่สูง 701.58 ตร.กม. ทั้งนี้ในการหาแผนที่ความเสี่ยงต่อประชกรนั้น แสดงให้เห็นว่า จาก 720 ตำบล ตำบลที่มีความเสี่ยงสูงถึง 290 ตำบล โดย 103 แห่งในนั้นมีความเสี่ยงสูงมาก โดยเขตเทศบาลเมืองบนเกาะราปูราปู (Rapu-rapu) ทั้งหมดมีความเสี่ยงสูงมาก ส่วนสำคัญของเขตเทศบาลตีวี (Tiwi), มาลินาว (Malinao), มานิตอ (Manito) และโยเวลีอาร์ (Jovellar) ก็มีความเสี่ยงที่สูงมากเช่นกัน ในขณะที่พื้นที่ทั้งหมดของเมืองเลกซปี (Legazpi) มีความเสี่ยงที่ต่ำมากที่จะเกิดดินถล่ม จากผลการศึกษาแสดงให้เห็นถึงความจำเป็นในการลดความเสี่ยงจากภัยพิบัติที่สามารถนำไปใช้ในการวางแผนจัดการเพื่อลดผลกระทบของแผ่นดินถล่มในเขตที่มีความเสี่ยงได้

สาขาวิชาคณิตศาสตร์และภูมิสารสนเทศ
ปีการศึกษา 2566

ลายมือชื่อนักศึกษา Kwardaa
ลายมือชื่ออาจารย์ที่ปรึกษา Tanghoun Sukarnipal
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม Luafin

KARL MALCOLM N. CORDOVA : LANDSLIDE SUSCEPTIBILITY MAPPING OF
ALBAY PROVINCE PHILIPPINES USING DEEP LEARNING ALGORITHMS. THESIS
ADVISOR : TANAKORN SRITARAPIPAT, Ph.D. THESIS CO-ADVISOR : ASST. PROF.
PANTIP PIYATADSANANON, Ph.D. 145 PP.

Keyword: LANDSLIDE SUSCEPTIBILITY, LANDSLIDE INVENTORY MAPPING, DEEP LEARNING, CNN-2D, MLP

The province of Albay regularly experiences landslides due to its geographic location in the southeastern part of the island of Luzon, making it prone to natural disasters. Being located in the country's eastern seaboard, the province is regularly hit by typhoons entering the Philippines. Since the province is situated in the Pacific Ring of Fire, Albay is also tectonically active as characterized by the presence of active faults traversing the area and being the location of Mayon Volcano, which is the most active volcano in the Philippines.

Despite these circumstances which necessitate the need for landslide susceptibility mapping, there is still a lack of comprehensive studies in this field due to the difficulty of preparing and obtaining a landslide inventory. To overcome these challenges, the study utilized two remote sensing techniques to identify landslide features, namely PSI (persistent scatterer interferometry)-based hotspot analysis, and clustering of spectral indices.

The first technique; PSI was used to identify locations where landslides are highly likely to occur based on their stability threshold. PSI was performed on a stack of 62 ascending and 65 descending Sentinel-1 SAR images dated from July 30, 2017 to December 31, 2020, extracting a total of 277,525 PSI points measuring V_{LOS} (line-of-sight velocity) with a stability threshold of 11 mm./yr. V_{LOS} measurements were projected along the direction of the steepest slope to measure slope deformation velocity (V_{SLOPE}). Results of Getis-Ord G_i^* hotspot analysis showed that out of 871 points with a 99% degree of confidence, 356 points have a V_{SLOPE} of $>33\text{mm./yr.}$ and are designated as landslide points.

Another technique used for landslide inventory mapping is the extraction of landslide features from the Gaussian clustering of spectral indices. In this study, two spectral indices (NDVI and BSI) supplemented with slope raster were used as input data in a Gaussian clustering algorithm to semi-automatically extract confirmed landslide features on clusters by comparing pre-landslide and post-landslide satellite images across three different dates, all of which were taken during the aftermath of typhoons. A total of 200 landslide features were successfully extracted through the manual post-processing of Gaussian clustering output features.

The landslide inventory data were split into 70% training data used as input in deep learning algorithms for landslide susceptibility mapping, and 30% testing data for accuracy assessment. Both deep CNN-2D and MLP showed high overall accuracy at 92% and 89%, respectively. Comparison with conventional machine learning algorithms showed that RF can perform on par with MLP with an overall accuracy of 86% followed closely by SVM at 85% and with LR showing the lowest accuracy at 79%.

Assessment of landslide susceptibility in the province shows that a total land area of 990.92 sq.km. out of 2500 sq.km. was classified as having very high susceptibility and 701.58 sq.km. at high susceptibility. Population risk exposure mapping also showed that 290 out of 720 subdistricts are at high risk with 103 of them at very high risk. The entire island municipality of Rapu-rapu is at very high risk. Significant parts of the municipalities of Tiwi, Malinao, Manito, and Jovellar are also at very high risk while the entirety of the city of Legazpi has very low risk to landslides. The results of the study imply the need of robust disaster risk reduction and management planning to mitigate the risk of landslides in vulnerable municipalities.

School of Mathematical Sciences
and Geoinformatics
Academic Year 2023

Student's Signature *Mucardana*
Advisor's Signature *Tanahuan Suberapikat*
Co-advisor's Signature *afufo*

ACKNOWLEDGEMENTS

First and foremost, I would like to thank all of the faculty of the School of Geoinformatics who have been an integral part of honing my abilities during my academic journey. I would like to give thanks to the ajarns who did their utmost best in helping me learn core subjects when I was first starting my studies in Thailand despite the complications encountered during the height of the pandemic namely Assoc. Prof. Dr. Suwit Ongsomwang, Dr. Siripon Kamontum, and Dr. Songkot Dasananda. I would like to especially thank my adviser Dr. Tanakorn Sritarapipat for his understanding and support on the duration of my research and my thesis co-adviser Asst. Prof. Dr. Pantip Piyatadsananon for her insightful comments.

I would also like to express my thanks to all of friends and seniors for their patience and willingness in giving a hand whenever I am struggling with my academics or needing someone to reach out to.

To all of my friends from different programs and those who currently reside in the Philippines particularly my fellow alumni from Partido State University's BS Geology program, I am very grateful for your presence and support at the times I needed it.

And to my parents, Josemari V. Cordova and Eunice C. Novio who themselves are also ajarns, for financially supporting my studies in Thailand and supported my decision to enroll in graduate school in the first place.

Karl Malcolm Cordova

CONTENTS

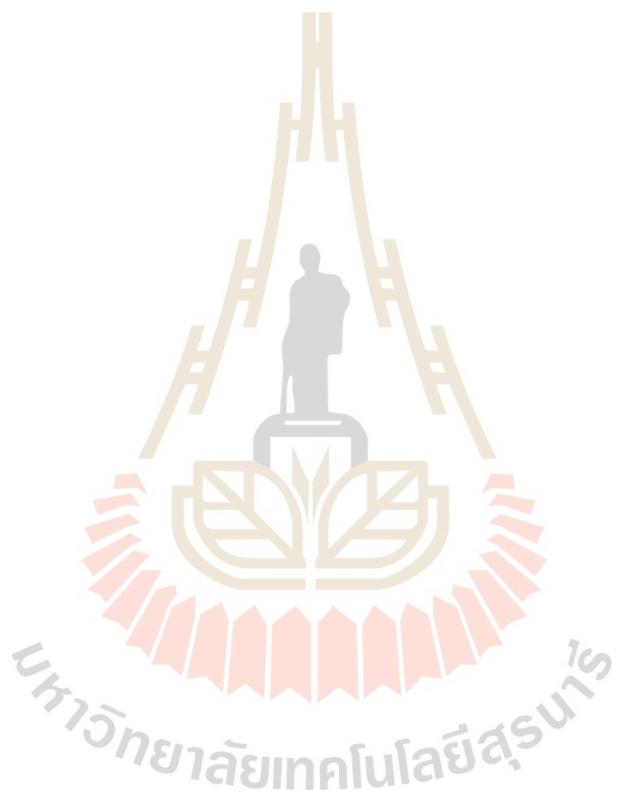
	Page
ABSTRACT IN THAI.....	I
ABSTRACT IN ENGLISH.....	III
ACKNOWLEDGEMENTS.....	V
CONTENTS.....	VI
LIST OF TABLES.....	IX
LIST OF FIGURES.....	X
LIST OF ABBREVIATIONS.....	XIII
CHAPTER	
I INTRODUCTION.....	1
1.1 Background of the Study.....	1
1.2 Research Objectives.....	2
1.3 Types of Landslide.....	2
1.4 Study Area.....	8
II LITERATURE REVIEW.....	11
2.1 Landslide Inventory Mapping.....	11
2.1.1 Overview of landslide inventory mapping.....	11
2.1.2 Detection and interpretation of slow-moving landslides using PSI.....	11
2.1.3 Extraction of landslide features from spectral indices.....	12
2.2 Landslide susceptibility mapping using deep learning algorithms.....	13
2.3 Landslide susceptibility mapping using conventional machine learning algorithms.....	14
2.4 Landslide vulnerability and risk assessment.....	15
III METHODOLOGY.....	17
3.1 Research Procedure.....	17
3.1.1 Scope and Limitations.....	17
3.2 Research Methodology.....	19

CONTENTS (Continued)

	Page
3.3 Data Acquisition and Preprocessing.....	21
3.4 Landslide Inventory Mapping.....	29
3.4.1 Persistent Scatterer Interferometry.....	29
3.4.2 Hotspot Analysis of PSI points.....	32
3.4.3 Gaussian clustering of spectral indices.....	33
3.5 Conventional Machine Learning Algorithms.....	36
3.5.1 Logistic regression.....	36
3.5.2 Random forest.....	37
3.5.3 Support vector machines.....	39
3.6 Deep Learning Algorithms.....	40
3.6.1 Multilayer perceptrons.....	40
3.6.2 Convolutional Neural Networks.....	41
3.7 Accuracy Assessment.....	43
3.8 Landslide Vulnerability and Risk Mapping.....	44
IV RESULTS.....	46
4.1 Landslide Inventory Mapping.....	46
4.1.1 PSI points processing.....	46
4.1.2 Landslide hotspot mapping of PSI points.....	51
4.1.3 Landslides extracted from spectral indices.....	55
4.2 Landslide Susceptibility Mapping.....	69
4.2.1 Optimal selection of landslide influencing factors.....	70
4.2.2 Landslide Susceptibility Mapping using Deep Learning Algorithms.....	71
4.2.3 Benchmark comparisons and accuracy assessment.....	79
4.3 Vulnerability and Risk Assessment.....	85
V CONCLUSION.....	91
5.1 Recommendations.....	92
REFERENCES.....	95
APPENDIX.....	105

CONTENTS (Continued)

	Page
CURRICULUM VITAE.....	145



LIST OF TABLES

Table	Page
3.1 Data types and sources.....	21
3.2 Selected landslide influencing factors.....	23
3.3 Parameters for calculating coefficient C	31
4.1 StaMPS parameters in MATLAB.....	46
4.2 Summary of acquired ascending and descending Sentinel-1 InSAR data.....	49
4.3 Summary of VSlope points.....	49
4.4 Hotspot analysis summary.....	53
4.5 Mutual information scores (IG Ratio) of each influencing factor.....	71
4.6 Specifications of the fully connected layers used by both MLP and CNN-2D.....	73
4.7 Specifications of CNN-2D model used in the study.....	74
4.8 Classification report of CNN-2D.....	78
4.9 Classification report of MLP.....	78
4.10 Area of landslide susceptibility classes.....	79
4.11 Classification report of logistic regression.....	80
4.12 Classification report of random forest including the parameters used.....	80
4.13 Classification report of SVM including the parameters used.....	80
4.14 Benchmark comparison metrics for all models used.....	81
4.15 Economic parameters for each city and municipality in Albay.....	86
4.16 Resiliency scores for each city and municipality.....	87

LIST OF FIGURES

Figure	Page
1.1 Schematic diagram of rockfall.....	3
1.2 Roadside rockfall at Sagñay-Tiwi Road, Patitinan, Camarines Sur.....	3
1.3 Schematic diagram of topple.....	4
1.4 Collapsed road at Malipot, Albay toppled due to erosion.....	4
1.5 Rotational and translational landslide.....	5
1.6 Translation landslide at Itogon, Benguet Province.....	5
1.7 Diagram of debris flow.....	6
1.8 Lahar flow from Mt. Mayon in the aftermath of Typhoon Goni.....	7
1.9 Diagram of debris avalanche.....	7
1.10 Debris avalanche landslide in Itogon, Benguet.....	8
1.11 Administrative boundaries of Albay Province.....	8
1.12 Bicol Volcanic Arc and regional structures.....	9
1.13 Active fault lines of Albay.....	10
3.1 Conceptual framework of landslide inventory and susceptibility mapping.....	20
3.2 Conceptual framework of landslide vulnerability and risk assessment.....	20
3.3 Elevation map of Albay in meters.....	23
3.4 Slope gradient map of Albay.....	24
3.5 Slope aspect of Albay.....	24
3.6 Buffer showing distance to active faults of Albay.....	25
3.7 Topographic wetness index of Albay.....	25
3.8 NDVI of Albay.....	26
3.9 Flow direction map of Albay.....	26
3.10 Terrain roughness index of Albay.....	27
3.11 Planform curvature of Albay.....	27
3.12 Profile curvature of Albay.....	28
3.13 Stream distance map of Albay.....	28

LIST OF FIGURES (Continued)

Figure	Page
3.14 Clustering boundary of K-means.....	34
3.15 Clustering boundary of Gaussian mixtures.....	35
3.16 Logistic regression curve.....	37
3.17 Schematic diagram of a random forest model.....	38
3.18 Hyperparameter separating two classes.....	39
3.19 RBF kernel separating non-linear classes in 3-d space.....	40
3.20 Structure of a typical MLP.....	41
3.21 Generalized CNN architecture.....	42
4.1 Resulting ascending VLOS points.....	47
4.2 Resulting descending VLOS points.....	48
4.3 Resulting VSlope points after spatial joining.....	50
4.4 Spatial autocorrelation results.....	51
4.5 Hotspot analysis results.....	52
4.6 Clusters located within Mayon Volcano.....	53
4.7 PSI-derived landslide inventory.....	54
4.8 Aerial view of villages buried by lahar flow in Legazpi City, Albay.....	56
4.9 Training points for November 2006 landslides.....	57
4.10 Silhouette score of GMM clusters extracted from November 2006 landslides.....	58
4.11 Pre-landslide and post-landslide Landsat-5 image of Guinobatan.....	58
4.12 Pre-landslide and post-landslide Landsat-5 image of Sto. Domingo.....	59
4.13 Pre-landslide and post-landslide Landsat-5 image of Daraga.....	59
4.14 Pre-landslide and post-landslide Landsat-5 image of Camalig.....	59
4.15 Aftermath of TD Usman in Tiwi, Albay dated Dec. 30, 2018.....	60
4.16 Training points for December 2018 Tiwi landslides.....	61
4.17 Silhouette score of clusters extracted from December 2018 Tiwi landslides.....	62
4.18 Sentinel-2 pre-landslide image of Tiwi, Albay dated February 9, 2018.....	62
4.19 Sentinel-2 post-landslide image of Tiwi, Albay dated January 15, 2019.....	63
4.20 100-feet chasm overlooking Bulawan River in Malilipot, Albay.....	64

LIST OF FIGURES (Continued)

Figure	Page
4.21 Training points for 2020 Malilipot landslides.....	65
4.22 Silhouette score of GMM clusters extracted from 2020 Malilipot landslides.....	66
4.23 Sentinel-2 pre-landslide image of Malilipot dated April 10, 2017.....	66
4.24 Sentinel-2 post-landslide image of Malilipot dated May 9, 2021.....	67
4.25 Landslide inventory map used in the study including non-landslide points.....	69
4.26 Feature importance plot.....	71
4.27 Schematic diagram of deep MLP used in the study.....	73
4.28 Schematic diagram of the CNN-2D model.....	75
4.29 Loss and accuracy metrics of CNN-2D in every epoch.....	75
4.30 Loss and accuracy metrics of Deep MLP in every epoch.....	76
4.31 Landslide susceptibility map of Albay using CNN-2D optimized with RMSprop..	77
4.32 Landslide susceptibility map of Albay with Deep MLP.....	77
4.33 ROC plots of deep learning and machine learning algorithms.....	81
4.34 Landslide susceptibility map of Albay using logistic regression.....	84
4.35 Landslide susceptibility map of Albay using random forest.....	84
4.36 Landslide susceptibility map of Albay using SVM.....	85
4.37 Social vulnerability map of Albay.....	88
4.38 Population map of Albay per subdistrict.....	88
4.39 Landslide risk exposure map of Albay.....	89

LIST OF ABBREVIATIONS

AUC	=	Area under ROC curve
BSI	=	Bare Soil Index
CNN	=	Convolutional Neural Networks
DEM	=	Digital Elevation Model
GMM	=	Gaussian mixture model
LOS	=	Line-of-sight
LR	=	Logistic regression
MGB	=	Mines and Geosciences Bureau
MLP	=	Multilayer Perceptron
NDRRMC	=	National Disaster Risk Reduction and Management Council
NDVI	=	Normalized Difference Vegetation Index
PAGASA	=	Philippine Atmospheric Geophysical and Astronomical Services Administration
PHIVOLCS	=	Philippine Institute of Volcanology and Seismology
PSI	=	Persistent Scatterer Interferometry
RF	=	Random Forest
ROC	=	Receiver Operating Characteristic
SAR	=	Synthetic Aperture Radar
SVM	=	Support Vector Machine
TWI	=	Topographic Wetness Index
TRI	=	Terrain roughness index
VRA	=	Vulnerability and Risk Assessment

CHAPTER I

INTRODUCTION

1.1 Background of the Study

Landslides are one of the most common types of geohazard, which refers to the downward movement of soil due to slope failure driven by gravity. Several factors, physical, geological, and hydrologic characteristics, influence the rate at which landslide events occur or potentially occur. Froude and Petley (2018) found that 55997 people were killed by landslides over 12 years between 2004 and 2016 in 4862 landslide events, with 75% of these events having occurred in Asia. In the same study, the Philippines alone accounts for 46% of rainfall-induced landslides in Southeast Asia, 42% of which were triggered by typhoons. The most recent notable landslide event in the province occurred on November 1, 2020 during the onslaught of Typhoon Goni where at least 14 people were reported dead after lahar flows from the Mayon Volcano ravaged nearby villages. The aftermath of the disaster left thousands of homes buried with damages estimated to be at least \$369 million, according to the NDRRMC.

The province of Albay regularly experiences landslides, the most common of which are flow-type landslides: debris flow (specifically, lahar flows) and debris avalanches. Landslide susceptibility mapping is crucial to vulnerable areas to mitigate the future impacts of a reoccurrence. Due to its dynamic environment being geologically active and the inevitable effect of climate change further intensifying typhoons, conducting landslide susceptibility mapping in the province of Albay is of utmost importance.

The mountainous terrain, complex (mostly volcanic) geology, and geographical location of the province of Albay in the eastern seaboard of the Philippines make it vulnerable to combined natural disasters with devastating effects on infrastructure and public safety. The most recent notable landslide event in the province occurred on November 1, 2020 during the onslaught of Typhoon Goni where at least 14 people were reported dead after lahar flows from the Mayon Volcano ravaged nearby villages.

The aftermath of the disaster left thousands of homes buried with damages estimated to be at least \$369 million, according to the NDRRMC (National Disaster Risk Reduction and Management Council).

Despite these circumstances, there is still a relative lack of comprehensive research focused on assessing the landslide susceptibility of the province in part due to the lack of available landslide inventory data in the area. The most recent study conducted in the province of Albay to assess landslide vulnerability was the use of hexagonal binning techniques by Abante (2021), which was only able to identify 376 sq.km of highly vulnerable areas which was only located inside the 12 km. radius of Mayon Volcano.

1.2 Research objectives

The objectives of this research were:

- I. To create a landslide inventory map of the study area using:
 - a. hotspot clustering of slope velocity points measured from PSI (persistent scatterer interferometry) analysis of SAR (synthetic aperture radar) data, and;
 - b. visual identification of landslide features from Gaussian clustering of spectral indices (BSI, NDVI).
- II. Utilize the produced landslide inventory map to create a landslide susceptibility map using deep learning algorithms (CNN-2D and MLP).
- III. Assess the performance of deep learning algorithms for landslide susceptibility mapping by comparing its ROC and overall accuracy with conventional ML algorithms (LR, RF, SVM)
- IV. Create a vulnerability and population risk exposure map of Albay to identify the number of subdistricts at high risk of landslides.

1.3 Types of Landslides

According to the USGS Landslide Handbook, landslides can be generally classified based on the type of material they carry: **earth** (fine-sized rocks or soil particles, or both), and **debris** (larger and coarser materials). For clarification purposes,

landslides are classified further according to their displacement as defined by Highland and Bobrowsky (2008):

A. Fall

Abrupt, downward movements of rock or earth, or both, that detach from steep slopes or cliffs. The falling material usually strikes the lower slope at angles less than the angle of fall, causing bouncing. The falling mass may break on impact, begin rolling on steeper slopes, and continue until the terrain flattens.



Figure 1.1 Schematic diagram of rockfall.



Figure 1.2 Roadside rockfall at Sagñay-Tiwi Road, Patitinan, Camarines Sur (Photo from Isarog Radio Broadcasting).

B. Topple

A topple is recognized as the forward rotation out of a slope of a mass of soil or rock around a point or axis below the center of gravity of the displaced mass. Toppling is sometimes driven by gravity exerted by the weight of material upslope from the displaced mass.



Figure 1.3 Schematic diagram of a topple.



Figure 1.4 Collapsed road at Malilipot, Albay toppled due to erosion dated November 2020 (Photo from The Philippine Examiner).

C. Slides

A slide is a downslope movement of a soil or rock mass occurring on rupture surfaces or relatively thin zones of intense shear strain. It is further classified into two types: **rotational**, if the surface of rupture is curved upward (spoon-shaped) and the slide movement is more or less rotational about an axis that is parallel to the contour of the slope, and **translational**, if the mass moves out, or down and outward, along a relatively planar surface with little rotational movement or backward tilting.

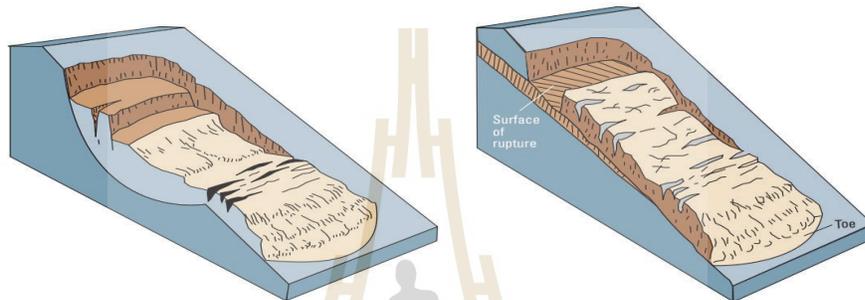


Figure 1.5 Rotational landslide (left) and translational landslide (right).



Figure 1.6 Translation landslide at Itogon, Benguet Province, Sept 27, 2018. (Photo from Associated Press).

D. Flow

A flow is a spatially continuous movement in which shear surfaces are short-lived, closely spaced, and usually not preserved. The component velocities in the displacing mass of a flow resemble those in a viscous liquid. Often, there is a gradation of change from slides to flows, depending on the water content, mobility, and evolution.

Debris flow

A form of rapid mass movement in which loose soil, rock, and sometimes organic matter combine with water to form a slurry that flows downslope. They have been informally and inappropriately called "mudslides" due to the large quantity of fine material that may be present in the flow. Debris flows carrying tephra and volcanic materials are called "lahar" flows.

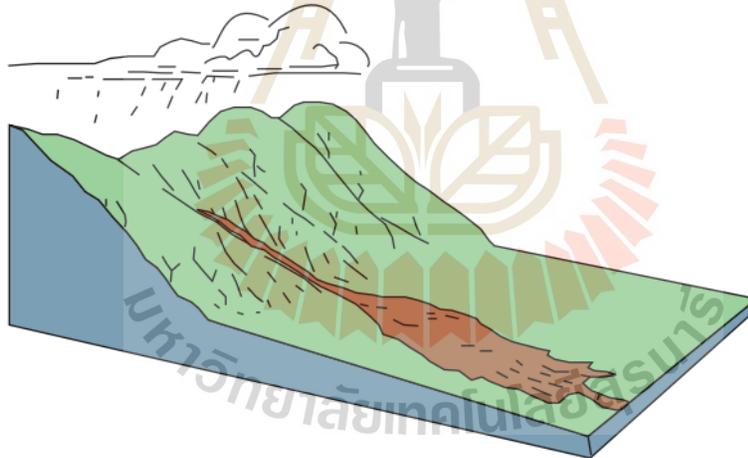


Figure 1.7 Diagram of a debris flow.



Figure 1.8 Lahar flow from Mt. Mayon burying homes in Guinobatan, Albay in the aftermath of Typhoon Goni dated November 2020 (Photo from GMA Network).

Debris avalanche

Debris avalanches are essentially large, extremely rapid, often open-slope flows formed when an unstable slope collapses, and the resulting fragmented debris is rapidly transported away from the slope.



Figure 1.9 Diagram of a debris avalanche.



Figure 1.10 Debris avalanche landslide in Itogon, Benguet dated January 2018 (Photo from Reuters).

1.4 Study Area

The province of Albay is located in the southeastern part of the largest island in the Philippines, Luzon, covering a total land area of 2575.77 sq.km. The province is divided into three cities and 15 municipalities, further divided into three cities and 15 municipalities into 720 subdistricts, with a total population of 1,374,768 as of 2020.

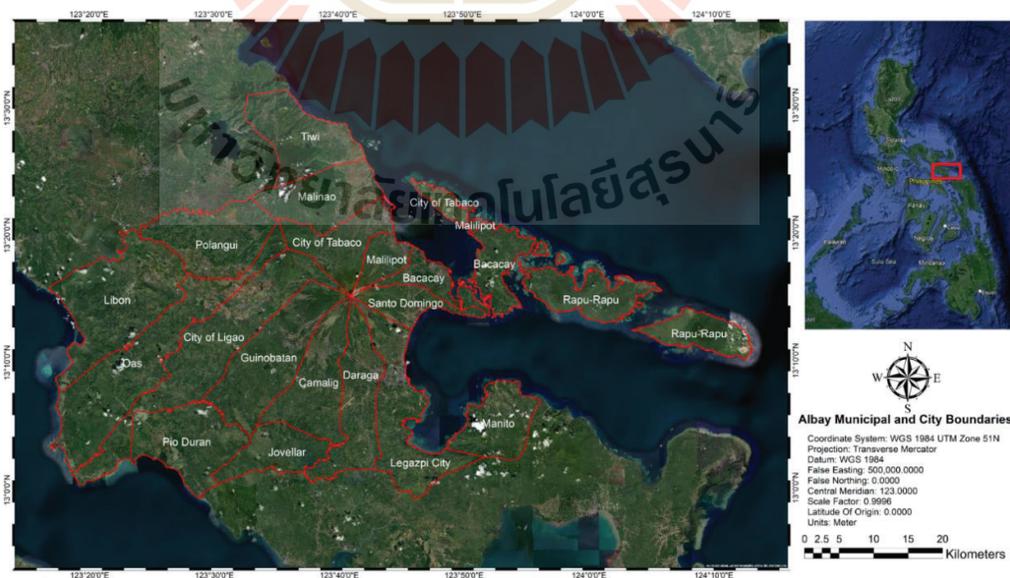


Figure 1.11 Administrative boundaries of Albay Province.

Due to its location being located in the eastern seaboard of the central Philippines, the province of Albay is prone to typhoons. The active Legazpi Lineament and Linao Fault traverse directly through the province, making it prone to earthquakes. In addition, the iconic Mayon Volcano, a stratovolcano with a near-perfect cone shape and the most active volcano in the Philippines is also located in the province.

Another volcano classified as potentially active is Mount Masaraga, which serves as the province's crucial geothermal energy hub. According to a 2013 report from the Human Development Network, the province of Albay suffers from a combined risk of natural disasters. An average of 20 typhoons hit the province each year, two destructive, and as many as 86 baranggays (villages) within three cities and municipalities are at risk from volcanic hazards. It is estimated that around 73% of the province is vulnerable to landslides, with 127 villages at high risk.

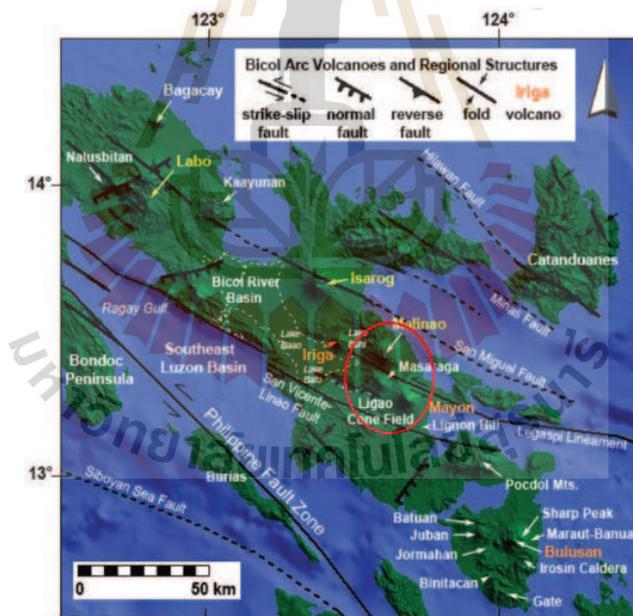


Figure 1.12. Bicol Volcanic Arc and regional structures with Masaraga, Mayon and Malinao highlighted. Obtained from Minimo and Lagmay (2016).



Figure 1.13 Active fault lines of Albay obtained from PHIVOLCS.

CHAPTER II

LITERATURE REVIEW

2.1 Landslide Inventory Mapping

2.1.1 Overview of landslide inventory mapping

Gathering landslide points is the first step in conducting landslide susceptibility mapping. According to Guzzetti et al. (2012), the quality of a landslide inventory depends on its accuracy and the type and certainty of the information provided. The easiest and most straightforward method involves gathering the location of previous landslide events from various sources, either from previous studies in the area or through an existing landslide inventory database. Geological field mapping is a common method in studying landslide features, but due to dynamic natural and human processes, there is considerable difficulty in this approach. Moreover, the in-situ visual investigation does not take into account the occurrences of historical landslides.

In the absence of historical data, NASA provides a freely available Global Landslide Catalog as part of the Cooperative Open Online Landslide Repository (COOLR) project (<https://gpm.nasa.gov/landslides/data.html>), where data is provided mostly by concerned professionals and citizen scientists, supplemented with online news articles (Juang et al., 2019).

2.1.2 Detection and interpretation of slow-moving landslides using PSI

PSI has been successfully conducted over the past decade to detect slow-moving ground motion. Persistent scatterer interferometry (PSI) is a differential InSAR technique that estimates the line-of-sight (LOS) yearly displacement rate from a sparse grid of processed radar points collected from multiple SAR images. Lu et al. (2014) stated that PSI processing helps overcome the disadvantages of temporal decorrelation and atmospheric noises exhibited by DInSAR methods by decoupling height and deformation over long time-series analysis. The results of PSI points, however, are highly dependent on the number of SAR images selected, the coherence of the image,

and the time interval between scenes. In the case of highly vegetated areas, PSI tends to exhibit noise and low density due to a lack of stable signal reflectors, such as buildings and barren land. To this end, it is necessary to adjust coherence manually and weeding parameters accordingly to account for noisy pixels while gathering as many PS points as possible.

A study by Cigna et al. (2013) analyzed the reliability of PSI in confirming previous landslide inventories and mapping the extent of potential landslides combined with in situ mapping for validation. A later study by Lu et al. (2014) described how PSI helped detect anomalous clusters of slow-moving landslides over mountainous and hilly areas. Using hotspot analysis to analyze and interpret areas where slope failures are likely to occur, their study detected new slow-moving landslides not present in previous landslide inventories by using ancillary topographic maps and optical images.

PSI has also been used for the refinement of existing landslide susceptibility maps. Ciampalini et al. (2016) described how the integration of PSI into existing landslide susceptibility maps can increase the degree of susceptibility over certain areas. Their findings also indicate how PSI successfully increased the reliability of new landslide susceptibility maps by detecting slow-moving landslides that typically occur in urban areas.

A more recent study by Cigna et al. (2021) aimed to comprehensively quantify the accuracy of PSI derived from Sentinel-1 SAR images by comparing their displacement velocity from GNSS and ground levelling data. Their results found a significantly high R-squared of 0.95 compared with geodetic levelling data with a relative error of only 20% for targets subsiding 15mm/year, which significantly decreases to 16% for targets subsiding faster than 50mm/year.

2.1.3 Extraction of landslide features from spectral indices

Spectral indices derived from multispectral satellite imagery provide an avenue to easily identify land cover and the state of vegetation in an area. In mountainous areas where slope failures are prevalent, spectral indices aid in identifying landslide features. One of the most spectral indices used is the normalized difference vegetation index (NDVI), which quantifies the health and amount of vegetation over the area

coverage. Yang, Wang and Shi (2013) highlighted the efficacy of NDVI in detecting landslides after the occurrence of events that induce rapid downslope movements, such as earthquakes, by identifying areas exhibiting sudden drops in NDVI values.

The heuristic approach in visually interpreting landslide features using NDVI is further highlighted by Fiorucci et al. (2019). By integrating stereoscopic multispectral images maximizing the use of radiometric and terrain information, the study was able to identify additional landslides than what was previously identified in existing landslide inventory maps due to better recognition of narrow, channeled landslides.

While still not commonly used for landslide detection, bare soil index (BSI) has also been successful in visual interpretation of landslide features due to its ability to discern sparsely vegetated and barren areas from dense canopies. Ariza et al. (2021) demonstrated the potential of bare soil index derived from Sentinel-2 images by successfully detecting 62% of landslides present compared to high-resolution SPOT-7 imagery.

2.2 Landslide susceptibility mapping using deep learning algorithms

Advancements in computational hardware increased the scalability of deep learning algorithms allowing for an increase in predictive performance. Given a set of input data with conditioning factors serving as independent variables, deep learning algorithms calculate the weights of input features in hidden layers in order to extract meaningful information. Park and Lek (2016) highlighted that aside from this inherent advantage, optimized deep learning models have the ability to derive solutions directly from the input data due to their adaptive nature. There are also relatively few parameters to consider in deep learning models allowing for streamlined deployment of models.

A 2019 study by Wang, Fang, and Hong using deep convolutional neural networks (CNN) found how convolutional neural networks, specifically CNN-2D models can be applied for regional analysis of landslide susceptibility. The study concluded how CNN-2D can slightly outperform other convolutional neural networks such as CNN-1D, CNN-3D, and LeNet5 in terms of overall accuracy achieving a score of 77.63%, 4% higher than other deep learning models and 7% higher than SVM during assessment.

A similar study conducted by Yi et al. (2020) on the use of convolutional neural networks for landslide susceptibility mapping on a regional scale found that the model significantly outperformed standard neural networks and logistic regression in terms of both training and validation accuracy metrics. Using multi-scale fusion input data to effectively expand the receptive field and improve feature extraction, CNN model obtained a training accuracy of 97%, 16% higher than standard neural networks with a high validation accuracy of 88% which is also a 6% increase compared with other models.

Subsequent studies conducted by Tien Bui et al. (2020) focusing on multilayer perceptrons found that deep MLPs outperform conventional machine learning algorithms and even shallow neural networks by a considerable margin achieving an overall accuracy of 90.53% and a high sensitivity of 95%. Comparison with a standard shallow neural network showed a 7% increase in sensitivity and 3% increase in overall accuracy. While there was a considerable difference in terms of predictive performance, it was noted that this method comes at the expense of heavier computational load and long process of fine-tuning hyperparameters to achieve high performance metrics; a trade-off that future researchers must consider.

The characteristic of convolutional neural networks to extract spatial patterns which are more representative of real-world landslide phenomena was further emphasized in a study by Azarafza et al. (2021). By not being constrained into a single pixel, the model was able to extract meaningful information contributing to a high accuracy score of 90.9%, outperforming other models by at least 5%.

2.3 Landslide susceptibility mapping using conventional machine learning algorithms

Landslide susceptibility mapping involves combining multiple influencing factors as raster of different formats and scales. This often presents problems in some lower-end hardware especially on studies conducted on a provincial or regional-scale. Deep learning models mostly require specialized hardware to run efficiently due to the complexities in its hyperparameters whereas conventional machine learning algorithms can be used as a relatively lightweight approach that can run on most entry-

level hardware. Conventional machine learning algorithms treat each input data as individual pixels allowing it to be trained faster at the expense of not being able to analyze for meaningful spatial patterns. Despite this obvious limitation, these algorithms still remain a viable option for binary classification of landslides that can provide a relatively high level of accuracy.

A study by Tien Bui et al. (2020) showed how SVM can perform on par with MLP in terms of sensitivity and overall accuracy with a score of 86.94% and 81%, respectively. The scores are only 1% lower than that observed from the MLP.

A 2022 study by Hussain et al. showed how logistic regression, one of the most common methods for landslide susceptibility mapping, can still perform with relatively high accuracy achieving an AUC score of 85.61%, a 10% increase as opposed to frequency ratio model which is another common method for mapping landslides.

2.4 Landslide vulnerability and risk assessment

Vulnerability and risk assessment are crucial components for drafting disaster risk reduction and management strategies. In areas where landslides are common or highly likely to occur, VRA is utilized to quantitatively determine the potential socioeconomic impact of disasters in order to identify areas that need higher priority for disaster mitigation.

According to Fell, et al. (2008), while there is an increasing need for quantitative risk management, there is no unifying procedure for risk mapping among countries leading to different accuracy and reliability of produced maps as a variety of input data and methodology is used similar to a statement described by Westen et al. (2008) regarding landslide susceptibility mapping. The study further emphasized that in the context of landslide zonation studies, the definition of hazard, susceptibility, and risk are used interchangeably. However, advances in GIS and remote sensing technology made it possible to utilize several tools to quantify risk and compare them with other hazards and risks present in an area.

In order to identify areas at risk of landslide, Arrogante-Funes et al. (2021) stated that landslide susceptibility and vulnerability mapping must be conducted as these procedures can help in finding variables that influence landslides. But as is the case

with landslide susceptibility and risk mapping, there is also no standard procedure in how vulnerability mapping can be conducted but suggested that socioeconomic factors must be considered. Subsequent studies by Singh and Kanungo (2021) and Wang, et al. (2021) stated that physical vulnerability can be integrated for landslide risk assessment to serve as an indicator of potential degree of damage to properties.

Risk mapping combines the quantified hazard and vulnerability classes in order to identify expected economic damages and loss of human life as a result of exposure to landslides. This procedure is a disaster mitigation strategy that can help policymakers identify areas that need higher priority in order to minimize losses. However, as hazard maps rely on dynamic spatio-temporal data the resulting maps can vary annually over certain periods of time as opposed to using static input data as is the case with landslide susceptibility mapping. Since landslide susceptibility mapping also integrates geophysical parameters such as annual precipitation, temperature, and NDVI which may be considered as spatio-temporal data, Arrogante-Funes, et al. (2021), and Ram and Gupta (2022) suggests the use of resulting landslide susceptibility maps as a substitute for a landslide hazard map even though the definition of terminologies are technically different.



CHAPTER III

METHODOLOGY

3.1 Research Procedure

3.1.1 Scope and Limitations

This section provides a summary of the scope of the study which is as follows:

1) For preparing a historical landslide inventory using spectral indices, the satellite imagery from Landsat and Sentinel-2 platforms will cover various time periods. The selection of suitable satellite imagery will be based on the following factors: exact date when the landslide was reported; clarity of the location (eg. cloud cover) at the date of the satellite image, and visibility of the landslide features.

Bare soil index and NDVI will be extracted from multispectral satellite images and stacked together with a slope raster. Landslide features will be visually identified and the values of the point locations are to be used as input in a Gaussian clustering model. Clusters showing landslide features will be manually extracted from the resulting cluster output. The locations of these clusters on post-landslide imagery are to be visually compared with pre-landslide satellite imagery in order to ensure that the features are not pre-existing prior to the reported occurrence of landslides in the area.

2) For preparing PSI-based landslide inventory data, ascending and descending data from Sentinel-1A/1B platforms on single VV polarization over a 3-year time period (2017-2020) will be prepared in order to calculate the average slope displacement rate of the area. The resulting persistent scatterer (PS) points will be merged and calculated for their standard deviation which will be designated as the stability threshold.

To reduce the number of points, incident points will be aggregated into a 100 m x 100 m grid cell and calculated for their average values and average distances to the eight nearest neighboring features. The resulting output will be calculated using a

Getis-Ord G_i^* hotspot analysis to identify red and blue hotspots with 99% confidence. The degree of susceptibility will be calculated from the significant points based on the number of standard deviations of their average vertical slope displacement from the designated stability threshold.

3) Aggregated points derived from PSI and spectral indices will be used as input dependent variables. The values of their location will be extracted from the raster of landslide causative factors serving as independent variables. To facilitate binary classification of landslide susceptibility, an equal number of non-landslide points will be randomly generated. The resulting data will be divided into 70% training data as input and 30% test data for accuracy assessment.

4) The training data will be used to train deep learning models which will create the landslide susceptibility maps of the area. The models will be assessed for their accuracy. Comparative performance with conventional machine learning algorithms will also be conducted for benchmarking purposes.

5) On the vulnerability and risk assessment stage (VRA), vulnerability mapping will only consider social vulnerability quantified from economic parameters and resiliency scores of each municipality and city. Risk mapping will be calculated based on the percentage of population in each subdistrict exposed to landslides.

The following points describe the limitations encountered during the conduct of the study and how these were addressed:

1) Due to the geographical location and climate of the study area, there is increased difficulty in visually identifying the location of landslide features. For this purpose, only three major landslide events encompassing various dates will be identified from multispectral satellite imagery.

2) The pre-existing problems concerning the identification of landslide features from spectral indices extend to PSI analysis. In addition to the aforementioned, PSI derived from C-band SAR platforms show mixed performance in vegetated areas. This characteristic and atmospheric noise caused by surrounding ocean causes incoherence in many of the SAR images which has to be addressed manually through weeding of noisy PS points.

3) Aside from noise and coherence problems, another tradeoff to consider is the large amount of storage required to accommodate SAR images which in turn, also requires high computational power and longer processing times over repeated experimentations.

4) For extracting landslide features from spectral indices, the study will not fully employ change detection and the total landslide area as GMM clustering is a semi-automatic algorithm which is highly reliant on manually extracting relevant clusters.

5) Due to absence of high resolution images and lack of auxiliary data regarding properties, vulnerability and risk maps will not include infrastructures and accompanying economical costs of potential landslide disasters. Additionally, due to the lack of dynamic data needed for conducting a comprehensive landslide hazard mapping, the study will directly use the landslide susceptibility map created as a substitute for a hazard map during vulnerability and risk assessment stage.

3.2 Research Methodology

The methodology encompasses the general flow of the case study which is mainly separated into three parts: landslide inventory mapping, landslide susceptibility mapping, and landslide vulnerability and risk assessment (VRA).

In this report, this section is divided into five parts:

(1) Data acquisition and preprocessing – for landslide inventory mapping, Sentinel-1 SAR data is acquired to get PSI points, and spectral indices will be extracted from multispectral satellite imagery from

(2) Landslide inventory mapping techniques, specifically hotspot analysis of persistent scatterer interferometry (PSI) points, and semi-automatic extraction of landslide features from Gaussian clustering of spectral indices;

(3) Utilization and optimization of deep learning models, specifically CNN-2D (two-dimensional convolutional neural networks) and MLP (multilayer perceptron), and;

(4) Benchmark comparison of deep learning models with conventional machine learning algorithms (random forest, SVM, logistic regression) in terms of ROC score and overall accuracy.

(5) Vulnerability and risk assessment (VRA) to identify subdistricts exposed to landslide occurrences in terms of percentage of population.

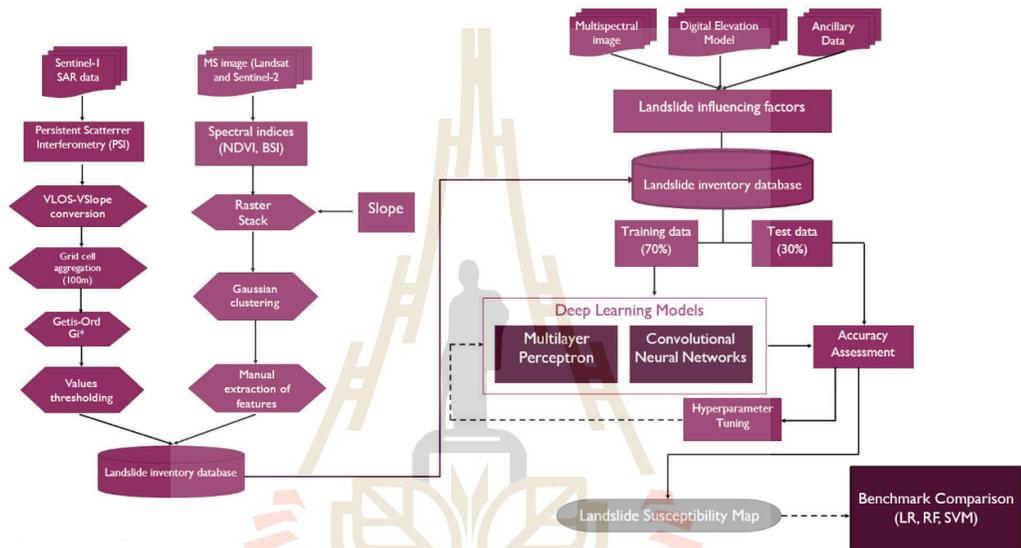


Figure 3.1 Conceptual framework of landslide inventory and susceptibility mapping.

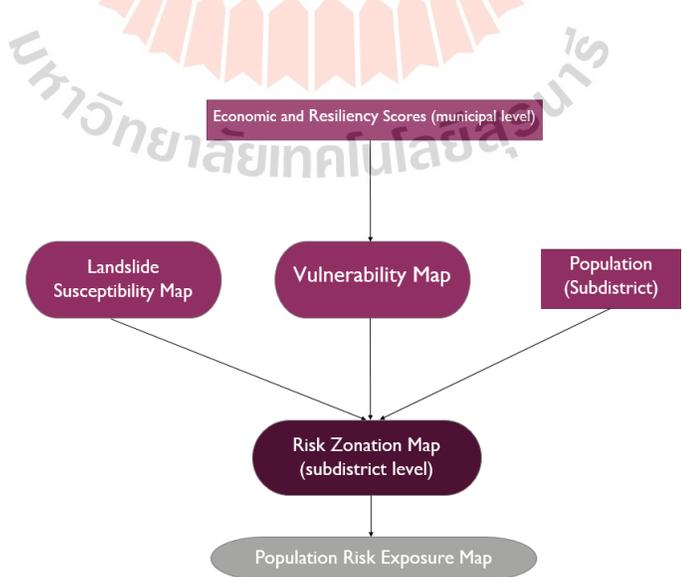


Figure 3.2 Conceptual framework of landslide vulnerability and risk assessment.

3.3 Data acquisition and Preprocessing

Moderate resolution multispectral satellite imagery from Sentinel and Landsat is to be prepared in advance. Existing geologic data and weather data are also to be collected from relevant government agencies. Cloud masking was applied on satellite images to avoid potentially undesirable results and then converted into null values prior to scaling.

To reduce training time, all input data with the exception of categorical variables are to be standardized to normally distribute data by reducing mean to 0 and variance to 1 as described in the following equation:

$$z = \frac{x - \mu}{\sigma} \quad (3.1)$$

Where, x is the original value, μ is the mean, and σ is the standard deviation.

The data will be divided into 70% training as input for deep learning algorithms and 30% test sets for accuracy assessment. The data to be used in the study and their sources are listed in Table 3.1.

Table 3.1 Data type and sources.

Data type	Source
Sentinel-1 SAR SLC IW: Path 69 <i>ascending</i> (2017-2020) Path 134 <i>descending</i> (2015-2020)	Copernicus Open Access Hub
Sentinel-2 multispectral image	Copernicus Open Access Hub
Landsat 5 ETM	EarthExplorer
ALOS-PALSAR DEM	Copernicus Open Access Hub
Geologic map	Mines and Geosciences Bureau (MGB)
Fault map	Philippine Institute of Volcanology and Seismology (PHIVOLCS)
Population data	Philippine Statistics Authority (PSA)

Selection of landslide factors

Currently, there are no established standards for a suitable landslide influencing factor. The selection of the environmental factors used in the susceptibility assessment depends on the type of landslide, the type of terrain, and the availability of existing data and resources (Westen et al., 2008). Truong et al. (2018) recommended the use of topographic features, land cover type, lithology, soil, hydrological features, and distance to active fault zones for places located within active fault zones. Topographic features include the slope and its aspect, altitude, and shape. Hydrological features include river networks, TWI (topographic wetness index), and SPI (stream power index). The aforementioned topographic and hydrological features can be readily extracted and quantified from DEM (digital elevation models)

In this study, landslide factors were selected based on their influence in inducing downslope soil movement. In general, slopes with a higher degree of the gradient are more susceptible to landslides. Differences in slope aspect influence the amount of precipitation and solar radiation, while curvature reflects the terrain's complexity and topography (Zhang et al., 2019). Soil and lithology effectively influence the occurrence of landslides due to differences in shear strength, porosity, density, and particle sizes. Seismic energy released by faults can effectively trigger landslides in steep areas with weak soil foundations. Hydrological influences soil movements as sediments are suspended and can travel in the same direction as water. Additionally, moisture can also affect soil cohesion through saturation. The NDVI reflects both the amount and condition of vegetation. Areas with higher vegetation are generally at lower risk of landslides as root cohesion effectively strengthens the underlying soil and prevents it from being carried and transported by erosional agents. A list of landslide influencing factors is summarized according to their category in Table 3.2.

Table 3.2 Selected landslide influencing factors.

Category	Factors
Physical	● Elevation
	● Slope
	● Aspect
	● Curvature
	● Topographic ruggedness index
	● Normalized difference vegetation index (NDVI)
	● Fault distance
Hydrological	● Topographic wetness index(TWI)
	● Stream distance

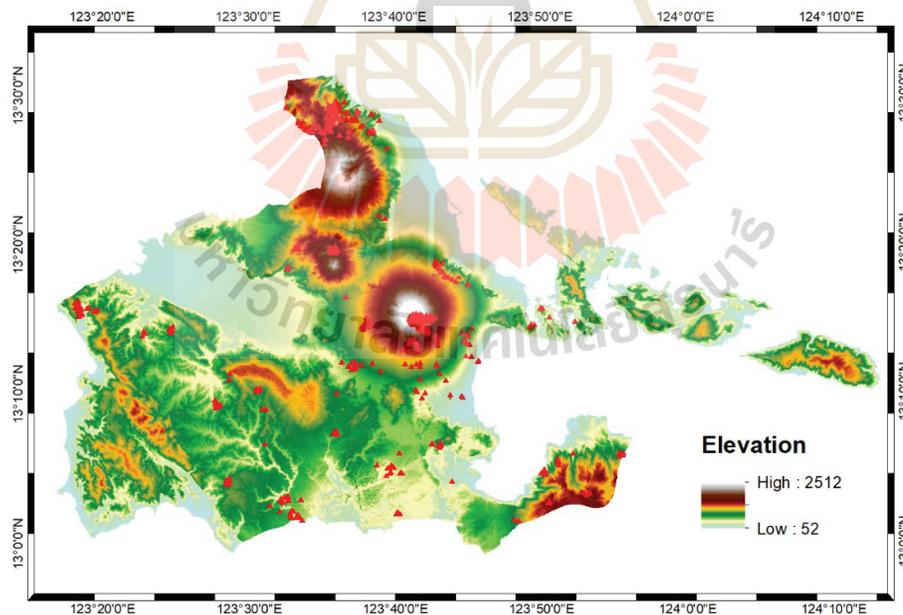


Figure 3.3 Elevation map of Albay in meters.

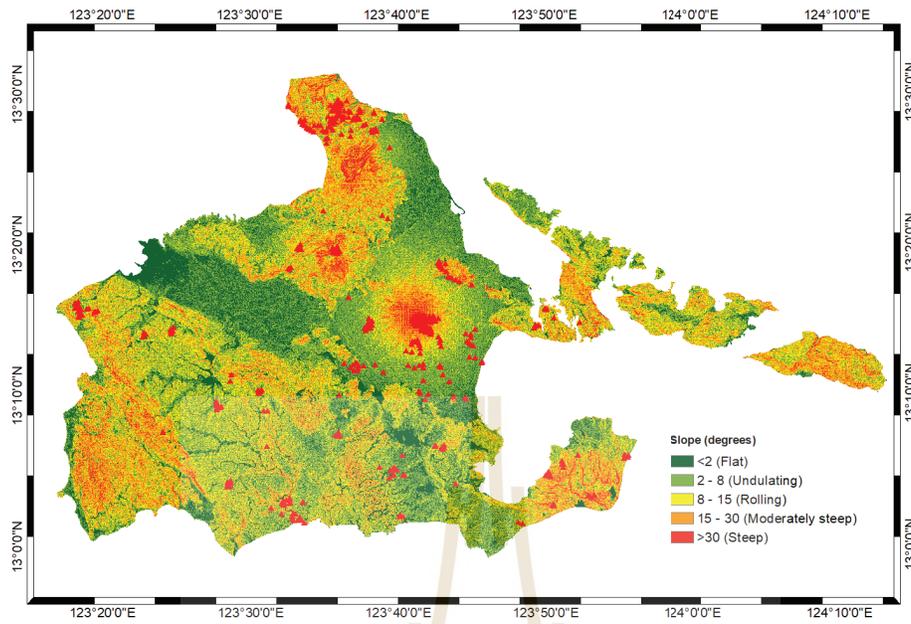


Figure 3.4 Slope gradient map of Albay.

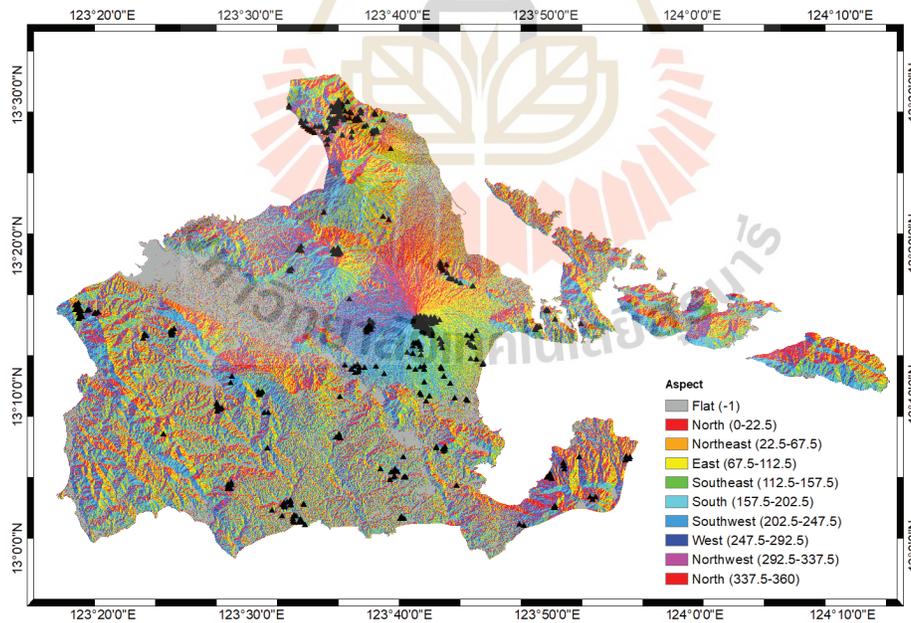


Figure 3.5 Slope aspect of Albay.

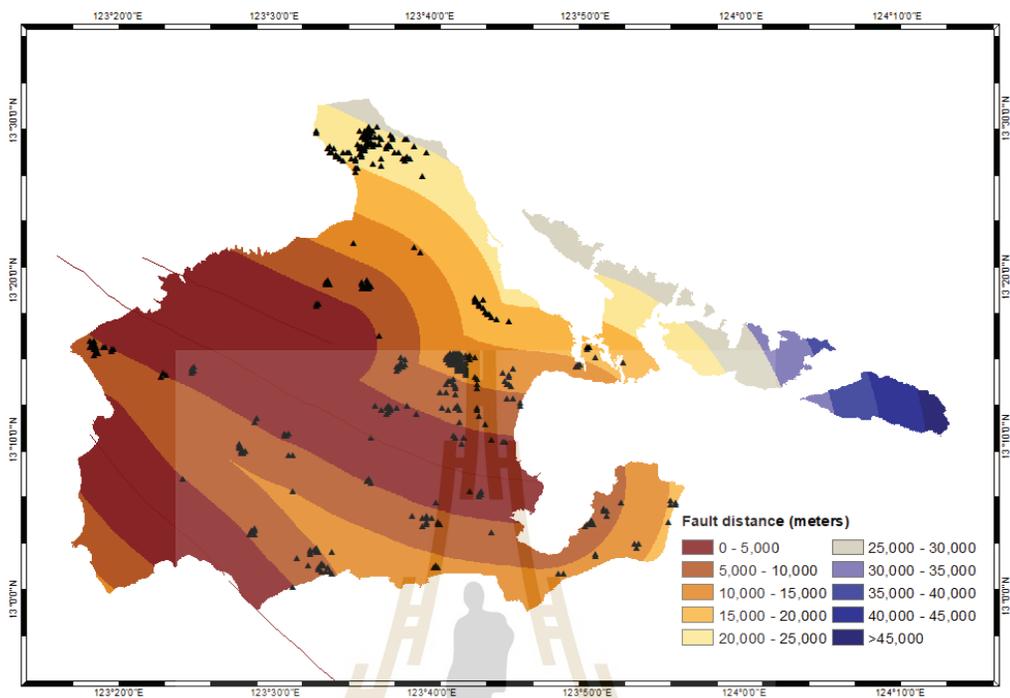


Figure 3.6 Buffer showing distance to active faults of Albay.

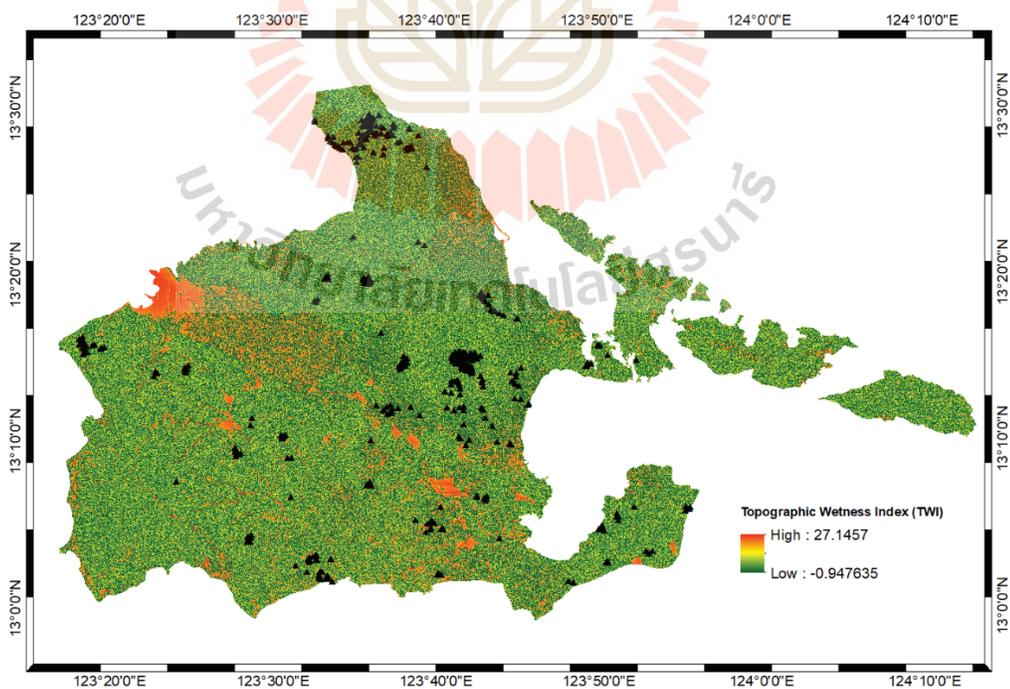


Figure 3.7 Topographic wetness index of Albay.

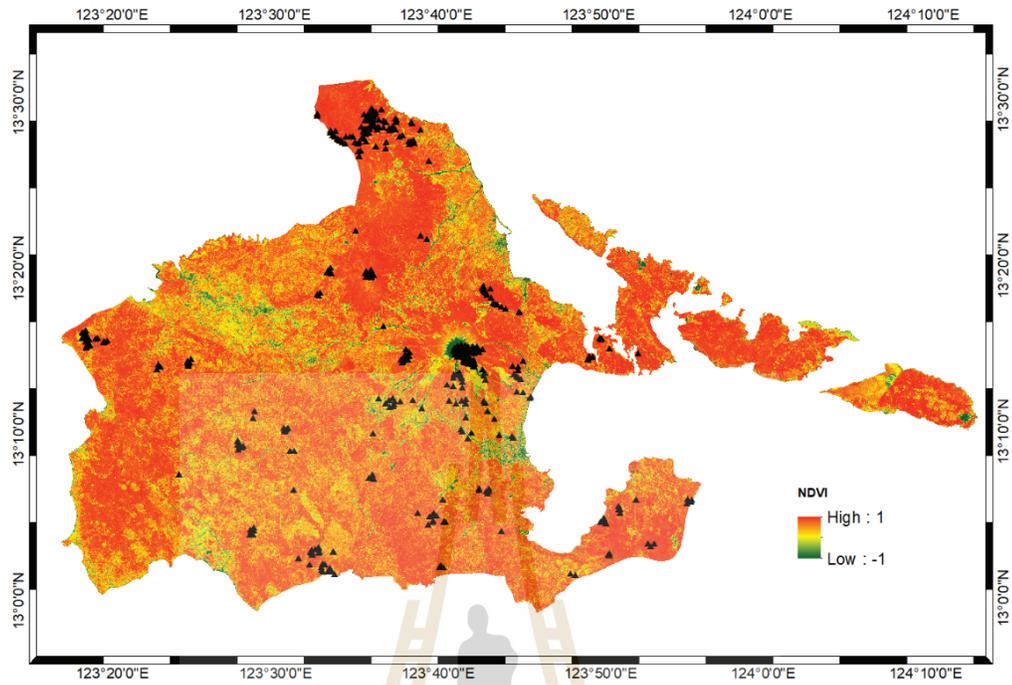


Figure 3.8 NDVI of Albay.

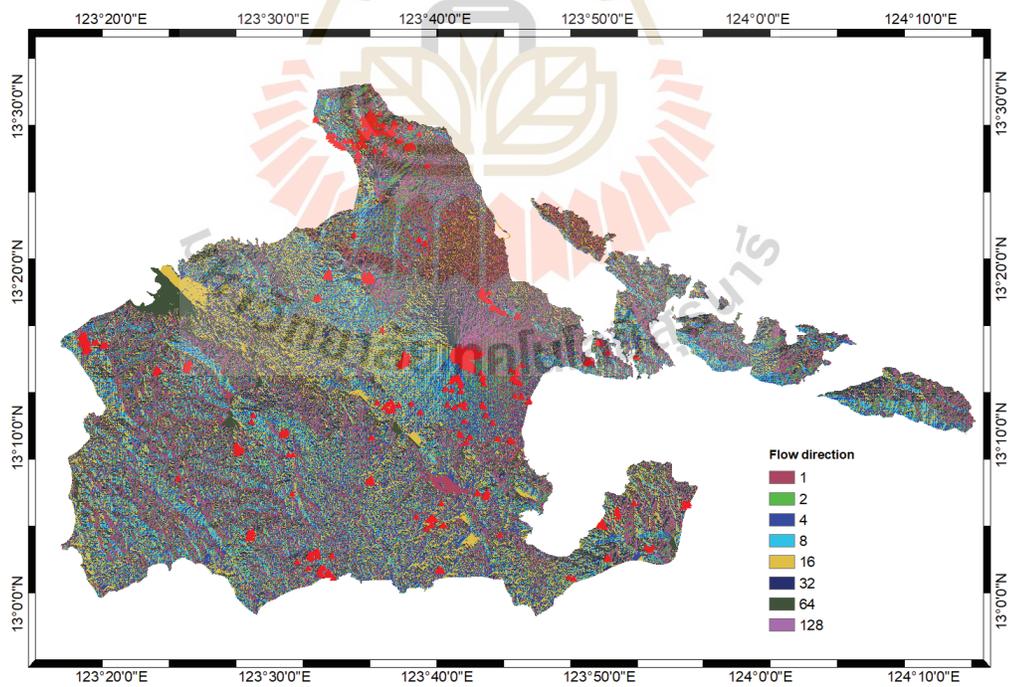


Figure 3.9 Flow direction map of Albay.

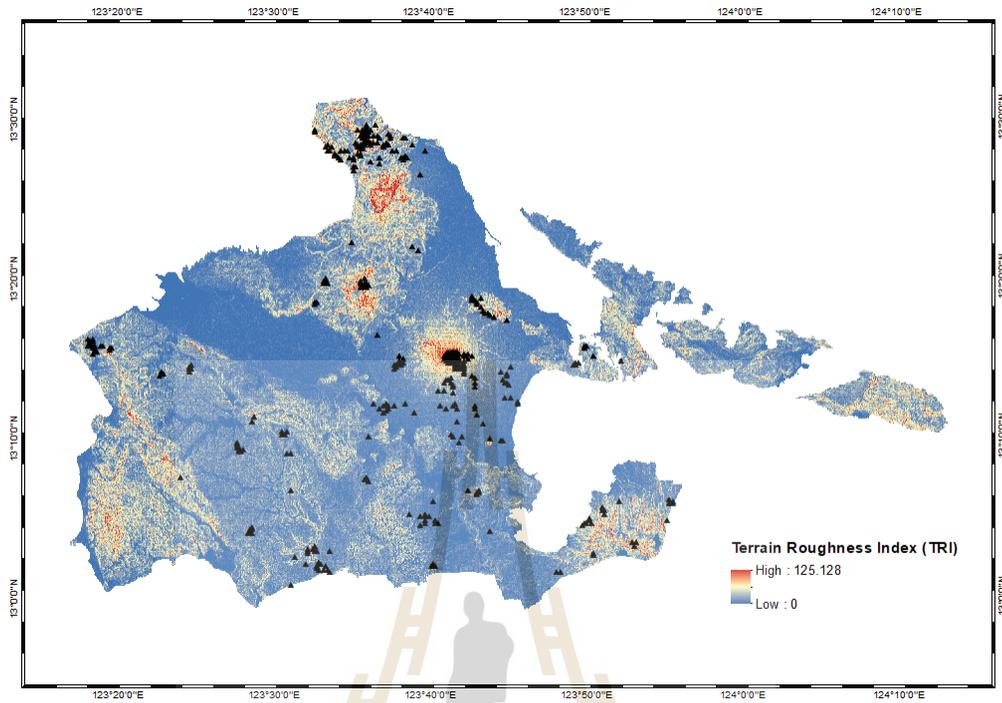


Figure 3.10 Terrain roughness index of Albay.

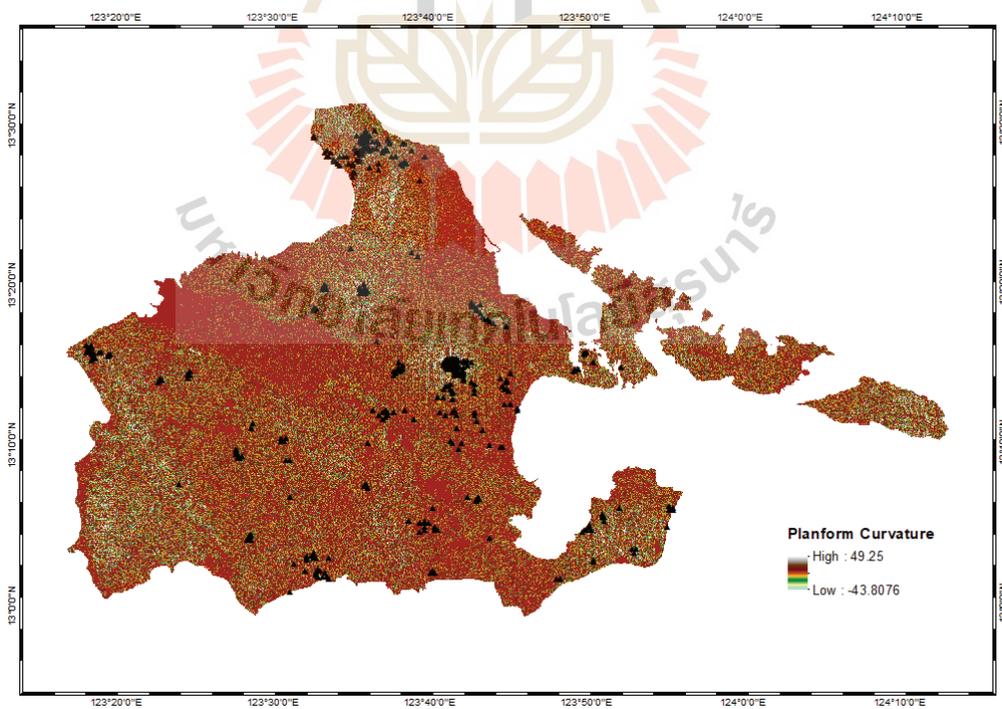


Figure 3.11 Planform curvature of Albay.

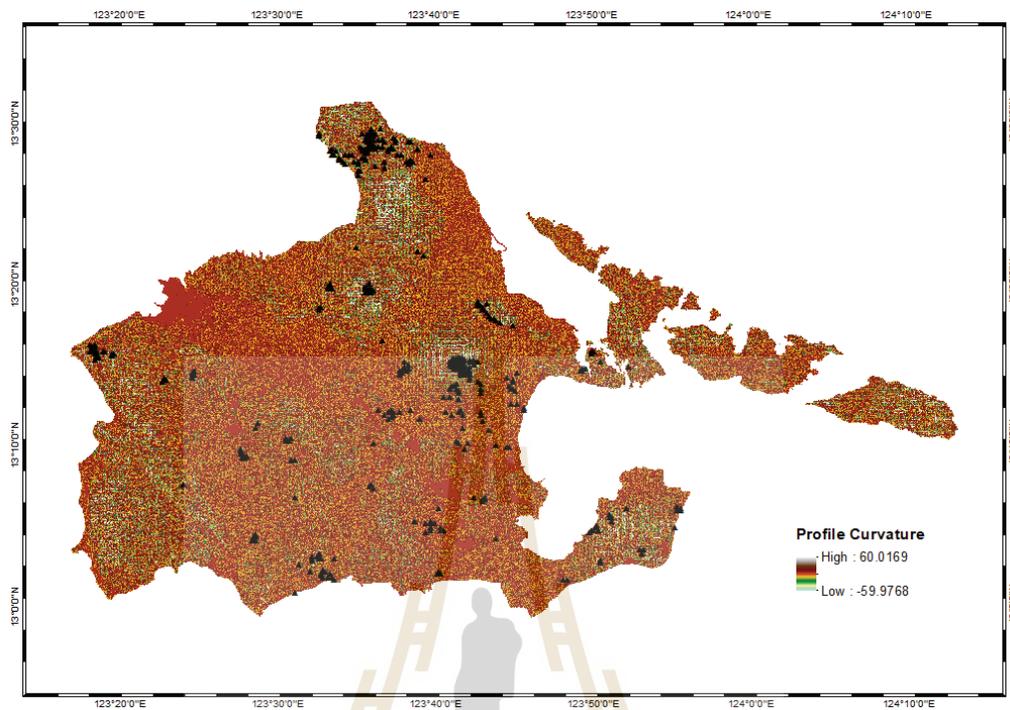


Figure 3.12 Profile curvature map of Albay.

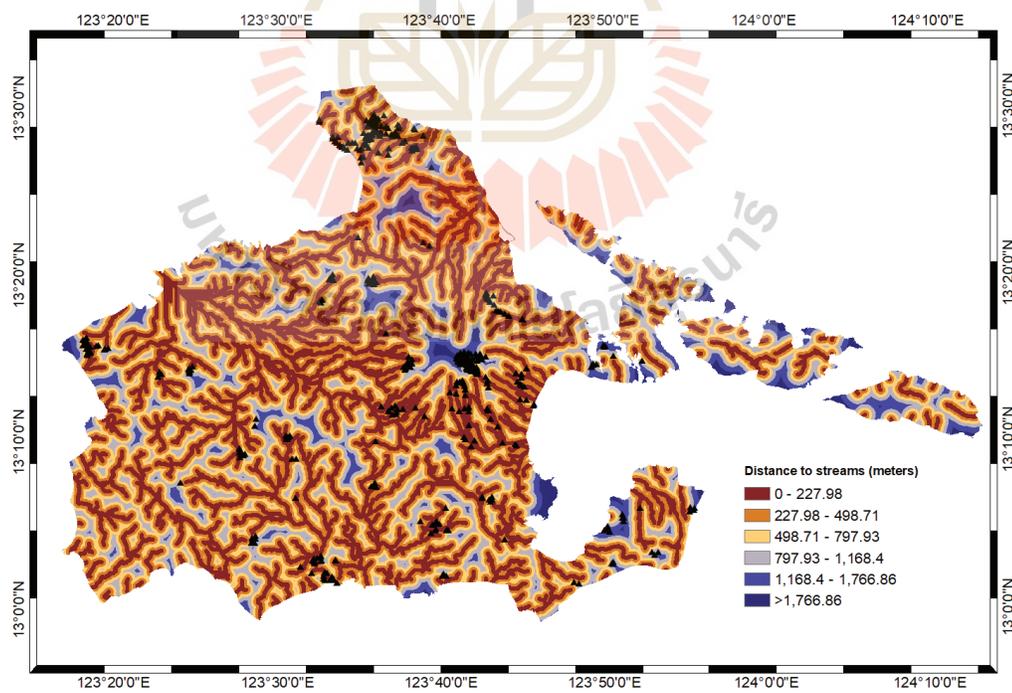


Figure 3.13 Stream distance map of Albay.

3.4 Landslide Inventory Mapping

Gathering landslide points is the first step in conducting landslide susceptibility mapping. According to Guzzetti et al. (2012), the quality of a landslide inventory depends on its accuracy and the type and certainty of the information provided. The easiest and most straightforward method involves gathering the location of previous landslide events from various sources, either from previous studies in the area or through an existing landslide inventory database. Geological field mapping is a common method in studying landslide features, but due to dynamic natural and human processes, there is considerable difficulty in this approach. Moreover, the in-situ visual investigation does not take into account the occurrences of historical landslides. To account for previous landslide events, an existing landslide inventory detailing the location and description of the feature can be used, preferably from official government sources. In the absence of historical data, NASA provides a freely available Global Landslide Catalog as part of the Cooperative Open Online Landslide Repository (COOLR) project, where data is provided mostly by concerned professionals and citizen scientists, supplemented with online news articles (Juang et al., 2019).

To fill in the gaps of landslide inventory at the local level, this study will utilize remote sensing techniques namely, persistent scatterer interferometry, and spectral indices to create a landslide inventory.

3.4.1 Persistent Scatterer Interferometry

Persistent scatterer interferometry, hereafter referred to as PSI, is a powerful tool for monitoring landslide displacement. It offers a synoptic view that can be repeated at different time intervals and at various scales (Tofani et al., 2013). PSI is commonly applied on multiple SAR images in conjunction with in situ monitoring to provide a coherent interpretation of landslide movements and account for the revisit time of SAR platforms.

For PSI processing, the technique is conducted via a two-part approach as described by Mancini et al. (2021):

- (i). Processing of acquired SAR images and formation of interferogram using ESA's SNAP (Sentinel Application Platform) software packages.

(ii) Processing of interferograms using StaMPS (Stanford Method of Persistent Scatterers) developed by Fomelis et al. (2018), implemented in MATLAB.

Interferogram generation done via ESA SNAP's Sentinel-1 toolbox starts with the automatic selection of a master image. All acquired SAR images are split based on the bursts, sub-swath and polarization needed for the study. For mapping landslides, VV polarization is typically used as it is sensitive to vertical displacement and downslope movement. Orbital corrections are then applied to the image splits and co-registered via backgeocoding. During the coregistration stage, master-slave image stacks are generated and undergo spectral diversity enhancement to improve coherence. Debursting is applied to remove redundant scan lines from different bursts and sub-swaths. Subsetting may be done at this stage to process the desired area of interest for interferogram generation. Topographic phase removal is applied to the interferogram subset with SRTM DEM to account for topographic variations resulting from the temporal difference of master-slave pairs before being exported to StaMPS format.

During StaMPS processing, candidate persistent scatterer (PS) points are selected from every interferogram via phase noise estimation and noisy PS points are weeded. Phase correction and unwrapping are done to preserve information and merge patches generated during PS point selection. Spatially-correlated noises are then estimated to account for DEM errors mapped into radar coordinates. Processed PS points are refined further to correct atmospheric noise, automatically done using Toolbox for Reducing Atmospheric Noise (TRAIN) as part of the StaMPS package in MATLAB. The parameters used during StaMPS operation are based on the recommendation of Hoser (2018) for PSI-based landslide detection.

Because SAR platforms operate with a LOS direction tilted concerning the vertical direction, coupled with the incidence angle being very small (30-45° for Sentinel-1A/B), sensors are more sensitive to vertical deformation. To compensate for this, combining data from ascension and descension orbits is necessary to extract both horizontal (in the east-west direction) components of the movement and, consequently, the actual vector of displacement (Tofani et al., 2013) in three dimensions.

To quantitatively analyze a large number of PSI points, PSI points need to be merged and aggregated into a grid cell size of 100m. In order to exclude PSI points reflected from built-up areas, PSI points with an average slope angle of $<5^\circ$ are removed. This is to ensure that no PSI points are going to be calculated in flat areas where slope failures are highly unlikely to occur.

The resulting V_{LOS} points are converted into V_{SLOPE} which projects data along the direction of the steepest slope using the formula proposed by Notti et al. (2014):

$$C = (nlos * \cos(S) * \sin(A - 1.571)) + (elos * (-1 * \cos(S) * \cos(A - 1.571)) + hlos * \sin(S)) \quad (3.2)$$

Where, $hlos$, $nlos$, and $elos$ are the directional cosines of the LOS platform calculated as:

$$\begin{aligned} hlos &= \cos(\alpha); \\ nlos &= \cos(1.571 - \alpha) * \cos(\eta); \\ elos &= \cos(1.571 - \alpha) * \cos(\omega); \\ \eta &= 3.142 - \theta; \quad \omega = 4.712 - \theta. \end{aligned} \quad (3.3)$$

The parameters used for calculating C is listed in Table 3.

Table 3.3 Parameters for calculating coefficient C .

Symbol	Description
S	Slope (radian)
A	Aspect
C	V_{LOS} and V_{SLOPE} ratio
θ	LOS azimuth (radian)
α	LOS incident angle
η	Ellipticity
ω	Altitude

Finally, V_{SLOPE} can now be calculated as:

$$V_{\text{SLOPE}} = V_{\text{LOS}} / C \quad (3.4)$$

Since landslides move downslope, V_{SLOPE} with positive values are removed. Positive values indicate displacement towards the satellite sensor. Negative values measure displacement moving away from the satellite sensor indicating downslope movement. V_{SLOPE} points are then converted into absolute value. The remaining V_{SLOPE} from ascending and descending platforms are aggregated and spatially joined to a 100m grid cell in order to further reduce the number of points and decrease computational load. The grid cells are calculated for their average V_{SLOPE} to account for incident points.

Although the method is still uncommon, the use of PSI-derived landslide inventory successfully quantifies landslide susceptibility. In a study conducted by Piacentini et al. (2015) on mapping slow-moving landslides in a medium-scale area, the landslide susceptibility map produced using PSI-derived data showed an AUC of 0.95 using a weight-of-evidence model. PSI has also proven useful in refining and updating existing landslide susceptibility maps, as demonstrated by the 2016 study of Ciampailini et al., in which the incorporation of PSI data led to an increase in high susceptibility areas due to the prevalence of slow-moving landslides.

3.4.2 Hotspot Analysis of PSI points

In order to accurately pinpoint areas with a high degree of landslide susceptibility with high confidence, hotspot mapping analysis will be conducted on the resulting V_{SLOPE} calculated from the time-series displacement of ascending and descending points similar to the approach adopted by Lu et al. (2012, 2019) and Liu et al. (2020).

Getis-Ord Hotspot Analysis returns z-score which describes intensity of clusters. The resulting z-score also corresponds to the level of confidence of calculated points or the p-value. The formula for Getis-Ord is written as:

$$G_i^*(d) = \frac{\sum x_j - x_i - n_{ij} * \bar{x}^*}{\sqrt{s^* \{[(n \times n_{ij}) - n_{ij}^2] / (n-1)\}}} \quad (3.5)$$

Where, n_{ij} is the total number of V_{SLOPE} points in location i given j neighbors within a scale distance d , x is the surface displacements of V_{SLOPE} points, \bar{x}^* is the mean value of displacements, and s^* is the standard deviation of displacements.

A fixed distance band is designated which is calculated from the average distance of a single point to eight neighboring features using the average nearest neighbor algorithm. The resulting red and blue hotspots are to be evaluated further to only select locations with p-value < 0.01, signifying 99% degree of confidence.

Finally, to extract points with a high degree of susceptibility a slope stability threshold is designated which is calculated from the standard deviation of the merged PSI points. Points with 4σ standard deviation from the stability threshold will be classified as potential landslide points as modified from the approach of Ciampalini, et al. (2016) originally used for the refinement of existing landslide susceptibility maps.

3.4.3 Gaussian clustering of spectral indices

Using data from multi-spectral satellite platforms such as Landsat 8 and Sentinel, spectral indices can be created which are sensitive to soil and vegetation, making them ideal for monitoring soil movement.

To differentiate barren land from other classes, bare soil index can be used. This index is based on the combination of the normalized difference vegetation index and normalized difference built-up index in order to enhance the contrast of bare soils (Diek et al., 2017). A recent study by Ariza et al. (2021) proposed the use of a bare soil index to detect traces of soil movement by combining blue, red, near-infrared and shortwave infrared bands from Sentinel-2 to observe variations in soil moisture described in the following equation:

$$BSI = \frac{(SWIR + Red) - (NIR + Blue)}{(SWIR + Red) + (NIR + Blue)} \quad (3.6)$$

Another index commonly used is the normalized difference vegetation index which is useful for monitoring changes in vegetation and forest cover. Yang et al. (2019) described how this index could be used for temporal analysis of slope movement in densely vegetated mountainous areas. The index can be described in the following equation:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (3.7)$$

To streamline the process of compiling landslide inventories through spectral indices, the study will use a Gaussian mixture model to divide mapped spectral indices and slope values into separate clusters using well-known landslide event locations as input.

Gaussian clustering is a probabilistic model that assumes samples are generated from k Gaussian distributions with unknown parameters (Wang, Azzari, and Lobell, 2019). K-means algorithm is used to initialize the Gaussian clustering by partitioning samples into clusters and assigning them to the nearest centroid using squared Euclidean distance. The cluster centroids are assigned according to the means of samples. K-means typically choose a centroid that will minimize the sum of squares within a cluster as illustrated in the following:

$$\text{minimize } C_1, \dots, C_k \left\{ \sum_{k=1}^K \frac{1}{|C_k|} \sum_{i, i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \right\} \quad (3.8)$$

Where, $|C_k|$ denotes the number of observations in the k cluster.

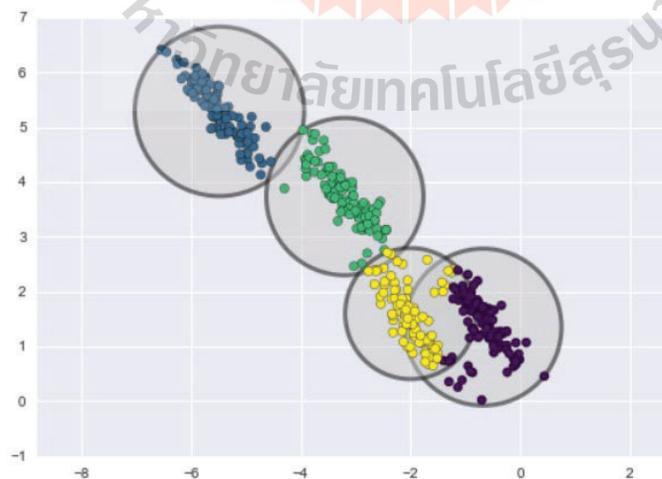


Figure 3.14 Clustering boundary of K-means (Photo from Github).

However, unlike K-means which performs a hard assignment on the clusters, Gaussian clustering implements an expectation-maximization algorithm by calculating the probability of samples belonging to each Gaussian and estimating the means and covariances of the clusters from assigned samples until it converges to a local optimum. Gaussian clustering is then calculated as:

$$\log p(\mathbf{X}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{i=1}^N \log \left(\sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \right) \quad (3.9)$$

$$\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$$

Where, $\mathcal{N}(\mathbf{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ are the derived Gaussian densities from cluster component k , parameter $\boldsymbol{\mu}_k$ is the mean of component, covariance $\boldsymbol{\Sigma}_k$, and the mixing coefficient $\boldsymbol{\pi}_k$ which is the estimate of the density of each Gaussian component.

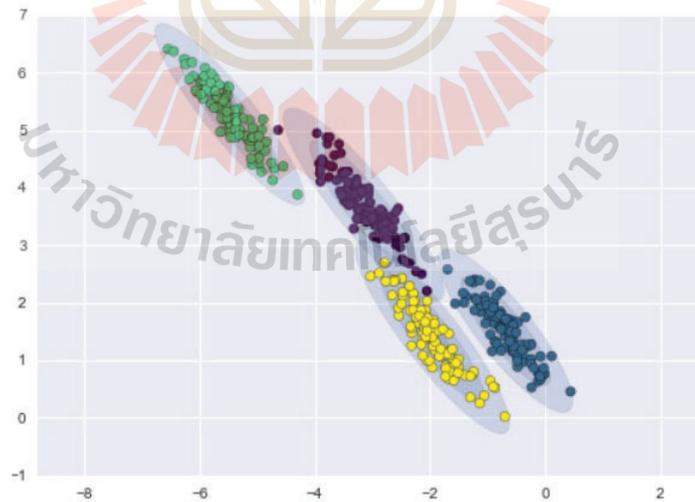


Figure 3.15 Clustering boundary of Gaussian mixtures (Photo from Github).

To assess quality of clusters, a silhouette coefficient is calculated as:

$$s = \frac{b - a}{\max(a, b)} \quad (3.10)$$

Where, a is the mean distance between a sample and all other points in the same class and b is the mean distance between a sample and all other points in the next nearest cluster.

The parameter optimal number of clusters k can then be determined to obtain predicted Gaussian clusters. As the underlying assumption of GMM clustering is tied to normal distribution of data, the resulting decision boundary is elliptical which allows the model to readily accommodate non-linear data by considering covariances. Its probabilistic nature also allows the model to optimize the likelihood of one data point belonging to a certain cluster allowing for better quantitative fitness of clusters.

The resulting clusters are visually evaluated from the original multi-spectral imagery in order to only extract landslide features and ensure the quality of the resulting landslide inventory. This is to determine whether the barren features are pre-existing or recurring prior to the landslide events such as the case with dry river beds, construction sites, dirt roads, and fallow lands.

3.5 Conventional Machine Learning Algorithms

For benchmark comparison purposes, this study will utilize conventional machine learning algorithms typically used for landslide susceptibility mapping, namely: logistic regression, random forest, and support vector machines. Their performance is compared with deep learning models. A brief summary of each conventional machine learning algorithm used in the study is described in the following sections.

3.5.1 Logistic Regression

Logistic regression is a modified form of generalized linear models used for binary classification tasks despite its name. A logistic regression model uses a logistic (sigmoid) function, which transforms a linear regression line into an S-shaped curve to predict outputs whose values are as close to but never exceed 0 and 1. The logistic

curve allows the model to obtain a broader range of variables. The equation can be written as:

$$p(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (3.11)$$

Where sigmoid constant e is the base natural logarithm equal to 2.718, and β_0 is the intercept of the coefficients of independent variables $\beta_1 x$.

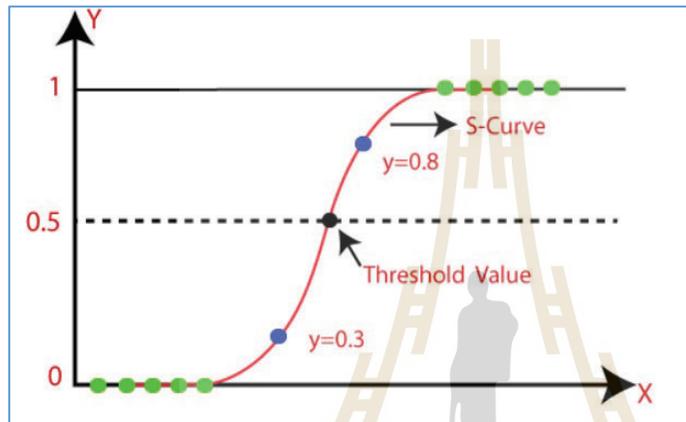


Figure 3.16 Logistic regression curve (Photo from Analytics India).

3.5.2 Random Forest

In a typical tree-based model, binary recursive splitting is applied to grow a large tree on given training data. It is composed of a root node where the first split will occur and then branches out into several decision nodes, stopping at the leaf node or terminal node, which will predict the outcome. The splitting of nodes is randomly determined based on the values of specific attributes. The splitting of nodes is controlled by the cost function Gini.

$$H(Q_m) = \sum_k p_{mk}(1 - p_{mk}) \quad (3.12)$$

Where, p_{mk} is the proportion of training observations of k class in node m .

The Gini index measures the impurity of a given node by measuring the total variance across all features in a given class. A Gini index of 0 indicates that the node is

pure. For binary tasks, the Gini index takes on a small value. The Gini index takes on a small value for binary tasks if the proportion of training observations is close to 0 or 1. A random forest model is a tree-based model which combines multiple decision trees and averages their output to improve predictive performance and reduce overfitting. A random forest uses a large number of unpruned random decision trees that their results are combined using an un-weighted majority of class votes (Rokach, 2016). To find the optimal number of decision trees to be created, bootstrap aggregation (OOB estimation) is implemented. Here, each new tree is fit from a bootstrap sample of the training observations. Random forest model can be written using the following formula:

$$f(\mathbf{a}) = \int \left[\sum_{Z_i} F(Z_i|\mathbf{a})P(Z_i|\mathbf{a}) \right] P_{\mathbf{a}}d(\mathbf{a}) \quad (3.13)$$

Where, $P(Z_i|\mathbf{a})$ and $F(Z_i|\mathbf{a})$ are the probability and relative frequency of the predicted output classification Z_i of the input variable \mathbf{a} , respectively, and $P_{\mathbf{a}}d(\mathbf{a})$ is the probability distribution of the input variable \mathbf{a} .

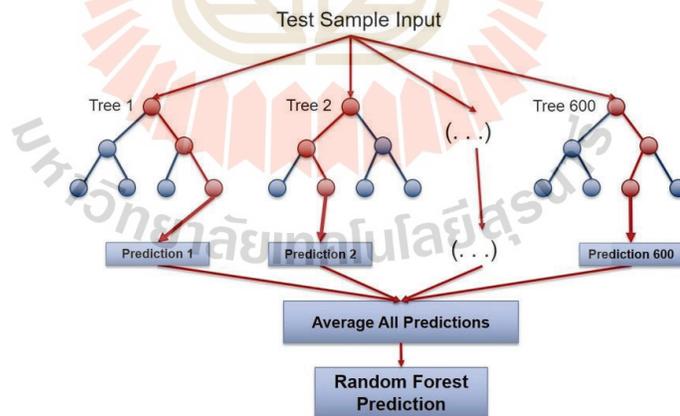


Figure 3.17 Schematic diagram of a random forest model (Photo from Medium).

According to Abidin et al. (2020), the ability of a random forest to outputs of many decision trees makes the model highly robust and cancel out biases making overfitting less of a problem albeit with a trade-off of being computationally intensive.

3.5.3 Support Vector Machine

Support vector machine (SVM) algorithm creates an optimum linear hyperplane that separates data patterns. For non-linear data, a radial basis function (RBF) kernel is used to convert the data into a linearly separable format in a high-dimensional feature space.

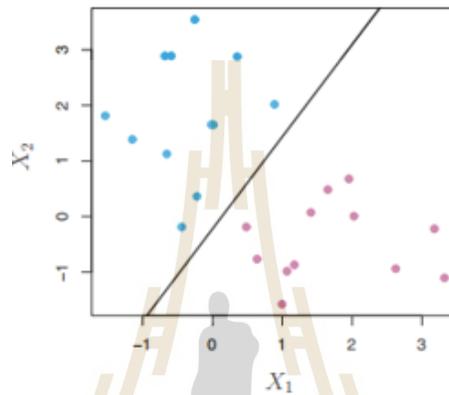


Figure 3.18 Hyperplane separating two classes obtained from James et al. (2013).

In the implementation of SVM, the parameters C and gamma must be considered. The parameter C , common to all SVM kernels, trades off the misclassification of training examples against the simplicity of the decision surface. Gamma defines how much influence a single training example has. A low C makes the decision surface smooth, while a high C aims at classifying all training examples correctly. Larger gamma means samples must be closer in distance to be affected and vice versa. The non-linear RBF kernel SVM can be written in the following equation:

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma(\mathbf{x}_i - \mathbf{x}_j)^2} \quad (3.14)$$

Where, γ is a positive constant of the training vectors \mathbf{x}_i and \mathbf{x}_j , and e is the natural logarithm. RBF kernel classifies data by finding similarities between training and test observations in terms of Euclidean distance. As SVM RBF kernel is a non-linear model, the transformed space is infinite-dimensional making the coefficients of each variable to become unrelated. As such, the relationship of features is only implied.

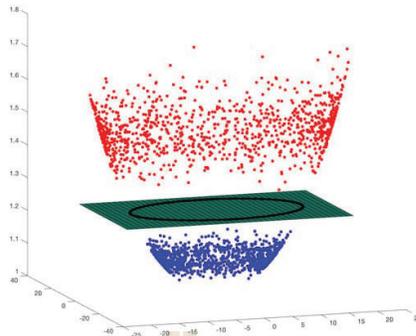


Figure 3.19 RBF kernel separating non-linear data in 3d space (Photo from Learn OpenCV).

3.6 Deep Learning Algorithms

Deep learning algorithms are the general form of an artificial neural network, which is also a subset of machine learning. The basic structure of a neural network comprises one input layer, a hidden layer, and an output layer. Any artificial neural network with 3 or more hidden layers are generally considered as a deep learning model (Emmert-Streib et al., 2020). After one forward pass, neural networks make use of backpropagation which automatically adjusts the weights and biases of hidden layers through reverse matrix multiplication to calculate the optimal gradient of a given loss function. Deep learning models to be used in this research are described below.

3.6.1 Multilayer Perceptron

Multilayer perceptron (MLP), also referred to as artificial neural network (ANN) is an interconnected network of simple computational elements, where the elements are named neurons (Mo et al., 2017).

For binary classification tasks, the MLP consists of the following:

- i. an input layer of neurons with values X_1, \dots, X_i ;
- ii. hidden layers which compute the weighted summation of extracted features from inputs defined as W_1, \dots, W_n , and a predefined offset or threshold value bias b ;

iii. output layer $\mathcal{Y}_1, \dots, \mathcal{Y}_i$ with values ranging from 0 to 1.

A typical MLP can be summarized in the following equation:

$$\sum_{i=0}^n w_i x_i + b \quad (3.15)$$

Where, $w_i x_i$ is the summed product of weights w and input data i , and b is the bias or offset value.

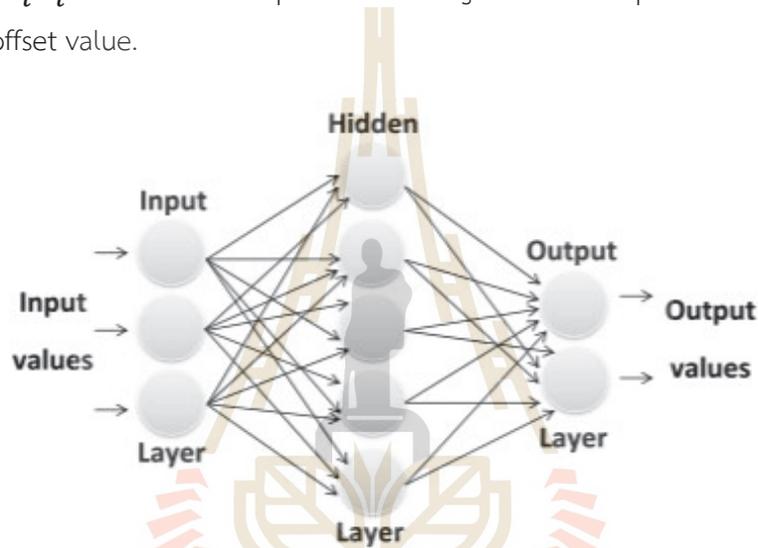


Figure 3.20 Structure of a typical MLP (Abderrahim, Chellali and Hamou, 2016).

For landslide susceptibility mapping, input layers represent the given landslide conditioning factors. The output layer is a continuous binary value ranging from 0 to 1, representing non-landslide and landslide classes.

3.6.2 Convolutional Neural Network

A convolutional neural network (CNN) is a multilayer feed-forward neural network designed specifically to process large-scale images or sensory data in the form of multiple arrays by considering local and global stationary properties (Hu et al., 2020). Convolutional neural networks (CNN) are similar to the standard MLP in that both have an input layer, multiple hidden layers of weight and biases, and an output layer.

This type of neural network has been found to excel at extracting meaningful local features based on their shared-weights architecture and space invariance

characteristics (Hu et al., 2018). The key difference between CNN and the standard MLP is its use of convoluted filter layers consisting of several convolutional units to extract features from the input layer. Another key component is the use of a pooling layer to reduce the feature's dimensionality, reducing the number of parameters and decreasing computational load (Wang, Fang, and Hong, 2019).

Koushik (2016) described a typical CNN model where layers are filter maps, and each layer can be written as a sum of convolutions of the previous layer using the equation:

$$x_j(u, k_j) = \rho \left(\sum_k (x_{j-1}(\cdot, k) * W_{j,k_j}(\cdot, k)) (u) \right) \quad (3.16)$$

Where, x_j is the subsequent layer of input data x_j , W_j is the convolutional filter of kernel size k , ρ is the activation function, u is the offset value, and $*$ is the discrete convolutional operator.

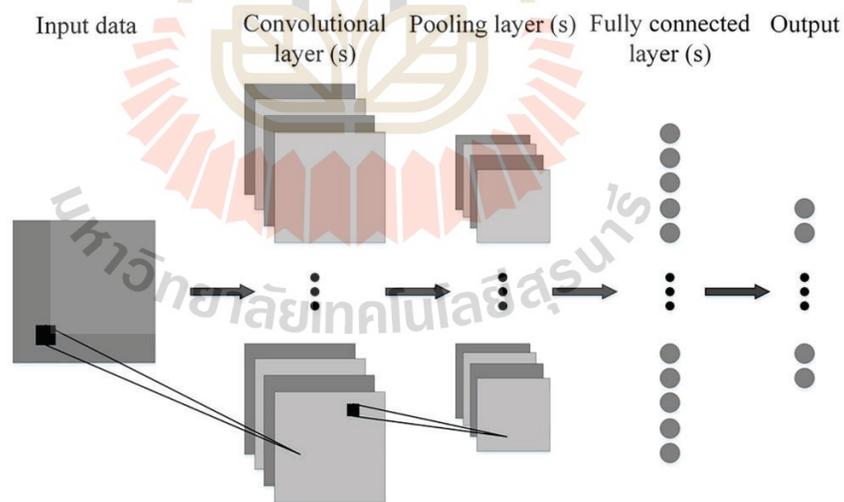


Figure 3.21 Generalized CNN architecture (Wang, Fang and Hong, 2019).

3.7 Accuracy Assessment

For benchmark comparison of deep learning algorithms, accuracy assessment is conducted using performance metrics available in the Scikit-learn library. Receiver operating characteristic (ROC) curve, which represents sensitivity as a function of the false positive rate (Chen et al., 2018) and is quantitatively measured by the AUC (Area under ROC curve) is used to visually compare the predictive performance of algorithms.

Since landslide susceptibility mapping is a binary classification task, precise statistical measures are used for accuracy assessment of produced maps. To identify the ratio of correctly identified classes over the number of predicted classes, precision of predicted positive and negative classes are measured using the following formula:

$$\frac{TP}{TP+FP} \text{ and } \frac{TN}{TN+FN} \quad (3.17)$$

Where, TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

To assess how an algorithm can correctly classify samples belonging to a specific class, recall is assessed. For binary classification tasks, two types of recall; sensitivity and specificity are assessed. Sensitivity is the recall of the positive class or the performance of an algorithm in classifying true positives or the landslide classes, described as:

$$\frac{TP}{TP + FN} \quad (3.18)$$

Specificity is the recall of the negative class which measures how well an algorithm performs in classifying true negative or non-landslide classes, described in the following formula:

$$\frac{TN}{FP + TN} \quad (3.19)$$

For a precise summary of the performance of models used for landslide susceptibility mapping, overall accuracy of all models are assessed which takes into account and the true and false prediction of all classes using the formula:

$$\frac{TP + TN}{TP + FP + TN + FN} \quad (3.20)$$

3.8 Landslide Vulnerability and Risk Mapping

Vulnerability and risk assessment is conducted after landslide susceptibility mapping to quantify the socioeconomic impact of potential landslide disasters. Vulnerability is the estimated degree of damages to an element in the event of a landslide occurrence (Lu et al., 2014, Wang et al., 2021). According to Abdulwahid and Pradhan (2017), vulnerability is typically used to describe physical, social, and economic indicators which quantifies the effect of landslides in an area. Vulnerability indicators which mostly rely on published information are generally assessed empirically (Fell et al., 2008).

Due to the lack of high resolution satellite data and information regarding the costs of infrastructures, the study assessed social vulnerability on a municipal scale using a formula modified from Guillard-Gonçalves and Zêzere (2018) and Gonzalez et al. (2020) described as:

$$V = \frac{E + R}{2} \quad (3.21)$$

Where, E is the indicator for economic parameters and R is the resiliency scores of municipalities to landslides.

Risk is the expected losses that will be incurred of an exposed element in the event of a disaster which includes the probability of a disaster and the magnitude of losses (Wang, et al., 2021). The exposed element considers the monetary values of properties at risk or the number of human lives (Lu, et al., 2014). Risk mapping typically

combines hazard values derived from dynamic spatial data and vulnerability maps. In this study however, landslide susceptibility map will be used in place of a hazard map as carried out by Ram and Gupta (2021) combined with exposed elements similar to a method by Lu et al. (2014). The modified landslide risk formula is then described as:

$$R = S \times V \times E \quad (3.22)$$

Where, S is landslide susceptibility, V is the vulnerability of municipalities to disasters, and E is the elements exposed to landslides described as the number of population per subdistrict in the study.

To identify subdistricts that are at risk of landslides and create a comprehensible output risk exposure map showing the percentage of population exposed to landslides, mean zonal statistics is also applied and then divided over the actual population of each subdistrict as applied by de Almeida et al. (2016) described in the following formula:

$$\text{Risk Exposure} = \frac{R}{P} \times 100 \quad (3.23)$$

Where, R is the average number of population exposed to landslides in a subdistrict and P is the number of population in each subdistrict.

CHAPTER IV

RESULTS

4.1 Landslide Inventory Mapping

The initial results of landslide inventory mapping highlight the utilization of PSI V_{LOS} measurements by identifying landslide points based on their V_{SLOPE} (vertical slope velocity). Another highlight of this section is the application of Gaussian clustering on spectral indices to extract three confirmed landslide events, which are further visually verified by a pre-landslide image and cross-verification through news reports.

4.1.1 PSI points processing

Results of StaMPS analysis in MATLAB using the parameters suggested by Lazecky (2011) and Hoser (2018) for landslide detection yielded 163,627 points for ascending images and 113,898 points for descending images, totaling 277,525 PS points. The maximum topographic error is modified from the default 20m. to 30m. since input SAR data are projected using a 30m SRTM digital elevation model. Parameters in StaMPS modified from the default settings are described in Table 4.1.

Table 4.1 StaMPS parameters as modified from Lazecky (2011) and Hoser (2018).

Parameters	Description	Default	Used
max_topo_err	Maximum uncorrelated DEM error.	20	30
clap_win	CLAP (Combined Low-Pass and High Pass) filter window	32	16
unwrap_grid_size	Resample grid spacing	200	100
unwrap_time_win	Smoothing window (in days) for estimating phase noise distribution for each pair of neighbouring pixels.	730	365
unwrap_gold_alpha	Value of α for Goldstein filter	0.8	5

Table 4.1 (Continued).

Parameters	Description	Default	Used
scla_deramp	Phase ramp estimation.	'n'	'y'
unwrap_gold_n_win	Goldstein filter window size	32	8
weed_standard_dev	Threshold standard deviation for dropping noisy pixels.	1	1.5
scn_time_win	Spatio-temporal filter window of atmospheric noise (in days)	365	50

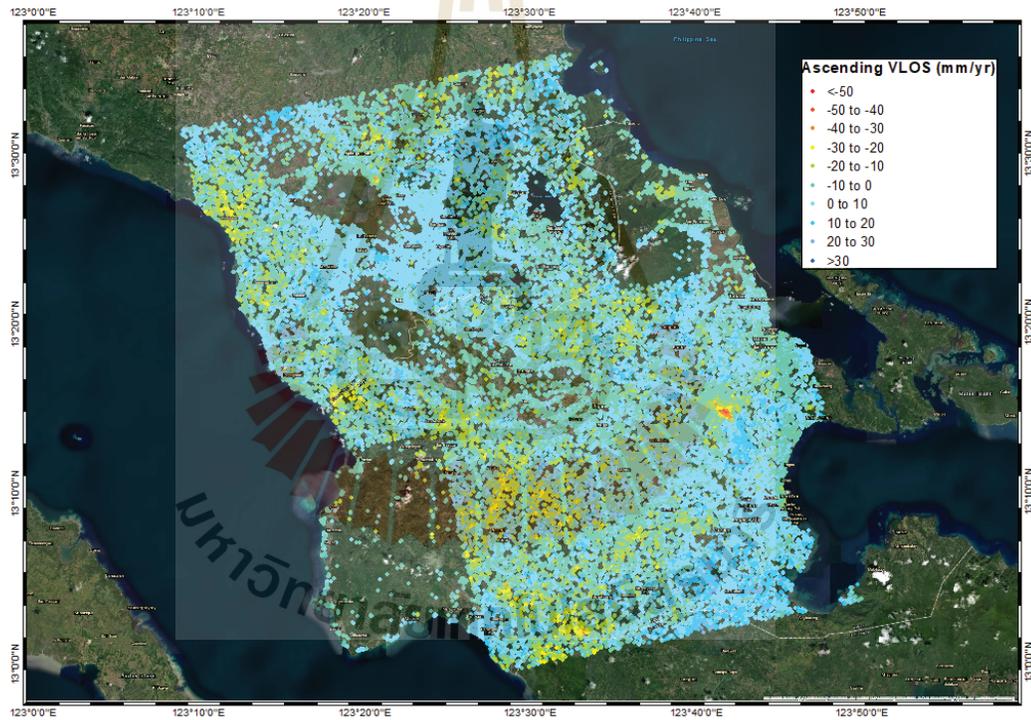


Figure 4.1 Resulting ascending VLOS points.

In order to extract V_{SLOPE} from the initial V_{LOS} of ascending and descending points, V_{LOS} is projected along the steepest slope following the procedure proposed by Notti et al. (2014) which considers the directional cosines of the V_{LOS} calculated from the incidence angle and azimuth of the SAR platform. V_{LOS} points located in slopes with $<5^\circ$ gradients are excluded from the calculation as this signifies reflectors located in flat areas where slope failure is highly unlikely to occur. As V_{SLOPE} values

signifying subsidence and potential slope failures are negative, positive values are discarded as this would signify landslides moving upslope, which is highly unlikely in nature.

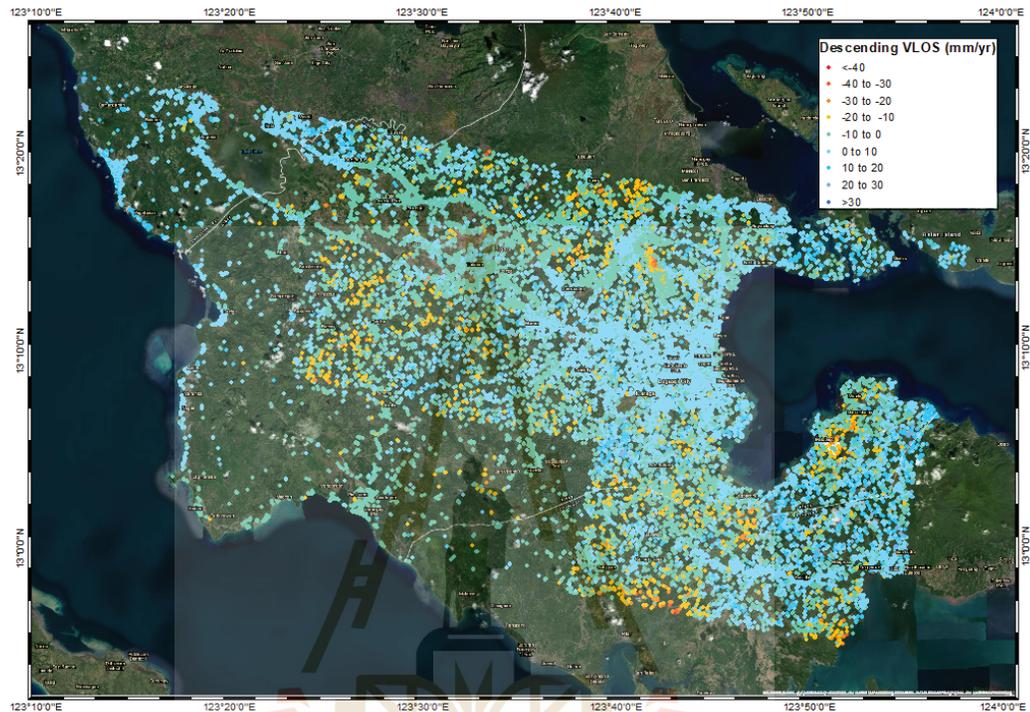


Figure 4.2 Resulting descending VLOS points.

As there are no ground levelling and geodetic measurements conducted to accurately determine the stability threshold of gathered PS points, a stability threshold of 11mm/yr. is designated to assess the degree of susceptibility of V_{SLOPE} . The stability threshold is calculated from the combined standard deviation of ascending and descending V_{LOS} points similar to the approach of Ciampalini et al. (2016). The designated stability threshold is also close to the 10mm/yr value determined by Righini et al. (2012) to discriminate significant slow-moving landslides accounting for the underestimation of slope movement by PSI. All positive V_{SLOPE} are excluded as this signifies either terrain uplift or landslides moving upwards which are impossible to happen in the real world.

To further reduce the number of points and account for incidence, the merged V_{SLOPE} points are spatially joined and aggregated to a 100m grid cell and calculated for

the average V_{SLOPE} in each individual cell. Grid cells with no count and those outside the boundaries of the study area are excluded. A 100 meter ring buffer is constructed outside the boundaries of the study area to account for the uncertainty in locations caused by spatial joining and phase shifting of PS points. A summary of the resulting V_{LOS} (ascending and descending) and V_{SLOPE} are described in Tables 4.2 and 4.3, respectively.

Table 4.2 Summary of acquired ascending and descending Sentinel-1 InSAR data.

	Ascending	Descending
Temporal range	10/12/2017 12/31/2020	– 07/30/2017 06/08/2020
Master image date	01/23/2019	01/09/2019
No. of images	62 images	65 images
No. of PS points	163,627 points	113,898 points
Min. V_{LOS} (mm/yr)	-59.13	-40.45
Max. V_{LOS} (mm/yr)	32.76	39.21
Stability threshold (mm/yr)	± 7.14	± 4.74
LOS Azimuth (θ)	-0.21	-2.93

Table 4.3 Summary of VSlope points.

No. of V_{SLOPE} features	57,438 points
<i>After spatial joining</i>	13,152 cells (100m grid)
Max. V_{SLOPE} (mm/yr.)	232.26
<i>After spatial joining</i>	226.31
Mean V_{SLOPE} (mm/yr.)	18.46
<i>After spatial joining</i>	18.34
Stability threshold (mm/yr.)	0 to 11

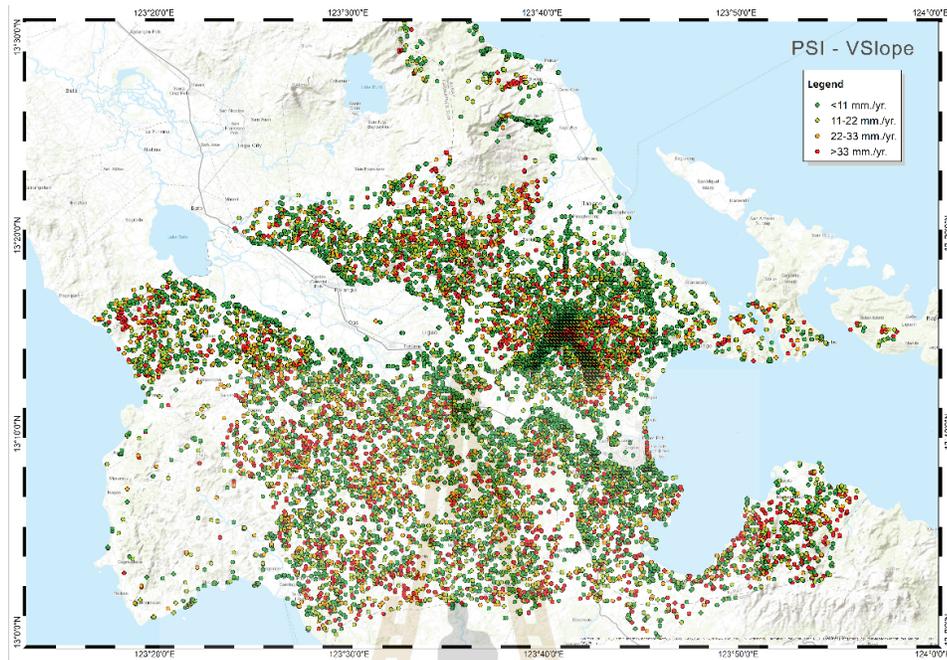


Figure 4.3 Resulting VSlope points after spatial joining.

The maximum V_{SLOPE} prior to spatial joining is located on the slopes of Mayon Volcano. Aggregation of V_{SLOPE} drastically reduced the maximum V_{SLOPE} . However, it was observed that all V_{SLOPE} exceeding 160 mm/yr^{-1} prior to aggregation are clustered within the permanent danger zone of Mayon Volcano where lahar flows and volcanic debris regularly accumulate during periodic eruptions. A mean V_{SLOPE} of around 18 mm/yr^{-1} indicates that slow-moving landslides are prevalent in the area. The similar mean V_{SLOPE} before and after spatial joining also indicates that aggregation of points did not significantly affect the value of V_{SLOPE} as points very close to each other typically have similar velocities.

A. Spatial Autocorrelation Analysis (Moran's I)

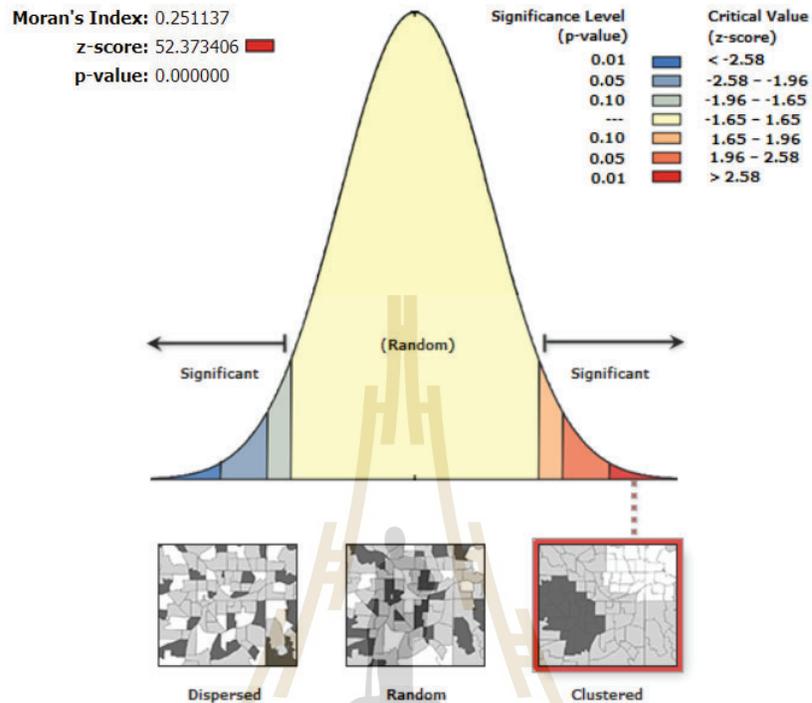


Figure 4.4 Spatial autocorrelation results.

Aggregated PSI points are calculated for their average distance to eight neighboring features to determine the distance band for spatial autocorrelation and hotspot analysis. The average distance threshold to eight neighboring features is determined to be 540 meters. Results of Moran's I shows that a p-value of <0.01 and a z-score of >2.58 indicates that V_{SLOPE} points are significantly clustered.

4.1.2 Landslide hotspot mapping of PSI points

Despite the precise millimeter measurements of PSI points, their quality is still highly dependent on the coherence, local terrain, topography, and quality of the digital elevation models to be used. Projections and grid resampling also introduce variations in the locations and actual values of PSI points which further necessitates in situ measurements which may be costly in terms of time and resources considering the large study area.

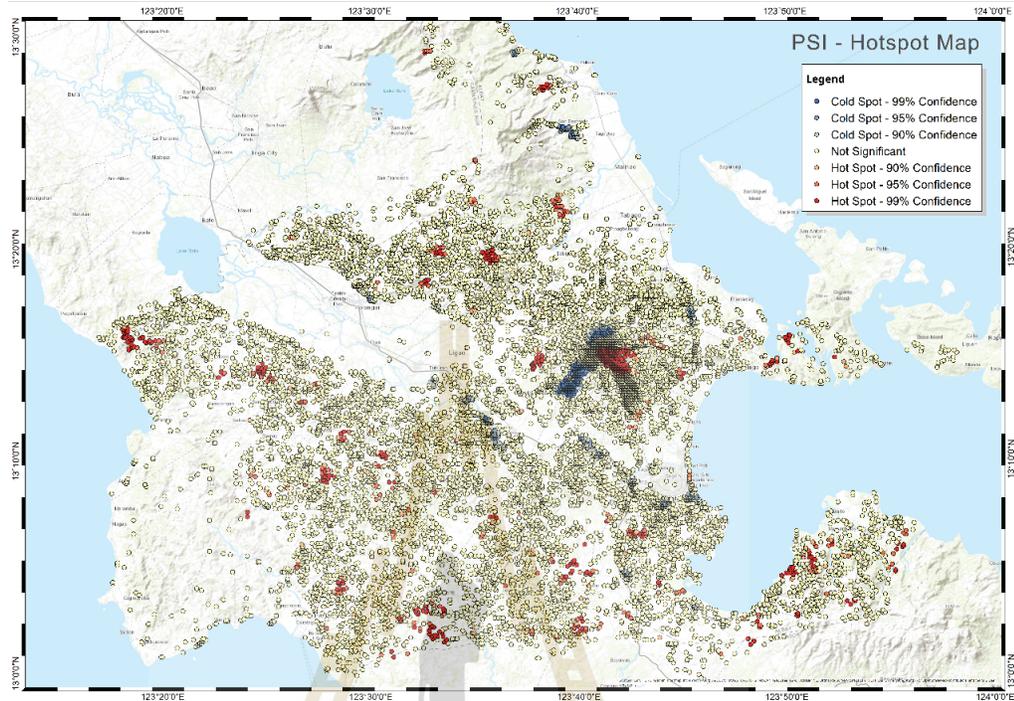


Figure 4.5 Hotspot analysis results.

To determine the significance of V_{SLOPE} points, Getis-Ord G_i^* hotspot analysis is conducted which creates red and blue hotspots and classifies their confidence level. Red hotspots indicate clustering of high V_{SLOPE} values, while blue hotspots indicate clustering of low V_{SLOPE} values. A fixed distance band of 540 m. is applied to the points with V_{SLOPE} as the weighting factor. In order to produce a reliable landslide inventory map derived from V_{SLOPE} , only red hotspots and blue hotspots with a 99% confidence level were accepted.

A high concentration of V_{SLOPE} was observed within the radius of active Mayon Volcano. This is further emphasized by a high concentration of hotspots and coldspots near the cone upon closer inspection. Mayon Volcano regularly experiences phreatic eruption. During typhoons and eruptions, volcanic debris and lahar flow also regularly accumulate within the designated 12km. permanent danger zone.

Finally, in order to discriminate critical areas from the extracted hotspot points, a stability threshold of 11mm/yr. derived from the standard deviation of ascending and descending V_{LOS} were used to classify the points. V_{SLOPE} points 4σ standard deviations from the stability threshold were designated as having very high degrees of susceptibility. This procedure is modified by Ciampalini et al. (2016) in classifying V_{SLOPE} points for landslide susceptibility.

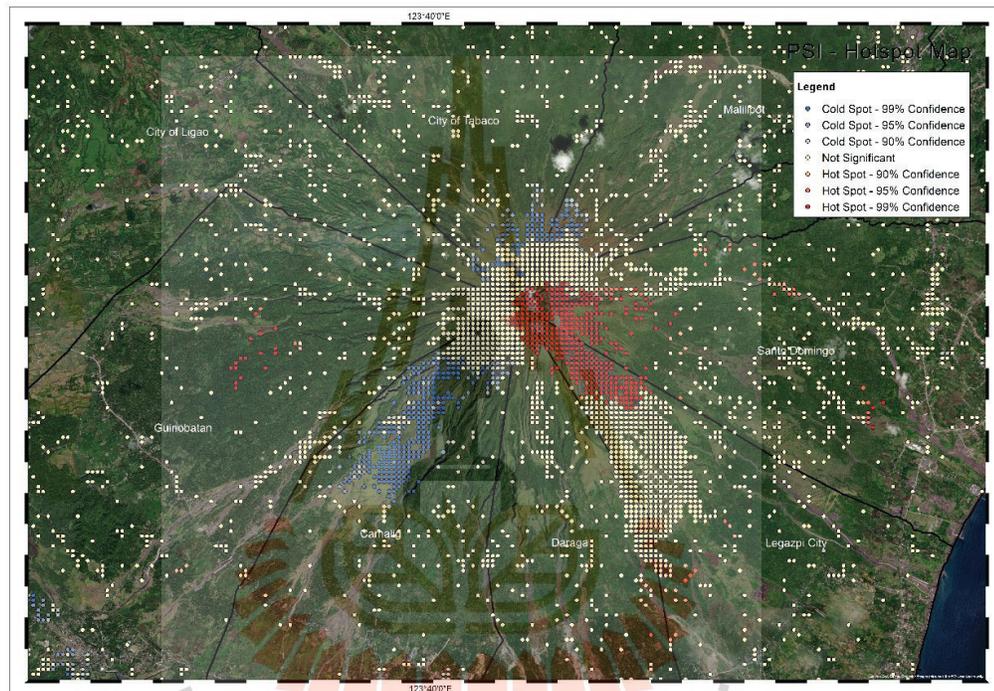


Figure 4.6 Clusters located within Mayon Volcano.

Table 4.4 Hotspot analysis summary.

99% confidence - Hot spots	591 points
- Cold spots	280 points
95% confidence - Hot spots	220 points
- Cold spots	216 points
90% confidence - Hot spots	213 points
- Cold spots	213 points
Insignificant	11419 points

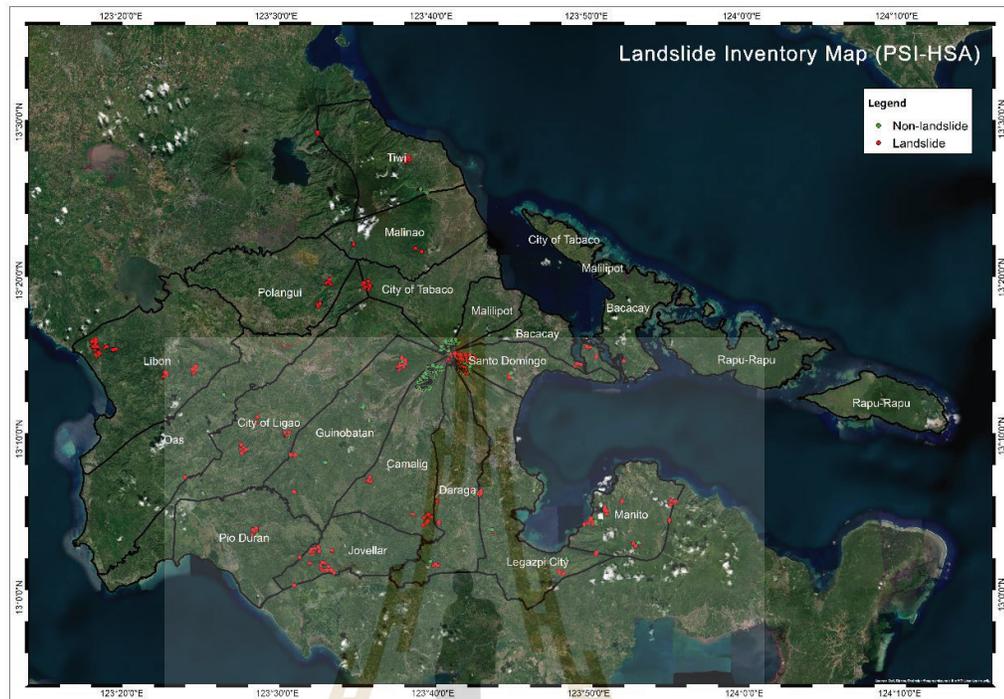


Figure 4.7 PSI-derived landslide inventory. Green points signify non-landslide points.

Out of 891 V_{SLOPE} points with 99% confidence level, 356 points were designated as having very high degrees of susceptibility indicating the potential for landslide and slope failure. 154 of these points are located within the permanent danger zone of Mayon Volcano. However, 11419 points have been considered insignificant as false discovery rate correction was applied to account for spatial dependence and reduce the rate of false positive results. Another possibility is the factor of a relatively short fixed distance band of 540 m. which was the average distance of one feature from the nearest neighboring feature. The aggregation of incident points may have also contributed to a relatively high number of insignificant points. As the mean of incident points is obtained after aggregation, this process inadvertently eliminates local patterns but may have introduced high variations in values between neighboring features leading to FDR considering them as insignificant.

To accommodate binary landslide susceptibility classification, points below 11mm/yr V_{SLOPE} at a 99% confidence level are designated as non-landslide classes due to their extremely slow velocity. These points are classified as having low susceptibility

degrees as their displacement rates are below the designated stability threshold. In total, there are 335 non-landslide points derived directly from hotspot analysis of PSI-derived V_{SLOPE} points.

4.1.3 Landslides extracted from spectral indices

Features extracted from spectral indices represent location where landslides have been confirmed by news reports and visual interpretation of multi-spectral satellite imagery. To extract landslide features efficiently, a Gaussian mixture model (GMM) clustering algorithm was applied on a stack of NDVI, bare soil index, and slope rasters. Due to the geography of the area, most satellite images have significant cloud covers all-year round which is prevalent during the typhoon season which starts from July to December. Frequent phreatic eruptions of Mayon Volcano also makes it difficult to identify eruption-induced landslides as smoke covers significant portions of the area near the permanent danger zone where most slope failures occur. This difficulty is further complicated by rapid vegetation regrowth due to frequent rainfall in the tropical humid environment of the study area. As such, only three major landslide events were considered in the study which were validated from archived news articles and visual interpretation of high-resolution satellite imagery. Visual interpretation was carefully conducted in order to not extract pre-existing features similar to landslides that might be incorrectly identified by the clustering model such as barren construction sites, dry river beds, fallow lands, and to some extent, built-up areas.

Using GMM clustering on a raster stack of spectral indices, the AUC showed relatively high predictive performance. However, as a semi-automatic algorithm, landslide features had to be extracted manually by removing other clusters and pre-existing features that were classified as belonging to the landslide clusters such as built-up areas, fallow lands, and dry river beds. Two classes indicating landslide and non-landslide points were used for clustering. The choice of the optimal number of clusters is arbitrary and varies across the three major landslide events. The final number of cluster components was chosen considering the AUC score of the probability of predicted clusters and the ease of visually separating landslide and non-landslide features in a GIS environment. A total of 85 confirmed landslide points were extracted across three major landslide events.

This section describes the three major landslide events recorded and used for creating landslide inventory maps derived from spectral indices. All landslide events were rainfall-induced because of typhoons.

A. November 2006 Typhoon Durian (local name: Reming)



Figure 4.8 Aerial view of villages buried by lahar flow in Legazpi City, Albay (Photo from Bulatlat, dated December 2006).

On November 30, 2006 supertyphoon Reming hit Albay triggering lahar flows and landslides coming from Mayon Volcano. According to PAGASA Weather Advisory, the typhoon packed a maximum sustained winds of 190 km/h near the center with a gustiness of 225 km/h (Orense and Ikeda, 2007). In a statement given by then-PAGASA administrator Graciano Yumul Jr. to the Philippine Star, 466 mm. of rainfall was recorded on a single day during the onslaught of Reming (Echemineda, 2006). In another statement by then-PHIVOLCS head Renato Solidum, the typhoon brought the heaviest single-day volume of rainfall recorded in almost 40 years during that time. Following the termination of retrieval operations on December 15, 2006 in another report by Philippine Star, the Provincial Disaster Coordinating Council (PDCC) of Albay recorded at least 546 killed that were retrieved and 243 missing presumed dead bringing the total number of dead to at least 658 (Dematera, 2006). A 2006 report by OCHA indicated that more than 1100 people were killed in Bicol Region alone. An

estimated P5 billion (USD 108 million) in damages to infrastructure, property, and agriculture were recorded. Lahar flows buried villages located in the foot of Mayon Volcano in the municipalities of Camalig, Daraga, Guinobatan, and Sto. Domingo.

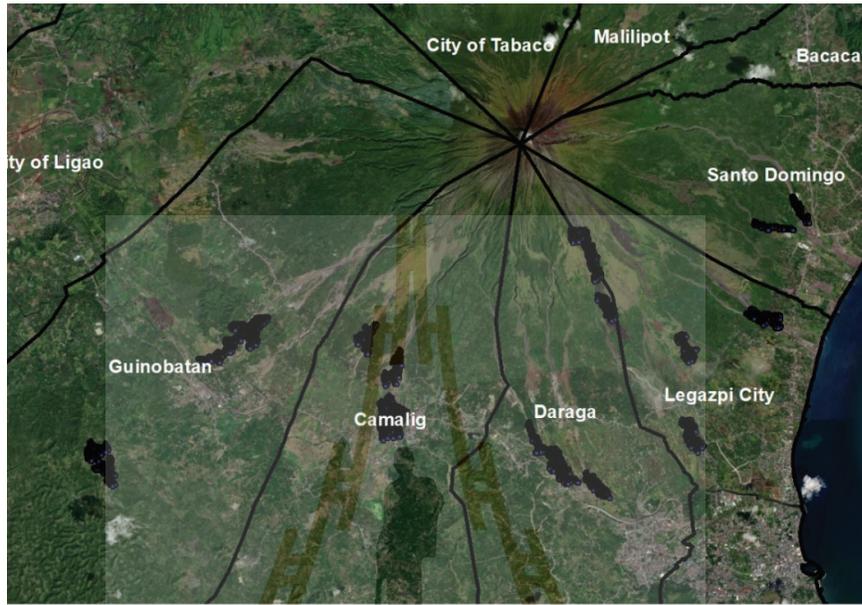


Figure 4.9 Training points for November 2006 landslides overlain on high-resolution satellite imagery.

A Landsat 5 TM+ multispectral image dated April 11, 2007 was used as the source image for extracting spectral indices. In order to not interfere with the NDVI and BSI, clouds were masked out. To aid in discerning landslide features from pre-existing features, another image dated August 30, 2006 was used as a pre-landslide image guide. 4207 points designated as landslide features were used as input for the GMM clustering algorithm. Using GMM, a total of 57 features were extracted from 14 clusters with a silhouette score of 69%. Since landslide features are to be manually extracted, the score is sufficient to discern landslide and non-landslide clusters in the area. The optimal choice of clusters is arbitrary and will highly depend on the ease of discrimination among clusters in the final output while considering the silhouette coefficient. 14 clusters were enough to visually discern landslides from non-landslides and other related pre-existing features such as fallow land and dirt roads without significant change in the silhouette coefficient. During an 8-month time period in the

pre-landslide and post-landslide satellite imagery, it was observed that most of the areas affected by lahar flows were located near the foot of the volcano well within the established permanent danger zone.

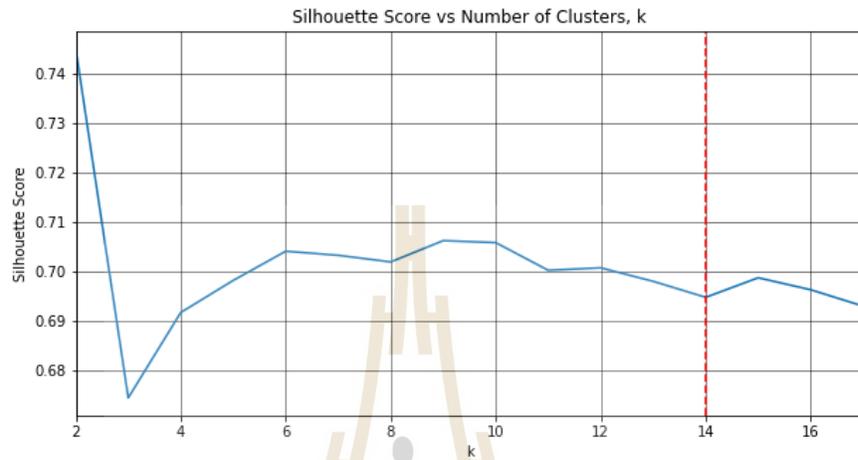


Figure 4.10 Silhouette score of GMM clusters extracted from November 2006 landslides.



Figure 4.11 Cloud-masked pre-landslide (left) and post-landslide (right) Landsat-5 image of Guinobatan.



Figure 4.12 Cloud-masked pre-landslide (left) and post-landslide (right) Landsat-5 image of Sto. Domingo.

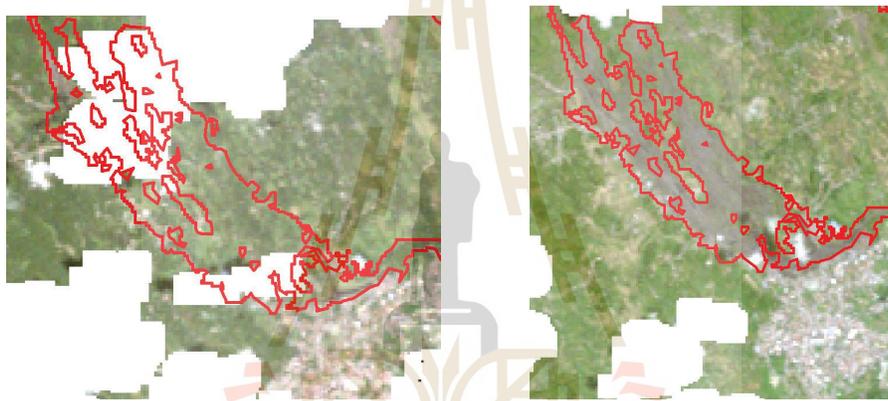


Figure 4.13 Cloud-masked pre-landslide (left) and post-landslide (right) Landsat-5 image of Daraga.

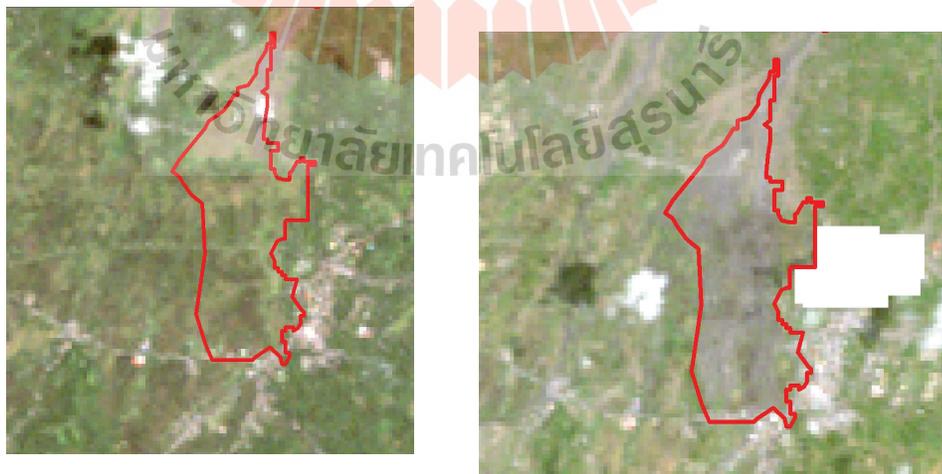


Figure 4.14 Cloud-masked pre-landslide (left) and post-landslide (right) Landsat-5 image of Camalig.

B. December 2018 Tiwi, Albay landslides (Tropical Depression Usman)



Figure 4.15 Aftermath of TD Usman in Tiwi, Albay dated Dec. 30, 2018 (Photo from ABS-CBN News).

A relatively weak tropical cyclone formed on December 25, 2018, with a maximum sustained wind of 45 km/h, entered the Philippine Area of Responsibility (PAR). In a statement issued by PAGASA to the Philippine Daily Inquirer on December 27, TD Usman slowed down from 15 km/h to 10 km/h. At its peak on December 28-29, TD Usman reached a maximum sustained wind of 55km/h near the center and gustiness of 65 km/h (Ramos, 2018). During the same 2-day period, 460.1 mm of accumulated rainfall was recorded by PAGASA in its synoptic station in Legazpi City, Albay. The slow velocity of the tropical depression combined with heavy rainfall caused multiple landslides in the province of Albay as well as surrounding regions. 3 people reportedly died after being buried by landslides in Legazpi City, according to a news report by Rappler (Barcia 2018). In Tiwi, Albay, one of the areas struck with devastating landslides, 11 people were reportedly killed (Magsino, 2018). On January 2, 2019, the chief of Bicol Regional Disaster Risk Reduction and Management Council

(RDRRMC) Claudia Yucot ordered the evacuation of residents in one village after discovering large cracks in mountainous areas (Barcia, 2019). A total of 126 people reportedly died in the aftermath of TD Usman with about 100 being recorded in Bicol Region with total damages to agriculture and infrastructure estimated to be at P4.2 billion or USD 85 million (Viray, 2019).

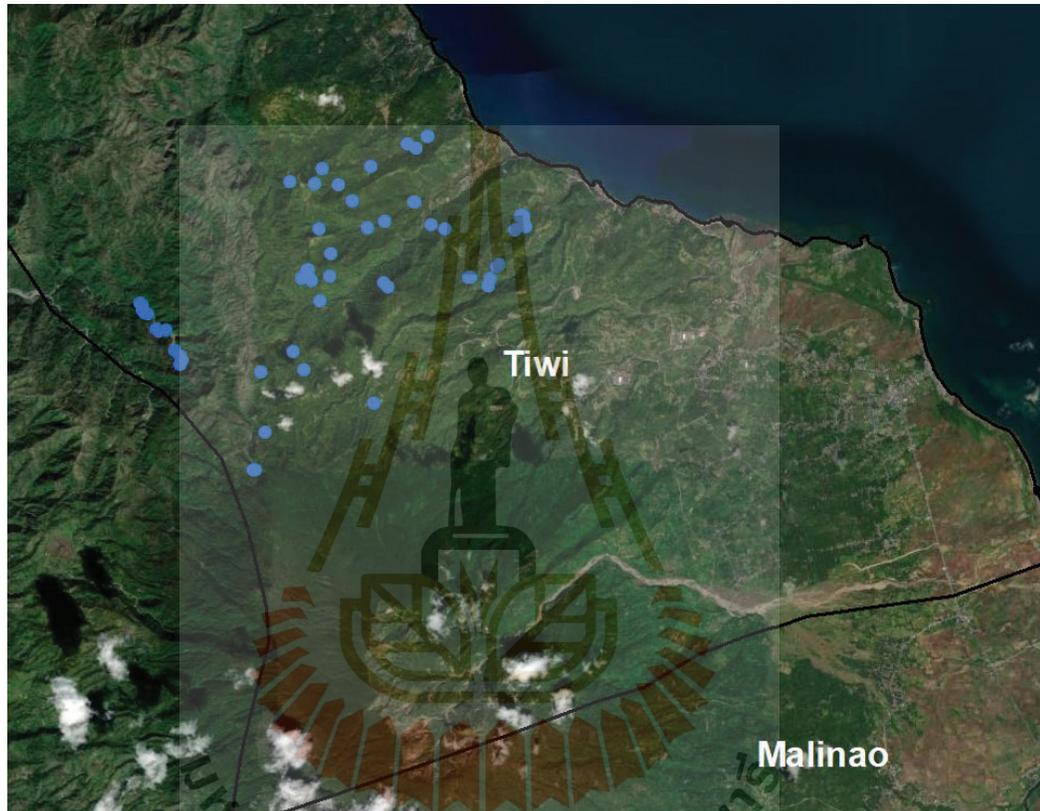


Figure 4.16 Training points for December 2018 Tiwi landslides overlain on high-resolution satellite imagery from ArcMap.

A Sentinel-2 multispectral satellite imagery dated January 15, 2019, was used to extract landslide features in Tiwi, Albay. A pre-landslide image of the area dated February 9, 2018 was used as a guide in determining pre-existing features from landslide features. Despite a minimal temporal difference of only 17 days from the date of collapse, the sizes of landslides are relatively small making it difficult to extract input features for the GMM clustering algorithm. This difficulty is reflected in a low silhouette coefficient of 55% in 10 clusters. However, 126 features were successfully extracted from 62 points after post-processing of clusters.

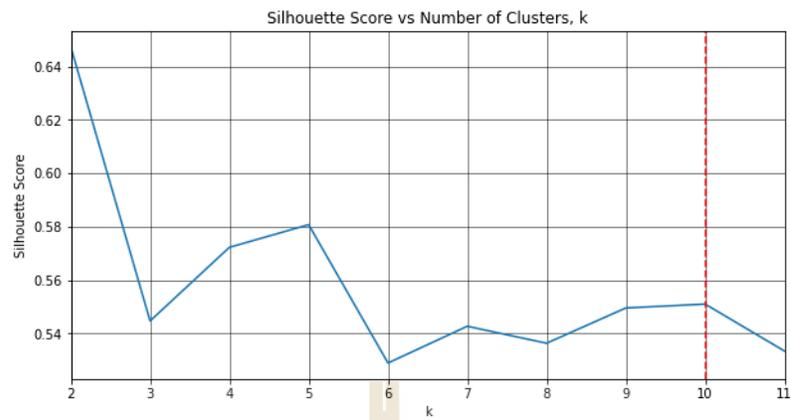


Figure 4.17 Silhouette score of GMM clusters extracted from December 2018 Tiwi landslides.

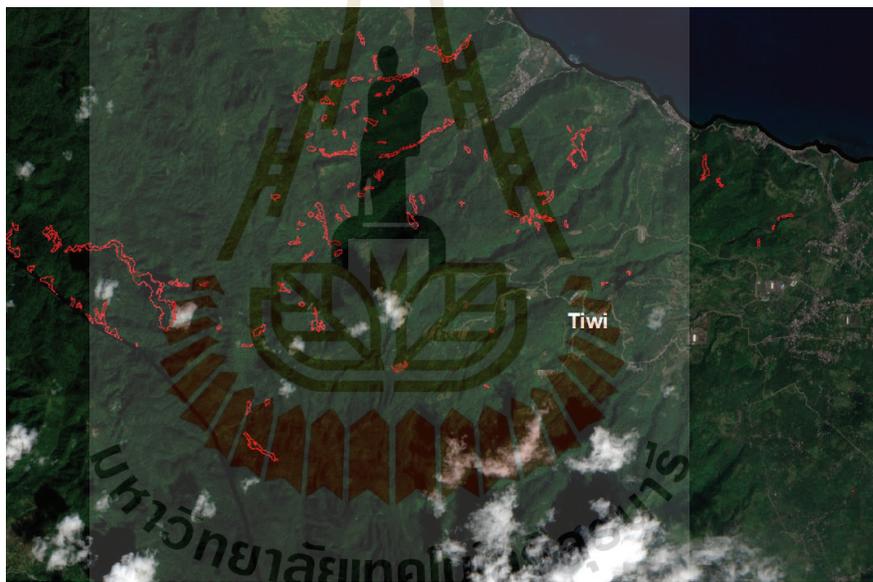


Figure 4.18 Sentinel-2 pre-landslide image of Tiwi, Albay dated February 9, 2018.

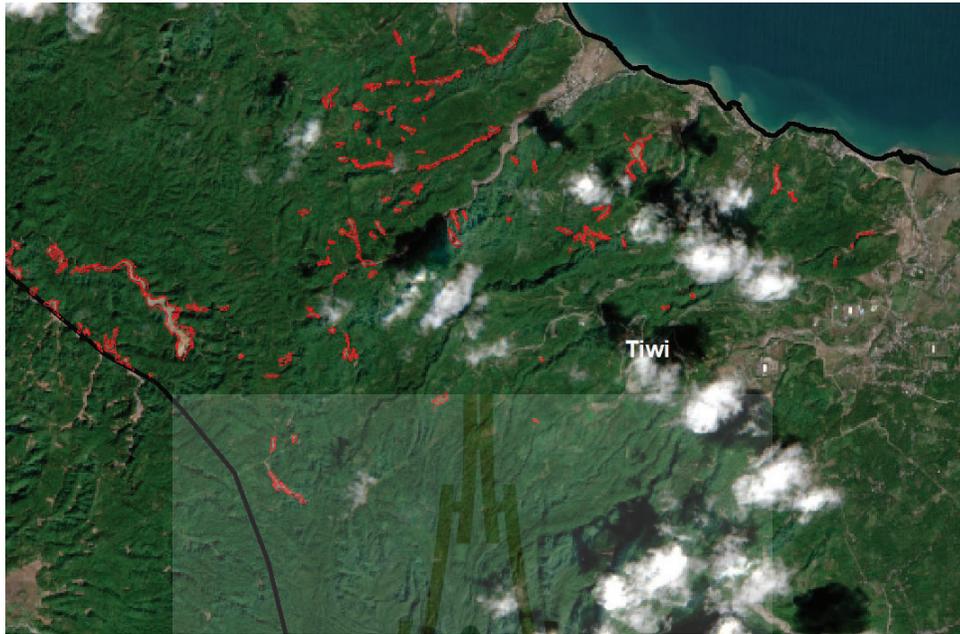


Figure 4.19 Sentinel-2 post-landslide image of Tiwi, Albay dated January 15, 2019.

C. November 2020 typhoons

Typhoon Goni (local name: Rolly) is a supertyphoon which first made landfall on November 1, 2020 with the maximum sustained wind of 225 km/h and gustiness of 310 km/h (Inquirer, 2020). According to PAGASA, the typhoon was the strongest worldwide in 2020 (Marquez, 2020). The typhoon displaced more than 316 thousand people or 87,000 families in the Bicol region according to the Office of Civil Defense, 116,000 of which are in Albay (Mier-Manjares, 2020). In a statement issued by NDRRMC to CNN Philippines (2020) on November 6, 22 people died, 13 of which are in Albay. Damage to infrastructure and property was estimated to be at around P14 billion (\$285 million). 77% of Albay's electrical grid was also damaged according to APEC (Albay Power and Electric Corporation) communication head Lesley Capus (Mier-Manjares, 2020). 98% of the grid was restored as of January 22, 2021 in a statement issued by National Electrification Administration (NEA) to BusinessWorld (Yang, 2021).

On the aftermath of Typhoon Goni, another supertyphoon codenamed Vamco (local name: Ulysses) hit Albay on November 10, 2020 with a recorded maximum sustained winds of 155 km/h and gustiness of 215 km/h with total damages exceeding P20 billion (USD 408 million) with 101 people confirmed dead (NDRRMC, 2021).

Among the hardest hit was the town of Malilipot, Albay already reeling from the destruction of Typhoon Rolly. Successive typhoons triggered landslides in the municipality – one of which is a prominent 100-foot chasm overlooking Bulawan River, displaced 65 families (Luces, 2020).



Figure 4.20 100-foot chasm overlooking Bulawan River in Malilipot, Albay dated November 10, 2020 (Photo from Manila Bulletin).

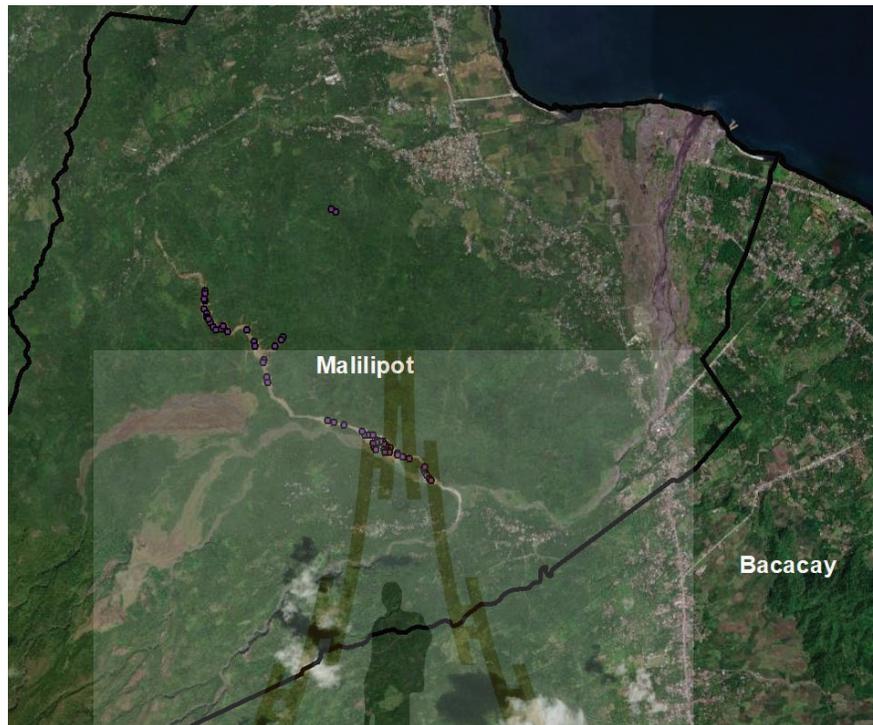


Figure 4.21 Training points for 2020 Malilipot landslides overlain on high-resolution satellite imagery.

Using spectral indices derived from Sentinel-2 satellite imagery dated May 29, 2021, a total of 17 landslide features was visually identified and extracted from 11 clusters with a silhouette coefficient of 58%. Despite a 6-month temporal difference from the date of the landslide, the extent of the landslide meant that it was not affected by vegetation regrowth during that time period. However, due to its smaller size relative to the locality, only 59 landslide points were used as input data. Despite this problem, the clarity of landslide features made it relatively easy to extract landslide points and visually discern landslides from non-landslide features.

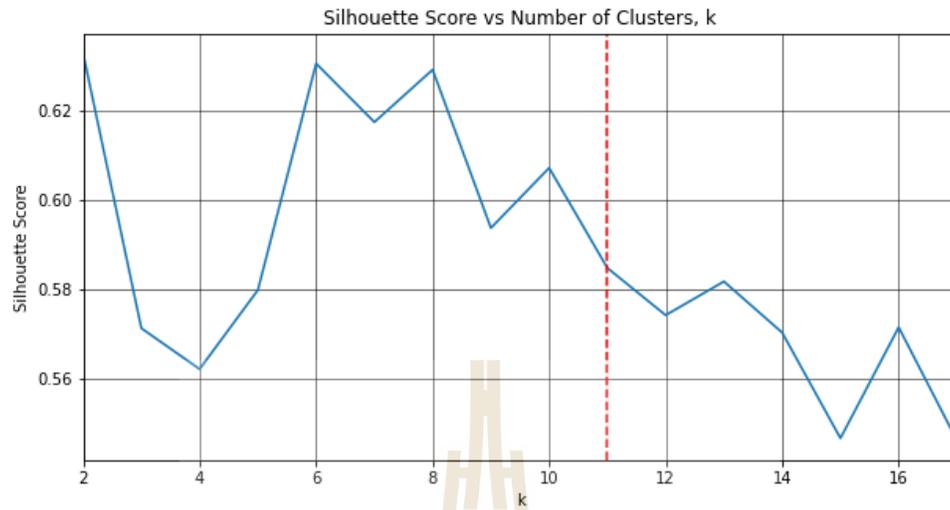


Figure 4.22 Silhouette score of GMM clusters extracted from 2020 Malilipot landslides.



Figure 4.23 Sentinel-2 pre-landslide image of Malilipot dated April 10, 2017.

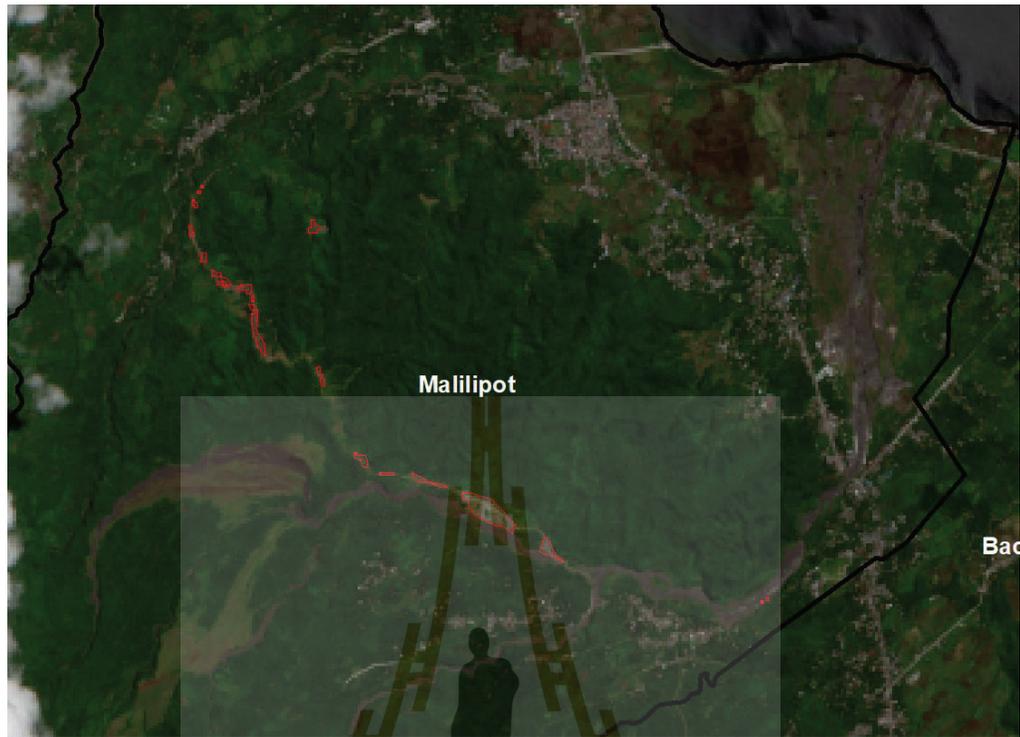


Figure 4.24 Sentinel-2 post-landslide image of Malilipot dated May 9, 2021.

During landslide inventory mapping, 200 landslide features were extracted directly from spectral indices using GMM clustering algorithms. Despite the difficulty of extracting landslides in small localities due to atmospheric factors, online news sources became an indispensable source of information to minimize the temporal range during the acquisition of multispectral satellite images. While there is also considerable difficulty in discerning landslide features from pre-existing features, this method combining a semi-automatic clustering algorithm with elements of citizen science yielded a fair amount of samples in an area where landslide inventory is inaccessible to the public.

There were 356 landslide points extracted from PSI using a combined technique of V_{SLOPE} calculation and hotspot mapping. Moreover, 332 non-landslide points were also extracted from the results of hotspot mapping. The totals of 556 landslide points and 332 non-landslide that are ready to be analyzed for binary landslide susceptibility classification.

Despite landslide points being only limited to mainland Albay and does not include the outer-lying island municipality of Rapu-Rapu and other areas within the jurisdiction of Bacacay, Malilipot, and Tabaco, the large sample size provides enough input for machine learning and deep learning models in order to classify the whole province.

To control the quality of the landslide susceptibility maps, 200 non-landslide points located in Mayon Volcano were omitted. To compensate for excluded points, non-landslide points were also arbitrarily created in the outer-lying islands and in areas where landslides are less likely to occur such as urban settlements located in flatlands within the coastline. Non-landslide points are also created in areas with <5 slope gradient. The figure indicates non-landslide points extracted from $PSI V_{SLOPE}$ points with 1σ stability threshold and points arbitrarily designated based on the fact that no significant slope movement have been made by both PSI and multi-spectral satellite imagery.

As a post-processing method, the results of the landslide inventory are further aggregated into 25×25 grid cells with a pixel size of 10m in order to be accommodated by deep learning algorithms. Aside from eliminating incident points, this process will also be highly beneficial for deep learning algorithms which considers the value of neighboring pixels during training. In order to avoid any error during training and prediction-related stemming from datasets, all input data with NaN values are further excluded. A total of 113 non-landslide points and 378 landslide points was used as input data giving a 1:3 ratio of samples. Despite the class imbalance, the ratio of the input landslide points are in line with studies done by Thabtah et al. (2020) and Starovoirov and Golub (2020) who stated that it will final the performance of binary classifiers. The final landslide inventory map including non-landslide points are illustrated in Figure 4.25.

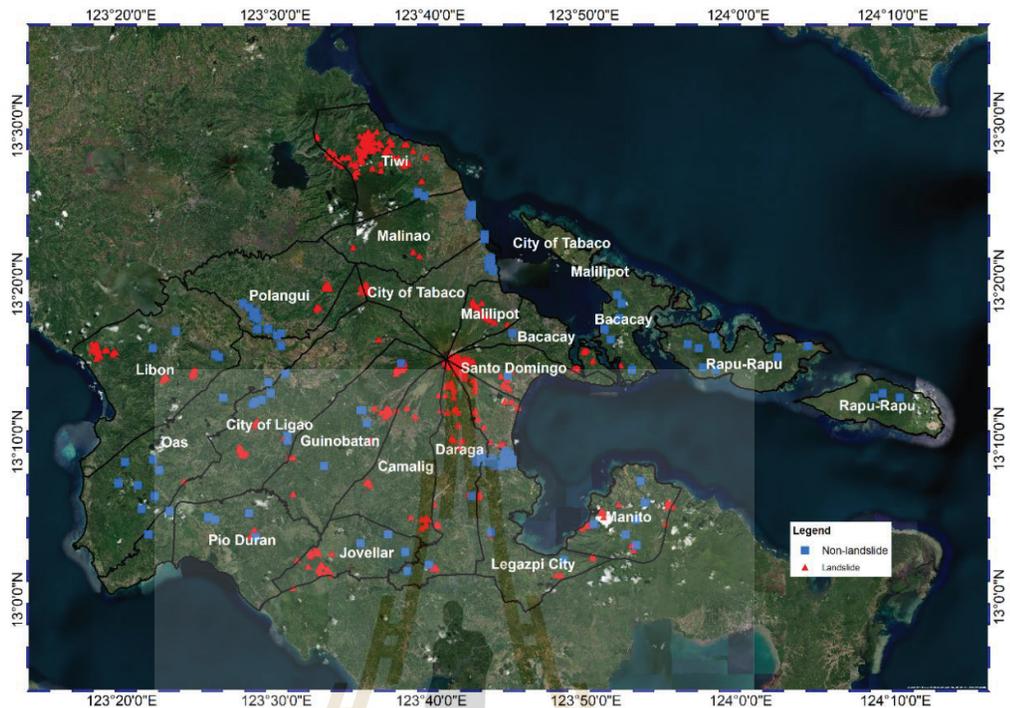


Figure 4.25 Landslide inventory map used in the study including non-landslide points.

4.2 Landslide Susceptibility Mapping

This section highlights the use of deep learning algorithms, namely multilayer perceptrons and two-dimensional convolutional neural networks (CNN-2D) for binary classification of landslide susceptibility mapping. Landslide data with NaN pixel values are excluded prior to training to control the quality of input data. The remaining data are split into 70% training and 30% validation data. The training data are normalized in order to reduce training time and improve predictive performance. In order to be accommodated by univariate statistical tests and conventional machine learning algorithms during benchmark comparisons, datasets were duplicated and calculated for their mean considering the values of neighboring pixels.

Prior to prediction, the study area is split into 12 equally sized overlapping tiles with a shape of 2500x2500 pixels to reduce computational load. The predicted tiles are merged together into a single mosaic. The means of edge pixels are calculated to account for overlap.

All landslide susceptibility maps were further post-processed using a 3x3 mean low-pass filter to smooth out edges and eliminate noisy pixels. Resulting maps are further classified using geometric intervals and discretized into 5 levels (very low, low, moderate, high, and very high).

4.2.1 Optimal selection of landslide influencing factors

Landslide influencing factors were selected based on their predictive capability using the information gain (IG) ratio. The results showed that stream distance and profile curvature have 0 IG and were thus eliminated from the input data. The lithology of the area is predominantly composed of non-porous and non-permeable volcanic igneous rocks which are highly resistant to weathering, hence making it difficult for water to flow. While significant portions are also composed of calcareous sedimentary rocks, these have varying degrees of porosity and permeability due to its structure of interlocking crystals and varying texture sizes making it generally impervious to flow. While some locations may have larger pore sizes which will aid in the formation of underground cavities and karst landforms, the variations make it less ideal to form streams and groundwater reservoirs. Profile curvature has 0 IG score despite planform curvature having a relatively higher degree of importance. This is due to the fact that most of the landslide points were extracted from PSI which records backscatter intensity perpendicular to the satellite sensor, hence recording velocity perpendicular to the direction of the slope. Profile curvature measures values parallel to the maximum slope direction which explains why it did not exhibit any degree of importance.

Fault distance and elevation have the highest IG scores at 0.14 and 0.136, respectively. Since the Legazpi Lineament and Linao faults are seismically active, many PSI V_{SLOPE} points have also been recorded within the faultlines. Additionally, landslides were previously recorded in elevated areas characterized by steep slopes and rugged topography such as in the municipalities of Malilipot and Tiwi, and within the radius of Mayon Volcano which is reflected in its high degree of importance.

The results of the univariate statistical analysis are illustrated in Figure 4.26 and the values described in Table 8.

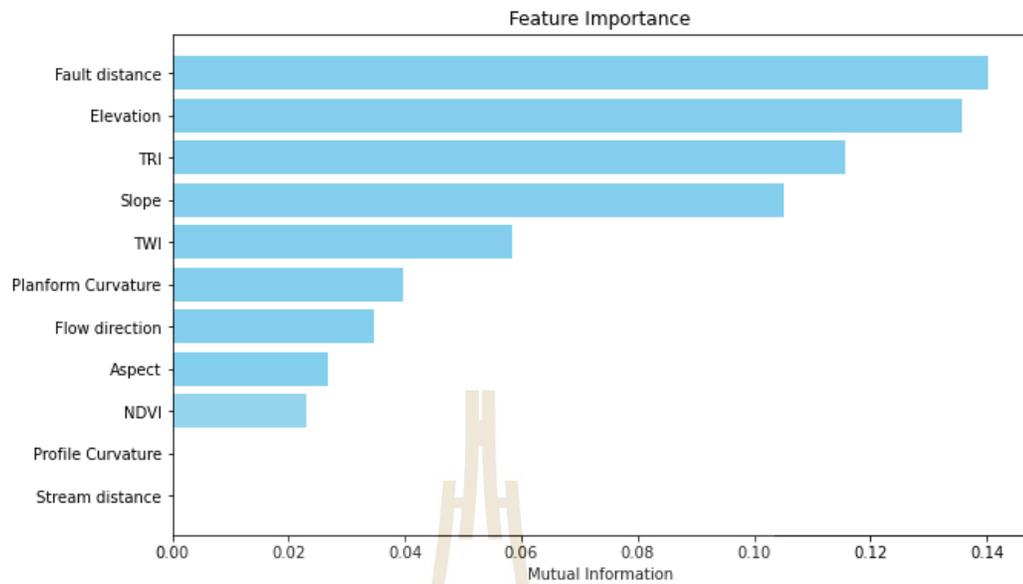


Figure 4.26 Feature importance plot showing mutual information of each influencing factor.

Table 4.5 Mutual information scores (IG Ratio) of each influencing factor.

Landslide influencing factors	Mutual Information (IG Ratio)
Fault distance	0.14
Elevation	0.136
TRI	0.116
Slope	0.105
TWI	0.058
Planform curvature	0.04
Flow direction	0.035
Aspect	0.027
NDVI	0.023
Profile curvature	0.0
Stream distance	0.0

4.2.2 Landslide Susceptibility Mapping using Deep Learning Algorithms

The high predictive performance of the two deep learning algorithms used in the study shows their viability for creating accurate landslide susceptibility maps. Both

algorithms are optimized using RMSprop (root-mean square propagation), an extension of stochastic gradient descent (SGD) first proposed by Hinton (2012) which estimates the moving average of squared gradients.

Both deep learning algorithms share the same hyperparameters with regards to the learning rate, activation function, optimizer, and number of neurons in fully connected layers. However, in the case of CNN-2D, three convolutional filters and average pooling layers were used to extract features from the input data. The choice of average pooling layers over max-pooling layers was because binary classification has probabilistic outputs. Max-pooling layers excel in detecting edges by extracting the maximum values of feature maps which is useful for hard classification tasks such as LULC classification but may eliminate crucial values in probabilistic tasks. Since landslide susceptibility mapping is a binary classification task, the information provided by the output is highly dependent on the predicted probabilistic values and as such, it needs information considering the values of all pixels in the feature maps. Each convolutional filter was regularized using L2 regularization or weight decay to smooth out the gradients and penalize large weights to mitigate overfitting, improving predictive performance.

To further reduce training time, early-stopping callback routines were implemented on both models to stop training when the accuracy did not increase after 35 iterations. To avoid the possibility of yielding a poor output due to early convergence, a warm start of 20 epochs was also implemented. This method allows the learning rate to gradually increase becoming constant after a number of iterations. Both models are created with Tensorflow version 2.10. The models are compiled and trained on a server computer running on Intel i9-9900KF 3.6 GHz CPU (up to 5 GHz overclocking), 4-slot 2.6 Ghz DDR4 RAM totaling 64GB, and NVIDIA GeForce RTX 2070 Super 8GB with 2560 CUDA cores, 320 tensor cores, and 40 raytracing acceleration cores. The architecture of both MLP and CNN-2D are illustrated in Figures 4.27 and 4.28. The description of the deep MLP used in the study are described in Table 4.6. A detailed description of convolutional filters, and their hyperparameters is given in Table 4.7. The fully connected layers in the CNN-2D model uses the same hyperparameters described in the MLP model.

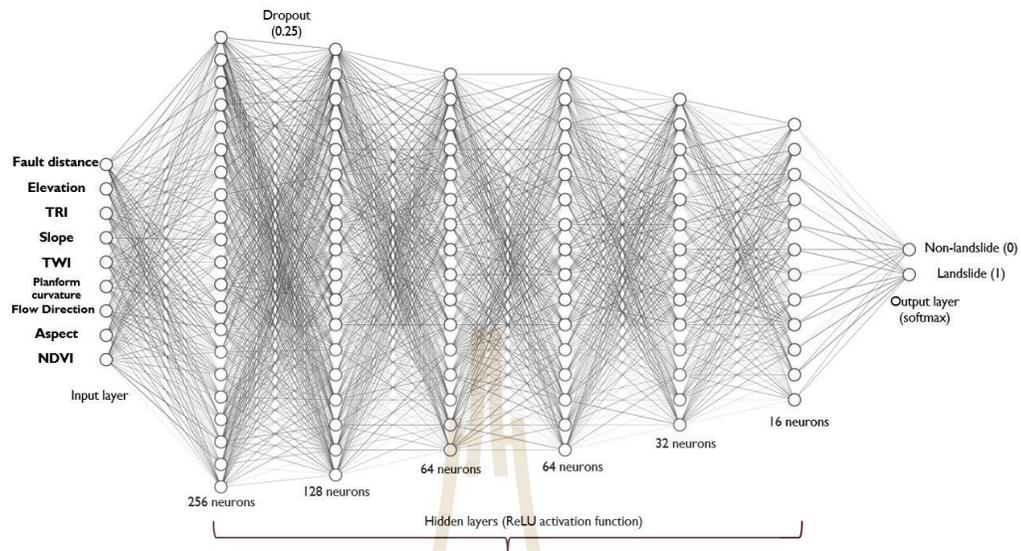


Figure 4.27 Schematic diagram of the deep MLP used in the study.

Table 4.6 Specifications of the fully connected layers used by both MLP and CNN-2D.

Hidden layers:

Dense layer 1 – 256 neurons (dropout = 0.25)

Dense layer 2 – 128 neurons

Dense layer 3 and 4 – 64 neurons

Dense layer 5 – 32 neurons

Dense layer 6 – 16 neurons

Dense layer activation function - ReLU

Output layer:

2 neurons

Softmax activation function

Learning rate = 0.0003

Warm-start enabled – 20 epochs

Optimization function = RMSprop

Early-stopping callback enabled – 35 iterations

Table 4.7 Specifications of CNN-2D model used in the study.

7x7 convolutional filter:

- 32 neurons
- Padding = 'same'
- L2 regularizer = 0.01
- ReLU activation function
- Batch normalization enabled

2x2 Average Pooling (strides = 2, same padding)

5x5 convolutional filter:

- 32 neurons
- Padding = 'same'
- L2 regularizer = 0.01
- ReLU activation function
- Batch normalization enabled

2x2 Average Pooling (strides = 2, same padding)

3x3 convolutional filter:

- 64 neurons
- Padding = 'same'
- L2 regularizer = 0.01
- ReLU activation function
- Batch normalization enabled

2x2 Average Pooling (strides = 1, same padding)

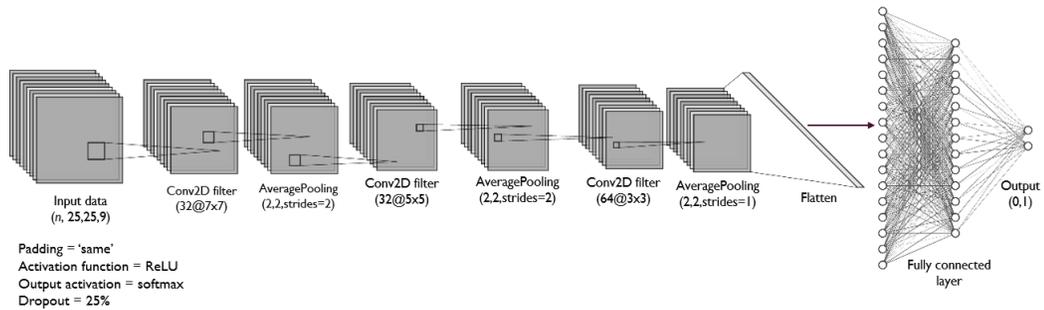


Figure 4.28 Schematic diagram of the CNN-2D model.

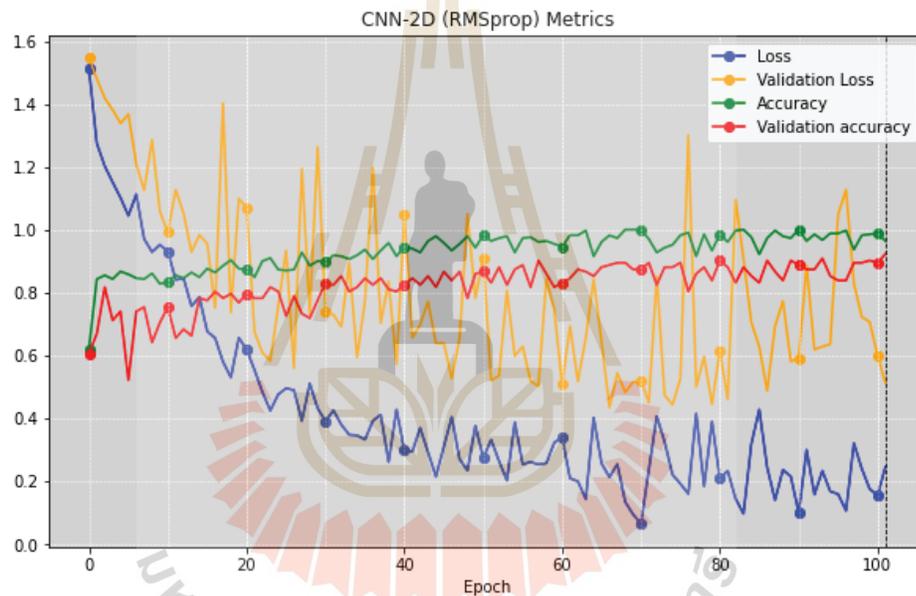


Figure 4.29 Loss and accuracy metrics of CNN-2D in every epoch.

Analysis metrics of CNN-2D shows that the model converged after 102 epochs due to the implementation of early-stopping callbacks. Considerable oscillations were observed in the loss and validation loss as shown in Figure 4.29. This may be attributed to a large variation of values in each sample as convolutional neural networks consider the value of neighboring pixels. However, the accuracy and validation accuracy remained relatively stable throughout training indicating that its overall predictive performance is not truly affected by variations in the training data.

Although MLPs do not consider neighboring pixels to identify spatial patterns as the model treats data as a 1-dimensional vector, similar trends in loss and validation loss were observed in the metrics of the MLP model which converged at 55 epochs. In addition to the learning rate which may have been too fast for both CNN-2D and MLP causing the loss values to oscillate, the similarity may be attributed to the input data as well. The original 25x25 grid size input data was converted to a 1x1 pixel containing the average values of all landslide influencing factors. During the process, large variations in the input values likely occurred which was reflected in the MLP metrics plot as shown in Figure 4.30.

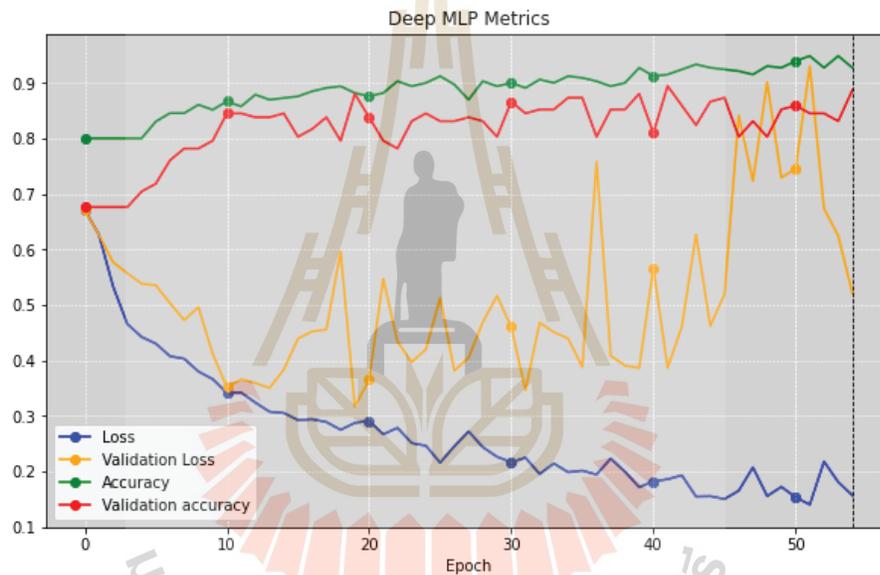


Figure 4.30 Loss and accuracy metrics of Deep MLP in every epoch.

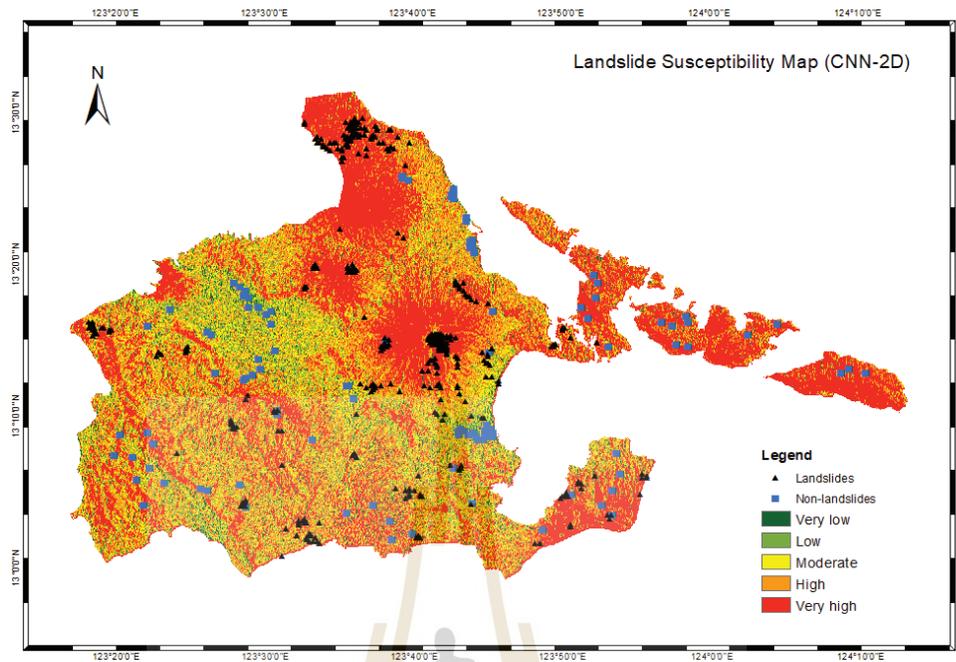


Figure 4.31 Landslide susceptibility map of Albay using CNN-2D optimized with RMSprop.

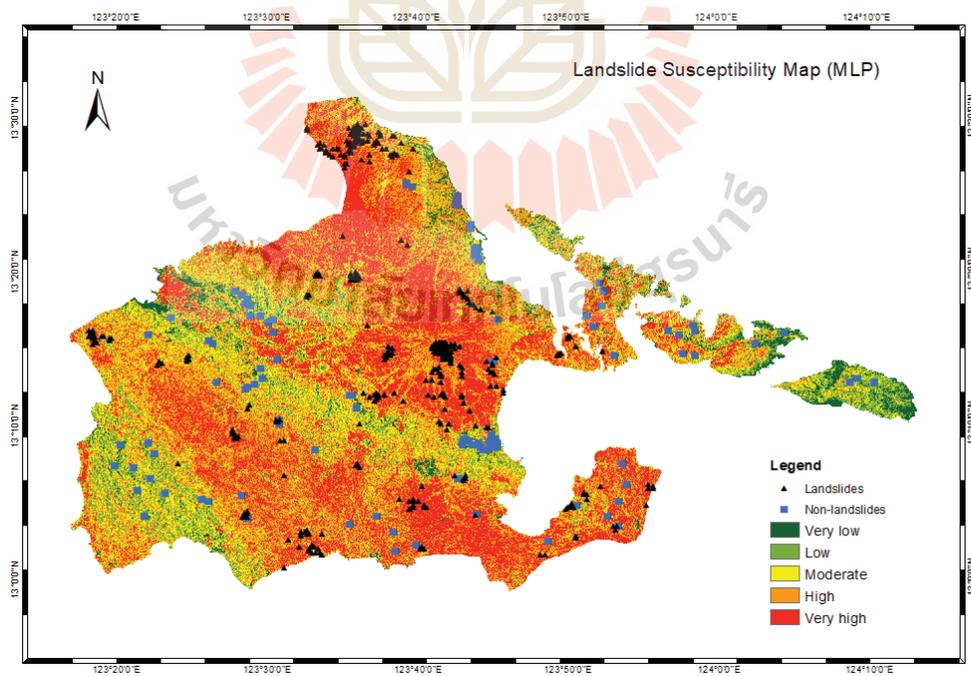


Figure 4.32 Landslide susceptibility map of Albay with Deep MLP.

The overall accuracy of 89% and 92% of MLP and CNN-2D, respectively shows that both models have comparably good results. While both models exhibit similar sensitivity in predicting landslide classes, CNN-2D shows higher sensitivity in predicting non-landslide points with a score of 93% as opposed to 84% shown by MLP. This contrast is further emphasized in Figures 68 and 69 wherein the CNN-2D model detected patterns of very high landslide susceptibility in the island municipality of Rapu-rapu located in the eastern part of the province where non-landslide points were designated. In the mountainous southwestern portions of the province, CNN-2D was also able to detect similar patterns of susceptibility that the MLP did not detect. As opposed to MLP which functions similarly with conventional machine learning algorithms in that it only considers individual pixels, CNN-2D considers the values of neighboring pixels during training and prediction. Specificity shows similar results in that both models can correctly identify true negative results at a relatively similar rate, but with CNN-2D outperforming MLP by 3 and 4% in landslide and non-landslide classes, respectively. The resulting output maps were classified using natural breaks to better illustrate the differences in values in each susceptibility class.

Table 4.8 Classification report of CNN-2D.

Classes	Precision	Recall	F-1 score
Non-landslide	93%	83%	87%
Landslide	92%	97%	94%
<i>Weighted average</i>	92%	90%	91%
<i>Overall model accuracy - 92%</i>			

Table 4.9 Classification report of MLP.

Classes	Precision	Recall	F-1 score
Non-landslide	84%	80%	82%
Landslide	91%	93%	92%

Table 4.9 (Continued).

<i>Weighted average</i>	89%	89%	89%
<i>Overall model accuracy - 89%</i>			

While the output susceptibility maps differ in their composition, both models are on par with one another based on their predictive performance with the difference being that CNN-2D can classify non-landslide points more accurately as exhibited by its high sensitivity.

A total land area of 2500 sq. km. was successfully classified by CNN-2D. Predicted probability values were classified using geometric interval and then reclassified into five discrete classes showing the degree of susceptibility. Most of the area are classified as having high to very high susceptibility to landslides. A total area of 990.92 sq. km. were classified as having very high susceptibility to landslides making it the highest observed susceptibility class followed by high susceptibility class totaling to 701.58 sq. km. Only 73.8 sq. km. of the province were classified as having very low susceptibility, most of it being located in low-lying urban areas near the coastline where landslides are highly unlikely to occur followed by 246.77 sq. km. of low susceptibility areas.

Table 4.10 Area of landslide susceptibility classes.

Class	Area (sq.km.)
Very Low	73.8
Low	246.77
Moderate	487.17
High	701.58
Very high	990.92

4.2.3 Benchmark comparisons and accuracy assessment

Conventional machine learning algorithms namely logistic regression, random forest, and support vector machines were used for benchmark comparisons. For

random forest and SVM, a grid-search cross-validation method was employed obtain the optimal number of parameters. Among the machine learning algorithms used, random forest had the highest accuracy with a score of 86% followed closely by SVM at 85%. Logistic regression had the lowest overall accuracy at 79%. Similar to CNN-2D models, all machine learning algorithms exhibited high sensitivity in classifying non-landslide points.

Table 4.11 Classification report of logistic regression.

Classes	Precision	Recall	F-1 score
Non-landslide	90%	39%	55%
Landslide	77%	98%	86%
<i>Weighted average</i>	81%	79%	76%
<i>Overall model accuracy - 79%</i>			

Table 4.12 Classification report of random forest including the parameters used.

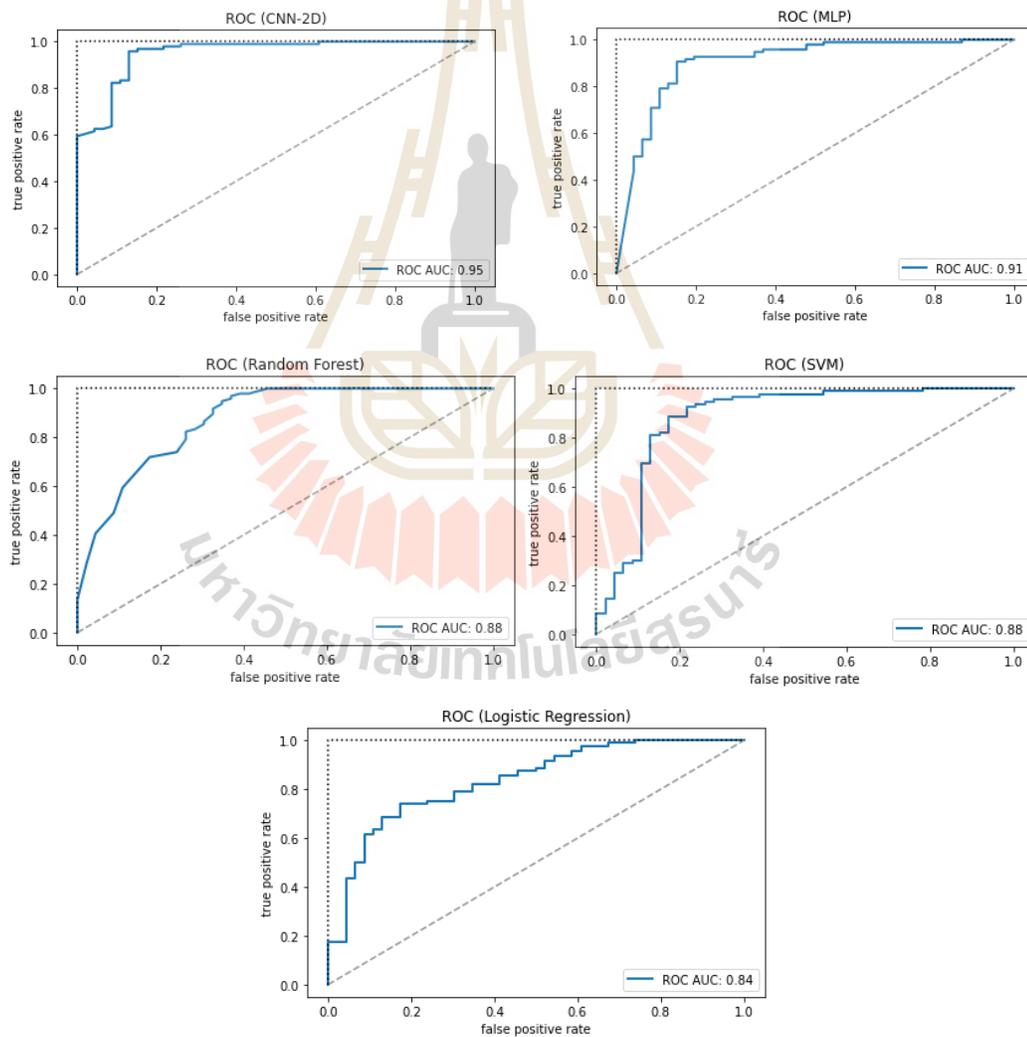
Classes	Precision	Recall	F-1 score
Non-landslide	93%	61%	74%
Landslide	84%	98%	90%
<i>Weighted average</i>	87%	86%	85%
<i>Overall model accuracy - 86%</i>			
<i>Parameters: 50 trees, gini criterion</i>			

Table 4.13 Classification report of SVM including the parameters used.

Classes	Precision	Recall	F-1 score
Non-landslide	93%	59%	72%
Landslide	83%	98%	90%
<i>Weighted average</i>	86%	85%	84%
<i>Overall model accuracy - 85%</i>			
<i>Parameters: C = 10, Gamma = 0.01, Kernel = radial basis function (RBF)</i>			

Table 4.14 Benchmark comparison metrics for all models used.

Model	Sensitivity	Specificity	AUC score	Overall accuracy
CNN-2D	97%	83%	95%	92%
MLP	93%	80%	91%	89%
LR	98%	39%	84%	76%
RFC	98%	61%	88%	86%
SVM	98%	59%	88%	85%



Figures 4.33 ROC plots of deep learning (CNN-2D, MLP) and conventional machine learning algorithms (Random forest, SVM, logistic regression).

Based on the resulting landslide susceptibility maps, both random forest and SVM showed comparably similar results. All resulting maps showed similar patterns of landslide susceptibility in the southwestern areas of the province similar to CNN-2D. The landslide susceptibility map created by logistic regression detected similar patterns of very high susceptibility in Rapu-rapu similar to CNN-2D. This is in part due to the fact that logistic regression also uses weight decay to penalize large weights similar to how the convolutional filters used in this study were regularized. Random forest model emphasized very low degree of susceptibility in areas with low slope gradient. Although the impact of fault distance in the predicted output has also been emphasized in the model in the same area, the low slope gradient in the area makes it less likely for landslides to occur. However, periodic earthquakes may trigger landslides in the surrounding mountainous areas as shown by the contrasting susceptibility classes in the predicted output.

All models predicted very high degrees of susceptibility in the western portion of Tiwi where heavy landslides were recorded in 2019 following successive typhoons. Very high degrees of susceptibility were also observed in the mountainous portion of the municipality of Manito bordering the neighboring Sorsogon province. In both SVM and CNN-2D, the center of Mayon Volcano is classified as having a moderate degree of susceptibility in contrast to other models which classified the entirety of Mayon Volcano as having very high degrees of susceptibility. It might be attributed to the fact that despite being active with periodic eruptions, the cone remains relatively stable and has not yet collapsed into a caldera retaining its near-perfect shape.

All models showed a high degree of sensitivity indicating that all models have a high rate of identifying landslide classes during accuracy assessment. All three conventional machine learning models have a very high true positive rate of 98% which is marginally higher than CNN-2D at 97%. MLP has a comparably lower score at 93%. However, all conventional machine learning models also have a significantly lower rate of true negatives with RFC and SVM scoring a specificity of 61% and 59%, respectively. LR has the lowest specificity score at 39%. Both deep learning models were able to identify true negatives at a significantly higher rate at 80% and 83% for MLP and CNN-2D, respectively.

Further analysis of the model metrics shows that CNN-2D outperforms all of the models. With a ROC score of 95% and overall accuracy of 92%, it indicates that CNN-2D is better at classifying landslides and distinguishing between non-landslide areas. This performance is due to the fact that the model excels at extracting meaningful features and analyzing spatial patterns by making use of convolutional filters. This is opposed to MLP and conventional machine learning algorithms which generally treat each input data as an individual pixel which will not allow the models to extract meaningful spatial context. However, MLP results also show close predictive performance despite not being able to detect the same hidden patterns shown by the landslide susceptibility map produced by CNN-2D. An AUC of 91% and overall accuracy of 89% indicates that MLP can be used for binary classification of landslides as an alternative to conventional machine learning algorithms due to its comparably higher predictive performance. It can also be used to substitute for CNN-2D which may be computationally intensive on most hardware.

Among the conventional machine learning algorithms, random forest and SVM show comparably similar results with overall accuracy of 86% and 85%, respectively. While both models have lower accuracy compared to deep learning models, the AUC score of 88% in both models show that their performance is not far behind from MLP. This is because both models are robust to multicollinearity. Random forest makes use of bootstrap sampling due to its method of splitting features into different sets and then picking the best feature among the samples. On the other hand, SVM models fitted with non-linear kernels; particularly RBF kernel as used in the study, avoids multicollinearity by transforming features into a higher dimensional space effectively decorrelating input variables. Logistic regression recorded the lowest score in every metric despite being commonly used for binary classification tasks. Multicollinearity is a common problem which affects the predictive performance of linear models as these models operate on an assumption that input data are independent of one another. The presence of L2 regularization or weight decay in logistic regression model further affected predictive performance by heavily penalizing higher weights observed in data with high variance.

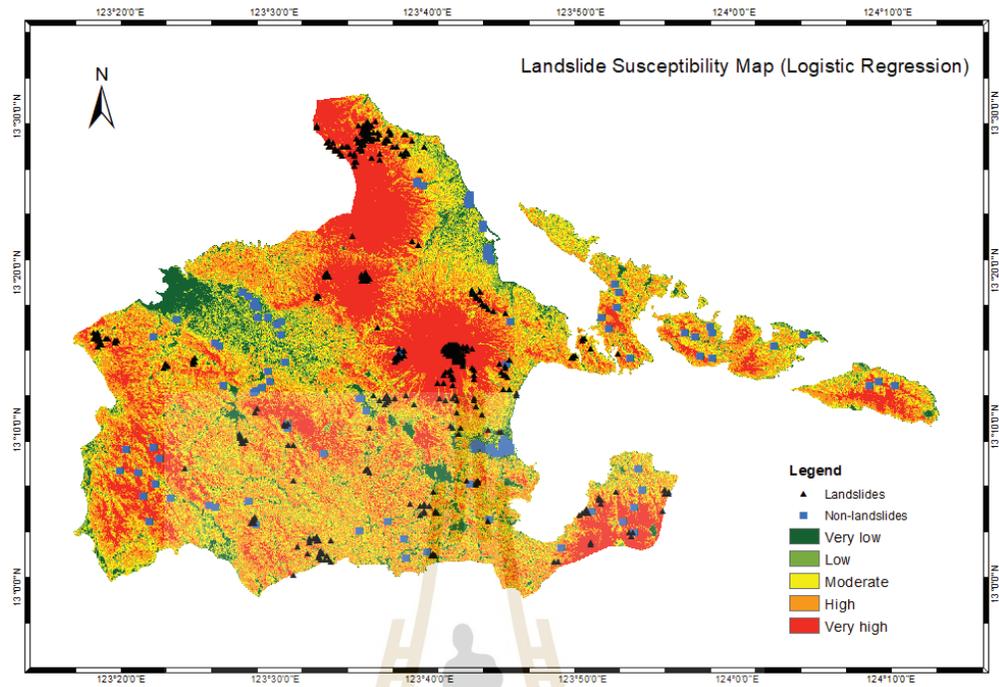


Figure 4.34 Landslide susceptibility map of Albay using logistic regression.

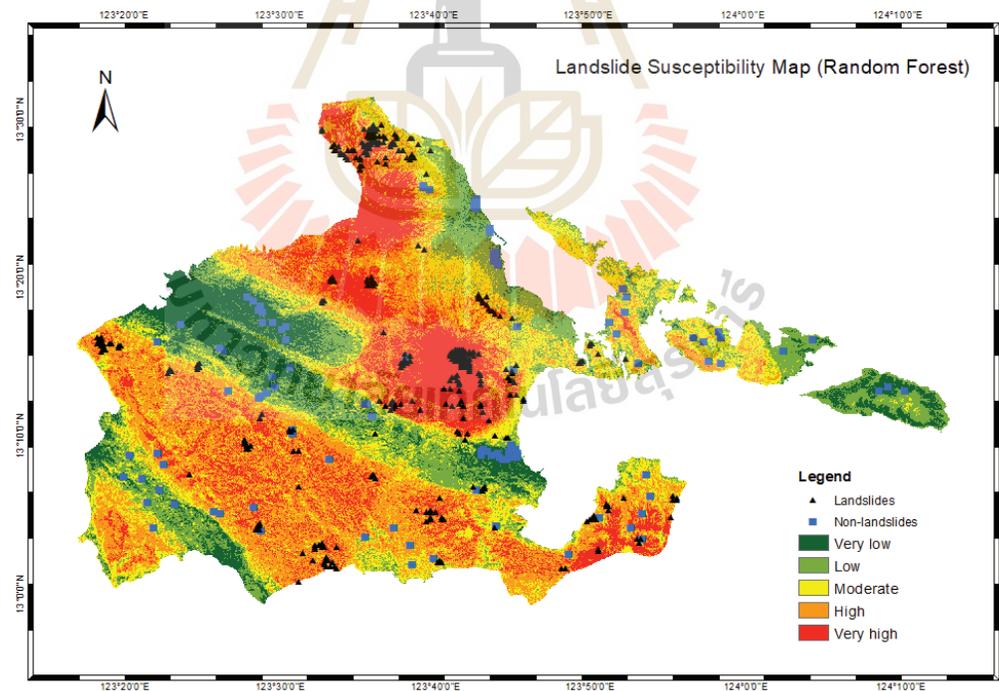


Figure 4.35 Landslide susceptibility map of Albay using random forest.

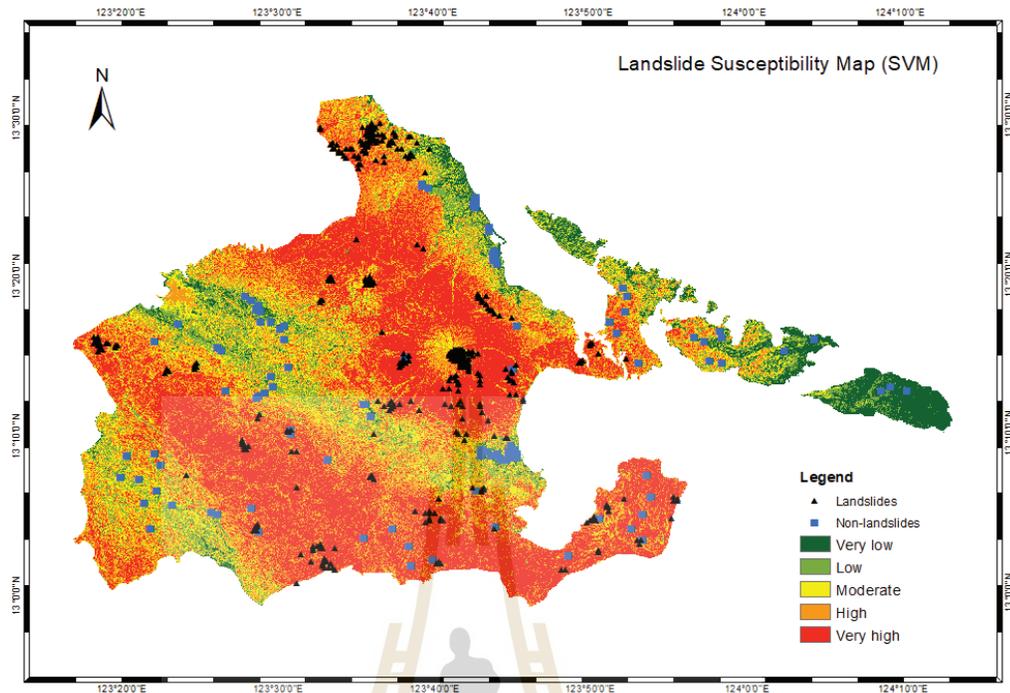


Figure 4.36 Landslide susceptibility map of Albay using SVM.

4.3 Vulnerability and Risk Assessment

In order to analyze the potential impact of landslides, vulnerability and risk assessment is conducted at the municipal level and subdistrict level, respectively. Geometric interval was used in order to balance the distribution of classes and to highlight changes between the median and extreme values in the resulting vulnerability and risk maps.

Vulnerability factors for each municipality in this study were divided into two indicators: economic parameters and resiliency. Economic parameters cover 2018 annual income (millions) obtained from the Department of Finance and poverty incidence data obtained from Philippine Statistics Authority. The annual economic income was used as an economic parameter as areas with higher economic activity can typically fund the necessary policies and equipment that can mitigate disasters. Poverty incidence, while seemingly tied to economic income, displays the general economic status and well-being of the citizens living in the area. Areas with a lower

incidence of poverty means that people living in the area are being given adequate welfare and social services which can help with their readiness to potential disasters allowing them to easily recover from the aftermath.

Due to lack of high resolution satellite data and information regarding costs of property and materials in infrastructures, resiliency values used data obtained from the 2022 annual ranking Cities and Municipalities Competitive Index of the Department of Trade and Industry detailing critical infrastructures.

Table 4.15 Economic parameters for each city and municipality in Albay.

City/Municipality	Annual income (Php millions)	Poverty incidence (percentage)
Bacacay	216.37	31.5
Camalig	233.95	21.1
Daraga	473.66	13.5
Guinobatan	288.39	25.6
Jovellar	96.53	37.5
Legazpi City	1368.08	13.8
Libon	260.07	37.4
Ligao City	844.09	23.1
Malilipot	183.51	26.1
Malinao	160.12	33.7
Manito	138.24	37.2
Oas	235.97	34.2
Pio Duran	171.59	37.5
Polangui	292.42	21.3
Rapu-rapu	150.38	42.7
Sto. Domingo	138.52	23
Tabaco	711.83	19.8
Tiwi	239.58	23.6

Table 4.16 Resiliency scores for each city and municipality.

City/Municipality	Healthcare infrastructure	Emergency infrastructure	Early Warning System	Utilities
Bacacay	0.2694	0.3768	1.0012	1.5043
Camalig	0.2134	0.3535	1.0037	1.517
Daraga	0.2816	0.3696	1.0025	1.4998
Guinobatan	0.0325	0.3554	1	1.0004
Jovellar	0.0133	0.3769	0.004	1.5001
Legazpi City	1.3211	0.645	1.0884	1.6553
Libon	0.101	0.4363	1.0012	0.9999
Ligao City	0.1053	0.6261	1.0071	1.8338
Malilipot	0.0084	0.3607	1.0356	1.2508
Malinao	0.0188	0.3724	1.0059	1.2509
Manito	0.0495	0.3608	1.0119	1.4982
Oas	0.0047	0.3492	1.0658	1.5008
Pio Duran	0.1609	0.4119	1.0237	1.5
Polangui	0.1969	1.5039	1.005	1.5039
Rapu-rapu	0.0848	0.3372	1.004	1.4995
Sto. Domingo	0.0214	0.3287	1.002	1.5006
Tabaco	0.1434	0.4108	1.0342	1.4997
Tiwi	0.0712	0.3665	1.0074	0.0036

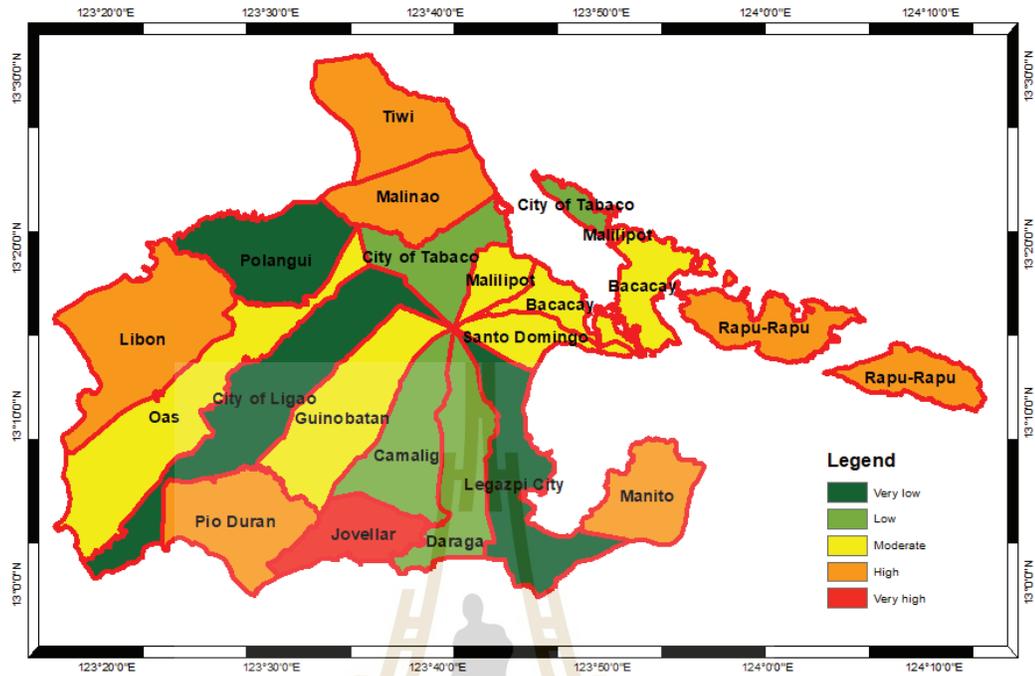


Figure 4.37 Social vulnerability map of Albay.

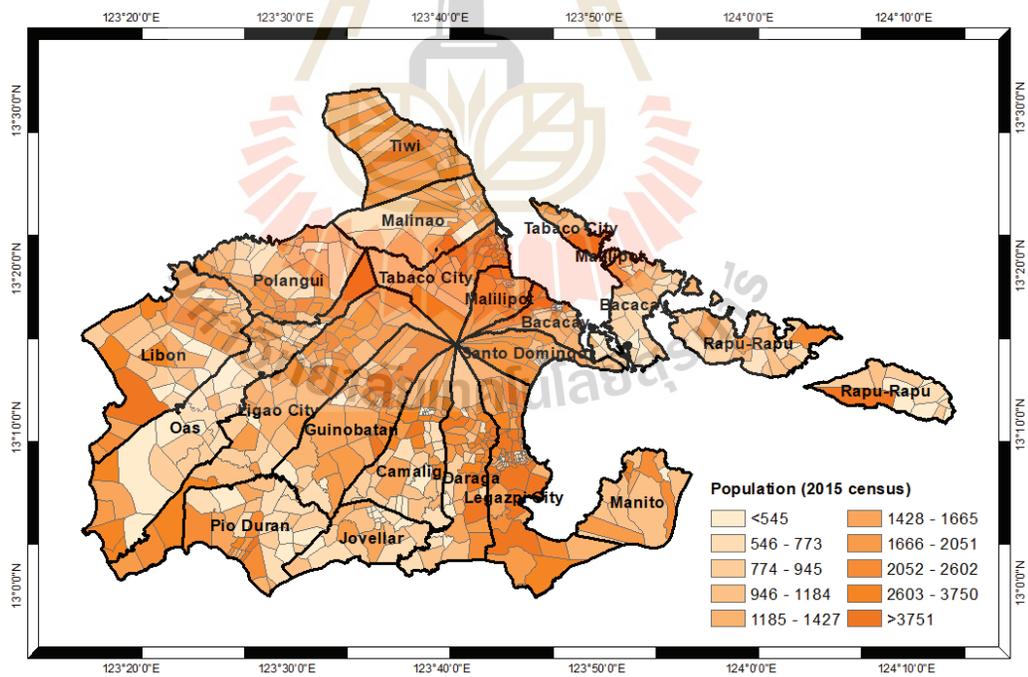


Figure 4.38 Population map of Albay per subdistrict.

The results show that Jovellar has a very high degree of vulnerability. Six other municipalities are also highly vulnerable. Only the capital city of Legazpi, Ligao City, and Polangui have a very low degree of vulnerability attributed to their high resiliency scores and high annual revenue.

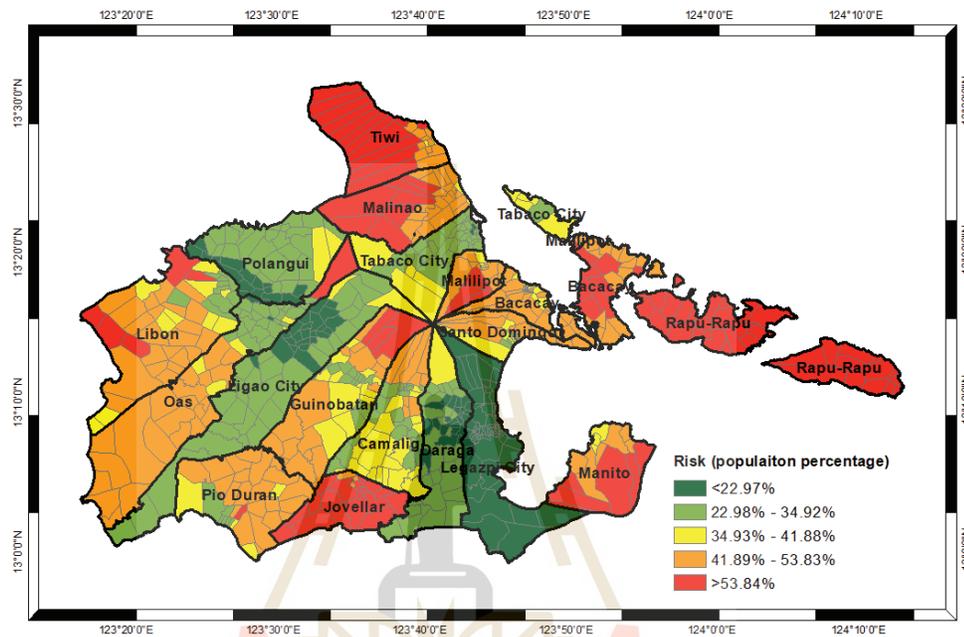


Figure 4.39 Landslide risk exposure map of Albay.

The landslide risk exposure map is generated by combining the results of the classified landslide susceptibility map and the social vulnerability map with the population of each individual subdistrict as per the 2015 census. No subdistrict population census data for 2020 and beyond for the province of Albay is freely available at the time of writing. The resulting landslide risk map is expressed in terms of the percentage of population exposed to landslides in terms of percentage in each subdistrict. Out of 720 subdistricts, 290 subdistricts have been classified as high risk with 103 subdistricts classified as very high risk. In the northern part of the province, the western portions of the municipalities of Tiwi and Malinao is classified as very high risk with more than half of the population in the subdistricts at risk of landslides. A single subdistrict in the municipality of Oas bordering Malinao, Tabaco City, Polangui, and Ligao City is classified as very high risk.

Significant portions of the municipalities of Oas, Libon, Guinobatan, and Camalig are classified as high risk. Nearly the entirety of Manito, Malilipot, Bacacay, and Sto. Domingo are at high to very high risk. The majority of the municipality of Jovellar is at very high risk while the entirety of the island municipality of Rapu-rapu is at very high risk to landslides. The aforementioned locations are highly susceptible to landslides with a moderate to high social vulnerability which may not be enough to fund disaster mitigation measures thereby putting significant number of lives at risk. The entirety of Rapu-rapu being at very high risk is a cause for major concern. While the municipality is relatively sparsely populated, its high vulnerability to disaster and its remote location from the mainland will exacerbate the consequences posed by a potential landslide disaster putting entire populations at risk.

Subdistricts in the municipality of Malilipot, Guinobatan, Sto. Domingo, and Camalig within the radius of Mayon Volcano are also classified as high to very high risk due to the inherent combined hazards posed by being located in near proximity to an active volcano. Subdistricts belonging to the cities of Ligao and Tabaco, and the municipality of Daraga also within the proximity of Mayon Volcano has moderate risk to landslides while the rest of the aforementioned areas have low risk exposure. This is mainly due to their low social vulnerability to landslide disasters owing to their high resiliency made possible by their high economic activity which can readily fund disaster mitigation strategies.

The entirety of the capital city of Legazpi is at very low risk despite its close proximity to high risk areas and high population. Being the center of commerce and governance in the province means that the city has enough funds for disaster risk mitigation measures. Additionally, minimizing the risk posed by disasters in the city will ensure that the whole province will not be paralyzed economically and politically as key local government offices, large establishments, and transportation hubs are located in the area.

CHAPTER V

CONCLUSION

The results of the study highlights the use of persistent scatterer interferometry and the semi-automatic extraction of landslide features from multispectral satellite imagery. These techniques provided a viable solution in compensating for the lack of landslide inventory data which is highly crucial in the province of Albay as it is prone to landslides that are further aggravated by other natural disasters.

By projecting the extracted time-series velocity points from SAR imagery to the direction of the steepest slope, a total of 13,152 PS points showing downward slope velocity (V_{SLOPE}) indicative of landslides were extracted after post-processing with the stability threshold set at 11 mm./yr. It has been found that the maximum V_{SLOPE} in the province of Albay is at 226.31 mm./yr with the average at 18.34 mm./yr. Using PSI-based hotspot analysis on the V_{SLOPE} points showed that the measured slope displacement are significantly clustered given a p-value of <0.01 and z-score of >2.58. A total of 356 landslide points showing 99% degree of confidence with a velocity of 33 mm./yr. given a stability threshold of 40 were successfully extracted.

Combining three significant landslide events that occurred in the province of Albay from 2007, 2018, and 2020 the study was also able to successfully extract landslide features using semi-automatic Gaussian clustering of BSI and NDVI which corresponds to the post-landslide satellite imagery which are then combined with slope values. A total of 200 landslide features have been extracted successfully with a silhouette coefficient ranging from 0.55-0.69. Post-processing of landslide points in order to be accommodated by the deep learning models as well as to reduce incident points gave a final inventory of 378 landslide points while 113 non-landslide points were manually created.

During landslide susceptibility mapping, both CNN-2D and MLP outperformed conventional machine learning algorithms in all metrics with overall accuracy of 92% and 89%, respectively. Random forest and SVM both with an AUC score of 88% shows

comparable performance with MLP with an AUC score of 91%. While also being outperformed by deep learning models, random forest in particular showed an overall on par with MLP at 86%. However, the significantly low rate of specificity for conventional machine learning models in contrast with their high sensitivity suggests that these models have a tendency to overestimate the rate at which landslide occurs which contributed to their significantly lower AUC and overall accuracy.

The results also suggested that despite the non-linear parameters used in the conventional machine learning algorithms (except LR), these models cannot find hidden patterns that will help improve predictive performance unlike in the case of deep learning models which utilizes a very high number of hyperparameters to arrive at an optimal solution at the cost of computational resources. In turn, conventional machine learning algorithms had a higher rate of overfitting which may be exacerbated by the inclusion of larger input data.

However, RF and SVM can still perform relatively on par with deep learning models in terms of AUC and overall accuracy. Given major adjustments and further optimization in order to improve sensitivity and specificity, both models may also have the potential to perform similarly or even outperform deep learning models. These observations also prove that conventional machine learning algorithms can still be used as an alternative to deep learning models which are generally computationally expensive and will require specific hardware for the architectures to run smoothly and allow for easier hyperparameter tuning.

More than half of the province have been classified as having high to very high susceptibility to landslides totaling to 1600 sq.km out of 2500 sq.km. that was successfully classified. Additionally, 290 subdistricts have been classified as having high to very high risk to landslides. The results imply the need for comprehensive disaster risk reduction and contingency plan to mitigate the impact of landslides on exposed populations, especially in poorer, vulnerable municipalities.

5.1 Recommendations

The study highlights the potential use of PSI to detect possible landslides based on their vertical slope velocity. While the use of hotspot clustering analysis may

provide statistically significant results and reduce the number of points, it is important to note that this approach and other similar techniques must be conducted in tandem with field verification to test their reliability such as using different for conceptualizing spatial relationships or using optimized hot-spot analysis as suggested by Lu et al. (2019) and Hakim et al. (2021). A corrected range index may also be used as also suggested by Notti et al. (2014) and Solari et al. (2019) to which also factors in the effect of LULC in estimating ground motion.

As the values recorded in $PSI_{V_{SLOPE}}$ points only have a limited temporal coverage, time-series analysis may be conducted as well in order to predict the dynamics and future trends of potential slope failures in the area. This is crucial in areas such as the province of Albay which regularly experiences a combination of natural disasters. Additionally, while the SAR images are not affected by weather patterns per se, their coherence and in turn, the backscatter intensity may be affected, leading to erroneous measurements or negative coherence in the images. Future studies using PSI must consider the factors mentioned above to refine its use for locating landslides.

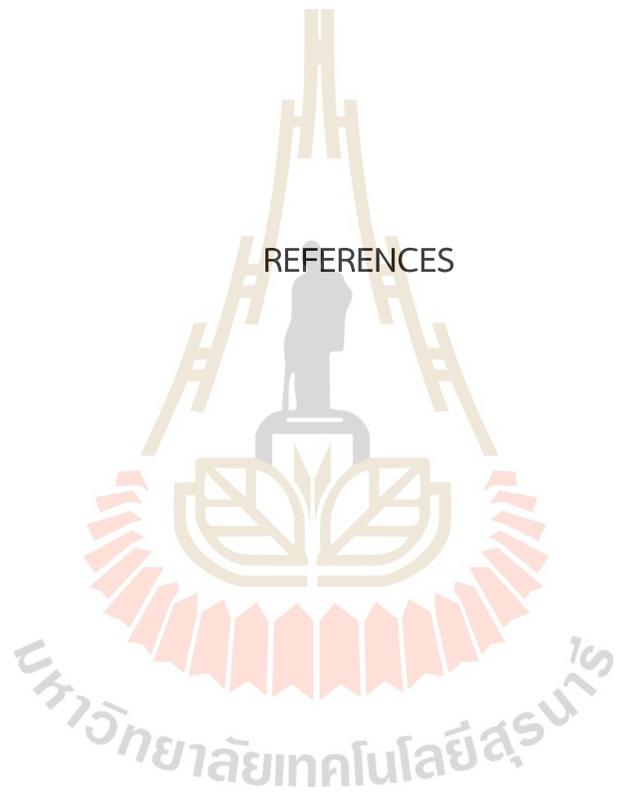
The use of multi-spectral satellite imagery for confirming locations of previous landslides in tandem with archival reports, while relatively effective in cases where data is unavailable, is a time-consuming process requiring manual cross-verifications with written articles in addition to side-by-side comparison of pre- and post- satellite imagery. This may also prove difficult in areas with large cloud covers, particularly in archipelagic countries with a large temporal gap between good quality satellite images. In the context of landslide mapping, this presents a problem as vegetation will have time to regrow in previously barren areas that suffered slope failures.

While Gaussian mixture modeling as an unsupervised learning method may expedite the process of pinpointing landslides, the approach used in the study still requires a prior knowledge to predict the probabilities of input data belonging to certain clusters. Furthermore, post-processing is required in order to extract landslide clusters and must also be cross-checked with pre-landslide satellite imagery in order to reduce the risk of misclassification. Future studies may opt to use high-resolution satellite imagery for easier manual extraction of landslides.

Additionally, while landslide susceptibility mapping using deep learning algorithms showed results with high accuracy, further validation methods may also be used in all models in order to assess how the models can perform under different conditions. Cross-validation of all chosen machine learning and deep learning algorithms may be conducted as suggested by Jiang and Chen (2016) and Shirzadi et al. (2017) to analyze how each machine learning and deep learning models can perform given independent datasets.

The lack of a comprehensive landslide hazard map is also another aspect that needs to be addressed as the probability may change over time due to its reliance on dynamic spatiotemporal data. The inclusion of annual rainfall as suggested by Wang et al. (2021) or seismicity and other natural disasters as recommended by Corominas et al. (2014) can be used in order to better analyze how landslide patterns can change over time. The results of the risk and vulnerability map are still highly subjective and are subject to change due to lack of spatial data concerning property costs. Data such as road accessibility, capacity of evacuation centers, locations of critical infrastructure, urban parcels, plantations, and registered business establishments, may also be considered in order to estimate the potential socio-economic impact of landslides.

REFERENCES



REFERENCES

- Abidin, N. H. Z., Remli, M. A., Ali, N. M., Phon, D. N. E., Yusoff, N., Adli, H. K., and Busalim, A. H. (2020). Improving intelligent personality prediction using Myers-Briggs type indicator and random forest classifier. *International Journal of Advanced Computer Science and Applications*, 11(11). doi: 10.14569/IJACSA.2020.0111125
- Abdulwahid, W. M., and Pradhan, B. (2016). Landslide vulnerability and risk assessment for multi-hazard scenarios using airborne laser scanning data (LiDAR). *Landslides*, 14(3), 1057-1076. doi: 10.1007/s10346-016-0744-0
- Agarap, A. (2018). Deep Learning using Rectified Linear Units (ReLU). ArXiv: 1803.08375
- Ariza, A., Davila, N.A., Kemper, H., and Kemper, G. (2021). Landslide detection in Central America using the differential bare soil index. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(2021), 679-684. doi:10.5194/isprs-archives-XLIII-B3-2021-679-2021
- Aslan, G., Foumelis, M., Raucoules, D., De Michele, M., Bernardie, S., and Cakir, Z. (2020). Landslide Mapping and Monitoring Using Persistent Scatterer Interferometry (PSI) Technique in the French Alps. *Remote Sensing*, 12(8). doi: 10.3390/RS12081305
- Ayer, T., Chhatwal, J., Alagoz, O., Kahn, C. E., Jr, Woods, R. W., and Burnside, E. S. (2010). Informatics in radiology: comparison of logistic regression and artificial neural network models in breast cancer risk estimation. *Radiographics*, 30(1), 13–22. doi:10.1148/rg.301095057
- Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., and Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. *Scientific reports*, 11(1), 24112. doi:10.1038/s41598-021-03585-1

- Barcia, R. B. (2019, January 2). Residents in 2 areas in Tiwi, Albay asked to leave due to land cracks. *Rappler*. <https://www.rappler.com/nation/220155-tiwi-areas-cracks-usman-january-2-2019/>
- Bonnabel, S. (2013). Stochastic gradient descent on riemannian manifolds. *IEEE Transactions on Automatic Control*, 58(9), 2217-2229. doi: 10.1109/TAC.2013.2254619
- Boukhater, C., Dakroub, O., Lahoud, F., and Awad, M. (2014). An intelligent and fair GA carpooling scheduler as a social solution for greener transportation. In MELECON 2014 - 2014 17th IEEE Mediterranean Electrotechnical Conference, 182-186. doi: 10.1109/MELCON.2014.6820528
- Bragagnolo, L., da Silva, R. V., and Grzybowski, J. M. V. (2020). Landslide susceptibility mapping with r. landslide: A free open-source GIS-integrated tool based on Artificial Neural Networks. *Environmental Modelling & Software*, 123, 104565. doi: 10.1016/j.envsoft.2019.104565
- Chen, W., Peng, J., Hong, H., Shahabi, H., Pradhan, B., Liu, J., Zhu, A., Pei, X., and Duan, Z. (2018). Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. *Science of the Total Environment*, 626(2018). 1121-1135 doi: 10.1016/j.scitotenv.2018.01.124
- Chollet, F. (2018). *Deep Learning with Python*. Manning Publications
- Ciampalini, A., Raspini, F., Lagomarsino, D., Catani, F., and Casagli, N. (2016). Landslide susceptibility map refinement using PSInSAR data. *Remote Sensing of Environment*, 184, 302-315. doi: 10.1016/j.rse.2016.07.018
- CNN Philippines Staff (2020, November 6). Rolly leaves 22 dead, causes P14-B damage to infrastructure and agriculture. *CNN Philippines*. <https://www.cnnphilippines.com/news/2020/11/6/Rolly-damage-death-toll.html>
- Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopolou, S., and Smith, J. T. (2014). Recommendations for the quantitative analysis of landslide risk. *Bulletin of Engineering Geology and the Environment*, 73, 209-263. doi: 10.1007/s10064-013-0538-8
- Dematera, C. (2006, December 15). Retrieval of 'Reming' victims called off. *The Philipp-*

- ine Star*. <https://www.philstar.com/headlines/2006/12/15/375103/retrieval-145reming146-victims-called-off>
- de Almeida, L. Q., Welle, T., and Birkmann, J. (2016). Disaster risk indicators in Brazil: A proposal based on the world risk index. *International Journal of Disaster Risk Reduction*, 17, 251-272. doi: 10.1016/j.ijdrr.2016.04.007
- DOST-PAGASA (2018). *Tropical Cyclone File Report: Tropical Depression Usman*. https://pubfiles.pagasa.dost.gov.ph/tamss/weather/tc_summary/TD_Usman_2018.pdf
- Echemineda, P. (2006, December 3). Pagasa admits it cannot measure storm rain intensity. *The Philippine Star*. <https://www.philstar.com/headlines/2006/12/03/373045/pagasa-admits-it-cannot-measure-storm-rain-intensity>
- Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., and Dehmer, M. (2020). An Introductory Review of Deep Learning for Prediction Models With Big Data. *Frontiers in Artificial Intelligence*, 3(2020), 1-23. doi: 10.3389/frai.2020.00004
- Espinosa, A. (2013). Geography and public planning: Albay and disaster risk management. *Manila: Human Development Network*.
- Fang, Z., Wang, Y., Peng, L., and Hong, H. (2020). Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Computers and Geosciences*, 139. doi: 10.1016/j.cageo.2020.104470
- Fayne, J. V., Roberts-Pierel, J., Rumsey, A. C., and Kirschbaum, D. (2019). Automated Satellite-Based Landslide Identification Product for Nepal. *Earth Interactions*, 23(3), 1-21. doi: 10.1175/EI-D-17-0022.1
- Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., and Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Engineering Geology*, 102(3-4), 85-98. doi: 10.1016/j.enggeo.2008.03.022
- Fiorucci, F., Ardizzone, F., Mondini, A. C., Viero, A., and Guzzetti, F. (2019). Visual interpretation of stereoscopic NDVI satellite images to map rainfall-induced landslides. *Landslides* 16, 165–174. doi: 10.1007/s10346-018-1069-y

- Froude, M. J., and Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161-2181. doi: 10.5194/nhess-18-2161-2018
- Gaidzik, K., and Ramírez-Herrera, M. T. (2021). The importance of input data on landslide susceptibility mapping. *Scientific reports*, 11(1), 19334. doi: 10.1038/s41598-021-98830-y
- Guillard-Gonçalves, C., and Zêzere, J. L. (2018). Combining social vulnerability and physical vulnerability to analyse landslide risk at the municipal scale. *Geosciences*, 8(8), 294. doi: 10.3390/geosciences8080294
- Guzetti, F., Mondini, A. C., Cardinali, C., Fiorucci, F., Santangelo, M., and Chang, K. (2012). Landslide inventory maps: New tools for an old problem. *Earth Science Reviews*, 112(1-2), 42–66. doi:10.1016/j.earscirev.2012.02.001
- He, M., Li, B., and Chen, H. (2017). Multi-scale 3D deep convolutional neural network for hyperspectral image classification. In 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 3904-3908, doi: 10.1109/ICIP.2017.8297014.
- Highland, L. M., and Bobrowsky, P. (2008). *The landslide handbook-A guide to understanding landslides* (No. 1325). US Geological Survey.
- Hinton, G. (2012) Lecture 6e rmsprop: Divide the Gradient by a Running Average of Its Recent Magnitude. COURSERA: Neural Networks for Machine Learning, 4, 26-31.
- Höser, T. (2018). *Analysing the capabilities and limitations of InSAR using Sentinel-1 data for landslide detection and monitoring* (Doctoral dissertation, Department of Geography, University of Bonn). doi:10.13140/RG.2.2.35085.59362
- Hu, Y., Zhang, Q., Zhang, Y., and Yan, H. (2018). A deep convolution neural network method for land cover mapping: A case study of Qinhuangdao, China. *Remote Sensing*, 10(12), 1-17. doi: 10.3390/rs10122053
- Hussain, S., Hongxing, S., Ali, M., Sajjad, M. M., Ali, M., Afzal, Z., and Ali, S. (2022). Optimized landslide susceptibility mapping and modelling using PS-InSAR technique: A case study of Chitral valley, Northern Pakistan. *Geocarto International*, 37(18), 5227-5248. doi:10.1080/10106049.2021.1914750
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to Statistical Learning*. New York: Springer.

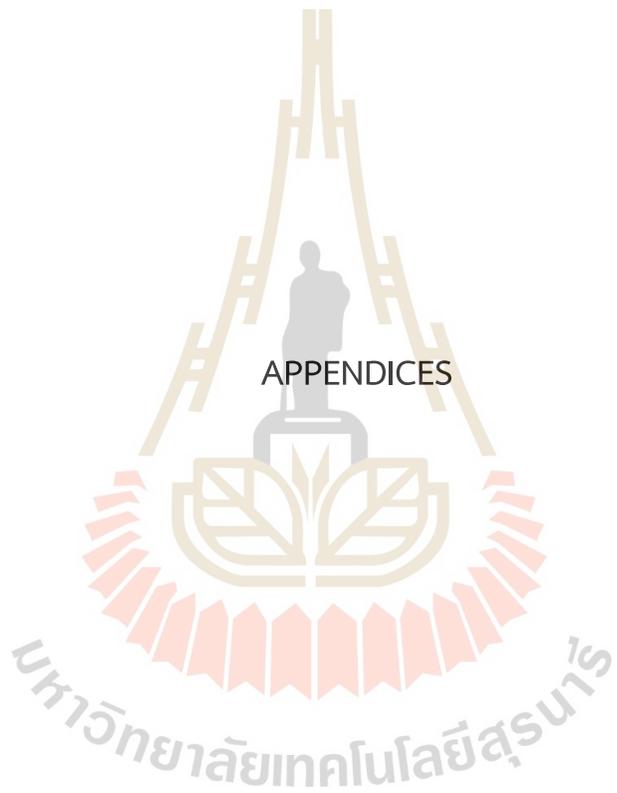
- Jiang, P., and Chen, J. (2016) Displacement prediction of landslide based on generalized regression neural networks with K-fold cross-validation. *Neurocomputing*, 198, 40-47. doi: 10.1016/j.neucom.2015.08.118
- Juang, C. S., Stanley, T. A., and Kirschbaum, D. B. (2019). Using citizen science to expand the global map of landslides: Introducing the Cooperative Open Online Landslide Repository (COOLR). *PloS one*, 14(7), doi: 10.1371/journal.pone.e0218657.
- Kamilaris, F., and Prenafeta-Boldu, F. X. (2018). Deep learning in agriculture: a survey. *Computer and Electronics in Agriculture*, 147, 70-90. doi: 10.1016/j.compag.2018.02.016
- Kingma, D., and Ba, J. (2014). Adam: A method for stochastic optimization. In 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. ArXiv: 1412.6980
- Koushik, J. (2016). Understanding Convolutional Neural Networks (Version 1). arXiv. doi:10.48550/ARXIV.1605.09081
- Lazecký, M. (2011). Monitoring of Terrain Relief Changes Using Synthetic Aperture Radar Interferometry.
- Lee, S., Hong, S., and Jung, H. (2017). A Support Vector Machine for Landslide Susceptibility Mapping in Gangwon Province, Korea. *Sustainability*, 9, 1-15. doi: 10.3390/SU9010048
- Liu, M., Cheng, D., and Yan, W. (2018). Classification of Alzheimer's Disease by Combination of Convolutional and Recurrent Neural Networks Using FDG-PET Images. *Frontiers in Neuroinformatics*, 12. doi: 10.3389/fninf.2018.00035.
- Liu, S., Segoni, S., Raspini, F., Yin, K., Zhou, C., Zhang, Y., and Casagli, N. (2020). Satellite InSAR as a New Tool for the Verification of Landslide Engineering Remedial Works at the Regional Scale: A Case Study in the Three Gorges Reservoir Area, China. *Applied Sciences*. 10(6435). doi: 10.3390/app10186435
- Lu, P., Casagli, N., Catani, F., and Tofani, V. (2012). Persistent Scatterers Interferometry Hotspot and Cluster Analysis (PSI-HCA) for detection of extremely slow-moving landslides. *International Journal of Remote Sensing*, 33(2), 466-489. doi: 10.1080/01431161.2010.536185

- Lu, P., Catani, F., Tofani, V., and Casagli, N. (2014). Quantitative hazard and risk assessment for slow-moving landslides from Persistent Scatterer Interferometry. *Landslides*, *11*, 685-696. doi: 10.1007/s10346-013-0432-2
- Lu, P., Bai, S., Tofani, V., and Casagli, N. (2019). Landslides detection through optimized hot spot analysis on persistent scatterers and distributed scatterers. *ISPRS Journal of Photogrammetry and Remote Sensing*. *156*, 147-159. doi: 10.1016/j.isprsjprs.2019.08.004
- Lucas, N. (2020, November 15). 65 families in Albay town to be relocated due to chasms created after typhoon Ulysses. *Manila Bulletin*. <https://mb.com.ph/2020/11/15/65-families-in-albay-town-to-be-relocated-due-to-chasms-created-after-typhoon-ulysses/>
- Magsino, D. (2018, December 30). Number of fatalities due to Usman in Tiwi, Albay climbs to 11. *GMA News*. <https://gmanetwork.com/news/topstories/regions/679872/number-of-fatalities-due-to-usman-in-tiwi-albay-climbs-to-11/story/>
- Marquez, C. (2020, October 31). Rolly is world's strongest so far in 2020 – Pagasa. *Philippine Daily Inquirer*. <https://newsinfo.inquirer.net/1354801/rolly-is-worlds-strongest-typhoon-so-far-in-2020-pagasa>
- Mier-Manjares, M. A. (2020, November 1). Over 87,000 families evacuated in Bicol provinces due to Typhoon Rolly. *Philippine Daily Inquirer*. <https://newsinfo.inquirer.net/1354915/over-87000-families-evacuated-in-bicol-provinces-due-to-rolly>
- Mier-Manjares, M. A. (2020, November 3). Power restoration in typhoon-hit Albay may take 2 months. *Philippine Daily Inquirer*. <https://newsinfo.inquirer.net/1355774/power-restoration-in-typhoon-hit-albay-may-take-2-months>
- Minimo, L. G., and Lagmay, A. M. F. A. (2016). 3D modeling of the Buhi debris avalanche deposit of Iriga Volcano, Philippines by integrating shallow-seismic reflection and geological data. *Journal of Volcanology and Geothermal Research*, *319*, 106-123. doi: 10.1016/j.jvolgeores.2016.03.002
- Mo, Y. J., Kim, J., Kim, J. K., Mohaisen, A., and Lee, W. (2017). Performance of deep learning computation with TensorFlow software library in GPU-capable multi-

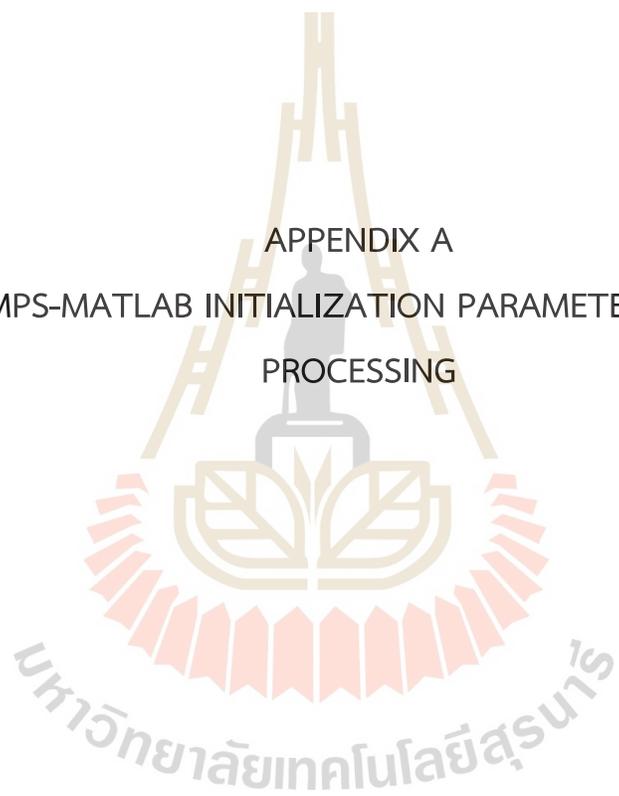
- core computing platforms. In 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN). 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN). IEEE. doi: 10.1109/icufn.2017.7993784
- National Disaster Risk Reduction and Management Council (2021). *Sitrep. No. 29 re Preparedness Measures and Effects for Typhoon “Ulysses” (I.N. VAMCO)*. https://ndrrmc.gov.ph/attachments/article/4138/SitRep_no_29_re_TY_Ulysses_a_s_of_13JAN2021.pdf
- Notti, D., Herrera, G., Bianchini, S., Meisina, C., García-Davalillo, J. C., and Zucca, F. (2014). A methodology for improving landslide PSI data analysis. *International Journal of Remote Sensing*, 35(6), 2186-2214. doi: 10.1080/01431161.2014.889864
- Orense, R. P., and Ikeda, M. (2007). Damage caused by typhoon-induced lahar flows from Mayon Volcano, Philippines. *Soils and foundations*, 47(6), 1123-1132.
- Pan, X., and Zhao, J. (2017) A central-point-enhanced convolutional neural network for high-resolution remote-sensing image classification, *International Journal of Remote Sensing*, 38(23), 6554-6581, doi: 10.1080/01431161.2017.1362131
- Park, S., Lee, C., and Lee S. (2018). Landslide Susceptibility Mapping and Comparison Using Decision Tree Models: A Case Study of Jumunjin Area, Korea. *Remote Sensing*, 10(10), 1545. doi: 10.3390/rs10101545
- Park, Y., and Lek, S. (2016). Artificial Neural Networks. In *Developments in Environmental Modelling* (pp. 123–140). Elsevier. doi: 10.1016/b978-0-444-63623-2.00007-4
- Piacentini, D., Devoto, S., Mantovani, M., Pasuto, A., Prampolini, M., and Soldati, M. (2015). Landslide susceptibility modeling assisted by Persistent Scatterers Interferometry (PSI): an example from the northwestern coast of Malta. *Natural Hazards*, 78(1), 681-697. doi: 10.1007/s11069-015-1740-8
- Ram, P., and Gupta, V. (2022) Landslide hazard, vulnerability, and risk assessment (HVRA), Mussoorie township, lesser himalaya, India. *Environ Dev Sustain* 24, 473–501 doi: 10.1007/s10668-021-01449-2
- Ramos, C. M. (2018, December 27). Tropical Depression ‘Usman’ slows down – Pagasa. *Philippine Daily Inquirer*. <https://newsinfo.inquirer.net/1066951/tropical-depression-usman-slows-down-as-it-continues-to-move-westward-pagasa>

- Shirzadi, A., Shahabi, H., Chapi, K., Bui, D. T., Pham, B. T., Shahedi, K., and Ahmad, B. B. (2017). A comparative study between popular statistical and machine learning methods for simulating volume of landslides. *Catena*, *157*, 213-226 doi: 10.1016/j.catena.2017.05.016
- Solari, L., Del Soldato, R., Montalti, R., Bianchini, S., Raspini, F., Thuegaz, P., Bertolo, D., Tofani, V., and Casagli, N. (2019). A Sentinel-1 based hot-spot analysis: landslide mapping in north-western Italy. *International Journal of Remote Sensing*, *40*(20), 7898-7921. doi: 10.1080/01431161.2019.1607612
- Starovoitov, V., and Golub, Y. (2020). Comparative study of quality estimation of binary classification. *Informatics*, *17*(1), 87–101. doi: 10.37661/1816-0301-2020-17-1-87-101
- Thabtah, F., Hammoud, S., Kamalov, F., and Gonsalves, A. (2020). Data imbalance in classification: Experimental evaluation. *Information Sciences*, *513*, 429-441. doi: 10.1016/j.ins.2019.11.004
- The Inquirer Staff (2020, November 2.) Super Typhoon Rolly batters Bicol, leaving at least 7 dead. *Philippine Daily Inquirer*. <https://newsinfo.inquirer.net/1355234/super-typhoon-rolly-batters-bicol-leaving-at-least-7-dead>
- Tien Bui, D., Tsangaratos, P., Nguyen, V. T., Liem, N. V., and Trinh, P. T. (2020). Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *Catena*, *188*, 104426. doi: 10.1016/j.catena.2019.104426
- Tofani, V., Raspini, F., Catani, F., and Casagli, N. (2013) Persistent Scatterer Interferometry (PSI) Technique for Landslide Characterization and Monitoring. *Remote Sensing*, *5*(3), 1045-1065. doi:10.3390/rs5031045
- Van Westen, C. J., Castellanos, E., and Kuriakose S. L. (2008). Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology*, *102*(3-4), 112-131. doi: 10.1016/j.enggeo.2008.03.010
- Viray, P. L. (2019, January 7). Agri, infra damage due to 'Usman' climbs to P4.2 billion. *The Philippine Star*. <https://www.philstar.com/headlines/2019/01/07/1883028/agri-infra-damage-due-usman-climbs-p42-billion>

- Wang, Y., Fang, Z., and Hong, H. (2019). Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Science of the Total Environment*, 666, 975-993. doi: 10.1016/j.scitotenv.2019.02.263
- Wang, Y., Wen, H., Sun, D., and Li, Y. (2021). Quantitative assessment of landslide risk based on susceptibility mapping using random forest and geodetector. *Remote Sensing*, 13(13), 2625. doi:10.3390/rs13132625
- Yang, A. Y. (2021, January 22). Power restored to 98% of Albay households after late-year typhoons. *BusinessWorld*. <https://www.bworldonline.com/economy/2021/01/22/340454/power-restored-to-98-of-albay-households-after-late-year-typhoons/>
- Yi, Y., Zhang, Z., Zhang, W., Jia, H., and Zhang, J. (2020). Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: A case study in Jiuzhaigou region. *Catena*, 195, 104851. doi:10.1016/j.catena.2020.104851
- Zhang, YD., Muhammad, K., and Tang, C. (2018). Twelve-layer deep convolutional neural network with stochastic pooling for tea category classification on GPU platform. *Multimedia Tools and Applications*, 77, 22821–22839. doi:10.1007/s11042-018-5765-3



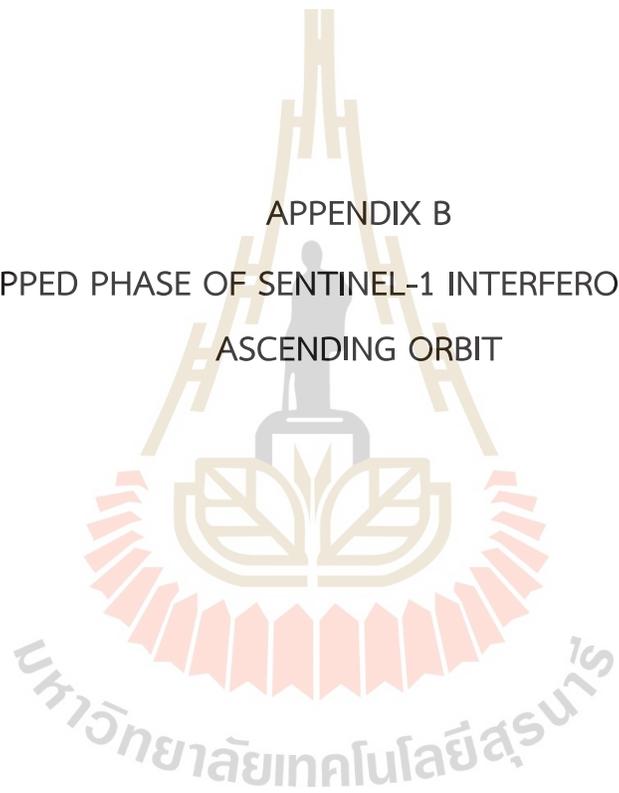
APPENDIX A
STAMPS-MATLAB INITIALIZATION PARAMETERS FOR PSI
PROCESSING



max_topo_err	30
quick_est_gamma_flag	'y'
select_reest_gamma_flag	'y'
filter_grid_size	'50'
filter_weighting	'P-square'
gamma_change_convergence	0.005
gamma_max_iterations	1
slc_osf	1
clap_win	16
clap_low_pass_wavelength	800
clap_alpha	1
clap_beta	0.3
select_method	'DENSITY'
density_rand	50
percent_rand	20
gamma_stdev_reject	0
weed_time_win	730
weed_max_noise	Inf
weed_standard_dev	1.5
weed_zero_elevation	'n'
weed_neighbours	'n'
unwrap_method	'3D'
unwrap_patch_phase	'n'
drop_ifg_index	[]
unwrap_la_error_flag	'y'
unwrap_spatial_cost_func_flag	'n'
unwrap_prefilter_flag	'y'
unwrap_grid_size	100
unwrap_gold_n_win	32
unwrap_alpha	8
unwrap_time_win	24
unwrap_gold_alpha	5
unwrap_hold_good_values	'n'

scla_drop_index	[]
scn_wavelength	100
scn_time_win	50
scn_deramp_ifg	[]
scn_kriging_flag	'n'
ref_lon	-Inf Inf
ref_lat	-Inf Inf
ref_centre_lonlat	0 0
ref_radius	Inf
ref_velocity	0
n_cores	1
plot_dem_posting	90
plot_scatterer_size	120
plot_pixels_scatterer	3
plot_color_scheme	'inflation'
shade_rel_angle	90 45
lonlat_offset	0 0
merge_resample_size	0
merge_standard_dev	Inf
scla_method	'L2'
scla_deramp	'y'
lambda	NaN
heading	NaN
subtr_tropo	'n'
tropo_method	'a_l'

APPENDIX B
WRAPPED PHASE OF SENTINEL-1 INTERFEROGRAMS ON
ASCENDING ORBIT



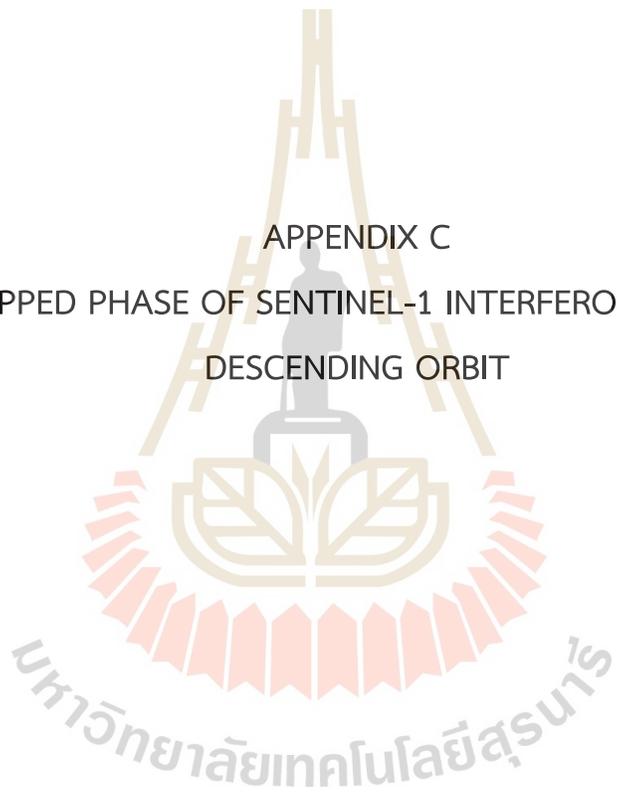








APPENDIX C
WRAPPED PHASE OF SENTINEL-1 INTERFEROGRAMS ON
DESCENDING ORBIT

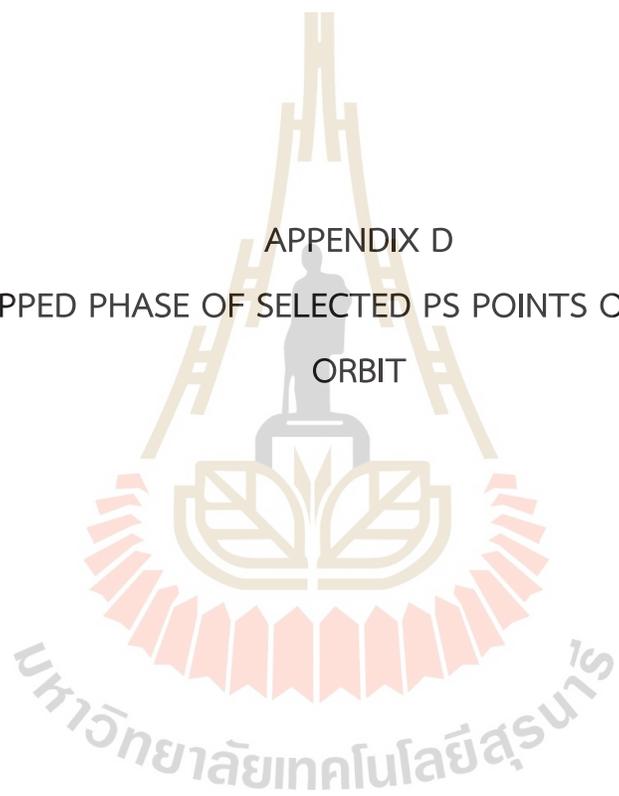




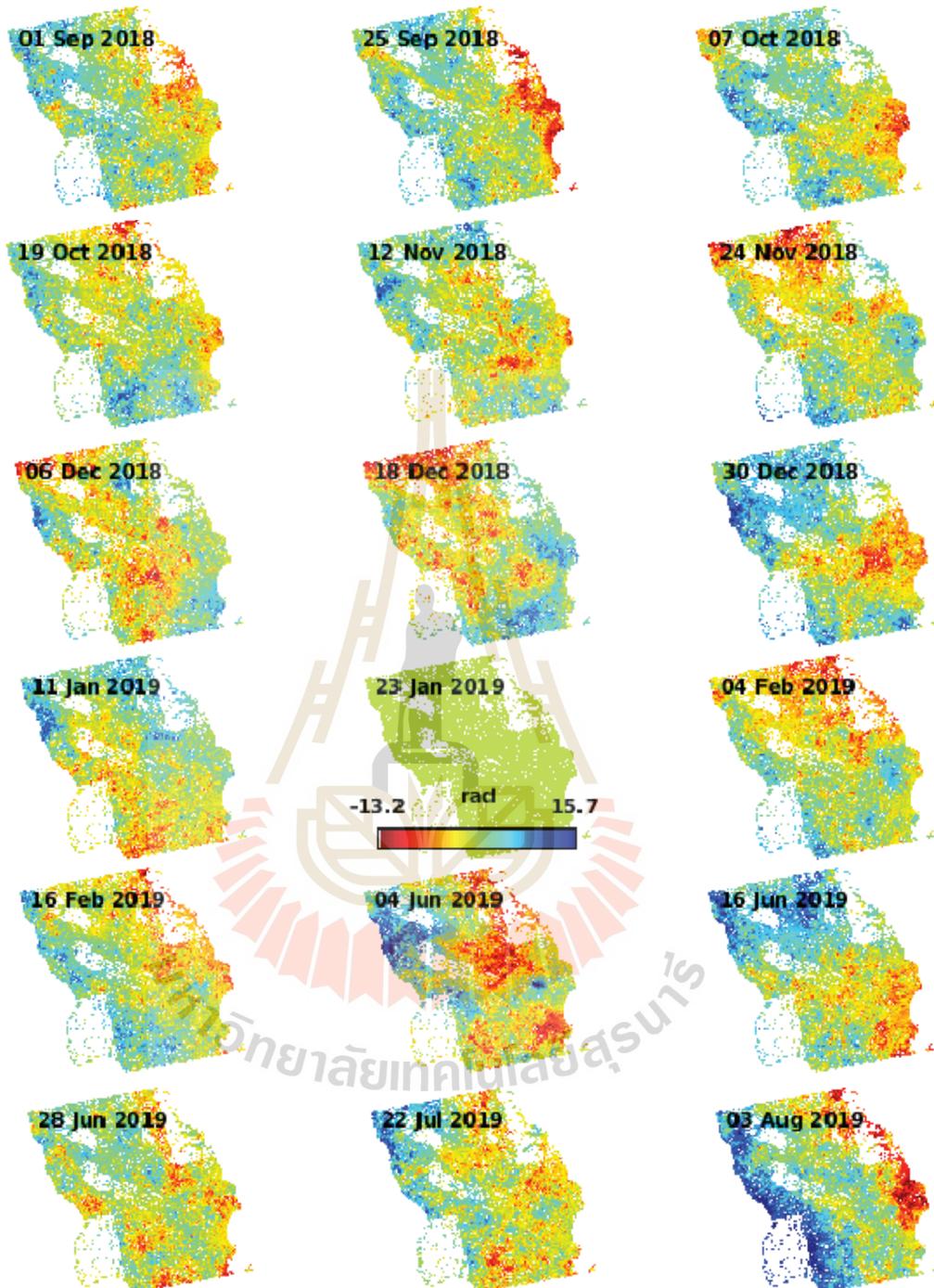


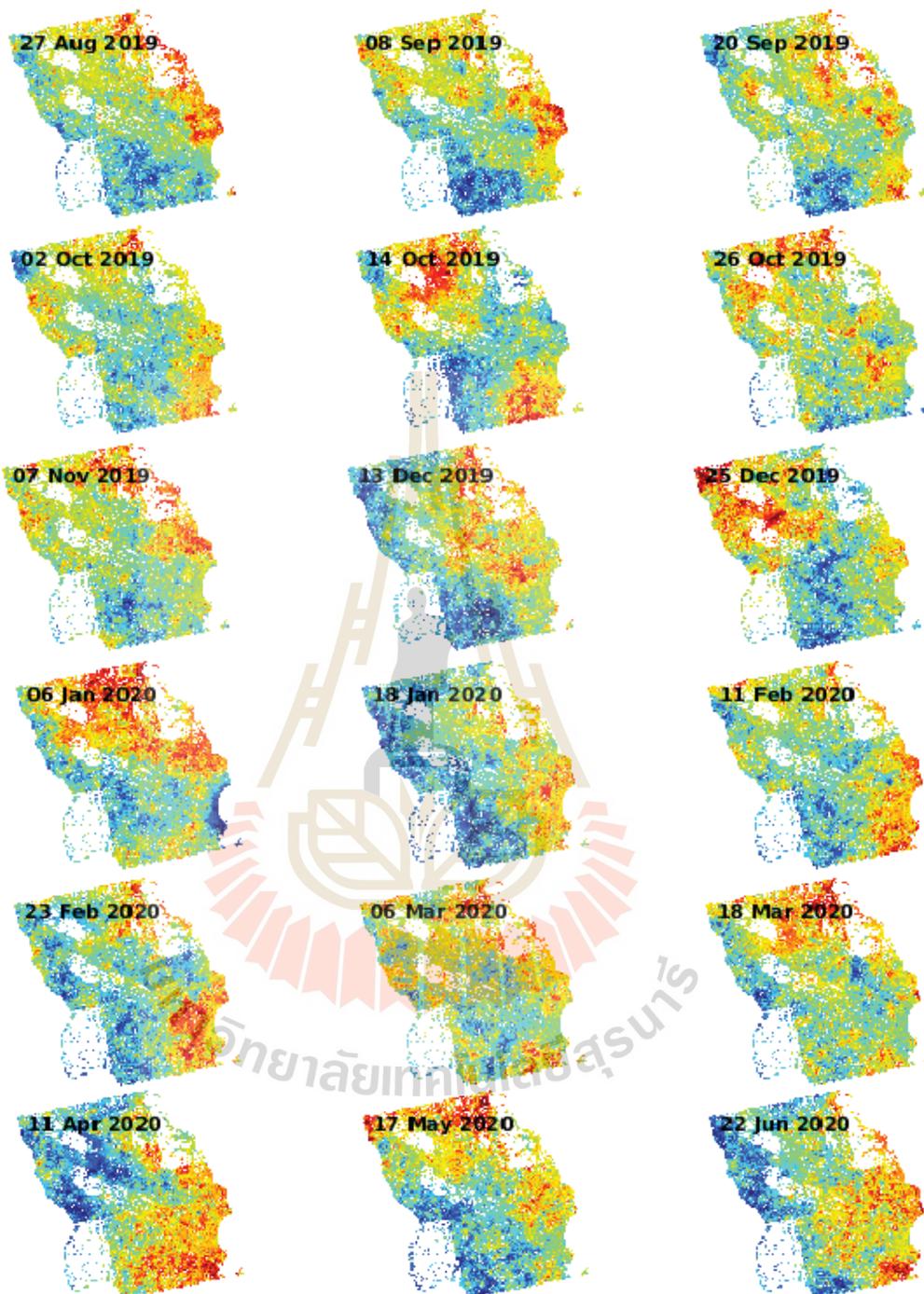


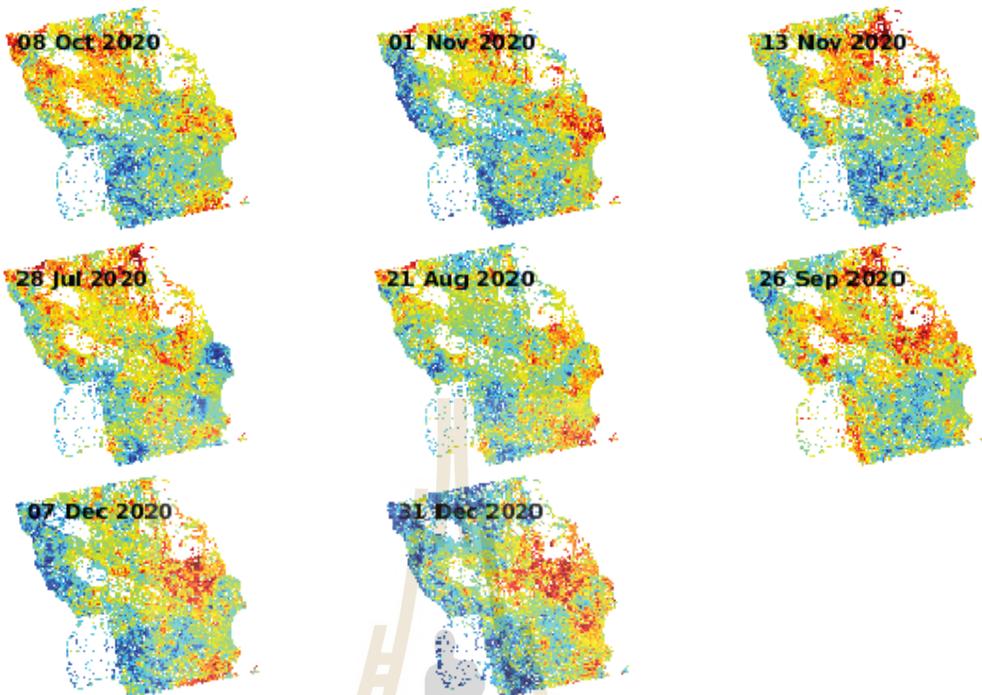
APPENDIX D
UNWRAPPED PHASE OF SELECTED PS POINTS ON ASCENDING
ORBIT



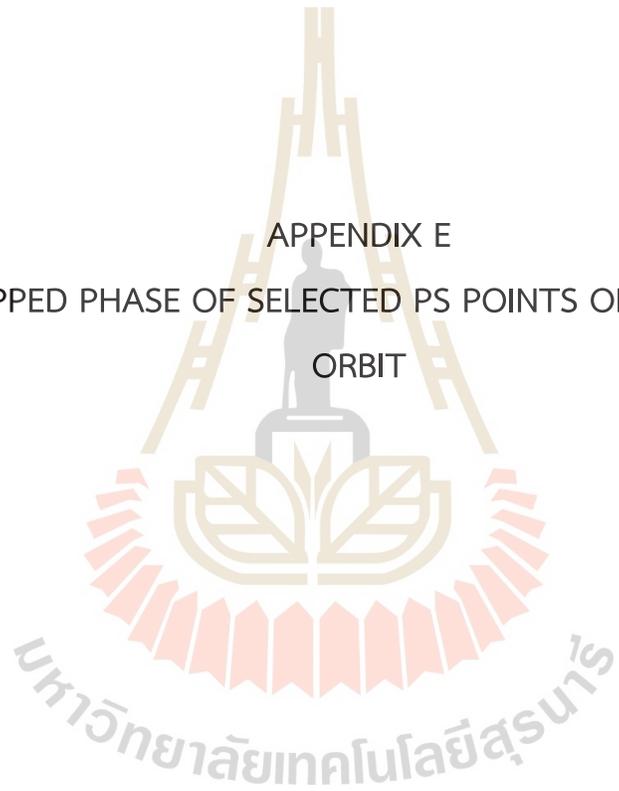




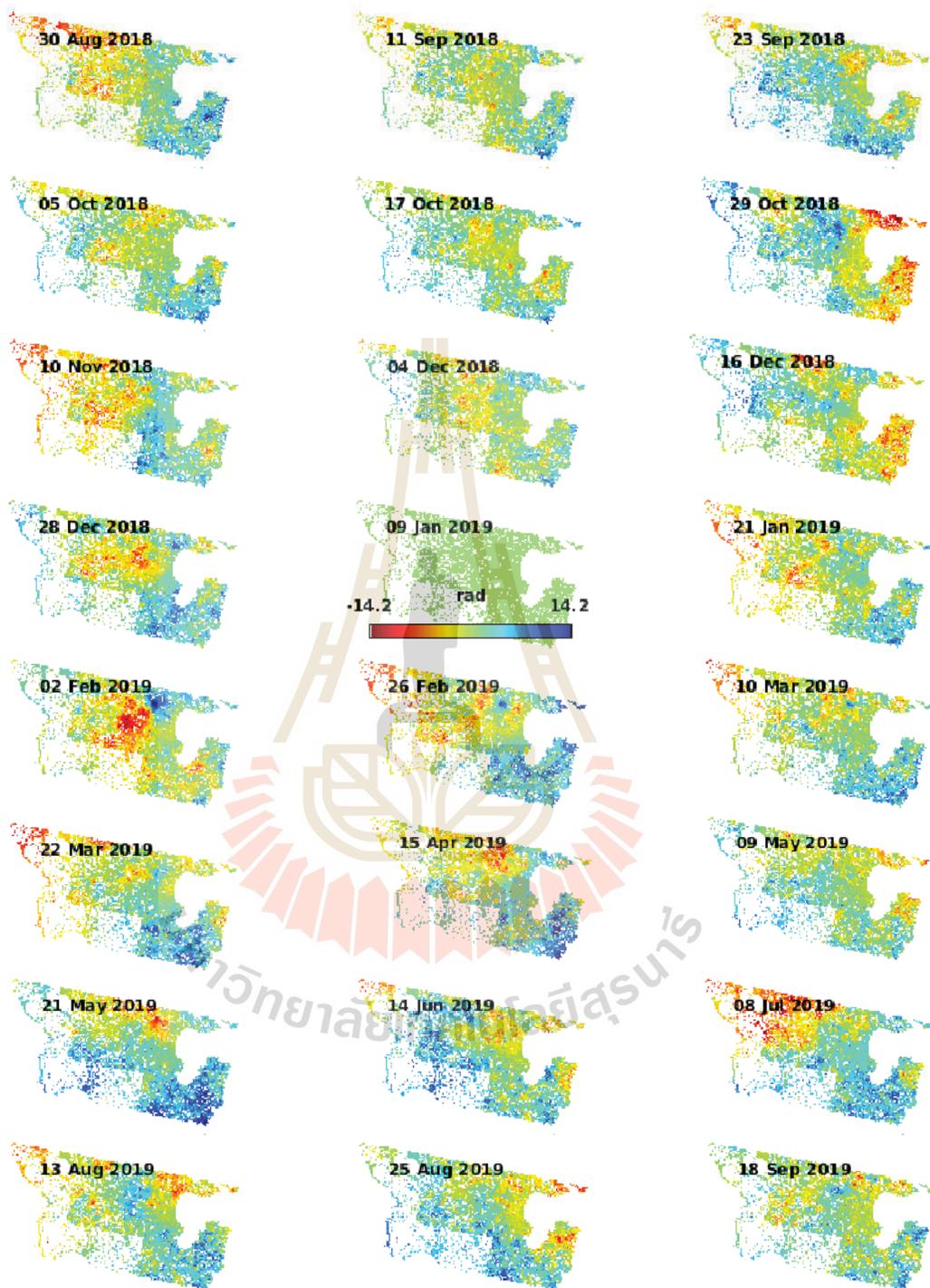




APPENDIX E
UNWRAPPED PHASE OF SELECTED PS POINTS ON DESCENDING
ORBIT

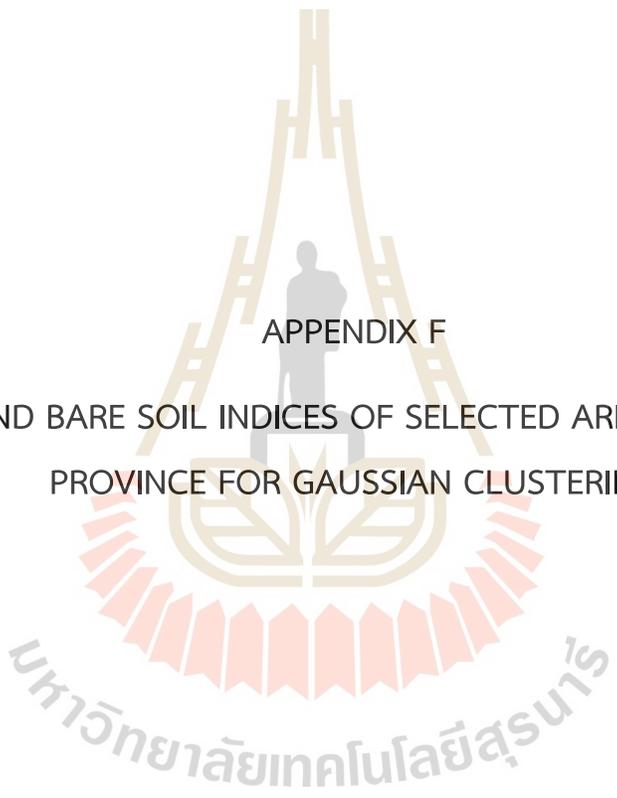








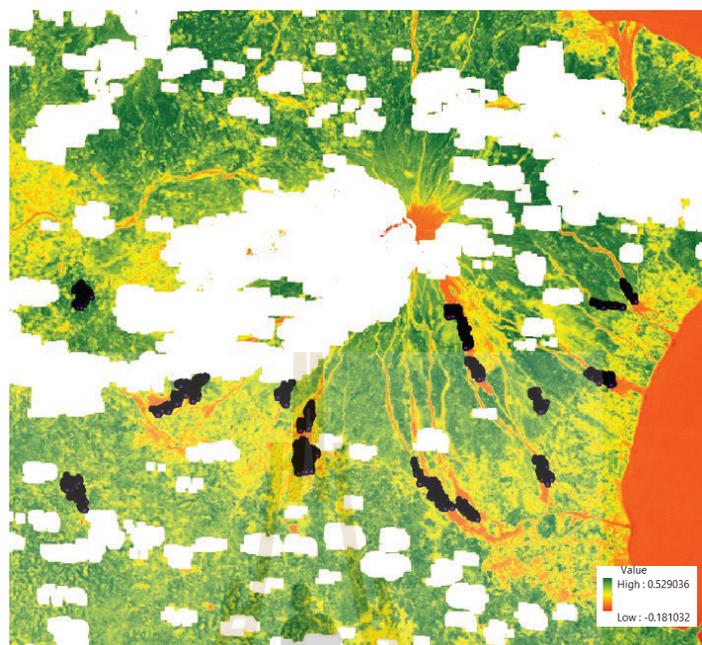
มหาวิทยาลัยเทคโนโลยีสุรนารี



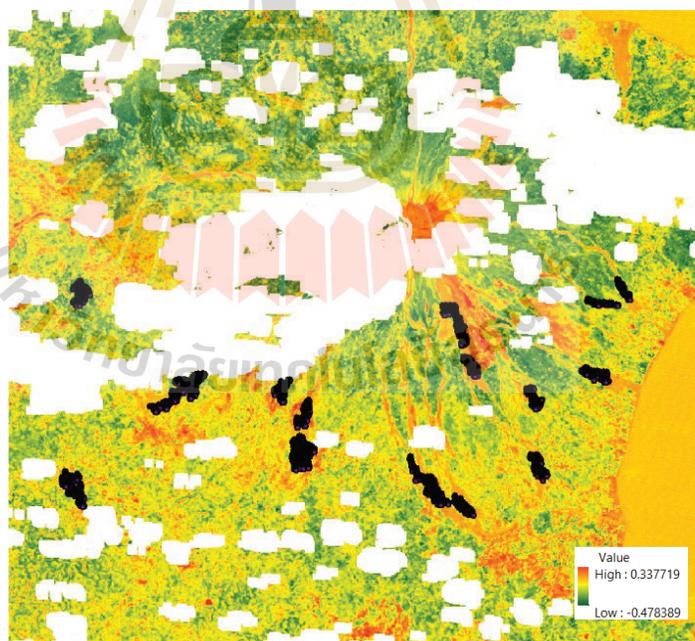
APPENDIX F

NDVI AND BARE SOIL INDICES OF SELECTED AREAS IN ALBAY
PROVINCE FOR GAUSSIAN CLUSTERING

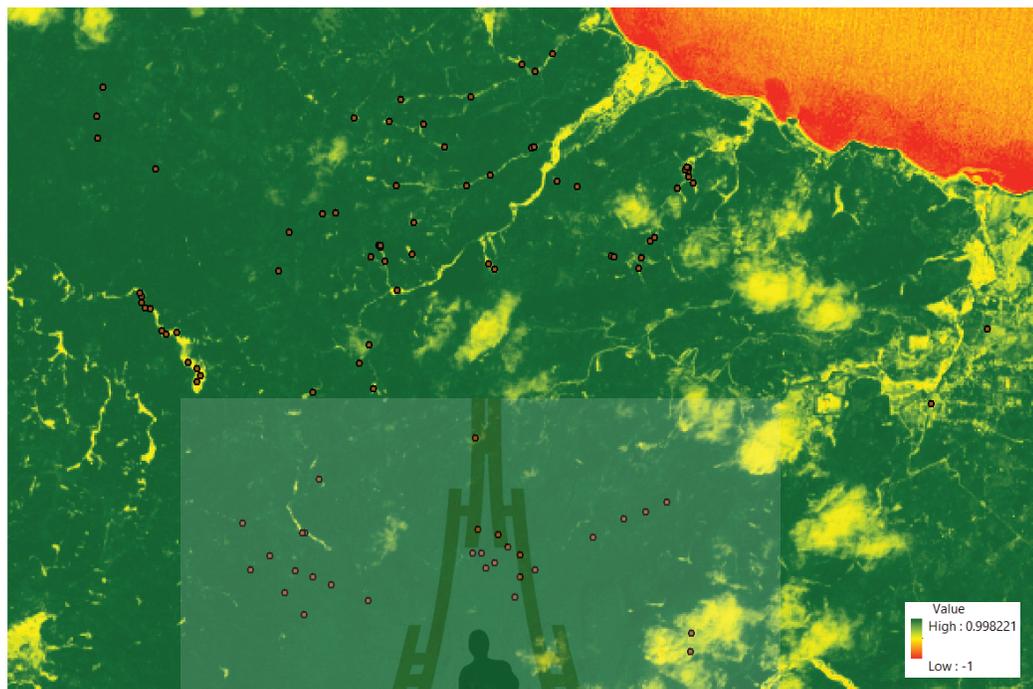
มหาวิทยาลัยเทคโนโลยีสุรนารี



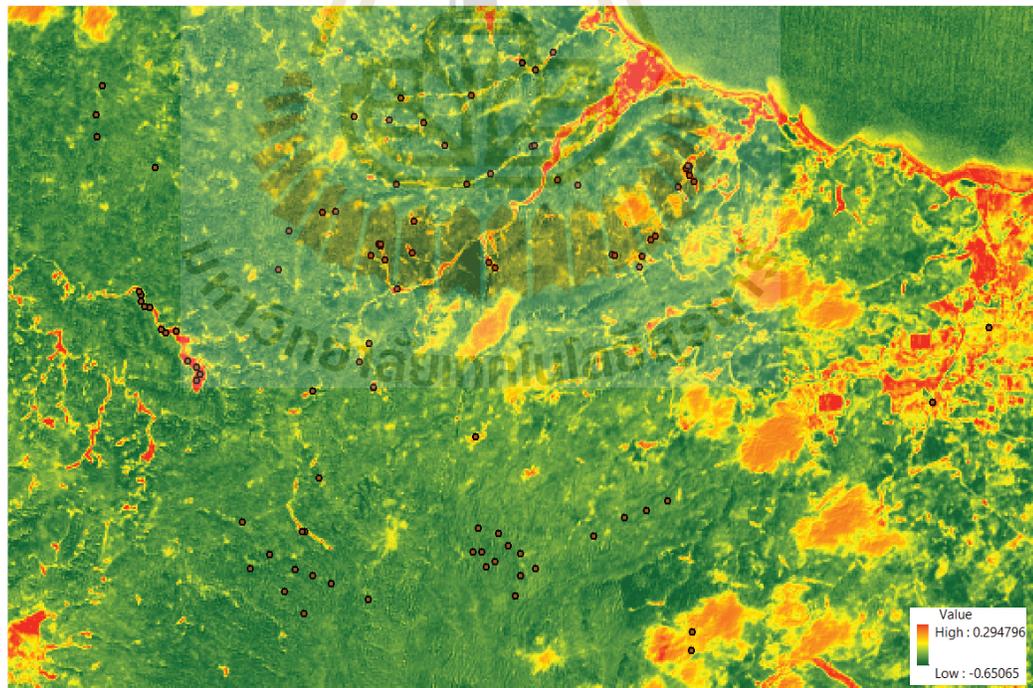
NDVI dated April 11, 2007



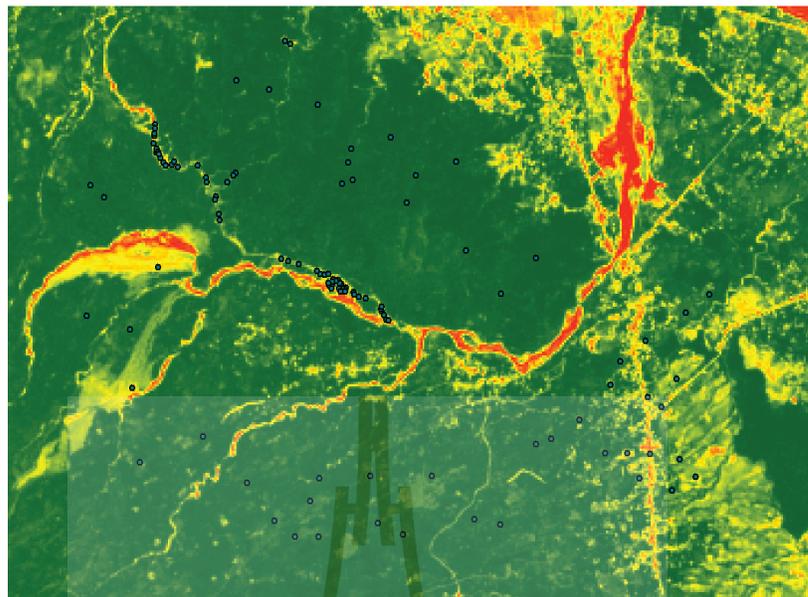
BSI dated April 11, 2007



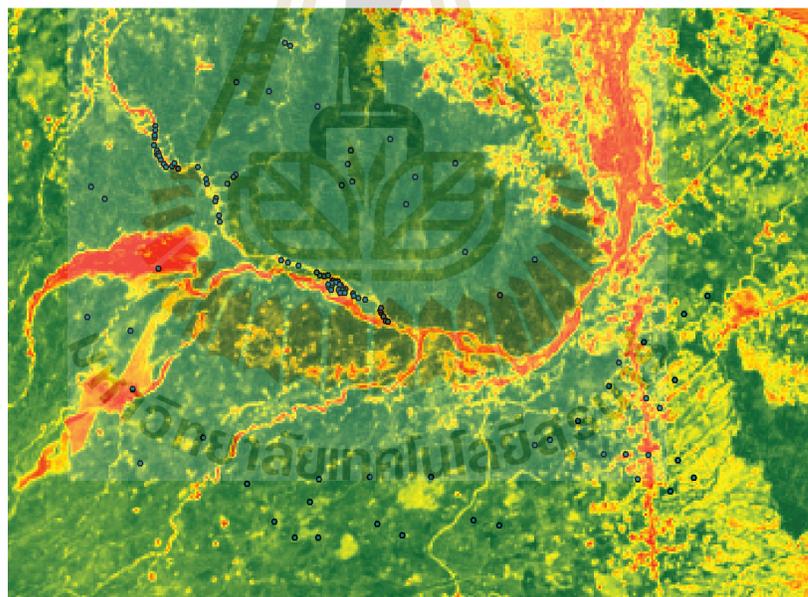
NDVI dated January 15, 2019



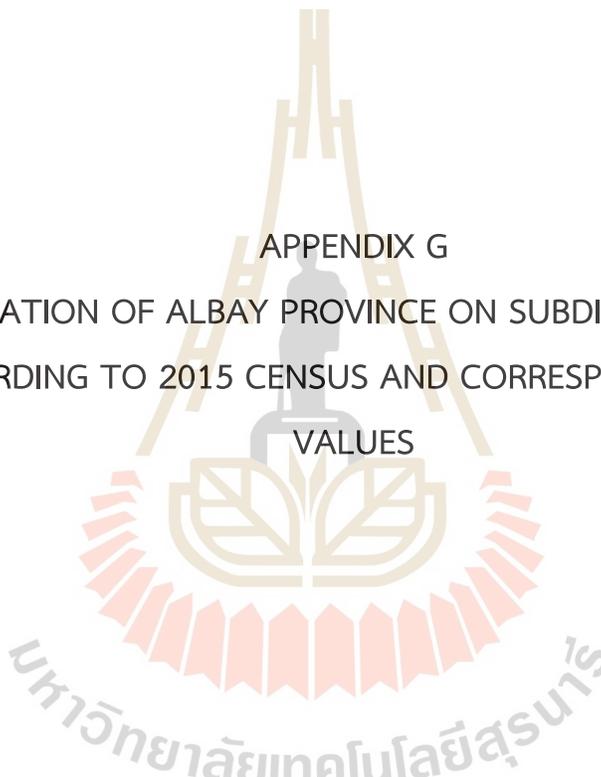
BSI dated January 15, 2019



NDVI dated May 29, 2021



BSI dated May 29, 2021



APPENDIX G
POPULATION OF ALBAY PROVINCE ON SUBDISTRICT LEVEL
ACCORDING TO 2015 CENSUS AND CORRESPONDING RISK
VALUES

Bacacay							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Baclayon	2703	1301.89	48.16%	Barangay 9	646	274.08	42.43%
Banao	1491	711.10	47.69%	Bariw	625	301.29	48.21%
Barangay 1	1308	522.90	39.98%	Basud	1746	833.80	47.75%
Barangay 10	773	408.07	52.79%	Bayandong	1650	795.57	48.22%
Barangay 11	197	87.57	44.45%	Bonga	7649	3960.57	51.78%
Barangay 12	298	84.65	28.41%	Buang	1337	660.16	49.38%
Barangay 13	1315	558.15	42.44%	Busdac	1182	622.83	52.69%
Barangay 14	1237	531.08	42.93%	Cabasan	2028	1108.39	54.65%
Barangay 2	285	125.78	44.13%	Cagbulacao	862	495.12	57.44%
Barangay 3	545	195.66	35.90%	Cagraray	703	378.65	53.86%
Barangay 4	274	91.95	33.56%	Cajogutan	1130	482.26	42.68%
Barangay 5	276	113.06	40.96%	Cawayan	1247	647.45	51.92%
Barangay 6	248	90.35	36.43%	Damacan	431	223.11	51.77%
Barangay 7	267	103.63	38.81%	Gubat Ilawod	1080	453.55	42.00%
Barangay 8	147	57.72	39.26%	Gubat Iraya	1159	581.95	50.21%
Hindi	3800	1671.31	43.98%	Pili Ilawod	1522	782.39	51.41%
Igang	2332	953.88	40.90%	Pili Iraya	997	420.36	42.16%
Langaton	765	406.15	53.09%	Pongco	1022	535.38	52.39%
Manaet	836	484.66	57.97%	San Pablo	1274	675.78	53.04%
Mapulang Daga	453	241.55	53.32%	San Pedro	1605	652.35	40.64%
Mataas	518	266.45	51.44%	Sogod	4552	1939.11	42.60%
Misibis	1007	529.00	52.53%	Sula	873	460.56	52.76%
Nahapunan	402	205.77	51.19%	Tambilagao	906	498.57	55.03%
Namanday	1482	766.75	51.74%	Tambongon	705	353.94	50.20%
Namantao	778	432.50	55.59%	Tanagan	1388	754.55	54.36%
Napao	1883	1016.20	53.97%	Pigcobohan	817	436.43	53.42%
Panarayon	1848	749.12	40.54%				

Camalig							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Anoling	964	474.73	49.25%	Libod	3233	1061.61	32.84%
Batigang	3389	937.45	27.66%	Ligban	714	314.99	44.12%
Bantonan	586	215.92	36.85%	Mabunga	137	53.85	39.31%
Barangay 1	1086	344.54	31.73%	Magogon	566	205.48	36.30%
Barangay 2	407	132.32	32.51%	Manawan	879	267.28	30.41%
Barangay 3	534	151.79	28.43%	Maninila	1085	354.12	32.64%

Barangay 4	316	97.12	30.74%	Mina	651	291.31	44.75%
Barangay 5	361	96.62	26.76%	Miti	926	329.09	35.54%
Barangay 6	385	106.65	27.70%	Palanog	3148	903.05	28.69%
Barangay 7	401	111.41	27.78%	Panoypoy	1304	481.48	36.92%
Bariw	2047	656.15	32.05%	Pariaan	863	322.41	37.36%
Binanderahan	583	241.99	41.51%	Quinartilan	924	300.67	32.54%
Binitayan	590	211.29	35.81%	Quirangay	2721	1267.05	46.57%
Bongabong	865	289.58	33.48%	Quitinday	259	107.64	41.56%
Cabañan	2820	1200.01	42.55%	Salugan	1750	736.01	42.06%
Cabranan Pequeño	852	329.82	38.71%	Solong	433	153.76	35.51%
Caguiba	1566	576.79	36.83%	Sua	1377	655.17	47.58%
Calabidongan	717	253.52	35.36%	Sumlang	2025	704.83	34.81%
Comun	1704	524.27	30.77%	Tagaytay	5652	2021.28	35.76%
Cotmon	2750	803.37	29.21%	Tagoytoy	620	250.28	40.37%
Del Rosario	842	260.17	30.90%	Taladong	1547	447.55	28.93%
Gapo	1665	608.66	36.56%	Taloto	998	349.09	34.98%
Gotob	635	277.74	43.74%	Taplacon	1181	442.37	37.46%
Ilawod	2807	873.32	31.11%	Tinago	2072	759.48	36.65%
Iluluan	1631	556.83	34.14%	Tumpa	1336	623.89	46.70%

Daraga							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Alcala	2935	729.20	24.84%	Kiwalo	1387	299.15	21.57%
Alobo	841	205.82	24.47%	Lacag	2533	647.19	25.55%
Anistag	11461	3117.45	27.20%	Mabini	673	152.86	22.71%
Bagumbayan	3416	784.50	22.97%	Malabog	4243	1023.88	24.13%
Balinad	2606	570.69	21.90%	Malobago	663	149.18	22.50%
Bañadero	1643	454.36	27.65%	Maopi	1197	296.99	24.81%
Bañag	2800	433.48	15.48%	Market Area Poblacion	2681	264.65	9.87%
Bascaran	4335	1111.66	25.64%	Maroroy	5518	957.06	17.34%
Bigao	1073	311.86	29.06%	Matnog	1772	552.45	31.18%
Binitayan	4447	655.82	14.75%	Mayon	1794	530.78	29.59%
Bongalon	1148	264.33	23.03%	Mi-Isi	1108	393.10	35.48%
Budio	288	94.74	32.90%	Nabasan	663	207.16	31.25%
Burgos	1227	300.41	24.48%	Namantao	1608	425.02	26.43%
Busay	2065	437.80	21.20%	Pandan	1859	419.76	22.58%
Canaron	513	155.23	30.26%	Peñafancia	3041	555.97	18.28%
Cullat	1492	296.15	19.85%	Sagpon	6587	803.97	12.21%
Dela Paz	909	198.74	21.86%	Salvacion	3648	941.05	25.80%
Dinoronan	691	185.26	26.81%	San Rafael	288	85.20	29.58%
Gabawan	2151	470.29	21.86%	San Ramon	1481	420.04	28.36%
Gapo	2062	439.62	21.32%	San Roque	4500	1110.72	24.68%

Ibaugan	621	207.02	33.34%	San Vicente Grande	1088	337.21	30.99%
Ilawod Area Poblacion	2243	206.69	9.21%	San Vicente Pequeño	336	91.21	27.15%
Inarado	1804	390.17	21.63%	Sipi	3750	521.92	13.92%
Kidaco	1373	333.51	24.29%	Tabon-Tabon	3124	668.01	21.38%
Kilicao	4328	900.51	20.81%	Tagas	6901	1404.95	20.36%
Kimantong	2144	480.74	22.42%	Talahib	643	183.74	28.58%
Kinawitan	451	110.63	24.53%	Villahermosa	2442	673.66	27.59%

Guinobatan							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Agpay	482	175.71	36.45%	Matobago	1867	821.13	43.98%
Balite	585	257.90	44.09%	Maninila	1482	705.85	47.63%
Banao	1074	420.18	39.12%	Mapaco	1460	514.71	35.25%
Batbat	1373	607.98	44.28%	Marcial O. Rañola	537	249.89	46.53%
Binogsacan Lower	1889	649.21	34.37%	Masarawag	3760	2184.86	58.11%
Binogsacan Upper	1131	364.45	32.22%	Mauraro	5980	2047.53	34.24%
Botolo	1442	675.67	46.86%	Minto	1726	737.81	42.75%
Bubulusan	1399	616.77	44.09%	Morera	2602	748.26	28.76%
Calzada	2787	858.07	30.79%	Muladbucad Grande	2079	1044.83	50.26%
Catomag	796	308.50	38.76%	Muladbucad Pequeño	2051	1136.34	55.40%
Doña Mercedes	1411	568.45	40.29%	Ongo	942	426.51	45.28%
Doña Tomasa	1352	747.17	55.26%	Palanas	427	203.94	47.76%
Ilawod	2209	659.81	29.87%	Poblacion	427	114.48	26.81%
Inamnan Grande	2177	759.76	34.90%	Pood	1336	629.66	47.13%
Inamnan Pequeño	1756	634.27	36.12%	Quibongbongan	2595	834.86	32.17%
Inascan	1481	774.73	52.31%	Quitago	2236	893.99	39.98%
Iraya	1771	507.31	28.65%	San Francisco	3302	1201.58	36.39%
Lomacao	2096	719.95	34.35%	San Jose	853	352.20	41.29%
Maguiron	2118	878.89	41.50%	San Rafael	3884	1459.87	37.59%
Maipon	3944	1689.06	42.83%	Sinungtan	1812	770.28	42.51%
Malabnig	1145	399.38	34.88%	Tandarora	1291	605.68	46.92%
Malipo	1760	707.56	40.20%	Travesia	3878	1312.82	33.85%

Jovellar							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Aurora Poblacion	96	57.78	60.18%	Mercado Poblacion	221	111.71	50.55%
Bagacay	985	574.55	58.33%	Plaza Poblacion	168	90.48	53.86%
Bautista	920	576.57	62.67%	Quitinday Poblacion	254	183.88	72.39%
Cabranan	1636	874.41	53.45%	Rizal Poblacion	360	203.90	56.64%
Calzada Poblacion	523	259.68	49.65%	Salvacion	400	245.19	61.30%
Del Rosario	679	425.33	62.64%	San Isidro	1020	643.86	63.12%
Estrella	513	280.87	54.75%	San Roque	1470	936.57	63.71%
Florista	790	475.95	60.25%	San Vicente	1976	1090.72	55.20%
Mabini Poblacion	636	405.39	63.74%	Sinagaran	570	323.40	56.74%
Magsaysay Pob	126	76.07	60.37%	Villa Paz	872	526.88	60.42%
Mamlad	1044	578.27	55.39%	White Deer Poblacion	752	482.89	64.21%
Maogog	1297	756.70	58.34%				

Legazpi City							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Bgy. 1 - Em's Barrio	3725	181.40	4.87%	Bgy. 42 - Rawis	8868	362.98	4.09%
Bgy. 10 - Cabugao	547	29.61	5.41%	Bgy. 43 - Tamaoyan	1642	74.13	4.51%
Bgy. 11 - Maoyod Poblacion	1110	58.24	5.25%	Bgy. 44 - Pawa	3469	181.63	5.24%
Bgy. 12 - Tula-Tula	2586	128.47	4.97%	Bgy. 45 - Dita	1791	85.52	4.77%
Bgy. 13 - Ilawod West Poblacion	721	39.77	5.52%	Bgy. 46 - San Joaquin	2260	100.73	4.46%
Bgy. 14 - Ilawod Poblacion	854	40.64	4.76%	Bgy. 47 - Arimbay	3753	184.47	4.92%

Bgy. 15 - Ilawod East Poblacion	2879	149.10	5.18%	Bgy. 48 - Bagong Abre	1627	85.16	5.23%
Bgy. 16 - Kawit-East Washington Dr	5372	289.07	5.38%	Bgy. 49 - Bigaa	6730	336.29	5.00%
Bgy. 17 - Rizal Sreet., Ilawod	2048	100.75	4.92%	Bgy. 5 - Sagmin Poblacion	1375	64.83	4.72%
Bgy. 18 - Cabagñan West	3856	203.24	5.27%	Bgy. 50 - Padang	1653	127.51	7.71%
Bgy. 19 - Cabagñan	1427	93.13	6.53%	Bgy. 51 - Buyuan	3895	233.80	6.00%
Bgy. 2 - Em's Barrio South	1820	109.67	6.03%	Bgy. 52 - Matanag	1895	130.29	6.88%
Bgy. 20 - Cabagñan East	641	42.69	6.66%	Bgy. 53 - Bonga	3503	253.11	7.23%
Bgy. 21 - Binanuahan West	917	53.04	5.78%	Bgy. 54 - Mabinit	1640	119.54	7.29%
Bgy. 22 - Binanuahan East	1619	68.76	4.25%	Bgy. 55 - Estanza	4237	336.87	7.95%
Bgy. 23 - Imperial Court Subd.	746	38.45	5.15%	Bgy. 56 - Taysan	4237	278.02	6.56%
Bgy. 24 - Rizal Street	641	39.44	6.15%	Bgy. 57 - Dapdap	2287	104.46	4.57%
Bgy. 25 - Lapu-Lapu	1398	74.27	5.31%	Bgy. 58 - Buragwis	4549	351.64	7.73%
Bgy. 26 - Dinagaan	798	43.00	5.39%	Bgy. 59 - Puro	4756	255.16	5.37%
Bgy. 27 - Victory Village South	1418	84.23	5.94%	Bgy. 6 - Bañadero Poblacion	1390	83.88	6.03%
Bgy. 28 - Victory Village North	2399	104.07	4.34%	Bgy. 60 - Lamba	1352	69.48	5.14%
Bgy. 29 - Sabang	1656	65.15	3.93%	Bgy. 61 - Maslog	4796	345.33	7.20%
Bgy. 3 - Em's Barrio East	900	37.45	4.16%	Bgy. 62 - Homapon	4832	337.06	6.98%
Bgy. 30 - Pigcale	1688	77.78	4.61%	Bgy. 63 - Mariawa	1664	103.94	6.25%
Bgy. 31 - Centro-Baybay	1415	86.57	6.12%	Bgy. 64 - Bagacay	1616	100.11	6.20%
Bgy. 33 - Pnr-Peñaranda St.-Iraya	2773	68.81	2.48%	Bgy. 65 - Imalnod	2146	133.59	6.23%
Bgy. 34 - Oro Site-Magallanes St.	1633	46.57	2.85%	Bgy. 66 - Banquerohan	6976	479.24	6.87%

Bgy. 35 - Tinago	375	16.74	4.46%	Bgy. 67 - Bariis	1812	144.01	7.95%
Bgy. 36 - Kapantawan	644	30.75	4.78%	Bgy. 68 - San Francisco	2479	177.50	7.16%
Bgy. 37 - Bitano	8559	348.61	4.07%	Bgy. 69 - Buenavista	1319	128.75	9.76%
Bgy. 38 - Gogon	5752	212.08	3.69%	Bgy. 7 - Baño	645	54.86	8.51%
Bgy. 39 - Bonot	3521	105.11	2.99%	Bgy. 70 - Cagbacong	2776	224.28	8.08%
Bgy. 4 - Sagpon Poblacion	953	66.57	6.99%	Bgy. 8 - Bagumbayan	3400	98.20	2.89%
Bgy. 40 - Cruzada	5853	191.58	3.27%	Bgy. 9 - Pinaric	1493	92.24	6.18%
Bgy. 41 - Bogtong	4753	244.49	5.14%	Bgy. 32 - San Roque	5632	322.56	5.73%

Libon							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Alongong	1529	709.30	46.39%	Rawis	2250	871.54	38.74%
Apud	1938	679.06	35.04%	Sagrada Familia	1666	737.72	44.28%
Bacolod	1450	451.48	31.14%	Salvacion	518	254.92	49.21%
Bariw	1297	560.94	43.25%	Sampongan	458	238.12	51.99%
Bonbon	3820	1193.55	31.24%	San Agustin	2517	1043.64	41.46%
Buga	3748	1521.00	40.58%	San Antonio	1519	749.42	49.34%
Bulusan	2169	938.96	43.29%	San Isidro	1883	611.68	32.48%
Burabod	2057	922.34	44.84%	San Jose	2982	1470.23	49.30%
Caguscoc	1051	570.51	54.28%	San Pascual	1229	575.33	46.81%
East Carisac	947	310.90	32.83%	San Ramon	764	397.17	51.99%
Harigue	876	438.16	50.02%	San Vicente	2989	1175.99	39.34%
Libtong	886	431.31	48.68%	Santa Cruz	1514	917.86	60.62%
Linao	1544	553.44	35.84%	Talin-Talin	1555	766.17	49.27%
Mabayawas	404	201.37	49.84%	Tambo	862	475.14	55.12%
Macabuños	2840	1598.24	56.28%	Villa Petrona	1497	560.70	37.46%
Magallang	512	211.75	41.36%	West Carisac	1053	313.09	29.73%
Malabiga	857	419.54	48.95%	Zone I	1486	523.02	35.20%
Marayag	966	321.56	33.29%	Zone II	1249	395.82	31.69%
Matara	894	421.12	47.11%	Zone III	870	229.86	26.42%
Molosbolos	1084	538.91	49.71%	Zone IV	1769	710.94	40.19%
Natasan	1033	473.16	45.80%	Zone V	1068	302.54	28.33%
Niño Jesus	342	147.60	43.16%	Zone VI	531	287.06	54.06%
Nogpo	2281	1033.09	45.29%	Zone VII	1635	488.75	29.89%
Pantao	6930	3622.18	52.27%				

Ligao City							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Abella	1047	256.45	24.49%	Mahaba	3560	745.26	20.93%
Allang	1472	416.71	28.31%	Malama	1864	512.28	27.48%
Amtic	1930	584.61	30.29%	Maonon	3042	849.84	27.94%
Bacong	1874	517.70	27.63%	Nabonton	2203	702.65	31.89%
Bagumbayan	2453	436.34	17.79%	Nasisi	1879	504.33	26.84%
Balanac	742	224.67	30.28%	Oma-Oma	1761	565.26	32.10%
Baligang	1463	526.87	36.01%	Palapas	1469	506.49	34.48%
Barayong	2438	536.15	21.99%	Pandan	4519	911.43	20.17%
Basag	2194	591.95	26.98%	Paulba	2464	778.62	31.60%
Batang	3672	757.69	20.63%	Paulog	2424	664.90	27.43%
Bay	1383	299.41	21.65%	Pinamaniquian	2088	515.47	24.69%
Binanowan	1208	310.98	25.74%	Pinit	2714	532.58	19.62%
Binatagan	3191	601.44	18.85%	Ranao-Ranao	1773	320.44	18.07%
Bobonsuran	659	133.77	20.30%	San Vicente	787	200.71	25.50%
Bonga	3196	860.71	26.93%	Santa Cruz	2192	346.70	15.82%
Busac	1229	390.65	31.79%	Tagpo	987	205.75	20.85%
Busay	1386	498.39	35.96%	Tambo	1525	515.72	33.82%
Cabarian	2569	675.48	26.29%	Tandarora	452	118.67	26.26%
Calzada	1717	298.53	17.39%	Tastas	2272	429.35	18.90%
Catburawan	1747	459.07	26.28%	Tinago	4800	955.16	19.90%
Cavasi	929	187.74	20.21%	Tinampo	2644	486.94	18.42%
Culliat	806	169.79	21.07%	Tiongson	944	261.90	27.74%
Dunao	760	148.42	19.53%	Tomolin	1458	286.28	19.64%
Francia	812	221.21	27.24%	Tuburan	5682	1091.80	19.22%
Guilid	3848	735.42	19.11%	Tula-Tula Grande	1825	449.91	24.65%
Herrera	2655	853.05	32.13%	Tula-Tula Pequeno	1224	314.68	25.71%
Layon	3152	705.44	22.38%	Tupas	1026	285.40	27.82%
Macalidong	1289	313.21	24.30%				

Malilipot							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Barangay I	2152	944.23	43.88%	San Antonio Santicon	791	415.64	52.55%
Barangay II	597	245.66	41.15%	San Antonio Sulong	794	455.35	57.35%

Barangay III	1296	542.88	41.89%	San Francisco	1565	658.03	42.05%
Barangay IV	1614	566.73	35.11%	San Isidro Ilawod	3705	1462.25	39.47%
Barangay V	1264	533.00	42.17%	San Isidro Iraya	4686	1987.28	42.41%
Binitayan	1834	804.29	43.85%	San Jose	5110	2264.57	44.32%
Calbayog	2094	1197.93	57.21%	San Roque	2277	1236.44	54.30%
Canaway	2599	1355.78	52.17%	Santa Cruz	1659	756.10	45.58%
Salvacion	1176	615.19	52.31%	Santa Teresa	2573	1401.49	54.47%

Malinao							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Awang	347	157.41	45.36%	Libod	1065	485.49	45.59%
Bagatangki	772	523.79	67.85%	Malolos	856	445.17	52.01%
Bagumbayan	788	362.35	45.98%	Matalipni	1939	1213.84	62.60%
Balading	3104	1484.63	47.83%	Ogob	1992	1048.33	52.63%
Balza	2084	941.34	45.17%	Pawa	1217	591.32	48.59%
Bariw	1524	676.18	44.37%	Payahan	1414	645.45	45.65%
Baybay	1730	638.17	36.89%	Poblacion	657	325.05	49.47%
Bulang	1028	680.38	66.18%	Quinarabasahan	821	523.99	63.82%
Burabod	837	393.22	46.98%	Santa Elena	1133	605.76	53.47%
Cabunturan	577	277.37	48.07%	Soa	1352	799.62	59.14%
Comun	1772	748.80	42.26%	Sugcad	1784	1122.37	62.91%
Diario	889	436.48	49.10%	Tagoytoy	2323	1123.12	48.35%
Estancia	4293	1872.94	43.63%	Tanawan	1739	985.23	56.65%
Jonop	2024	903.05	44.62%	Tuliw	1761	826.02	46.91%
Labnig	3479	1536.15	44.15%				

Manito							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Balabagon	507	261.93	51.66%	Holugan	1199	578.99	48.29%
Balasbas	1581	760.36	48.09%	It-Ba	3588	1477.97	41.19%
Bamban	1010	621.09	61.49%	Malobago	958	589.85	61.57%
Buyo	4038	1528.86	37.86%	Manumbalay	998	527.31	52.84%
Cabacongan	945	605.80	64.11%	Nagotgot	2510	1508.70	60.11%
Cabit	1093	613.47	56.13%	Pawa	2297	1075.06	46.80%
Cawayan	1805	1177.48	65.23%	Tinapian	1165	610.93	52.44%
Cawit	1013	419.78	41.44%				

Oas							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Badbad	602	264.66	43.96%	Maramba	3301	1404.46	42.55%
Badian	2681	1266.11	47.23%	Matambo	489	151.02	30.88%
Bagsa	862	248.80	28.86%	Mayag	452	196.36	43.44%
Bagumbayan	879	215.61	24.53%	Mayao	1618	321.31	19.86%
Balogo	4236	2505.14	59.14%	Moroponros	357	166.18	46.55%
Banao	470	219.42	46.68%	Nagas	1631	767.59	47.06%
Bangiwon	910	434.49	47.75%	Obaliw-Rinas	1504	472.05	31.39%
Bogtong	367	172.69	47.06%	Pistola	1135	372.46	32.82%
Bongoran	1816	602.87	33.20%	Ramay	734	316.98	43.19%
Busac	2079	633.24	30.46%	Rizal	864	303.25	35.10%
Cadawag	485	231.91	47.82%	Saban	1908	525.83	27.56%
Cagmanaba	3305	1645.31	49.78%	San Agustin	1285	386.68	30.09%
Calaguimit	867	422.85	48.77%	San Antonio	939	418.25	44.54%
Calpi	1335	438.47	32.84%	San Isidro	1529	491.60	32.15%
Calzada	911	289.30	31.76%	San Jose	744	342.83	46.08%
Camagong	1167	411.96	35.30%	San Juan	1503	468.44	31.17%
Casinagan	617	273.61	44.35%	San Miguel	772	330.97	42.87%
Centro Poblacion	225	44.52	19.79%	San Pascual	841	324.10	38.54%
Coliat	723	345.51	47.79%	San Ramon	1710	491.90	28.77%
Del Rosario	733	302.99	41.34%	San Vicente	938	299.18	31.90%
Gumabao	793	223.72	28.21%	Tablon	940	409.03	43.51%
Ilaor Norte	1495	443.94	29.69%	Talisay	779	326.53	41.92%
Ilaor Sur	2202	634.35	28.81%	Talongog	1192	396.40	33.25%
Iraya Norte	1982	668.88	33.75%	Tapet	1813	905.06	49.92%
Iraya Sur	774	286.64	37.03%	Tobgon	1648	754.53	45.78%
Manga	2322	652.41	28.10%	Tobog	1312	391.36	29.83%
Maporong	1184	385.11	32.53%				

Pio Duran							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Agol	1135	479.20	42.22%	Lawinon	1280	584.14	45.64%
Alabangpuro	835	381.04	45.63%	Macasitas	508	216.51	42.62%
Banawan	3235	939.22	29.03%	Malapay	1087	453.45	41.72%

Barangay I	4070	1645.53	40.43%	Malidong	2062	913.92	44.32%
Barangay II	1796	486.19	27.07%	Mamlad	503	228.51	45.43%
Barangay III	986	303.58	30.79%	Marigondon	1562	481.39	30.82%
Barangay IV	1472	415.60	28.23%	Matanglad	681	322.13	47.30%
Barangay V	2362	781.09	33.07%	Nablangbulod	657	298.95	45.50%
Basicao Coastal	2300	884.35	38.45%	Oringon	667	307.69	46.13%
Basicao Interior	599	276.63	46.18%	Palapas	1181	523.17	44.30%
Binodegahan	2517	1248.96	49.62%	Panganiran	621	280.53	45.17%
Buenavista	1247	502.21	40.27%	Rawis	875	391.13	44.70%
Buyo	452	223.15	49.37%	Salvacion	621	294.08	47.36%
Caratagan	4536	2584.57	56.98%	Santo Cristo	451	201.67	44.72%
Cuyaoyao	1618	694.71	42.94%	Sukip	868	388.16	44.72%
Flores	1410	617.57	43.80%	Tibabo	662	330.25	49.89%
La Medalla	1837	793.59	43.20%				

Polangui							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Agos	3631	776.19	21.38%	Kinuartilan	594	166.86	28.09%
Alnay	2398	617.81	25.76%	La Medalla	855	238.77	27.93%
Alomon	1285	216.83	16.87%	La Purisima	572	139.18	24.33%
Amoguis	506	155.88	30.81%	Lanigay	3830	968.56	25.29%
Anopol	1155	395.02	34.20%	Lidong	1563	452.28	28.94%
Apad	1665	332.54	19.97%	Lourdes	733	206.67	28.19%
Balaba	1697	581.86	34.29%	Magpanambo	1342	359.89	26.82%
Balangibang	1716	270.76	15.78%	Magurang	3884	494.89	12.74%
Balinad	3256	891.02	27.37%	Matacon	4398	836.50	19.02%
Basud	3752	628.95	16.76%	Maynaga	1349	383.86	28.45%
Binagbangan	922	250.49	27.17%	Maysua	1004	375.52	37.40%
Buyo	1066	320.18	30.04%	Mendez	1025	172.69	16.85%
Centro Occidental	3578	671.75	18.77%	Napo	3448	650.44	18.86%
Centro Oriental	2393	362.31	15.14%	Pinagdapugan	969	245.23	25.31%
Cepres	1529	414.32	27.10%	Ponso	5039	1353.96	26.87%
Cotmon	652	237.86	36.48%	Salvacion	1208	301.96	25.00%
Cotnogan	1652	369.25	22.35%	San Roque	1687	457.65	27.13%
Danao	1131	363.31	32.12%	Santa Cruz	639	171.39	26.82%
Gabon	3381	638.92	18.90%	Santa Teresita	855	219.95	25.73%
Gamot	1424	389.82	27.38%	Santicon	2878	559.76	19.45%
Itaran	1902	488.92	25.71%	Sugcad	3986	993.02	24.91%
Kinale	2402	388.55	16.18%	Ubaliw	3270	550.95	16.85%

Rapu-Rapu							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Bagaobawan	769	490.66	63.81%	Linao	484	323.77	66.90%
Batan	1093	735.20	67.26%	Matobago	869	608.65	70.04%
Bilbao	1265	868.17	68.63%	Mananao	1178	815.05	69.19%
Binosawan	747	465.75	62.35%	Mancao	1146	771.50	67.32%
Bogtong	1072	790.12	73.71%	Manila	840	576.17	68.59%
Buenavista	820	555.14	67.70%	Masaga	517	321.86	62.26%
Buhatan	1016	649.60	63.94%	Morocborocan	929	635.64	68.42%
Calanaga	927	623.19	67.23%	Nagcalsot	899	599.58	66.69%
Caracaran	1387	972.37	70.11%	Pagcolbon	244	172.43	70.67%
Carogcog	421	312.90	74.32%	Poblacion	5840	4261.58	72.97%
Dap-Dap	1024	725.33	70.83%	Sagrada	723	498.29	68.92%
Gaba	836	536.15	64.13%	San Ramon	2230	1494.75	67.03%
Galicia	2020	1313.55	65.03%	Santa Barbara	198	144.57	73.01%
Guadalupe	346	243.92	70.50%	Tinocawan	753	536.73	71.28%
Hamorawon	1118	782.79	70.02%	Tinopan	713	494.71	69.38%
Lagundi	696	506.73	72.81%	Viga	543	350.19	64.49%
Liguan	958	657.80	68.66%	Villahermosa	2299	1556.27	67.69%

Sto. Domingo							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Alimsog	1082	521.72	48.22%	San Fernando	2193	1113.61	50.78%
Bagong San Roque	1410	588.78	41.76%	San Francisco Poblacion	654	217.05	33.19%
Buhatan	1215	575.84	47.39%	San Isidro	2843	938.40	33.01%
Calayucay	1663	846.28	50.89%	San Juan Poblacion	1072	411.41	38.38%
Del Rosario Poblacion	730	267.40	36.63%	San Pedro Poblacion	471	105.85	22.47%
Fidel Surtida	2780	1378.46	49.58%	San Rafael Poblacion	615	275.92	44.86%
Lidong	3076	1286.77	41.83%	San Roque	1636	783.63	47.90%
Market Site Poblacion	200	77.39	38.69%	San Vicente Poblacion	1356	461.97	34.07%

Nagsiya Poblacion	995	297.50	29.90%	Santa Misericordia	2714	1328.09	48.93%
Pandayan Poblacion	493	287.88	58.39%	Santo Domingo Poblacion	673	253.18	37.62%
Salvacion	1946	1008.24	51.81%	Santo Niño	1870	713.27	38.14%
San Andres	3280	1385.02	42.23%				

Tabaco City							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Agnas	1377	455.42	33.07%	Oson	1184	459.72	38.83%
Bacolod	1907	448.46	23.52%	Panal	3317	883.39	26.63%
Bangkilingan	4542	1203.77	26.50%	Pawa	4248	1079.66	25.42%
Bantayan	1645	532.34	32.36%	Pinagbobong	2992	805.05	26.91%
Baranghawon	4357	1157.18	26.56%	Quinale Cabasan	1536	382.74	24.92%
Basagan	1190	341.77	28.72%	Quinastillojan	1805	547.78	30.35%
Basud	2561	581.08	22.69%	Rawis	1516	560.15	36.95%
Bogñabong	2527	746.18	29.53%	Sagurong	2921	1032.76	35.36%
Bombon	1788	435.92	24.38%	Salvacion	2525	699.21	27.69%
Bonot	1188	360.77	30.37%	San Antonio	5393	1474.62	27.34%
Buang	2949	1176.27	39.89%	San Carlos	3859	1058.70	27.43%
Buhian	1599	537.04	33.59%	San Isidro	1306	436.55	33.43%
Cabagñan	2277	672.91	29.55%	San Juan	1521	320.18	21.05%
Cobo	3592	986.52	27.46%	San Lorenzo	7892	2091.50	26.50%
Comon	2497	751.72	30.10%	San Ramon	3408	929.64	27.28%
Cormidal	2187	579.00	26.47%	San Roque	5173	1325.95	25.63%
Divino Rostro	1257	321.22	25.55%	San Vicente	3891	1026.41	26.38%
Fatima	3923	950.20	24.22%	Santo Cristo	5530	1395.95	25.24%
Guinobat	1918	531.60	27.72%	Sua-Igot	1148	412.47	35.93%
Hacienda	5655	2000.56	35.38%	Tabiguian	1642	495.97	30.21%
Magapo	1063	386.82	36.39%	Tagas	2431	596.54	24.54%
Mariroc	5344	1542.14	28.86%	Tayhi	3389	899.68	26.55%
Matagbac	4895	1286.93	26.29%	Visita	1647	606.36	36.82%
Oras	1356	424.72	31.32%				

Tiwi							
Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage	Subdistrict	Population (2015)	Population at Risk	Risk Exposure Percentage
Bagumbayan	2315	1442.82	62.32%	Libtong	2304	1048.19	45.49%
Bariis	1092	682.41	62.49%	Matalibong	1397	885.72	63.40%
Baybay	2265	931.49	41.13%	Maynonong	824	529.20	64.22%
Belen	1406	752.23	53.50%	Mayong	1632	1039.23	63.68%
Biyong	1680	1003.49	59.73%	Misibis	2192	1394.11	63.60%
Bolo	1205	566.62	47.02%	Naga	4629	2294.49	49.57%
Cale	5557	3399.44	61.17%	Nagas	3206	1443.54	45.03%
Cararayan	2012	887.36	44.10%	Oyama	1289	534.55	41.47%
Coro-Coro	1747	774.26	44.32%	Putsan	1234	700.96	56.80%
Dap-Dap	891	572.82	64.29%	San Bernardo	1759	1124.20	63.91%
Gajo	1536	717.11	46.69%	Sogod	2103	1199.80	57.05%
Joroan	2741	1731.67	63.18%	Tigbi	3168	1411.47	44.55%
Libjo	2936	1326.11	45.17%				



CURRICULUM VITAE

Name Karl Malcolm N. Cordova

Date of Birth December 18, 1997

Place of Birth San Jose, Occidental Mindoro, Philippines

Education

Bachelor of Science in Geology - Partido State University (2014-2019)

Master of Science in Geoinformatics – Suranaree University of Technology
(2020-2024)

Publications

Cordova, K.M.N., Sritarapipat, T., & Piyatadsananon, P. (2021). Application of CNN on Landslide Susceptibility Analysis: Case Study on 2018 Hokkaido Eastern Iburi Earthquake. Paper presented at The 42nd Asian Conference on Remote Sensing (ACRS2021), Can Tho University, Can Tho City, Vietnam, 22nd-24th November 2021

มหาวิทยาลัยเทคโนโลยีสุรนารี