



UNDERGRADUATE THESIS

ESTIMATION OF LAND AREA BASED ON- SCREEN DIGITIZATION UTILIZING KARNEY'S ALGORITHM AND SUPPORT VECTOR REGRESSION

IVAN HERDIANTO
NPM 22081010047

THESIS ADVISORS

Dr. Rr. Ani Dijah Rahajoe, S.T.,M.Cs.
Achmad Junaidi, S.Kom., M.Kom.

**MINISTRY OF HIGHER EDUCATION, SCIENCE, AND TECHNOLOGY
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FACULTY OF COMPUTER SCIENCE
INFORMATICS STUDY PROGRAM
SURABAYA
2026**

APPROVAL SHEET

ESTIMATION OF LAND AREA BASED ON-SCREEN DIGITIZATION UTILIZING KARNEY'S ALGORITHM AND SUPPORT VECTOR REGRESSION

By:
IVAN HERDIANTO
NPM. 22081010047

Has been defended before, and accepted by, the Board of Assessors of the Thesis Examination of the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jawa Timur, on August 20, 2026:

Approved,

Dr. Rr. Ani Dijah Rahajoe, S.T., M.Cs.
NIP. 19730512 200501 2 003



(Advisor I)

Achmad Junaidi, S.Kom., M.Kom.
NIP. 197811102025211048



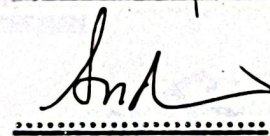
(Advisor II)

**Dr. Firza Prima Aditiawan, S.Kom.,
M.T.I., M.C.F., M.O.S.**
NIP. 19860523 202121 1 003



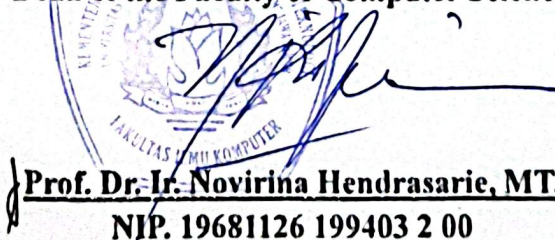
(Head Assessor)

Andreas Nugroho S., S.Kom., M.Kom.
NIP. 19900412 202406 1 003



(Assessor I)

Acknowledge by,
Dean of the Faculty of Computer Science



Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 00

APPROVAL SHEET

**ESTIMATION OF LAND AREA BASED ON ON-SCREEN DIGITIZATION
UTILIZING KARNEY'S ALGORITHM AND SUPPORT VECTOR
REGRESSION**

By:
IVAN HERDIANTO
NPM. 22081010047

Approved to proceed to the Thesis Examination



Approved by,

Coordinator of Informatics Study Program
Faculty of Computer Science

Dr. Intan Yuniar Purbasari, S.Kom. MSc.

NIP. 19800602 202521 2 029

STATEMENT OF ORIGINALITY

I am the undersigned:

Student Name : Ivan Herdianto
NPM : 22081010047
Degree Program : Bachelor (S1)
Study Program : Informatics
Faculty : Faculty of Computer Science

Hereby declares that this undergraduate thesis contains no part of any other scientific work that has been submitted to obtain an academic degree at any higher education institution. Furthermore, it does not contain any work or opinions previously written or published by others, except for those which are explicitly cited in this thesis and listed completely in references.

And I declare that this scientific document is free from elements of plagiarism. If in the future indications of plagiarism are found in this Thesis, I am willing to accept sanctions in accordance with the applicable laws and regulations.

Thus, I made this statement without any coercion from anyone and to be used as it should.



Surabaya, August 27, 2026
Declarant,



IVAN HERDIANTO
NPM. 22081010047

ABSTRACT

Student Name / NPM: Ivan Herdianto/22081010047
Thesis Title: Estimation Of Land Area Based On-Screen Digitization Utilizing Karney's Algorithm And Support Vector Regression
Supervisor: 1. Dr. Rr. Ani Dijah Rahajoe, S.T.,M.Cs.
2. Achmad Junaidi, S.Kom., M.Kom.

This research examines spatial data accuracy issues in *on-screen digitizing*, which is susceptible to human error and planar projection distortions. The objective is to compare the accuracy of the planar method with Karney's Algorithm (based on the WGS84 ellipsoid), test the effectiveness of Support Vector Regression (SVR) as a digitizing bias corrector, and assess its consistency across various land area stratum. The study was conducted using 400 land polygon samples in Surabaya, categorized into small, medium, and large land areas, utilizing data from the ATR/BPN BHUMI WebGIS as the official reference. The SVR model was then optimized using GridSearchCV and evaluated through MAPE, R^2 , MAE, RMSE metrics, as well as the Wilcoxon test. Geodesic computation evaluated across all 400 samples demonstrates that Karney's Algorithm is more accurate, achieving a MAPE of 6.89% compared to the planar method (8.41%), with a mean absolute difference of 45.69 m² between the two methods. The application of SVR significantly improved the overall accuracy, achieving a MAPE of 5.34% and an R^2 of 0.9997. In cross-stratum testing, SVR also demonstrated the best consistency, outperforming both Random Forest Regression and Ridge Regression models. It can be concluded that the integration of Karney's Algorithm and SVR yields an accurate, efficient, and adaptive method for estimating digitized land areas across varying geometric characteristics of land parcels.

Keywords: On-screen digitizing, Karney's Algorithm, Support Vector Regression, land area estimation, GIS.

ACKNOWLEDGEMENTS

Praise and gratitude the author expresses to Allah Subhanahu Wa Ta'ala for all His grace, guidance, and blessings, so that the author was given the ability and ease in completing this thesis entitled "Estimation Of Land Area Based On-Screen Digitization Utilizing Karney's Algorithm And Support Vector Regression". Prayers and peace be upon the Prophet Muhammad Shallallahu 'Alaihi Wasallam, along with his family and companions.

The author realizes that this thesis could not have been completed without the assistance, guidance, and support of many parties. Therefore, the author would like to express the deepest gratitude to:

1. The author's beloved parents, for endless prayers and immeasurable love, sacrifices, and teachings of hard work and responsibility that have become the compass of the author's life. This thesis is wholeheartedly dedicated to both of them.
2. Mrs. Dr. Rr. Ani Dijah Rahajoe, S.T., M.Cs. as Supervisor I and Mr. Achmad Junaidi, S.Kom., M.Kom. as Supervisor II, who with great patience and dedication have guided the author from the very beginning to the completion of this research.
3. All lecturers of the Informatics Study Program, UPN "Veteran" East Java, for the invaluable knowledge, insights, and experiences provided throughout the author's academic journey.
4. The great family of Informatics Class of 2022, "ZFORDDA". It has been a journey full of stories and memories that will never be forgotten.
5. Himpunan Mahasiswa Informatika (HIMATIFA), for broadening the author's collegiate experience through organizational engagement, leadership development, and meaningful collaborative endeavors.
6. Lelga Salfaura Chelsea Fonda, who has consistently accompanied the author through both the highs and lows of this research process. Your companionship has been a source of calm, encouragement, and a place to share stories that is beyond measure.

7. 510, Perunggu, The Band CAMINO, Nuwa, and Ilham Ispram, for the works that continuously accompanied the author through many long nights during the research process. Music is the language of the soul, and your works spoke at exactly the right moments.
8. Hafidz, Gusti, Habib, Daffa, Fazri and all friends who cannot be mentioned one by one, for the togetherness and support that made this research journey feel lighter and more enjoyable.
9. Last but not least, I wanna thank me. I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for never quitting, I wanna thank me for always being a giver, And tryna give more than I receive, I wanna thank me for tryna do more right than wrong, I wanna thank me for just being me at all times.

The author acknowledges that this thesis is far from perfect. Therefore, constructive criticism and suggestions from all parties are greatly appreciated for future improvement. It is hoped that this thesis may be beneficial to readers in general and contribute to the advancement of knowledge in particular.

Surabaya, August 20 2026

Author

TABLE OF CONTENT

ABSTRACT	vii
ACKNOWLEDGEMENTS.....	ix
TABLE OF CONTENT	xi
LIST OF FIGURES.....	xv
LIST OF TABLES	xvii
LIST OF SOURCE CODE	xix
CHAPTER I INTRODUCTION.....	1
1.1. Background.....	1
1.2. Problem Formulation	5
1.3. Research Limitation.....	5
1.4. Research Objectives.....	6
1.5. Research Benefits	6
CHAPTER II LITERATURE REVIEW	9
2.1. Previous Research.....	9
2.2. Geographic Information System.....	12
2.3. Georeferencing.....	12
2.4. On-Screen Digitizing	13
2.5. Remote Sensing and Satellite Imagery	14
2.6. Validity of BHUMI BPN WebGIS Data as a Legal Reference.....	14
2.7. Map Projections and Geometric Distortions.....	15
2.8. Stratified Random Sampling.....	15
2.9. Morphological Features for Geodesic Calculations.....	16
2.10. GridSearchCV and K-Fold Cross-Validation.....	16
2.11. Planar/Shoelace Method	17
2.12. Karney Algorithm (Geodesic Area Calculation).....	18
2.13. Support Vector Regression (SVR) in Spatial Analysis	19
2.14. <i>Ridge Regression</i> in Spatial Analysis.....	22
2.15. <i>Random Forest Regression</i> in Spatial Analysis	23
2.16. Evaluation Method.....	23

2.16.1.	R^2 , MAE, MAPE, dan RMSE	24
2.16.2.	Shapiro-Wilk Test (Normality Test)	25
2.16.3.	Wilcoxon Signed-Rank Test (Statistical Significance Test)	26
CHAPTER III METHODOLOGY		27
3.1.	Research Framework	27
3.2.	Data Collection	29
3.3.	Architectural Design	34
3.3.1.	Geodesic Computation and Feature Engineering	35
3.3.2.	Data Splitting and Pre-processing	37
3.3.3.	Model Training and Optimization Pipeline	39
3.3.4.	Application of Models for Comparative Evaluation	43
3.4.	Performance Evaluation Methodology	48
3.4.1.	Quantitative Evaluation Metrics	48
3.4.2.	Land-Stratum Analysis	51
3.4.3.	Statistical Significance Test	52
3.4.4.	Indicators of Success	53
CHAPTER IV RESULTS AND DISCUSSION		55
4.1.	Data Preparation	55
4.2.	Data Pre-Processing	56
4.2.1.	Coordinate Reference System Validation	56
4.2.2.	Geometric Value Validation	57
4.2.3.	Ground Truth Column Validation	57
4.2.4.	Remove Duplicates Based On ID and Geometry	58
4.2.5.	Data Preprocessing Results	59
4.3.	Test Results	60
4.3.1.	Geodetic Computation Process	60
4.3.2.	Calibration Error and Model Optimization Results	68
4.4.	Discussion on Model Evaluation	72
4.4.1.	Aggregate Performance Analysis	72
4.4.2.	Permutation Feature Importance Analysis	73
4.4.3.	Stability Analysis and Statistical Significance	75
4.4.4.	Land-Stratum Analysis	80

4.5.	User Interface Implementation	82
CHAPTER V CONCLUSIONS AND RECOMMENDATIONS		87
5.1.	Conclusions.....	87
5.2.	Recommendations.....	88
BIBLIOGRAPHY		89
APPENDIX.....		97

LIST OF FIGURES

Figure 3.1 Research Flowchart	27
Figure 3.2 Georeferencing Process	30
Figure 3.3 Georeferencing Result	30
Figure 3.4 Digitization Results	31
Figure 3.5 Enter Attribute Value	31
Figure 3.6 Vector Data Results.....	32
Figure 3.7 Land Parcel on the BHUMI ATR/BPN WebGIS	33
Figure 3.8 Geodesic Computing Architecture Diagram.....	35
Figure 3.9 SVR Model Architecture Diagram	40
Figure 3.10 Comparative Evaluation Pipeline Diagram	49
Figure 4.1 Land digitization results	55
Figure 4.2 Feature Contribution to Prediction per Model.....	74
Figure 4.3 Interactive Maps and On-Screen Digitizing	82
Figure 4.4 Implementation of Spatial Feature Extraction and SVR Prediction....	84
Figure 4.5 Visualization of Results and Metric Evaluation	85

LIST OF TABLES

Table 2.1 Comparison Matrix of Related Studies in Geographic Information Systems	10
Table 3.1 Sample Distribution by Land Stratum.....	34
Table 3.2 Morphological Features and Their Interpretation	36
Table 3.3 Dataset Splitting Scheme	37
Table 3.4 SVR Hyperparameter Search Space.....	39
Table 3.5 RF Hyperparameter Space	47
Table 3.6 Model Performance Evaluation Metrics.....	49
Table 3.7 Final Evaluation Metrics Results	51
Table 4.1 Preview of Coordinate Reference System validation results data.....	59
Table 4.2 Data Validation and Cleaning Results	59
Table 4.3 Statistical Results of the Difference Between Karney's Algorithm and the Planar Method	67
Table 4.4 Performance evaluation metrics for the SVR model on test data.....	69
Table 4.5 Ridge Regression Model Evaluation Metrics.....	70
Table 4.6 Evaluation Metrics for Random Forest Regressor Models	71
Table 4.7 Primary Scenario Model Evaluation Metrics	72
Table 4.8 Permutation Feature Importance Results	73
Table 4.9 Wilcoxon Signed-Rank Significance Test Results	75
Table 4.10 Model Evaluation Metrics for the 70:15:15 Scenario	76
Table 4.11 Model Evaluation Metrics for the 80:10:10 Scenario	77
Table 4.12 Model Evaluation Metrics for the 60:20:20 Scenario	78
Table 4.13 Cross-Scenario Model Stability	79
Table 4.14 Stability of Cross-Stratum Models.....	80

LIST OF SOURCE CODE

Source Code 4.1 Coordinate Reference System Validation	56
Source Code 4.2 Geometric Value Validation	57
Source Code 4.3 Ground Truth Column Validation	57
Source Code 4.4 Remove Duplicates Based On ID and Geometry	58
Source Code 4.5 Import Library For Geodetic Computation	61
Source Code 4.6 Initiate Geodetic Calculation	61
Source Code 4.7 Geodetic Computation Pipeline.....	62
Source Code 4.8 Coordinate Extraction.....	63
Source Code 4.9 Planar Method Computation.....	63
Source Code 4.10 Extraction Morphological Feature.....	64
Source Code 4.11 Validate Value of Morphological Extraction.....	64
Source Code 4.12 Processing Polygon Data	65
Source Code 4.13 Initiate Evaluation Metric	66
Source Code 4.14 Geodetic Computation Analysis Ouput	67
Source Code 4.15 Basemap Interface	83
Source Code 4.16 Initiate Model	84
Source Code 4.17 Calculation and Visualization of Evaluation Metrics	86