

博士 学位 論文

Doctoral Thesis

論文題目

Thesis Title

Evaluating Land Use Change from a Pre-disaster
Recovery Planning Perspective: A Model to Analyze
Post-disaster Settlement Locations

東北大学大学院工学研究科

Graduate School of Engineering,
TOHOKU UNIVERSITY

専攻/Department: Architecture and Building Science

学籍番号/ ID No: C1TD7004

氏名 /Name: MARTIN Garcia Fry

Evaluating Land Use Change from a Pre-disaster Recovery Planning Perspective: A Model to Analyze Post-disaster Settlement Locations

**Graduate School of Engineering
Tohoku University, Japan**

March 2024

Martin Garcia-Fry

氏名	Martin Garcia Fry
授与学位	博士（工学）
学位授与年月日	令和6年3月26日
学位授与の根拠法規	学位規則第4条第1項
研究科、専攻の名称	東北大学大学院工学研究科（博士課程）都市・建築学専攻
学位論文題目	Evaluating Land Use Change from a Pre-disaster Recovery Planning Perspective: A model to Analyze Post-disaster Settlement Locations (災害事前復興計画の観点からの土地利用変化の評価：災害後の定住地分析モデル)
指導教員	東北大学教授 村尾 修
論文審査委員	主査 東北大学教授 村尾 修 東北大学教授 姥浦 道生 東北大学教授 窪田 亜矢 東北大学教授 越村 俊一

論文内容要旨

Post-disaster recovery in the periphery of growing cities faces challenges and opportunities to manage urban growth, while shrinking cities offer a chance to explore non-traditional approaches to generate growth. A few cases of post-disaster resettlement have successfully located displaced households in urbanizing areas with access to livelihood opportunities. Post-disaster recovery has seen progress in the velocity and quality of infrastructure rehabilitation and housing stock reconstruction. However, most recovery programs have ‘failed’ to resettle households close to livelihood choices in ample spheres of life, as well as restored the ecological balance of affected areas, and sustained population growth as opposed to decline. Therefore, it remains unclear whether population, land cover, and mobility affect settlement location in the context of disaster recovery plans. What has not been done, and this thesis does, is to clarify long temporal relationships between population, land cover, and mobility associated to recovery after the 2011 Great East Japan Earthquake and the 2010 Merapi volcano eruptions, and demonstrate a model that evaluates resettlement sites to help local governments formulate disaster recovery plans.

When a disaster occurs, a local government is faced with the challenge of redeveloping devastated areas. So, I assumed that learning from disaster recovery experiences of the past in other countries could help inform future recovery plans. A growing population may be vital to regenerate an area that has been affected by a disaster. I also thought that population growth affects the physical environment which supports people’s activities. These activities are best reflected in land cover change. Therefore, if population trends and land cover changes can be analyzed based on past disaster recovery experiences, perhaps local governments can use the results as indicators for envisioning recovery plans for newly affected areas.

In Chapter 1, the background of this study is described including an overview of two different resettlement initiatives in periurban areas with over ten years of data. First, resettlement after the 2011 Great East Japan Earthquake in Tohoku,

Japan, is described due to relocation areas undergoing population decline. I used population data to determine whether an area was growing or struggling to grow. Given that most areas were found to struggle with growth, I thought there was no need to consider land cover. Second, resettlement after the 2010 Merapi volcano eruptions in Java, Indonesia, is reviewed due to relocation areas currently dealing with issues of population growth. Urbanizing regions experience a higher intensity of land use growth than depopulating regions. Therefore, I evaluated land change interactions and estimated causal effects of resettlement to help avoid land cover degradation after future resettlement initiatives. This chapter indicates the position, purpose, and structure of this study.

Chapter 2 reviews a variety of post-disaster resettlement cases over the past 20 years examining location effects on mobility, land cover, and population. Mobility infrastructure is shown to be essential for the socioeconomic improvement of relocated households. Land cover is threatened by large investments that do not consider side effects in urbanizing regions. Some of these cases have seen polluted or degraded land due to a lack of public utilities and needs assessments in the surrounding areas. Disasters have been known to cause population movement away from affected areas with rates of return based on the effectiveness of disaster resettlement locations, job availability, amenities, and mixed land uses. However, in developing countries where governance is lacking, informal populations are attracted to increased accessibility and new infrastructure causing additional development in unprepared land. Yet, isolated resettlement sites can have opposite effects, aggravating population loss in countries already facing regional population decline. These issues are addressed elucidating recovery needs amid overall growing or declining regional trends. This chapter also discusses methods and practical challenges in the field of land use and transport including micro data collection, survey methods, and micro simulation modeling issues. Then, it describes land cover change detection using remote sensing data and emphasizes the need for automated regional mapping models for disaster recovery applications. To render accurate mapping products, machine and deep learning data inefficiencies are also addressed. Lastly, the need for population and urban geography data analysis is discussed leading to population trend evaluations in growing and shrinking areas affected by the 2011 Great East Japan Earthquake. This chapter describes the scope and contents of this study underscoring research gaps and methodological constraints.

Chapter 3 utilizes population change census data (fertility, mortality, migration, and population counts) collected from 2000 to 2021 over rural and peri-urban municipal areas affected by the 2011 Great East Japan Earthquake. This chapter aims to find peri-urban features likely to drive population growth and sustain future resettled communities in a depopulating region. Tabular data was plotted to analyze population trends over 20 coastal-affected municipal areas to identify post-disturbance anomalies if any. Results saw exacerbated population trends after the disaster with no change in overall trends. Rising mortality rates and fertility loss were seen during the reconstruction process. I included geospatial data (slope, inundation height, ocean front length, distance to train stations, area, and density) to characterize urban features driving population

growth in peri-urban and rural areas. Logistic regression models with over 89% accuracy were applied detecting a decentralized regional structure with vehicular accessibility and built-up density driving growth in small peri-urban municipal areas. Then, I evaluated peri-urban population changes before and after the disaster over areas with high population densities living inside the tsunami inundation range at the time of the event to find resilient population growth drivers. Empirical analyses rendered interconnected built-up agglomerations with accessible road networks and farmlands as potential urban growth drivers. Based on these results, I proposed a resettlement area detection method for depopulating regions from the standpoint of local governments evaluating access to vital urban structures. This chapter clarifies the relationship between post-disaster population growth and internal land structures in the region.

Chapter 4 processed multisource remote sensing data over a rapidly urbanizing district located in Yogyakarta's rural region to explore land cover change in a context of urban growth. This Chapter evaluates land use growth prompted by post-disaster resettlement sites after the 2010 Merapi volcano eruptions. Here, the aim is to help local governments find candidate resettlement sites next to roads that sustain and guide land use growth. I photo-inspected pre-existing land cover samples using high-spatial resolution satellite imagery in Google Earth and vetted each sample against fused values of canopy structure from the Ice, Cloud, and land Elevation Satellite-2 and the Global Forest Canopy Height dataset to circumvent interpretation bias. Reference samples were held while the prevailing quality samples down streamed multitemporal Sentinel-2 satellite data and Digital Elevation Model data to estimate multitemporal amplitude metrics. These metrics were parsed in a canopy height vetting code to automatically supervise the feature space and vertical height of candidate training samples. A Random Forest model was learnt with a refined set training samples to predict 2016-2021 Sentinel-2 satellite images over Java, Indonesia. Mapping outcomes rendered 7 strata, 3 stratified, and 4 land cover change classes with 82% overall accuracy. Site-specific change detections showed that resettlement sites caused 16% forest loss and 18% cropland growth as a medium for 15% built-up growth (in district shares). This chapter proposed a Ground Cover Change model to help local governments find resettlement sites next to roads that sustain land use growth. This method is useful to monitor post-disaster investments and opens the accessibility of remotely-sensed image classification tasks to people with no extensive domain knowledge.

Based on Chapter 4's results, Chapter 5 surveys the local road network to find a candidate resettlement site for evaluation in Merapi affected areas. In particular, this chapter evaluates post-disaster mobility to clarify whether a land-use network with distributed livelihood options can complement resettled farmers in an urbanizing scenario. A survey questionnaire was administered to 42 Javanese-speaking subjects living in the selected resettlement site to collect census, sociodemographic, and mobility data. Then, a travel simulation was performed including density, land use, and accessibility measures to enable a comprehensive diagnosis of livelihoods. Results suggested a restrained access to livelihood options with 26% of the monthly income used for travel. I assumed that livelihood options nearby could significantly improve people's

economic growth. Therefore, I proposed a land use microsimulation method to deliver a predictive network of land use change. A second travel simulation was performed to measure travel and livelihood changes. In total, 30% utility savings and 63% livelihood diversifications were recorded in the simulation with farming consolidating as the primary employment and home-business entrepreneurship as the secondary employment. As a result, mobility and livelihoods were improved enhancing the settlement's location. This chapter developed a 'what if?' modeling system with seven data inputs and ten sub-models sorted into two integration processes to evaluate resettlement sites. Sub-models reported in this chapter saw more than 80% predictive accuracies but, six simulation assumptions were made based on theoretical logic.

Chapter 6 discusses findings of this study: Chapter 3 proved that vehicular accessibility and built-up density surrounded by farmlands played an important role in the stability of periurban population growth areas. Chapter 4 proved that resettlement sites caused minor land use growth effects (15-18% of the district's total). These findings can be used for the land use microsimulation model in Chapter 5. Then, making the model, I saw that a land use network generated urban growth along the network increasing transit and access to livelihood options, thus reducing travel expenses. This result is one example for this case only, and is based on the model's performance, but I think this method provides useful information to consider the image of future affected areas.

Findings also indicate that post-disaster population data saw no significant change in population trends after the 2011 Great East Japan Earthquake. This study then evaluated land cover data during the recovery process after the 2010 Merapi volcano eruptions and found small post-resettlement effects on land cover change with no significant trend alterations. Lastly, it evaluated a resettlement site in Merapi affected areas based on post-disaster mobility input data and found that settlement locations could be improved if livelihood options were located nearby. Therefore, a land use microsimulation method was developed to verify whether livelihoods were improved. The analysis of results indicated that mobility and livelihoods were improved. In conclusion, post-disaster population, land cover, and mobility relationships were clarified rendering a preliminary model system for urban planners in private, public, and federal institutions interested in evaluating candidate relocation sites for future disaster recovery plans.

This conclusion only refers to two cases. However, results were obtained by monitoring and evaluating two ten-year recovery cases and there has never been any research like this before. I expect to investigate other cases in the future and hopefully improve the model's accuracy. I believe that by accumulating this kind of research and evolving it into an automated tool that can be used on demand to contribute to the formulation of recovery plans, other local governments will be able to use it.

様式 2 の(5)

履歴書

ふりがな	マーティン ガルシア フライ	性別	生年月日
氏名	Martin Garcia Fry	男 M 女 F	昭和 62 年 9 月 15 日 or 平成 1987 年 9 月 15 日
本籍	現住所		
都道府県 ペルー	〒980-0845 宮城県仙台市青葉区立町 25-5-206 号		
<p>最終学歴 令和 6 年 3 月 26 日 東北大学大学院工学研究科（都市・建築学専攻）博士課程修了</p>			
<p>研究歴</p> <p>2020 年 10 月 01 日 Clarified land use and transport interactions after the 2010 2022 年 03 月 01 日 Merapi volcano eruptions</p> <p>2022 年 03 月 01 日 Explored land cover change prompted by post-disaster 2023 年 06 月 01 日 resettlement sites after the 2010 Merapi volcano eruptions</p> <p>2022 年 09 月 28 日 Evaluated land structures driving population growth after the 2023 年 11 月 01 日 2011 Great East Japan Earthquake.</p>			
<p>職歴</p> <p>2019 年 04 月 01 日 Administrative Assistant – International Research Institute of 2022 年 03 月 31 日 Disaster Science (Part-Time)</p> <p>2015 年 03 月 01 日 Architect – Jose Bentin Diez Canseco Arquitectos & 2015 年 10 月 31 日 Asociados S.R.L. (Full-Time)</p>			

様式 2 の (6)

論 文 目 錄

氏 名	Martin Garcia Fry					
博士論文						
Evaluating Land Use Change from a Pre-disaster Recovery Planning Perspective: A Model to Analyze Post-disaster Settlement Locations (災害事前復興計画の観点からの土地利用変化の評価：災害後の定住地分析モデル)						
		(1 冊)				
題名	公表の方法	公表の年月日				
Land-use microsimulation model for livelihood diversification after the 2010 Merapi volcano eruptions (第 5 章)	Transportation Research Part D: Transport & Environment Vol.104, No. 103189	令和 4 年 03 月 01 日				
参考論文 題 名	公表の方法	公表年月日	冊数			

Acknowledgements

This dissertation would not have been possible without the guidance and assistance of several people who contributed and assisted in the development and evolution of this study.

First and foremost, I offer mi sincere gratitude to my academic advisor, professor Osamu Murao, who accepted me as a student in his laboratory back in 2017 and who has devoted unsurmountable energy and time to my preparation, personal development, and to the overall improvement of my research. His central research in urban vulnerability and disaster risk has been a guide to my academic development. I have admired his passion for architecture, which I believe brought our interests closer and helped me deliver fruitful labor throughout these past years. I am also indebted to his kindness and honest suggestions enabling academic progress. In particular, professor Murao has devoted his time and effort into helping me understand the meaning and importance of research as well as encouraging my personal thrive. I have enjoyed working in professor Murao's laboratory and I hope that our paths will cross numerous times to continue our work in disaster science.

My academic fulfillment would not have been possible without the extensive assistance of Murao sensei's laboratory members. In particular, Kato-san who provided a wealth of professional advice, assistance, administrative work, and laughter that made me smile every single time. I would like to extend mi gratitude to her and to the students at Murao sensei's lab including Gao-san, Haruna-san, Yonemura-san, Kaneyama-san, Mizuki-san, Kyota-san, Yitong-san, Mufidatun-san, Lin-san, and past students of Murao Lab. I also want to commend Kimura-san for her kindness and willingness to assist in matters of the laboratory. In addition, Sugiyasu sensei also provided insurmountable knowledge and assistance throughout my years as a master's student in Murao Lab, thank you.

I would also like to extend my gratitude to professor Koshimura for the comments provided in early stages of my work as a doctoral student. Professor Koshimura has contributed many times to the development of my research activities and provided me with knowledge, tools, and specific

remote sensing tips which I found extremely useful. Further, I believe professor Koshimura has been very kind in sharing his resourceful knowledge in geospatial data and has understood my intentions thoroughly which I think has helped me improve my research. I would also like to thank professor Bruno Adriano for devoting his time and energy into revising my work and making detailed and proactive observations. I am extremely grateful for their support.

This dissertation would not have been possible without the help of professor Ubaura who guided this research into the right direction. His comments and questions have engaged my work in many ways for which I am most grateful. Professor Kubota also guided my work with a tremendous knowledge in the field of urban planning and for which I am incredibly blessed to have had as a co-reviewer. Although we did not have enough time to discuss matters of common interest, I believe her convictions and knowledge had a deep impression on my work.

Some of my colleagues and friends also deserve my gratitude given their enormous work towards bringing research objectives to completion. In particular, professor Bachri who was caring and industrious since the day we met. This dissertation owes him a great deal of appreciation for the priceless knowledge of Merapi affected areas and his contributions to this research. I am also indebted to Wirawan for collecting data on behalf of this study and for his willingness to improve academic standards. I also want to thank Luis Moya for his help through my early days as a researcher and his disposition to teach complex tasks in fruition of our research targets.

I would also like to express sincere gratitude for my wife, Gabriela, who provides relentless support and helps me navigate life in Japan on a daily basis. Her direct and unassuming character has been of great value to my work. I am also incredibly thankful for taking much of her time revising my work with a careful consideration on detail. She has been an emotional and mental pillar throughout this journey and for that, I am extremely grateful.

I would like to extend my gratitude to my mother and father for always being present, even at a long-distance in Peru, and for always encouraging me to follow my dreams. They both cultivated

health, social responsibility, and academic value at home, which has ultimately been the major drivers of my work through difficult and challenging times. I am very lucky to have an older sister who encouraged my work by believing in me, and a younger brother who always provided indefinite support through his charismatic character and companionship.

This study would not have been possible without the financial support I received from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan. As an embassy-recommended Monbukagakusho scholarship awardee, I am very grateful.

Table of Contents

Abstract.....	I
Resume.....	V
Publications.....	VI
Acknowledgements.....	VII
List of Tables.....	XIV
List of Figures.....	XV
CHAPTER 1 POST-DISASTER RESETTLEMENT AND RECOVERY PROBLEM.....	01
1.1 Introduction.....	01
1.2 Resettlement in Urbanizing Areas after the 2010 Merapi Volcano Eruptions.....	03
1.3 Resettlement in Shrinking Areas after the 2011 Great East Japan Earthquake.....	04
1.4 Research Question and Purpose.....	06
<i>References.....</i>	10
CHAPTER 2 POST-DISASTER RECOVERY RESEARCH AND DEVELOPMENT.....	12
2.1 Long-Term Implications of Resettlement.....	12
2.1.1 Mobility and Livelihoods.....	14
2.1.2 Land Cover.....	16
2.1.3 Population.....	18
2.2 Mobility Estimation.....	19
2.2.1 Travel Demand Modeling.....	21
2.2.2 Micro Data.....	22
2.2.3 Microsimulation Models.....	23
2.3 Land Cover Change Detection.....	24
2.3.1 Training Data Collection Methods.....	25
2.4 Population Change.....	27
2.4.1 Small Area Census and Geospatial Data.....	29
<i>References.....</i>	30
CHAPTER 3 EVALUATING URBAN GROWTH STRUCTURES FOR RESETTLEMENT AREA DETECTION AFTER THE 2011 GREAT EAST JAPAN EARTHQUAKE.....	36
3.1 Introduction.....	36
3.2 Study Area.....	38
3.2.1 The 2011 Great East Japan Earthquake, Tohoku, Japan.....	38
3.3 Data and Methods.....	39

3.3.1 Data Collection.....	39
3.3.2 Demographic Data.....	40
3.3.3 Geographic Data.....	44
3.4 Geographic Feature Integration.....	45
3.4.1 Peri-Urban and Rural Areas	45
3.4.2 Peri-Urban Municipal Area Analysis.....	46
3.5 Urban Structure Analysis.....	48
3.6 Discussion and Concluding Remarks.....	49
3.6.1 Post-Disaster Population Changes.....	49
3.6.2 Urban Structures.....	50
3.6.3 Limitations.....	51
3.6.4 Concluding Remarks.....	52
<i>References.....</i>	53
CHAPTER 4 AUTOMATIC SUPERVISION OF LAND COVER LABELS TO INCREMENT THE RELIABILITY OF TRAINING DATASETS FOR DISASTER RECOVERY CHANGE DETECTION.....	55
4.1 Introduction.....	55
4.2 Study Area and Methodology.....	58
4.2.1 The Cangkringan District, Yogyakarta, Java, Indonesia.....	58
4.2.2 Methodology.....	59
4.3 Data.....	61
4.3.1 Lidar.....	61
4.3.2 Land Cover.....	62
4.3.3 Digital Elevation Model.....	62
4.3.4 Sentinel-2.....	62
4.4 Ground Cover Change Model.....	63
4.4.1 Sample Allocations.....	63
4.4.2 Sample Supervisions.....	66
4.4.3 Model Optimization.....	67
4.4.4 Classifications and Accuracy Assessments.....	69
4.5 Evaluation.....	72
4.5.1 Mapping Outcomes.....	72
4.5.2 Land Interactions.....	73
4.5.3 Resettlement Areas of Influence.....	75
4.6 Discussion and Concluding Remarks.....	77

4.6.1 Automatic Sample Calibration and Model Optimization.....	77
4.6.2 Land Change Effects of Resettlement.....	77
4.6.3 Implications and Limitations.....	78
4.6.4 Concluding Remarks.....	79
<i>References</i>	81
<i>Supplementary Information</i>	85
CHAPTER 5 LAND USE MICROSIMULATION MODEL FOR LIVELIHOOD DIVERSIFICATION AFTER THE 2010 MERAPI VOLCANO ERUPTIONS.....	95
5.1 Introduction.....	95
5.2 The influence of Urban Form on Travel.....	97
5.2.1 Model of Microsimulation.....	98
5.3 Objective Area and Methodology.....	98
5.3.1 The Cangkringan Sub-District.....	98
5.3.2 Methodology.....	100
5.4 Transport Component.....	102
5.4.1 Data.....	102
5.4.2 Activity-Based Schedules.....	104
5.4.3 Travel Simulation.....	104
5.5 OLUTM Model (2019-2030).....	107
5.5.1 Geographic Environment.....	107
5.5.2 Socioeconomic Environment.....	111
5.5.3 Settlement Environment.....	117
5.5.4 Transport Component.....	120
5.6 Discussion and Concluding Remarks.....	122
5.6.1 Limitations and Assumptions.....	124
5.6.2 Concluding Remarks.....	124
<i>References</i>	126
<i>Supplementary Information</i>	130
CHAPTER 6 CONCLUSIONS.....	156
6.1 Main Summary.....	156
6.2 Key Take Aways.....	158
6.3 Limitations.....	160
6.4 Policy Implications.....	160

6.5 Concluding Remarks.....	162
<i>References</i>	164
【APPENDIX 1】 Survey Questionnaire.....	165
【APPENDIX 2】 User Guide.....	170

List of Tables

- Table 2.1. Post-disaster resettlement and recovery problems.
- Table 3.1. Selected municipalities sorted by levels of affected population.
- Table 4.1. Features used for image classifications.
- Table 4.2. Sample allocations.
- Table 4.3. Data calibration summary.
- Table 4.4. Model performance. Root mean square errors (RMSE) and 95th percentile standard errors in parenthesis.
- Table 4.5. Error quantification with two reference sample designs.
- Table 4.6. Confusion matrix of post-informed land cover change with error-adjusted areas expressed in percentages of area proportion with 95% confidence intervals.
- Table 4.7. Cropland area change (Ha)
- Table 4.8. 2016-2021 Mean intensity of change over total change.
- Table 5.1. Road network accessibility measures.
- Table 5.2. Travel tours in segments with distances, road widths, and route choices.
- Table 5.3. Road network types.
- Table 5.4. Local multipliers for tradable, non-tradable, and agriculture jobs in the Cangkringan district.
- Table 5.5. Sensitivity of job location choice variables.
- Table 5.6. Sensitivity of HH location choice variables.
- Table 5.7. Predictive capacity of the model.
- Table 5.8. Monthly travel expenditures.
- Table 5.9. Model assumptions.

List of Figures

- Figure 1.1. Disaster recovery planning cycle.
- Figure 1.2. Study Workflow.
- Figure 1.3. Model Structure.
- Figure 2.1. Detached resettlement site in central Sulawesi, Indonesia.
- Figure 2.2. Disaggregate urban travel demand modeling system.
- Figure 2.3. Pre-existing samples are traditionally discriminated using visual inspection techniques.
- Figure 2.4. Dynamics of mobility during the first 5 months. Movement is shown to decrease during the first 50 days.
- Figure 3.1. Study areas in the Tohoku region.
- Figure 3.2. Population changes divided into two four analysis groups.
- Figure 3.3. Mortality and fertility trends in Tohoku affected coastal areas.
- Figure 3.4. In-migration and out-migration trends in Tohoku affected coastal areas.
- Figure 3.5. Geospatial data collected to characterize population changes in Tohoku affected coastal areas.
- Figure 3.6. Population counts per feature unit increase in peri-urban (G1) and rural areas (G3).
- Figure 3.7. Population counts per feature unit increase in peri-urban areas before and after the 2011 GEJE.
- Figure 3.8. Urban structure analysis over Iwanuma municipal area.
- Figure 3.9. Proposed Method 1.
- Figure 4.1. Land cover classification of Java.
- Figure 4.2. Proposed method 2.
- Figure 4.3. Sample distributions.
- Figure 4.4. Feature contributions with SHAP values.
- Figure 4.5. The Special Region of Yogyakarta in 2019.
- Figure 4.6. Land change in rural and semi-urban sub-districts of the Cangkringan district.

- Figure 4.7. Mean transitions in 200-m tiles.
- Figure 4.8. Resettlement areas of influence.
- Figure 5.1. Study area in the Special Region of Yogyakarta.
- Figure 5.2. Model architecture.
- Figure 5.3. Layout of the OLUTM model (proposed method 3).
- Figure 5.4. The transport component (left).
- Figure 5.5. Model integration flow.
- Figure 5.6. Density map of the Cangkringan district in 2019.
- Figure 5.7. Built capacity of buildings sharing a side with intersecting roads.
- Figure 5.8. Land uses by 2030.
- Figure 5.9. Settlement simulation process.
- Figure 5.10. Vehicular congestion at peak hour transit (nodes 2,3 and 4).
- Figure 5.11. Comparative assessment of fuel consumption.

CHAPTER 1

POST-DISASTER RESETTLEMENT AND RECOVERY PROBLEM

In Chapter 1, the background of this study is described including an overview of two different resettlement initiatives in the rural periphery of urban areas with over ten years of data. First, resettlement after the 2010 Merapi volcano eruptions in Java, Indonesia, is reviewed due to relocation areas currently dealing with issues of population growth. Second, resettlement after the 2011 Great East Japan Earthquake in Tohoku, Japan, is described due to relocation areas undergoing population decline. This chapter indicates the position and purpose of this study.

1.1 Introduction

Rapid urbanization in the periphery of urban areas poses an increasingly urgent call for planners in the context of natural disaster risk. Urbanization, understood as an increase in the urban population and workforce—i.e., manufacturing as compared to agricultural workforce—has almost always involved land use change from non-urban to urban use in response to the need for space in existing settlement areas (Nuissl and Diedentop, 2021). Urban areas are growing faster than populations in most parts of the world (Seto et al. 2011; Angel et al. 2011), and in return, this causes a significant alteration of land cover change accompanied by changes in urban structure and urban form.

According to the Lewis dual-economy model (Lewis, 1954), rural populations migrate to urban areas seeking employment opportunities. However, not all migrant workers have access to formal employment, and may choose to be self-employed, or informally employed (Bederman and Adams, 1974). The visible outcome of this trend is the expansion of built-up areas (Seto et al. 2016), with spillover effects in high-risk, informal, and unprepared land. In addition, global targets such as biodiversity loss and food security have been linked to land use extent affecting carbon sources and

partial land cover degradation (Winkler et al. 2021). This makes urban growth a necessary implication of disaster recovery plans, and quantifying its dynamics, a critical task in supporting these debates.

Disaster recovery planning exercises hazard contingencies and community-based resilience in an overall urban recovery plan. Pre-disaster recovery plans, harmonized with other plans as shown in Figure 1, can provide a strong foundation for post-disaster reconstruction projects (FEMA, 2017). This is because planning can be applied before a disaster occurs to anticipate safety, environmental, and societal needs while preparing evacuation and relocation sites to efficiently mitigate and respond to future natural disaster events.

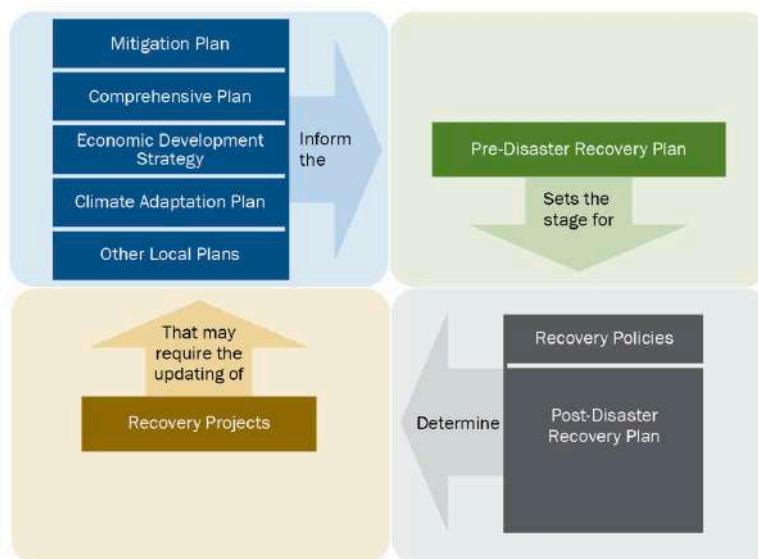


Figure 1.1. Disaster recovery planning cycle (FEMA, 2017)

There is also a need to address several pre-existing issues at the local level, which are particularly important to the communities living on that area. For example, urbanization can have severe implications on rural livelihoods and natural resources which account for the local identity and resiliency of people (Winkler et al. 2021). It can also overwhelm land owners in fragile conditions causing them to sell their land immediately after a catastrophic event (König et al. 2010). In contrast, shrinking cities experience population loss which leads to the marginalization of demographic groups that are more attached to land. This increases the risk of mortality among the elderly as a result of

isolation, and decreases birth rates due to young people's movement (Lima and Eischeid, 2017). Encouraging a town or city to make progress towards resolving these issues post-disaster, may prove too challenging. Therefore, an organized reconstruction and community planning process that is ready to be implemented after a disaster will ensure that the affected people are able to undertake an organized process and do not miss the opportunities to rebuild in a sustainable and more resilient way. This is why pre-disaster recovery plans are better suited to address pre-existing local needs, take advantage of available resources, and include opportunities to increase local resilience, sustainability, accessibility, and social equity (FEMA, 2017).

1.2 Resettlement in Urbanizing Areas after the 2010 Merapi Volcano Eruptions

On October 26, 2010, Indonesia's most active stratovolcano erupted thrice displacing thousands of people from their homes causing massive damage to land and property (Kuncoro et al. 2012). In response, REKOMPAK—a community-driven disaster reconstruction and resettlement program—relocated 2,608 households in safely-designated areas with a majority of houses built in 15 housing resettlement sites (huntraps) in the 5 villages that make up the Cangkringan sub-district of Sleman Regency, Yogyakarta, Java, Indonesia (Garcia-Fry and Murao, 2020).

This resettlement approach was first piloted in Banda Aceh after the 2004 Indian Ocean tsunami. A request was made by the national government of Indonesia to assign the World Bank as a trustee to manage a Multi Donor Fund (MDF), which pooled together a total of 654.5 million USD from 15 country donors. This amount of money under one management provided a suitable solution for beneficiaries by putting aside individual interest. By 2006, a 5.7 magnitude earthquake struck the island of Java and four years later in 2010, Mt. Merapi erupted evacuating over 400,000 people temporarily from danger zones (areas within 20 km from Mt. Merapi) (The World Bank, 2016). Since REKOMPAK still had people and resources on the ground, it was possible to mobilize a quick response and scale support through this program. In light of this, the Java Reconstruction Fund was

able to efficiently allocate its first block of aid while other support was being organized (The World Bank, 2012). By 2014, the program had finished the reconstruction of physical infrastructure including paved roads, housing, and public utilities with evacuation plans in place and some livelihood recovery programs organized. The affected livelihoods were based on dairy farming and agriculture, while tourism was the secondary employment with the standard minimum wages (Rindrasih, 2018).

Households of the largest resettlement site, Pagerjurang, currently believe that resettling has granted them safety from the volcano. However, resilience did not translate into secured livelihoods (Miller, 2018): Even though tourism became a main source of livelihood diversification, it did not promote choices of livelihood in ample spheres of life; middle-aged farmers and the elderly are too tired of commuting to the dairy farming cooperative and their land on Mt. Merapi; and, the Cangkringan district is undergoing rural urbanization which is shifting land-use structures that generate place-attachment. As a result, rural identities are at stake, risking tourist-attractive value (Umaya et al. 2020).

Some studies claim that rural income sources, such as small-holder dairy farming and agriculture or ecotourism, do not provide sufficient economic returns for the population at large, often causing people to sell their land for ad-hoc profits or reorient their livelihoods to off-farm jobs (König et al. 2010). Rural fragmentation in Yogyakarta's wider region is now dependent on how public and private investments will manage urban growth to sustain the environment.

1.3 Resettlement in Shrinking Areas after the 2011 Great East Japan Earthquake

On March 11, 2011, at 15:00 pm in the afternoon, a sudden and violent crustal deformation generated a 9.0 magnitude earthquake off the northeastern coast of Honshu, Japan (Hirose et al. 2011). The sudden event released a large-scale surge of energy triggering tsunami tidal waves which led to

massive destruction of the Pacific coastline and instant displacement of 470,000 people from their homes (Gusman et al. 2012).

Tanaka (2023) emphasizes three important features of the 2011 earthquake that should be recognized in relation to postwar Japan: First, Japan is one of the most advanced countries in disaster prevention, holding tsunami countermeasures that encompass “hard” (physical) and “soft” (intangible) facilities particularly enhanced in the northeastern coast which falls under the tsunami-prone zone designation; second, the Great East Japan earthquake caused the largest monetary damage in the postwar period and consequently, the largest financial investment for the reconstruction and recovery process. In contrast to the 1995 Great Hanshin-Awaji earthquake which caused damage of approximately 9.6 trillion Japanese yen (JPY, according to the National Land Agency) and received a reconstruction budget of 16.3 trillion JPY, the amount of damage caused by the Great East Japan earthquake was estimated in 16.9 trillion JPY (estimated by the Cabinet Office) and received a reconstruction budget of 32 trillion JPY for the subsequent 10-year period (Tanaka, 2023). Third, the 2011 disaster was the first major disaster event that Japan has faced since the era of population decline. Of the three features, this study examines the latter to characterize population growth areas for future resettlement.

The post-disaster reconstruction of housing and communities was undertaken at the core of all recovery plans with a notable 12.9 trillion JPY. Collective relocation projects (community-based reconstruction projects) consisted of relocating households to higher grounds or away from high-risk coastal areas. Various land readjustment projects were also carried out to raise land and ensure safer residential areas. However, several issues have been pointed out: First, the 460 billion JPY used for land readjustment projects have rendered approximately one third of the total land unused; second, collective relocation projects have not been able to move residents “collectively” to safer areas. In most cases, fewer than half of the original residents participated in the relocation projects causing a separation of family and neighborhood ties in the community. In total, only one third of the victims

rebuilt their homes using the three housing reconstruction support projects offered by the government (Tanaka, 2023). Third, post-disaster voluntary movement was reported in coastal and urban peripheral areas, such as Watari and Yamamoto towns, as a result of mobility obstructions (Muroi, 2022), community disintegration (Tanaka, 2023), and inconvenient life in temporary shelters (Miyasada and Maly, 2021). Fourth, relocation projects have resulted in the peripheral expansion of residential areas, the formation of low-density urban areas, and the creation of a decentralized regional structure based on the separation of jobs and residences (Tanaka, 2023). Therefore, the sustainability of settlements remains uncertain.

1.4 Research Question and Purpose

Is it possible to derive an appropriate reconstruction and recovery plan to resolve these issues from prior data (population, land cover) and surveys of relocating residents' travel routines?

The objective of this study is to clarify the relationship between preliminary population, land cover data and resettlement location, as well as mobility and livelihoods, and to develop a model that will contribute to the formulation of future recovery plans. I evaluated post-disaster land use change and clarified relationships between population, land cover, and mobility data associated to resettlement sites in areas affected by 2010 Merapi Volcanic Eruption and 2011 Great East Japan Earthquake. The two natural disaster cases are suitable candidates to better understand post-disaster recovery in growing and declining population contexts. They also offer reliable data for more than 10 years. The 2011 earthquake is important to explore population growth in a depopulating region because these areas have growing and declining population areas allowing an effective exploration of urban structures driving population growth. In contrast, population growth in urbanizing regions is partly a consequence of urbanization (i.e., employment shares and informal real estate transactions) which can cause uncertainty. Therefore, Chapter 3 selects the 2011 Great East Japan Earthquake affected areas to detect urban features associated with population growth and resettle future

populations in places with a high probability of growth. My assumption before conducting this research was that population growth (in a depopulating region) is likely to be a product of particular urban structures driving growth. Results proved that urban features such as peri-urban accessibility and aggregate built-up structures surrounded by farmlands play an important role in the stability of population growth areas. I also found that pre-disaster population data combined with geographic data can be used to locate resettlement areas near regional cities.

Based on these potential resettlement areas, Chapter 4 assumed that resettlement and infrastructure investments have significant land use change effects. Therefore, Chapter 4 selects Merapi affected areas because urbanizing regions experience a higher intensity of land use growth than depopulating regions. Therefore, the maximum causal effects of resettlement are estimated to prevent land cover degradation in developed and developing countries. Chapter 4 proved that resettlement sites in Merapi affected areas have caused minor land use growth effects (15-18% of the district's total) with no change in trends. These results can help local governments identify candidate settlement locations next to roads that sustain land use growth. The reason for estimating population data in a depopulating region and land cover data in an urbanizing region is because these data choices secure reliable estimates in each of the associated contexts:

1. Population growth in depopulating regions help to compare growing and declining areas to detect unique urban features driving population growth.
2. Land use growth in an urbanizing region helps to secure the maximum causal effects of resettlement with the potential to apply results as a surplus value in depopulating regions.

Chapter 5 evaluates post-disaster mobility in Merapi affected areas to clarify whether a land use network with distributed livelihood options can complement rural labor in a candidate resettlement site. This chapter assumed that resettlement and redevelopment will increase transit and access to livelihood opportunities. It selects Merapi affected areas because it offers the necessary setting to evaluate land use change in an urbanizing rural area. Application to growing populations in

depopulating regions can help to understand whether these land use networks will expedite growth. Chapter 5 also assumes that livelihood options near resettlement sites will decrease mobility costs. Therefore, a land use microsimulation method was proposed to deliver a predictive network of land use change. Results proved that a land use network with distributed livelihood options reduced travel expenditures by 30%. It also proved that rural livelihoods can be diversified, instead of replaced, if settlements are exposed to transit yielding home-business entrepreneurship (i.e., shops, small restaurants, hair salons, etc.). Figure 1.2 illustrates this study's workflow.

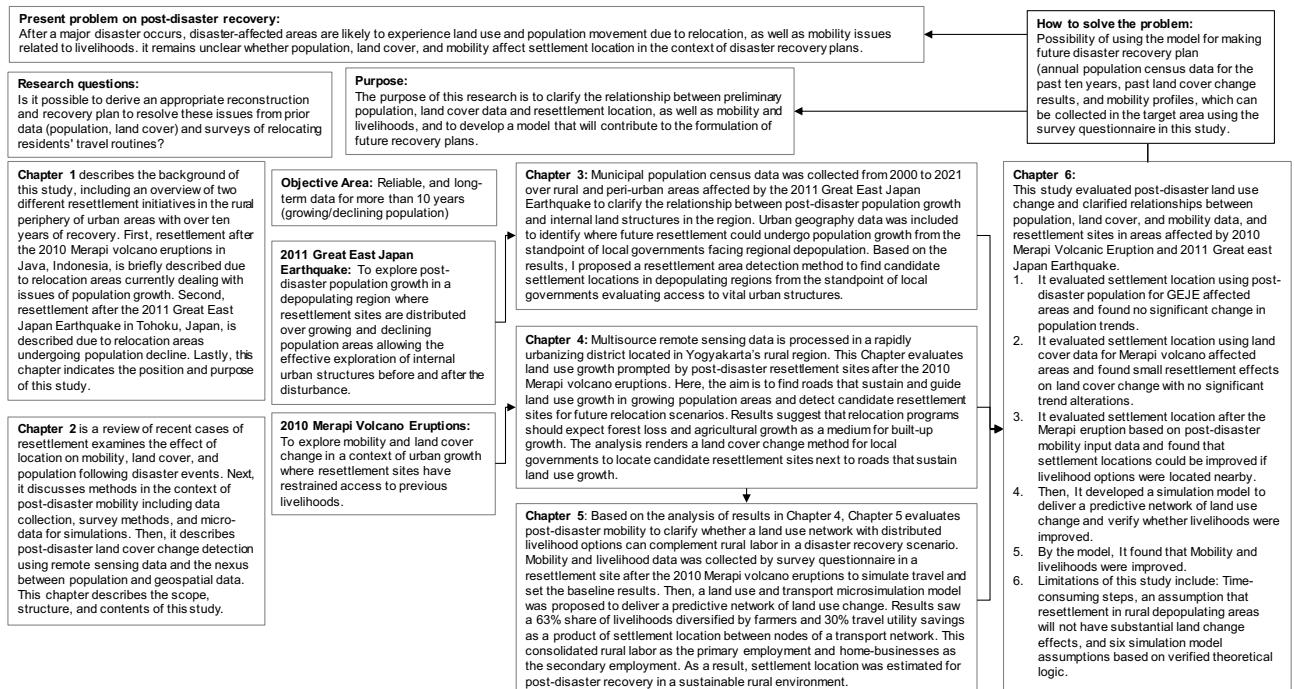


Figure 1.2. Study Workflow.

This study proposed a model in Figure 1.3 because Chapter 3 proved that pre-disaster population trends were unchanged in a depopulating context. The first method can be used to find urban structures that generate population growth in a depopulating region. Chapter 4 proved that resettlement sites do not change land cover trends in rural urbanizing areas where maximum effects can be expected. The second method can be applied to past resettlement cases in peri-urban areas

undergoing urban growth to help local governments find future relocation sites next to roads that sustain land use growth. Based on Chapter 4's result, Chapter 5 surveyed the local road network in Merapi affected areas and applied land use growth estimates to find an existing resettlement site. It evaluated mobility proving that increased accessibility can improve livelihood diversification. The third method can use mobility profiles from a target area where a land use network is required. Future applications of Chapter 5 can use Chapter 3 and Chapter 4's result in the rural periphery of a depopulating regional city if study areas are within the same ecozone.

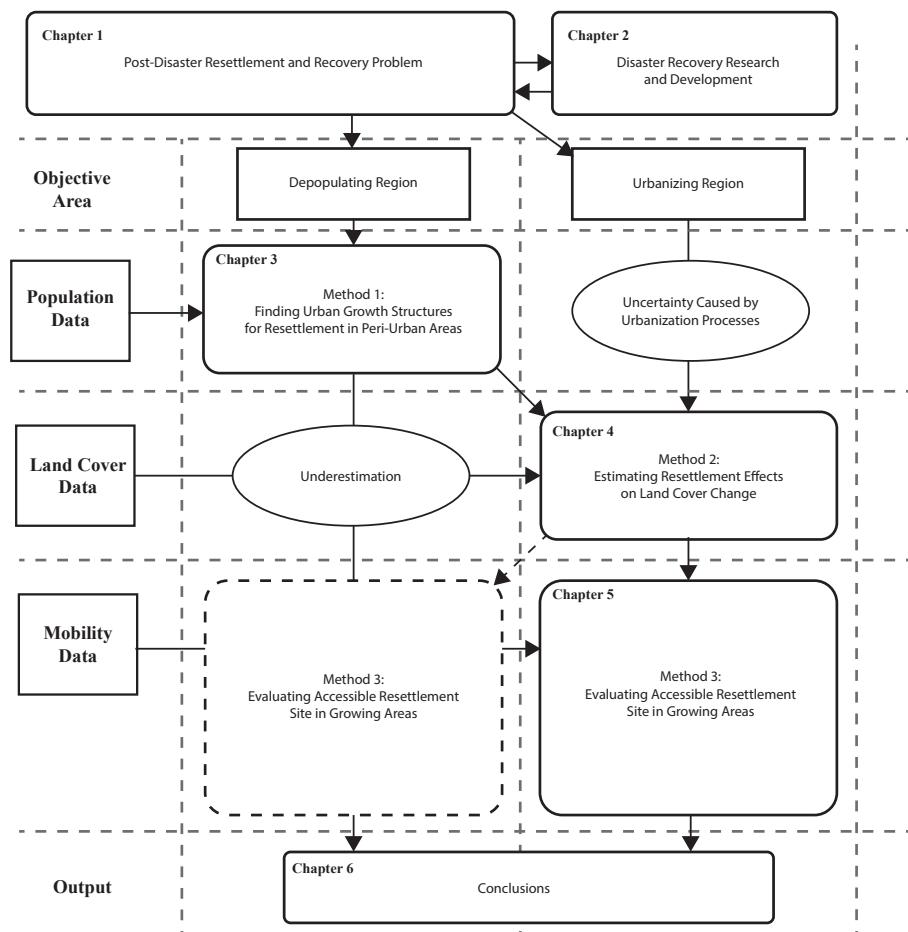


Figure 1.3. Model Structure.

References

- Nuissl, H., Siedentop, S. (2021). Urbanisation and Land Use Change. In: Weith, T., Barkmann, T., Gaasch, N., Rogga, S., Strauß, C., Zscheischler, J. (eds) Sustainable Land Management in a European Context. Human-Environment Interactions, vol 8. Springer, Cham. https://doi.org/10.1007/978-3-030-50841-8_5
- Seto K.C., Fragkias M, Güneralp B, Reilly M.K., 2011. A Meta-Analysis of Global Urban Land Expansion. PLOS ONE 6, e23777. <https://doi.org/10.1371/journal.pone.0023777>
- S. Angel, J. Parent, D. L. Civco, A. Blei, D. Potere, 2011. The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. Prog. Plann. 75, 53–107. <https://doi.org/10.1016/j.progress.2011.04.001>
- Lewis, W.A., 1954. Economic Development with Unlimited Supplies of Labour. The Manchester School 22, 139-191. <https://doi.org/10.1111/j.1467-9957.1954.tb00021.x>
- Bederman, S. H., & Adams, J. S. (1974). Job accessibility and underemployment. Annals of the Association of American Geographers, 64, 378–386.
<https://doi.org/10.1111/j.1467-8306.1974.tb00986.x>
- Seto, K.C., and Ramankutty, N., 2016. Hidden linkages between urbanization and food systems. Science 352 (6288) <http://www.doi.org/10.1126/science.aaf7439>
- Winkler, K., Fuchs, R., Rounsevell, M., et al. 2021. Global land use changes are four times greater than previously estimated. Nat. Commun. 12, 2501. <https://doi.org/10.1038/s41467-021-22702-2>
- FEMA, 2017. Pre-disaster recovery planning guide for local governments. FEMA publication FD 008-03. Accessed: November, 2023. <https://www.fema.gov/sites/default/files/2020-07/pre-disaster-recovery-planning-guide-local-governments.pdf>
- Kuncoro, W., Nugroho, D., Sarwani, M., 2012. Agricultural recovery action for area affected by 2010 Merapi volcano eruption, Indonesia. Int. J. Geoinf. 8, 1-8. <https://creativecity.gsc.osaka-cu.ac.jp/IJG/article/view/630>
- Garcia-Fry, M., and Murao, O., 2020. A post-disaster forecast analysis for relocation settlements using a combined model sequence. Proceedings of the 17th World Conference on Earthquake Engineering, 7d-001, 1-12.
- The World Bank, 2016. Community-based Settlement Rehabilitation and Reconstruction Project for Central and West Java and Yogyakarta Special Region. Report No. ICR00003662, pp. 5-10, 24, 25.
- The World Bank, 2012. Rekompak; Rebuilding Indonesia's Communities After Disasters. Jakarta: The MDF for Aceh and Nias and the JRF Secretariat, pp. 66-67, 72 -74, 158 -165.
- Rindrasih, E., 2018. Under the Volcano: Responses of a Community-Based Tourism Village to the 2010 Eruption of Mount Merapi Indonesia. Sustainability, 10, 1-17. <https://www.mdpi.com/2071-1050/10/5/1620>
- Miller, M.A., 2018. Reconfiguring rural aspirations through urban resettlement: Navigating futurity after the 2010 eruption of Mount Merapi, Indonesia. In: Bunnell, T., and Goh, D., (Ed.), Urban Asias: Essays on Futurity Past and Present, Jovis, Berlin, pp. 193–202. <https://urbanaspirationsasia.com/urban-asias/>
- Umaya, R., Hardjanto, Soekmadi, R., Sunito, S., 2020. Livelihood adaptation patterns of sub-villages communities in the slope of Merapi Volcano. IOP Conf. Ser: Earth Environ. Sci. 528, p. 1. <https://doi.org/10.1088/1755-1315/528/1/012020>
- König, H.J., Schuler, J., Suarma, U., McNeill, D., Imbernon, J., Damayanti, F., Dalimunthe, S.A., Uthes, S., Sartohadi, J., Helming, K., Morris, J., 2010. Assessing the Impact of Land Use Policy on Urban-Rural Sustainability Using the FoPIA Approach in Yogyakarta, Indonesia. Sustainability, 2, 1991-2009. <https://doi.org/10.3390/su2071991>
- Hirose, F., Miyaoka, K., Hayashimoto, N., Yamazaki, T., and Nakamura, M., 2011. Outline of the 2011 off the Pacific coast of Tohoku Earthquake (Mw 9.0)—Seismicity: foreshocks, mainshock, aftershocks, and induced activity—. Earth Planet Sp 63, 1. <https://doi.org/10.5047/eps.2011.05.019>

Gusman, A.R., Tanioka, Y., Sakai, S., Tsushima, H., 2012. Source model of the great 2011 Tohoku earthquake estimated from tsunami waveforms and crustal deformation data, *Earth Planet. Sci. Let.*, 341–344, pp. 234-242. <https://doi.org/10.1016/j.epsl.2012.06.006>

Tanaka, S., 2023. The 2011 Great East Japan Earthquake and Tsunami: The highest casualties and largest reconstruction funds—Characteristics of major disasters and future challenges in developed countries. In Special Issue: Experience of Great Disasters: Revisioning 3/11 and beyond. *Japanese Journal of Sociology* 32, 7-24. <https://doi.org/10.1111/ijjs.12147>

Abe, T., 2014. Population movement in the Tohoku region after the Great East Japan Earthquake disaster. *Science reports of the Tohoku University, 7th Series, Geography*, 60, 83-95. <https://tohoku.repo.nii.ac.jp/records/11754>

Muroi, K., 2022. Post-disaster reconstruction in the rural-urban fringe following the Great East Japan Earthquake. *Proceedings of the 13th AIWEST-DR 2021, E3S Web of Conferences* 340, 03001. <https://doi.org/10.1051/e3sconf/202234003001>

Miyasada, A., and Maly, E., 2021. Impacts of collective housing relocation in the Ogatsu area of Ishinomaki City after the 2011 Great East Japan Earthquake and Tsunami. *Proceedings of the 12th AIWEST-DR 2019 in IOP Conf. Ser.: Earth Environ. Sci.* 630 012014. <http://www.doi.org/10.1088/1755-1315/630/1/012014>

Garcia-Fry, M., Murao, O., Bachri, S., Moya, L.A., 2022. Land-Use microsimulation model for livelihood diversification after the 2010 Merapi volcano eruptions. *Transp. Res. Part D: Transp. Environ.* 104, 103189, p.18. <https://doi.org/10.1016/j.trd.2022.103189>

CHAPTER 2

POST-DISASTER RECOVERY RESEARCH AND DEVELOPMENT

Chapter 2 provides a review of post-disaster resettlement cases over the past 20 years examining location effects on mobility, land cover, and population. Next, it discusses methods and practical challenges in the field of land use and transport including data collection, survey methods, and micro data for simulations. Then, it describes land cover change detection using remote sensing data in the post-disaster context and the nexus between population and geospatial data. This chapter describes the scope and contents of this study.

2.1 Long-Term Implications of Resettlement

Permanent displacement is often a direct consequence of sudden-onset hazard events, or slow-onset climate change (Rigaud et al. 2018), which forces the migration of people to safer and more sustainable locations. While some natural disasters are unpredictable (e.g., earthquakes, landslides, and volcanic eruptions) others can be more predictable (e.g., floods, tornadoes, cyclones, and hurricanes). Through disaster prevention measures and a culture of safety, risks can be reduced. However, land, infrastructure, and the housing stock are more susceptible to damage often causing mass permanent displacement and livelihood loss.

A direct consequence of displacement is resettlement by reconstruction or relocation. There are at least six major factors influencing resettlement (Shi 2008): Political and religious, economic benefit, eco-environmental, engineering construction, warfare and natural disasters. Relocating communities can consume a lot of resources and time. This is sometimes not in accordance with the expectation of people who want to get back to their previous state. Also, many cases of resettlement have failed to relocate communities in places that provide choices of livelihood in ample spheres of life (e.g., see Table 2.1). Successful cases overcome this disadvantage through the availability of reliable infrastructure, mobility options, and proximity to new and previous livelihood options. These

cases also restore damaged ecosystems and maintain the ecological balance of resettlement areas enabling sustainable communities. In contrast, failed cases have not been able to sustain previous livelihoods and promote diverse livelihood options for the socioeconomic improvement of relocated people. Some of these cases have not been able to conserve the environment leading to health, security, and resource consumption issues.

Table 2.1. Post-disaster resettlement and recovery problems.

Case	Relevance	Problem	Reference
2018 Sulawesi Earthquake	After the 2018 Sulawesi earthquake, resettlement forced part-time farmers to choose fishing over farming because of the many kilometers required to reach their farmlands (Paripurno et al. 2022, p.4).	Mobility	Paripurno et al. 2022
2011 Great East Japan Earthquake	After the 2011 Great East Japan Earthquake, collective relocation projects resulted in the creation of a decentralized regional structure based on the separation of jobs and residences (Tanaka, 2023, p.16, Section 3.4, last paragraph).	Mobility	Tanaka 2023
2011 Great East Japan Earthquake	After the 2011 Great East Japan Earthquake, collective relocation projects resulted in the creation of low-density areas in an already depopulating region (Tanaka, 2023, p.16, Section 3.4, last paragraph).	Population	Tanaka 2023
2010 Merapi volcano eruptions	After the 2010 Merapi volcano eruptions, several households were relocated to safely designated areas, but to sustain their livelihoods they have to travel back to their farmlands or change their identities as farmers (Miller, 2018, p.202).	Mobility	Miller 2018
2010 Merapi volcano eruptions	Settlements are not physically embedded in the nearby provincial capital city and challenge the official administrative boundary of cities promoting the expansion of built-up areas (Miller, 2018, p.200)	Land Cover	Miller 2018
2009 L'Aquila Earthquake	After the 2009 L'Aquila Earthquake, 19 resettled communities were isolated and made dependent on private transport to reach valued destinations and amenities (Contreras et al. 2017, p. 80, Section 3, paragraph 1 & 3).	Mobility	Contreras et al. 2017
2009 L'Aquila Earthquake	Communities demonstrated a positive correlation between dissatisfaction and the distance and travel time to a place with employment.	Mobility	Contreras et al. 2013
2004 Indian Ocean Tsunami	Neuhuen relocation sites after the 2004 Indian Ocean tsunami underscore the importance of infrastructure and mobility tied to livelihood recovery (Sina et al. 2019, p.180).	Mobility	Sina et al. 2019
2004 Indian Ocean Tsunami	Most people working in the capital city, Banda Aceh, spent more than 30% of their income on travel. Some jobs were also location-reliant, such as farming and fishing industries, which meant that people had to commute to their previous villages for work (p. 181).	Mobility	Sina et al. 2019
2004 Indian Ocean Tsunami	In Arugama bay located on the eastern coast of Sri Lanka, relocated and low-income fishermen could not foresee any livelihood options and had to walk 2 kms daily to sustain their families (Ingram et al. 2006, p.609 Section 4.2).	Mobility	Ingram et al. 2006
2004 Indian Ocean Tsunami	These resettlement camps have threatened the environmental conservation of the area threatening people's health and sustainability (p. 610).	Land Cover	Ingram et al. 2006
2011 Christchurch Earthquake	Massive outflow of people to the outer areas of the city, except people over 50 years of age representing a cohort of the population that tends to stay due to age and land ownership as well as a deeper connection to place (King and Gurtner, 2021, section 6.4.1).	Population	King and Gurtner 2021
2005 Hurricane Katrina	Only 60% of the population returned to the city where most of the damage occurred (King and Gurtner, 2021, section 6.4.2). The other 40% were thinking of leaving before the event, but damage prompted their migration.	Population	King and Gurtner 2021
Cyclones Larry and Yasi, 2006 & 2011	An already declining population trend was exacerbated by the loss of land and property displacing most of the agricultural labor force (King and Gurtner, 2021, section 6.4.3).	Population	King and Gurtner 2021
The 2011 Great East Japan Earthquake	Iwanuma is a successful case of recovery given its growth after the disaster and has continued to experience that growth.	Population	Muroi 2022
Bangladesh Floods 1995-2015	Land use changes exacerbated vulnerability in disaster risk areas and annual weather-related disasters have forced people's movement to peri-urban areas causing unprecedented vulnerability and land cover loss.	Land Cover	Parven et al. 2022
2011 Great East Japan Earthquake	Newly designated disaster risk zones caused land cover change. All coastal-damaged areas changed from agricultural activities to uninhabitable grasslands.	Land Cover	Ishihara and Tadono 2017

2005 Hurricane Katrina	Post-Katrina land management indicated that built-up land disappeared and several structures were abandoned during the recovery phase. Land cover took over several roads and revegetated the affected areas.	Land Cover	Reif et al. 2011
------------------------	---	------------	------------------

This Chapter reviews different cases of resettlement examining the effects of location on mobility and livelihoods, land cover, and population change underscoring research gaps. Then, it describes how contents of this research are organized to enable local governments formulate pre-disaster recovery plans. Disaster management studies in the past twenty years (2000-2021) have mainly focused on preparedness and response from 2000 until 2014. Yet, recovery studies have gained momentum since 2016 (Wolbers et al. 2021). The priority of a response stage is to distribute relief and rescue operations, while recovery stages prioritize progress beyond the reconstruction of buildings and infrastructure (Contreras et al. 2016). A considerable amount of money from donors and governments is destined to finance the recovery process, but allocation and management are crucial to overcoming expectations of development endorsed by the Sendai Framework for Disaster Risk Reduction 2015-2030 (UNISDR, 2015).

2.1.1. Mobility and Livelihoods

Post-disaster recovery research focuses on people's lives and livelihoods including mental health and the redevelopment of affected areas in socioeconomic terms (Contreras et al. 2016). For example, resettlement after the 2004 Indian Ocean tsunami underscored the importance of livelihood recovery tied to infrastructure and mobility options. In fact, Sina et al. (2019) interviewed resettled households in Neuheun relocation sites located in Aceh Besar, Indonesia, and stated that "*infrastructure was the 'blood vessel' that facilitates their mobility and connection to the workplace, markets, administrative centers, schools, and health centers, and is essential for the socio-economic improvements of the relocated sites*" (p. 180). Another important issue is the distance from relocation sites to livelihood areas rendering people's inability to sustain travel and household costs. Neuheun villagers that worked in the capital city, Banda Aceh, spent more than 30% of their income on travel.

Also, location-reliant jobs, such as farming and fishing industries, forced people to commute to their previous villages for work (Sina et al. 2019).

In 2009, the government of L'Aquila, Italy, experienced a 5.9 magnitude earthquake causing the relocation of 19 communities to distant land making people dependent on private transport to access urban amenities (Contreras et al. 2017). Given that residential areas were separated from jobs and livelihood options, people's dissatisfaction was found correlated with travel distance and time to workplaces (Contreras et al. 2013). The 2010 Merapi volcano eruptions and the 2011 Great East Japan Earthquake also continued that trend separating residences and jobs (Miller 2018; Tanaka 2023). More recently, the 2018 Sulawesi earthquake in the Indonesian archipelago also relocated people in dispersed and seemingly isolated resettlement sites (see Figure 2.1). As a result, Paripurno et al. (2022) reported that part-time farmers had to choose fishing over farming given the many kilometers required to reach their land.



Figure 2.1. Detached resettlement site in central Sulawesi, Indonesia.

Source: (Rizki, Vun, and Bald, 2022/World Bank. License: CC BY NC-SA 4.0)

Previous events have shown that long distance commuting from home to work after disaster displacement is one of the top problems reported by the relocated workers (Oliver-Smith 1996). Past research focused on site, layout, housing and popular input as crucial issues in the determination of

success or failure in post-disaster resettlement (Oliver-Smith 1991). Yet, failure to address livelihood options for the socioeconomic improvement of relocated people at their new location determines the sustainability of relocation projects on the long-term.

Successful relocation projects often involve the identification of equivalent lost land near new sites or bringing new land into production through land recovery (Cernea 2021). Unfortunately, there is less empirical research on positive experiences reconstructing livelihoods than there is describing and deplored displacement's pathologies. Even though success is less frequent than failure, developing knowledge on the former is imperative for policy and practical purposes. One example is the case of land scarcity near Shuikou dam in China, which led project leaders to convert hillsides and uplands into agrarian terraces for horticulture and forested areas (Cernea 2021). Intensified agriculture increased pre-settlement income and generated livelihoods for 19,700 people (The World Bank 1996). In most cases, complementary strategies and diversification can benefit the rest of resettlers': Crop intensification and crop shifts, diversification of primary activities, use of project-created resources such as disaster mitigation infrastructure, new employment, etc. Another example of successful livelihood development was the case of communities in Saguling, Indonesia, who saved their former land's topsoil for increased fertility and yield near resettlement sites (Costa-Pierce 1998). Although these cases were not led by disasters, they illustrate the possibilities embedded in post-disaster resettlement.

2.1.2 Land Cover

Land cover change using remote sensing technologies in the post-disaster context is a field of research that is largely at bay. Monitoring studies focus on the urban fringe of developing cities to understand the dynamics of risk associated with informal growth patterns (Fekete 2022), climate change (Gibson et al. 2016), and inconsistent planning schemes (Wijaya 2018). Multitemporal time-series analyses of remote sensing data have been increasingly applied in these areas (Li et al. 2018;

Shaw and Das 2018). A number of disaster response studies have focused on pre- and post-disaster damage assessments using optical remote sensing (Koshimura et al. 2020), SAR images (Moya et al. 2020), UAV photogrammetry (Calantropio et al. 2021), and multimodal/temporal datasets (Adriano et al. 2021). However, very few studies examine the effects of resettlement on socioeconomic (land use) and environmental (land cover) change.

Monitoring the recovery process helps to understand factors of vulnerability and keep them under control to avoid reproducing urban conditions that may have contributed to the original disaster (Contreras et al. 2016). It also provides data useful to encourage the formulation of pre-disaster recovery plans and the improvement of existing ones. A comprehensive study of the 1995-2015 recurring floods in Bangladesh's deltas reported that land use change and annual disaster events have forced people's movement to peri-urban areas causing unprecedented vulnerable conditions and significant deforestation effects (Parven et al. 2022). This example portrays the importance of managing urban growth and the environmental conservation of rural areas in the periphery of urban centers. The following statement by Miller (2018) also reflects on the importance of settlement location after the 2010 Merapi volcano eruptions:

“Even though the post-disaster Huntap (permanent settlements) are not physically embedded in the nearby provincial capital city of Yogyakarta, they are constitutive of the sort of “thoroughgoing urbanisation” of Asia that social demographer Gavin Jones says is characteristic of Asia’s emergent future. Not only do “urbanised corridors” like Huntap Pagerjurang challenge the official administrative boundaries of cities that are conventionally cast as representative units of “the urban”, but they are also imbued with certain urban norms and values that pervade and transform Indonesian societal spaces” (p. 200).

The paradigm of growth, expressed in developing countries where most urbanization is presently occurring, is equally contrasted by slow growth and population decline in developed countries (King and Gurtner 2021). For example, after the 2011 Great East Japan Earthquake, newly designated disaster risk zones rendered uninhabitable coastal areas which underwent land cover change from agricultural activities to grasslands (Ishihara and Tadono 2017). Also, Reif et al. (2011) reported that built-up land disappeared and several structures were abandoned during the recovery phase following the 2005 Hurricane Katrina in the U.S. Land cover took over several roads and persistently re-vegetated the affected areas. Repurposing land and infrastructure can help to resize density or enhance opportunities for green urbanism, ecological restoration, and hazard mitigation (Hollander et al. 2009).

Another implication of rural resettlement is the type of resettlement, including centralized and scattered resettlement sites. Concentrated rural settlements can achieve sustainable development and improve the resilience of people after disaster events (Zhao et al. 2022). Scattered settlements can waste land, generate poor living conditions, and degrade the environment. Centralized resettlement is considered an effective way to utilize rural land, improve infrastructure, public services, and living conditions (Yi et al. 2012). However, scattered villages are the basic form of rural areas and the usual choice for disaster resettlement.

2.1.3. Population

Population movement after sudden-onset disaster events has been a recurring topic of research in recent decades (Yabe et al. 2020; Abe 2014). A comprehensive study monitoring post-disaster population movement in five countries pointed out that temporary migration can affect a household's decision to relocate (Yabe et al. 2020). After hurricane Katrina in 2005, only 60% of the population returned to the affected areas (King and Gurtner 2021). The other 40% had decided to leave before the event, but damage prompted their migration. A similar experience can be attributed to cyclones

Larry and Yasi in 2006 and 2011, respectively, damaging rural areas of Innisfail in North Queensland, Australia. The pre-disaster declining trend was exacerbated by the loss of land and property displacing most of the agricultural labor force (King and Gurtner 2021).

Christchurch in New Zealand was hit by a large earthquake in 2011 which generated a massive outflow of people to the outer areas of the city. This case resonates with the latter cases in terms of the demographic profile of those moving away from hard-hit areas. However, people over 50 years of age represented a large cohort of the population that stayed due to land ownership and a deep connection to the place (King and Gurtner 2021). Some accounts of younger households moving in search of employment has been reported (Muroi 2022; Yabe et al. 2020). The movement of young and educated households have three significant consequences according to Lima and Eischeid (2017): First, public revenue is shrunk lowering public investment capacity and private investment interest; second, it reduces the capacity of inhabitants with adequate resources to entrepreneur businesses that create jobs; and third, it declines birth rates even further due to young couples' movement. As a result, these consequences reduce job opportunities and fuel depopulation, thus creating a negative feedback cycle (Lima and Eischeid 2017).

Much less attention has been paid to the relationship between resettlement sites and population change after disaster events. Reconstruction in the aftermath of disasters faces changed demography with a shift in service and infrastructure needs (King and Gurtner 2021). Very few studies have explored these uncertainties to understand post-disaster population change (Wilson et al. 2022). Moreover, aggravated population loss has been reported after the 2011 Tohoku Earthquake due to the separation of residences, jobs, and amenities (Tanaka 2023; Muroi 2022; Miyasada and Maly 2021; Abe 2014). Therefore, it remains unclear whether population, land cover, and mobility affect settlement location in the context of post-disaster recovery plans.

2.2. Mobility Estimation

Transportation and employment are two major constituents of the urban environment. When combined, they provide access to an equitable range of livelihood opportunities. Several meanings of accessibility exist (see Litman 2021), but this study follows the importance of reducing vehicle dependence as a factor of neighborhood accessibility (Krizek 2003). This entails shorter distances in daily travel routines and mobility choices for specific travel destinations.

While transport satisfies the need for mobility owing to the separation of activities, urban form research examines physical characteristics that compose urban environments (shape, size, and configuration of its parts) (Wegener and Fuerst 2004). For example, residents and firms want to be close to each other, saving on travel utilities and time. This generates economies of scale driven by agglomeration effects (Levinson and Wu 2020). These effects commonly expand in nodes of the transport network, where local and inflowing resources meet profitably, gathering building densities, activities, and people (Hynynen 2006). Nodes are commonly formed in road intersections or open spaces, and are defined by the scale of attraction and employment shares in a traffic analysis zone (TAZ). Transport analysts use TAZ's to assign the expected growth and model trips in a network (Clifton et al. 2008). Figure 2.2 shows the general framework of decisions relevant to model urban travel demand demonstrating decisions and interactions considered most important (Ben-Akiva, Bowman, and Gopinath 1996).

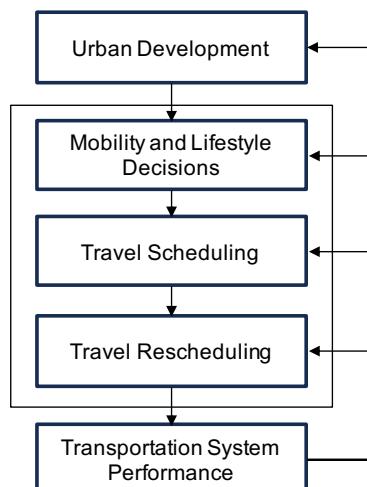


Figure 2.2. Disaggregate urban travel demand modeling system.

Source: (Adapted from Bowman and Ben-Akiva (2001))

This framework focuses on the household and individual decisions that influence travel such as mobility, lifestyle, activity, travel scheduling, and rescheduling (Bowman and Ben-Akiva 2001). It also shows the urban development process affecting individual decisions, and the interaction of all areas subject to the performance of the transportation system. Individual decisions can occur in different timeframes: (1) Mobility and lifestyles, reflected on residential location, employment, and vehicle ownership happen in irregular and infrequent annual timeframes; (2) activity and travel scheduling, is a more regular and frequent function that can happen in days or weeks. This includes particular activities, assignment of activities to household members, the order of activities performed, locations, times, and travel modes; and, (3) rescheduling, which occurs within a day in response to information or an unexpected disturbance which causes change in planned activities and travel schedules (Bowman and Ben-Akiva 2001). Urban development, on the other hand, will reflect the choices of governments, real estate developers, and firms which interact through policies, regulations, and locational needs influencing the decision of people and households, and together the urban form. All of these decisions affect the performance of a transportation system.

This study evaluates urban development in a TAZ with three urban form measures: (1) density, defined by Handy et al. (2002) as the amount of population, employment, or building square-meters (m^2) per unit of area; (2) land use, as the proximity of activities measured by the area of neighborhood shopping in a 400-meter radius, equivalent to the assumed median walking distance by Yang and Diez-Roux (2012); and, (3) accessibility, estimated in connectivity, road intersections, and block lengths (Krizek 2003).

2.2.1. Travel Demand Modeling

Bowman and Ben-Akiva (2001) used the “primary destination” approach, assuming that one activity and destination are identifiable as a travel tour generator, representing a start and end time for the tour. This method identifies work destinations as the motivation for trip generation from Monday to Saturday, from 5:00 am to 4:00 pm (Mohammadi and Taylor 2019). The tour aggregation method, attributed to geographer Hagerstrand (1970), Chapin (1974), and Jones (1977) is applied to gather “home-work-other-home” (hwoh) schedule of activities reflecting people’s daily travel routines. Modeling individual activities requires a large set of travel routines which can be generated from openly available datasets, existing transport databases, or administered from targeted field surveys in a study area.

While most countries do not have access to the whereabouts of people during daytime, some developed countries do have access to datasets of people’s work location and their residential addresses. Even if activity-based travel schedules are made available, several practical challenges still remain: First, human behavior functions in various locations at different points in time and involves the cost of movement from one place to another (Hagerstrand 1970); second, the household affects individual travel decisions, which may have implications on other members of the household as a whole (Chapin 1974); and third, decisions at one point in time are influenced by past and anticipated future events. This shapes behavior with responses exhibiting time lags or asymmetry (Goodwin et al. 1990). Furthermore, challenges in lifestyle and mode choice sector of this field have also brought modern difficulties for simple travel demand modeling. New alternatives like park-and-ride, car-sharing, or shared buses and remote employment derive new patterns that require new methods (Moeckel et al. 2003).

2.2.2. Micro Data

A new method developed to circumvent these issues is Monte Carlo microsimulation (Moeckel et al. 2003). Microsimulation models have been used to reproduce human behavior at the

individual level, i.e., how individuals select options following their perceptions, preferences, habits to constraints, and uncertainty. These models use micro data, but the collection of these data (single buildings, people's registration, firm's locations) is neither allowed or undesirable for privacy reasons. Therefore, these models work with synthetic micro data which can be retrieved from aggregate data sets.

This study followed a variant of the most prominent method to generate a synthetic population developed by Beckman et al. (1996) because this procedure is used for most integrated microsimulation models (Moeckel et al. 2003). Each household is represented by size, age, presence of children (under 18 years old), income of workers, and vehicle ownership. Employment requires individual jobs and employment sectors. The model accounts for jobs located in non-residential floorspace or residential units to account for home-based employment (Waddell 2003).

2.2.3. Microsimulation Models

Microsimulation has had applications in social science since the 60's (Orcutt et al. 1961), but spatial models had small impact even though they covered several areas of expertise (Moeckel et al. 2003). More recently, integrated urban models (land use and transport models) are complex systems in which transportation and land-use co-evolve in time. The range of models in this area vary from case to case, but it is argued that "complete" models include (Miller, E. 2018):

- Building stock supply subdivided by type (residential, commercial, etc.).
- Residential and employment location choice given building stock supply.
- Demographic change in the area of application.
- Business and firmographic change in the area of application.
- Residence-work and -school links as a result of market processes.
- Vehicle ownership.
- Daily travel per person in multi-modal travel simulations.

- Daily travel flow as a result of transportation network performance.
- Estimation of system impacts which include travel costs, energy, gas emissions, etc.

These processes rely on the individual decision of agents (persons, households, firms, etc.) as they go through their daily activities, collaboration between agents to achieve mutual utilities, and market demand/supply processes that interact with agents, compete with each other for resources, and exchange goods and services (Miller, E., 2018). Therefore, urban modelling systems directly depend on the relationship between travel demand and urban form, while development of the urban form and location choices depend on the accessibility provided by the transportation system (Wegener, 1995).

The Operational Land Use and Transport Microsimulation (OLUTM) model described in Chapter 5 of this study follows the prospect theory, developed by Kahneman and Tversky (1979), to simulate random decisions based on the potential value of loss or gains in a two-stage decision-making process. The models here maximize the utility behavior of agents and objects (jobs, buildings, land uses, etc.) with sociodemographic data found by Kitamura (2009) to affect travel behavior.

While many important studies have advanced the computational barriers of microscopic simulation (Miller and Salvini 2005; Waddell et al. 2003, Wagner and Wegener 2007, Zhu et al. 2018, Zondag et al. 2015), they mainly process findings from urban areas and do not address post-disaster mobility in rural areas (Garcia-Fry et al. 2022).

2.3. Land Cover Change Detection

Monitoring post-disaster land cover change is essential to understand causal effects of investments. Human activities and natural forces constantly change the surface of the Earth. To keep the state of Earth's environment in a sustainable state of change, it is necessary to use change detection techniques (Zhu et al. 2021). Change in land use at various scales has significant impacts on land cover (a crucial area of research for global climate change), environmental ecology, and natural

hazards (Liu et al. 2023). This body of knowledge enhances our understanding of global human and natural phenomena relationships as well as better resource management and land utilization.

Remote sensing is the primary source of data for land change detection. Remote sensing data supplies broad coverage over long periods of time. Two traditional methods to detect change exist: binary and semantic class segmentation. The former relates to change/no change distinctions between image pairs on the pixel (raster) level, while the latter compares multi-temporal image classifications or direct satellite images to reflect ‘from-to’ transitions.

In practical remote sensing applications, abrupt changes on the pixel level (e.g., change caused by human activity or natural disturbances) are difficult to estimate compared to synchronized change (e.g., vegetation growth and change in soil moisture). This is mainly due to the size of change in fractions of a pixel and the spectral magnitude of changes (Xu et al. 2022; Zhu et al. 2022). The predictive uncertainty of algorithms can also stream inconsistent time-series data (Zhao et al. 2019). This can be attributed to machine and deep learning models in the computer vision field suffering from insufficient and inefficient training labels (Huang et al. 2015). Manual labeling is hampered by the time-consuming process to supervise all samples, while domain expertise is limited to particular sites (Li et al. 2022). Another big challenge for remote sensing data learners is label noise (Burke et al. 2021). Given the variety and complexity of applications, datasets with perfect annotations rarely exist (Zhou et al. 2023).

2.3.1. Training Data Collection Methods

Researchers have explored various ways to overcome these issues using unsupervised, semi-supervised, and self-supervised learning models. Transfer learning and domain adaptation networks are unsupervised methods where a model is pre-trained on a source image to be applied on a target image elsewhere (Saha et al. 2022). The accuracy of these networks depends on the difference between source and target data. Therefore, weakly-supervised models offer an alternative solution

entailing: (i) incomplete supervision, where a subset of training labels are available; (ii) inexact supervision, where coarse-grained labels are given; and, (iii) inaccurate supervision, where the given labels are not always true on the ground (Yue et al. 2022). Active learning is another method that requires a small number of labeled samples to draw instances according to some measures (Desai and Ghose, 2022). Likewise, self-supervised models use unlabeled data to identify features and labeled samples to adapt the model for downstream tasks (Wang et al. 2022). However, the reliability of supervised labels can be contested due to interpretation biases (Stehman, 1999). Also, unsupervised labels carry multi-scale and high-dimensional land cover features which complicate their use.

One solution to this problem is using pre-existing land cover maps to generate reliable, cost-effective, and consistent national or continental land cover maps (Hermosilla et al. 2022). Many studies have used pre-existing land cover products to capitalize on training samples (Inglada et al. 2017; Wessels et al. 2016). However, source map sample errors can confuse the classifier and spread over target maps. Therefore, various quality control measures are suggested to filter and refine spurious samples. Ancillary data can provide descriptive information on factors such as topographic and hydrologic characteristics, climate, and land cover features which can allow the separation and detection of mislabeled samples. It does so by separating classes with similar spectral characteristics. Land cover features extracted from remote sensing spectral data can derive multitemporal statistical metrics. These metrics interpret phenological and morphological time-sequencing factors, or attributes of a given class (Figure 2.3). They retain information without regard to time or place, provided they represent a statistically valid representation of an area (Song et al. 2001).

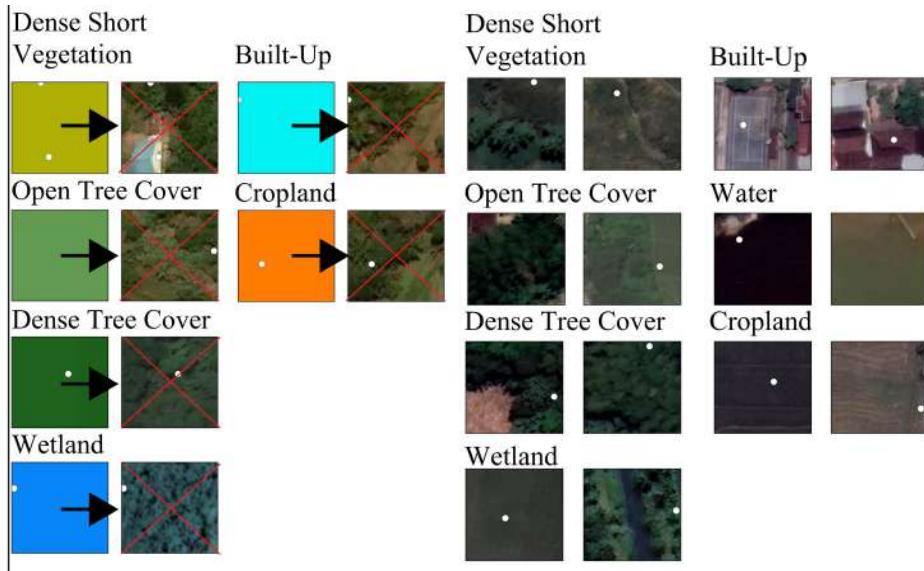


Figure 2.3. Pre-existing samples are traditionally discriminated using visual inspection techniques.

While multitemporal land cover features have been used to collect labels from remote sensing images (Tong et al. 2020), or to classify images directly (Potapov et al. 2020), they have not been retained from visually and vertically pre-inspected samples to automatically supervise the feature space and vertical height of candidate training samples.

2.4. Population Change

Post-disaster population change is a field of research that has surfaced over the two decades but has seen much of its attention focused on post-disaster displacement (King and Gurtner 2021). For example, Yabe et al. (2020) studied the effect of displacement across the U.S., Puerto Rico, and Japan using mobile phone GPS data (Figure 2.4) and found that five population displacement cases saw returnees approximately one month later. It also reported that people move across cities in search of employment opportunities or closer to family homes looking for stability and futurity.

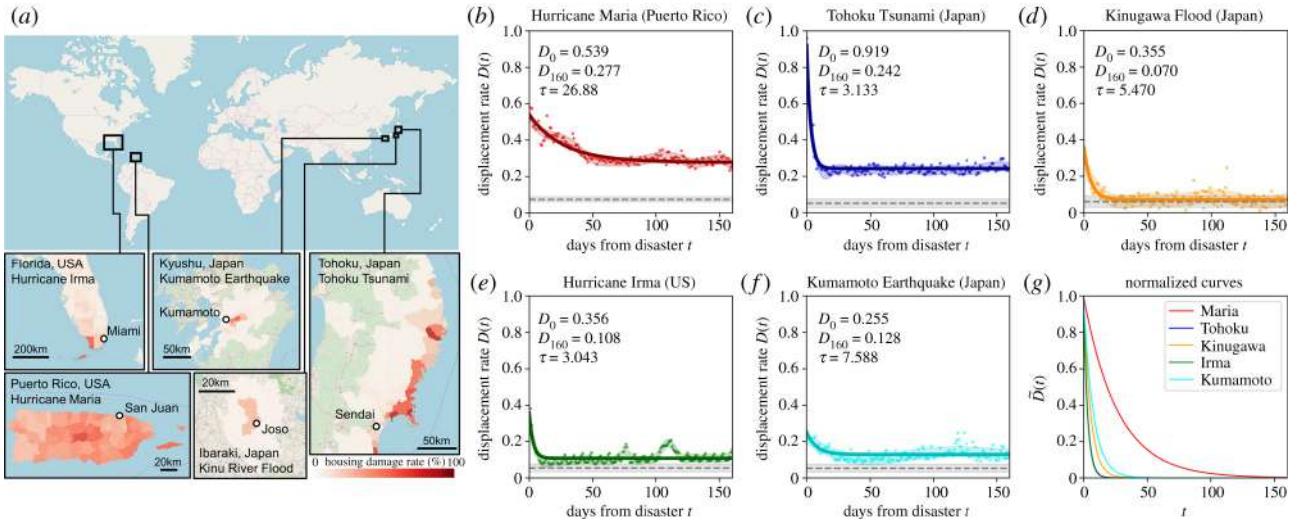


Figure 2.4. Dynamics of mobility during the first 5 months. Movement is shown to decrease during the first 50 days.

Source: (Permission requested, Yabe et al. 2020).

Demographers have studied disasters largely when the consequences occur at a scale potential to affect regional or national populations (Frankenberg, Laurito, and Thomas 2015). In addition to the obvious mortality effects of disasters, they can also have longer-term or indirect consequences that affect populations. Demography changes with the scale of data analysis and the different social layers which make some demographic groups more susceptible to disaster risk. Two recent studies by Frankenberg et al. (2023) and (2020) revealed that mortality rates increased over 10 years after the 2004 Indian Ocean Tsunami, but decreased in the subsequent 5 years. This was explained as a consequence of long-term resilience gained many years after the disaster event. Mortality has been studied by several social scientists and psychologists focused on the post-disaster context.

In contrast, fertility has been less studied over the past decades as reported in a recent review that examines post-disaster-related fertility studies (Lee et al. 2023). This study included a systematic review of natural disasters (e.g., earthquakes and weather-related disasters) and technological disasters (e.g., poisoning, explosions, radiation) (Lee et al. 2023). A total of 50 studies rendered negative fertility effects 5-10 years after natural disasters. It also states that after 10 years, most studies agreed that fertility rates returned to the median trend before the event.

While many studies have explored the demographic scene in post-disaster contexts, small area population studies are less known to explore these uncertainties through geospatial indicators or structural urban land use systems.

2.4.1. Small Area Census and Geospatial Data

Urban risk research has evolved around urban vulnerability, inequity and poverty, resilience, and similar trends (Hossain et al. 2017). Yet, post-disaster population growth in a depopulating context is less understood and opens the possibility for research avenues that explain relationships between urban features and population growing trends amid overall declining tendencies. One study by Gruebner et al. (2015) applies a correlation method to evaluate spatial vulnerability associated with post-disaster traumatic stress disorders following Hurricane Sandy in New York City. Another study utilizes mobile phone data and nighttime lights to find spatiotemporal relationships between urban, suburban, and rural counties after Hurricane Irma in the U.S. and reveal a hierarchical structure between areas to understand spatial recovery impacts (Park et al. 2024).

Urban recovery planners can benefit from studies that target relationships between population change and urban features in the post-disaster context. Therefore, the next chapter will explore the relationship between post-disaster population growth and geographic features to find urban growth structures for sustainable resettlement site detections. To the best of my knowledge, population growth and urban structures have not been explored in a depopulating context and require special attention to understand where future resettlement should take place.

References

- Rigaud, K. K., Sherbinin, A. d., Jones, B., Bergmann, J., Clement, V., Ober, K., Schewe, J., Adamo, S., McCusker, B., Heuser, S., Midgley, A., 2018. Groundswell: Preparing for Internal Climate Migration. Washington, DC: World Bank. <http://hdl.handle.net/10986/29461>
- Shi, G., 2008. Discussion on Resettlement Science. Proceedings of 16th IAHR-APD Congress and 3rd Symposium of IAHR-ISHS. IV, 1456-1462. Nanjing: Springer-Verlag Berlin Heidelberg.
- Wolbers, J., Kuipers S., and Arjen, B., 2021. A systematic review of 20 years of crisis and disaster research: Trends and progress. Risks Hazards Crisis Public Policy, 12, 374–392. <https://doi.org/10.1002/rhc3.12234>
- Contreras, D., Blaschke, T., Tiede, D., Jilge, M., 2016. Monitoring recovery after earthquakes through the integration of remote sensing, GIS, and ground observations: The case of L'aquila (Italy). Cart. Geog. Inf. Sci., 43, 115–133. <https://doi.org/10.1080/15230406.2015.1029520>
- UNISDR (United Nations International Strategy for Disaster Reduction), 2015. Sendai framework for disaster risk reduction 2015–2030. http://www.wcdrr.org/uploads/Sendai_Framework_for_Disaster_Risk_Reduction_2015-2030.pdf Accessed: Sep. 2022.
- Sina, D., Chang-Richards, A.Y., Wilkinson, S., Potangaroa, R., 2019. What does the future hold for relocated communities' post-disaster? Factors affecting livelihood resilience. International Journal of Disaster Risk Reduction 34, 173-183. <https://doi.org/10.1016/j.ijdrr.2018.11.015>
- Contreras, D., Blaschke, T., and Hodgson, M.E., 2017. Lack of spatial resilience in a recovery process: Case L'Aquila, Italy. Tech. Forecast. and Soc. Change 121, 76-88. <https://doi.org/10.1016/j.techfore.2016.12.010>
- Contreras, D., Blaschke, T., Kienberger, S., and Zeil, P., 2013. Spatial connectivity as a recovery process indicator: the L'Aquila earthquake. Technol. Forecast. Soc. Chang., 80, 1782-1803. <https://doi.org/10.1016/j.techfore.2012.12.001>
- Paripurno, E.T., Mahojwala, G., Prabaswara, G. & Khabibah, S.U., 2022, 'Agency-driven and community-driven impact in livelihood recovery: Beneficiaries stories', *Jàmbá: Journal of Disaster Risk Studies* 14(1), a1200. <https://doi.org/10.4102/jamba.v14i1.1200>
- Rizki, M.H., Vun, J., and Bald, A., 2022. Three years on: Four lessons learned from post-disaster recovery in Central Sulawesi, Indonesia. East Asia & Pacific on the Rise. The World Bank Blogs, Washington, D.C., United States. Accessed: December, 2023. <https://blogs.worldbank.org/eastasiapacific/three-years-four-lessons-learned-post-disaster-recovery-central-sulawesi-indonesia>
- Oliver-Smith, A., 1996. Anthropological Research on Hazards and Disasters. Annual Review of Anthropology 25, 303–328. <http://www.jstor.org/stable/2155829>
- Oliver-Smith, A., 1991. Successes and Failures in Post-Disaster Resettlement. Disasters 15, 12-24. <https://doi.org/10.1111/j.1467-7717.1991.tb00423.x>
- Cerneia, M.M., 2021. The Risks and Reconstruction Model for Resettling Displaced Populations. In: Koch-Weser, M., Guggenheim, S. (eds) Social Development in the World Bank. Springer, Cham. https://doi.org/10.1007/978-3-030-57426-0_16
- World Bank, 1996. Successful reservoir resettlement in China: Shuikou hydroelectric project. Washington, DC: World Bank. Accessed on January, 2024: <https://documents1.worldbank.org/curated/ar/118191468770421776/pdf/multi-page.pdf>
- Costa-Pierce, B. A., 1998. Constraints to the Sustainability of Cage Aquaculture for Resettlement From Hydropower Dams in Asia: An Indonesian Case Study. The Journal of Environment & Development, 7(4), 333-363. <https://doi.org/10.1177/107049659800700402>

- Fekete, A., 2022. Peri-urban growth into natural hazard-prone areas: mapping exposure transformation of the built environment in Nairobi and Nyeri, Kenya, from 1948 to today. *Nat Hazards*, 1-24. <https://doi.org/10.1007/s11069-022-05515-4>
- Gibson, T. D., et al., 2016. Pathways for transformation: Disaster risk management to enhance resilience to extreme events. *J. Extreme Events* 3, 1-23. <https://doi.org/10.1142/S2345737616710020>
- Wijaya, N., 2018. Disaster risk reduction and climate change adaptation integration into peri-urban development planning: A case study of Bandung metropolitan area, Indonesia. *IOP Conf. Ser.: Earth Environ. Sci.* 145, 012079. <http://www.doi.org/10.1088/1755-1315/145/1/01207>
- Li, X., Zhou, Y., Zhu, Z., Liang, L., Yu, B., and Cao, W., 2018. Mapping annual urban dynamics (1985–2015) using time series of Landsat data. *Remote Sens. Environ.* 216, 674-683. <https://doi.org/10.1016/j.rse.2018.07.030>
- Shaw, R. and Das, A., 2018. Identifying peri-urban growth in small and medium towns using GIS and remote sensing technique: A case study of English Bazar Urban Agglomeration, West Bengal, India. *The Egyptian J. Remote Sens. Space Sci.* 21, 159-172. <https://doi.org/10.1016/j.ejrs.2017.01.002>
- Koshimura, S., Moya, L., Mas, E., Bai, Y., 2020. Tsunami Damage Detection with Remote Sensing: A Review. *Geosci.* 10, 177. <https://doi.org/10.3390/geosciences10050177>
- Moya, L., Mas, E., Koshimura, S., 2020. Learning from the 2018 Western Japan Heavy Rains to Detect Floods during the 2019 Hagibis Typhoon. *Remote Sens.* 12, 2244. <https://doi.org/10.3390/rs12142244>
- Calantropio, A., Chiabriando, F., Codastefano, M., and Bourke, E., 2021. Deep learning for automatic damage assessment: Application in post-disaster scenarios using UAV data. *ISPRS annals of the Photogrammetry and Spatial Information Sciences V-1-2021 XXIV ISPRS Congress* 113-120. <https://doi.org/10.5194/isprs-annals-V-1-2021-113-2021>
- Adriano, B., Yokoya, N., Xia, J., Miura, H., Liu, W., Matsuoka, M., and Koshimura, S., 2021. Learning from multimodal and multitemporal earth observation data for building damage. *ISPRS J. Photogramm. and Remote Sens.* 175, 132-143. <https://doi.org/10.1016/j.isprsjprs.2021.02.016>
- Parven, A., Pal, I., Witayangkurn, A., Pramanik, M., Nagai, M., Miyazaki, H., Wuthisakkaroong, C., 2022. Impacts of disaster and land-use change on food security and adaptation: Evidence from the delta community in Bangladesh. *International Journal of disaster risk reduction*, 103119. <https://doi.org/10.1016/j.ijdrr.2022.103119>
- King, D., and Gurtner, Y., 2021. Land Use Planning for Demographic Change After Disasters in New Orleans, Christchurch and Innisfail. In: Karácsonyi, D., Taylor, A., Bird, D. (eds) *The Demography of Disasters*. Springer, Cham. https://doi.org/10.1007/978-3-030-49920-4_6
- Ishihara, M., and Tadono, T., 2017. Land cover changes induced by the great east Japan earthquake in 2011. *Sci Rep* 7, 45769. <https://doi.org/10.1038/srep45769>
- Reif, M., Macon, C.L., and Wozencraft J.M., 2011. Post-Katrina Land-Cover, Elevation, and Volume Change Assessment along the South Shore of Lake Pontchartrain, Louisiana, U.S.A. *Journal of Coastal Research* 62:30-39. https://doi.org/10.2112/SI_62_4
- Hollander, J. B., Pallagst, K., Schwarz, T., & Popper, F. J., 2009. Planning shrinking cities. *Progress in Planning*, 72(4):223–232. <https://ssrn.com/abstract=1616130>
- Ingram, J.C., Franco, G., Rio, C.R., Khazai, B., 2006. Post-disaster recovery dilemmas: challenges in balancing short-term and long-term needs for vulnerability reduction. *Environ. Sci. & Policy* 9(7-8), 607-613. <https://doi.org/10.1016/j.envsci.2006.07.006>
- Zhao, L., Zhou, S., Zhong, J., Ao, Y., Wang, Y., Wang, T., and Chen, Y., 2022. Rural Post-Earthquake Resettlement Mode Choices: Empirical Case Studies of Sichuan, China. *Frontiers in Public Health*, 10, 861497. <https://doi.org/10.3389/fpubh.2022.861497>

- Yi, P., Shen, L., Tan, C., Tan, D., and Wang, H., 2012. Critical determinant factors (CDFs) for developing concentrated rural settlement in post-disaster reconstruction: a China study. *Nat Hazards* 66, 355–73. <http://www.doi.org/10.1007/s11069-012-0488-7>
- Yabe, T., Tsubouchi, K., Fujiwara, N., Sekimoto, Y., Ukkusuri, S.V., 2020. Understanding post-disaster population recovery patterns. *J. R. Soc. Interface* 17: 20190532. <http://dx.doi.org/10.1098/rsif.2019.0532>
- Lima, M.F., and Eischeid, M.R., 2017. Shrinking cities: rethinking landscape in depopulating urban contexts, *Landscape Research*, 42:7, 691-698. <http://www.doi.org/10.1080/01426397.2017.1372167>
- Wilson, T., Grossman, I., Alexander, M., Rees, P., and Temple, J., 2022. Methods for Small Area Population Forecasts: State-of-the-Art and Research Needs. *Popul Res Policy Rev* 41, 865–898. <https://doi.org/10.1007/s11113-021-09671-6>
- Litman, T., 2021. Evaluating Accessibility for Transport Planning. *Victoria Transp. Policy Inst.* 1–62. <https://www.vtpi.org/access.pdf>
- Krizek, K.J., 2003. Operationalizing Neighborhood Accessibility for Land Use-Travel Behavior Research and Regional Modeling. *J. Plan. Edu. Res.* 22 (3), 270–287. <https://doi.org/10.1177/0739456X02250315>
- Wegener, M., Fuerst, F., 2004. Land-Use Transport Interaction: State of the Art. Integration of Transport and Land Use Planning (TRANSLAND) 2A. SSRN Electronic J. 1–119. <https://doi.org/10.2139/ssrn.1434678>
- Levinson, D.M., Wu, H., 2020. Towards a general theory of access. *J. Transp. Land Use* 13 (1), 129–158. <https://doi.org/10.5198/jtu.2020.1660>
- Hynynen, A., 2006. Node-Place-Model-A strategic tool for regional land use planning. *Nordisk arkitekturforskning - Nordic J. Arch. Res.* 18 (4), 21–29. <http://arkitekturforskning.net/na/article/view/164>
- Clifton, K., Ewing, R., Knaap, G., Song, Y., 2008. Quantitative analysis of urban form: A multidisciplinary review. *J. Urbanism: Int. Res. Placemaking Urban Sust.* 1 (1), 17–45. <https://doi.org/10.1080/17549170801903496>
- Ben-Akiva, M., Bowman, J.L., and Gopinath, D., 1996. Travel demand model system for the information era. *Transportation* 23, 241–266. <https://doi.org/10.1007/BF00165704>
- Bowman, J.L., Ben-Akiva, M.E., 2001. Activity-based disaggregate travel demand model system with activity schedules. *Transp. Res. Part A: Policy Pract.* 35 (1), 1–28. [https://doi.org/10.1016/S0965-8564\(99\)00043-9](https://doi.org/10.1016/S0965-8564(99)00043-9)
- Handy, S.L., Boarnet, M.G., Ewing, R., Killingsworth, R.E., 2002. How the built environment affects physical activity. *Am. J. Prev. Med.* 23 (2), 64–73. [https://doi.org/10.1016/S0749-3797\(02\)00475-0](https://doi.org/10.1016/S0749-3797(02)00475-0)
- Yang, Y., Diez-Roux, A.V., 2012. Walking Distance by Trip Purpose and Population Subgroups. *Am. J. Prev. Med.* 43(1), 11–19. <https://doi.org/10.1016/j.amepre.2012.03.015>
- Mohammadi, N., Taylor, J.E., 2019. Recurrent Mobility: Urban Conduits for Diffusion of Energy Efficiency. *Sci. Rep.* 9 (20247), 1–12. <https://doi.org/10.1038/s41598-019-56372-4>
- Hagerstrand, T., 1970. What about People in Regional Science? *Papers Reg. Sci.* 24 (1), 7–24. <https://doi.org/10.1111/j.1435-5597.1970.tb01464.x>
- Chapin, F.S., 1974. Human activity patterns in the city: Things people do in time and in space, Wiley- Interscience, New York, 507 pp. <https://www.abebooks.com/978047145639/Human-activity-patterns-city-Things-0471145637/plp>
- Jones, P.M., 1977. New Approaches to Understanding Travel Behaviour: The Human Activity Approach. Working Paper 28. Transp. Stud. Unit, University of Oxford, Oxford, England.

- Goodwin, P., Kitamura, R., and Meurs, H., 1990. Some Principles of Dynamic Analysis of Travel Behavior. In: Jones, P. (ed.), *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*, Aldershot: Avebury, England. <https://hdl.handle.net/2066/140605>
- Moeckel, R., Wegener, M., Spiekermann, K., 2003. Creating a Synthetic Population. In: 8th International Conference on Computers in Urban Planning and Urban Management (CUPUM), 1–18.
- Beckman, R.J., Baggerly, K.A., and McKay, M.D., 1996. Creating Synthetic Baseline Populations. *Transportation Research*, 30 (6), 415-429. [https://doi.org/10.1016/0965-8564\(96\)00004-3](https://doi.org/10.1016/0965-8564(96)00004-3)
- Waddell, P., Borning, A., Noth., M., Freier, N., Becke, M., Ulfarsson, G., 2003. Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim. *Netw. Spat. Econ.* 3, 43–67. <https://doi.org/10.1023/A:1022049000877>
- Orcutt, G.H., Greenberger, M., Rivlin, A., Korbel, J., 1961. *Microanalysis of Socioeconomic Systems: A Simulation Study*. New York: Harper and Row.
- Miller, E.J., 2018. The case for microsimulation frameworks for integrated urban models. *J. Transp. Land Use* 11 (1), 1025–1037. <https://doi.org/10.5198/jtlu.2018.1257>
- Wegener, M., 1995. Current and future land-use models. In: G. A. Shunk, P. L. Bass, C. A. Weatherby, & L. J. Engelke (Eds.), *Travel model improvement program land-use modeling conference proceedings*, 13–40. Washington, DC: Travel Model Improvement Program.
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2), 263–292. <https://doi.org/10.2307/1914185>
- Kitamura, R., 2009. Life-style and travel demand. *Transp.* 36, 679–710. <https://doi.org/10.1007/s11116-009-9244-6>
- Miller, E.J., and P.A., Salvini., (2005). ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems. *Netw. Spat. Econ.* 5, 217–234. <https://doi.org/10.1007/s11067-005-2630-5>
- Wagner, P., Wegener, M., 2007. Urban Land Use, Transport and Environment Models. *disP - Plan. Rev.* 43 (170), 45–56. <https://doi.org/10.1080/02513625.2007.10556988>
- Zhu, Y., Diao, M., Ferreira, J., Zegras, C., 2018. An integrated microsimulation approach to land-use and mobility modeling. *J. Transp. Land Use* 11 (1), 663–1659. <https://doi.org/10.5198/jtlu.2018.1186>
- Zondag, B., de Bok, M., Geurs, K.T., Molenwijk, E., 2015. Accessibility modeling and evaluation: The TIGRIS XL land-use and transport interaction model for the Netherlands. *Comp. Environ. Urb. Syst.* 49, 115–125. <https://doi.org/10.1016/j.compenvurbsys.2014.06.001>
- Zhu, Q., Guo, X., Deng, W., Shi, S., Guan, Q., Zhong, Y., Zhang, L., Li, D., 2022. Land-Use/Land-Cover change detection based on a Siamese global learning framework for high spatial resolution remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* 184, 63-78. <https://doi.org/10.1016/j.isprsjprs.2021.12.005>
- Liu, B., Song, W., Meng, Z., Liu, X., 2023. Review of Land Use Change Detection—A Method Combining Machine Learning and Bibliometric Analysis. *Land*, 12, 1050. <https://doi.org/10.3390/land12051050>
- Xu, C., Du, X., Yan, Z., Zhu, J., Xu, S., and Fan, X., 2022. VSDF: A variation-based spatiotemporal data fusion method. *Remote Sens. Environ.* 283, 113309. <https://doi.org/10.1016/j.rse.2022.113309>
- Zhu, Z., Qiu, S., and Ye, S., 2022. Remote sensing of land change: A multifaceted perspective. *Remote Sens. Environ.* 282, 113266. <https://doi.org/10.1016/j.rse.2022.113266>
- Zhao, K., et al., 2019. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote Sens. Environ.* 232, 111181. <https://doi.org/10.1016/j.rse.2019.04.034>

- Huang, X., Weng, C., Lu, Q., Feng, T., and Zhang, L. 2015. Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas. *Remote Sens.* 7, 16024-16044. <https://doi.org/10.3390/rs71215819>
- Li, Y., Zhou, Y., Zhang, Y., Zhong, Wang, L. J., Chen, J., 2022. DKDFN: Domain Knowledge-Guided deep collaborative fusion network for multimodal multitemporal remote sensing land cover classification. *ISPRS J. Photogramm. Remote Sens.* 186, 170-189. <https://doi.org/10.1016/j.isprsjprs.2022.02.013>
- Burke, M., Driscoll, A., Lobell, D. B., and Ermon, S., 2021. Using satellite imagery to understand and promote sustainable development. *Sci.* 371. <http://www.doi.org/10.1126/science.abe8628>
- Zhou, M., Li, D., Liao, K., and Lu, D., 2023. Integration of Landsat time-series vegetation indices improves consistency of change detection. *Int. J. of Digital Earth* 16, 1276-1299. <http://www.doi.org/10.1080/17538947.2023.2200040>
- Saha, S., Ebel, P., and Zhu, X.X., 2022. Self-Supervised multisensor change detection. *IEEE Transac. On Geosci. and Remote Sens.* 60, 4405710. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9538396>
- Yue, J., Fang, L., Ghamisi, P., Xie, J. L., Chanussot, J., and Plaza, A., 2022. Optical remote sensing image understanding with weak supervision: Concepts, methods, and perspectives. *IEEE Geosci. and Remote Sens. Mag.* 10, 250-269. <http://doi.org/10.1109/MGRS.2022.3161377>
- Desai, S. and Ghose, D., 2022. Active learning for improved semi-supervised semantic segmentation in satellite images: in 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 1485-1495. <http://www.doi.org/10.1109/WACV51458.2022.00155>
- Wang, Y., Albrecht, C. M., Braham, N. A. A., Mou, L., and Zhu, X. X., 2022. Self-supervised learning in remote sensing: A review: in *IEEE Geosci. and Remote Sens. Mag.* 10, 213-247. <http://www.doi.org/10.1109/MGRS.2022.3198244>
- Stehman, S. V., 1999. Basic probability sampling designs for thematic map accuracy assessment. *Int. J. Remote Sens.* 20, 2423-2441. <https://doi.org/10.1080/014311699212100>
- Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., 2022. Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes. *Remote Sens. Environ.* 268, 1-17. <https://doi.org/10.1016/j.rse.2021.112780>
- Inglaña, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sens.* 9, 95. <https://doi.org/10.3390/rs9010095>
- Wessels, K., Van Den Bergh, F., Roy, D., Salmon, B., Steenkamp, K., MacAlister, B., Swanepoel, 955 D., Jewitt, D., 2016. Rapid land cover map updates using change detection and robust random forest classifiers. *Remote Sens.* 8, 888. <https://doi.org/10.3390/rs8110888>
- Song, C., Woodcock, C. E., Seto, K. C., Pax-Lenney, M., Macomber, S. A. 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sens. Environ.* 75, 230–244. [https://doi.org/10.1016/S0034-4257\(00\)00169-3](https://doi.org/10.1016/S0034-4257(00)00169-3)
- Tong, X., Xia, G., Lu, Q., Shen, H., Li, S., You, S., and Zhang, L., 2020. Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sens. Environ.* 237, 1-20. <https://doi.org/10.1016/j.rse.2019.111322>
- Potapov, P., Hansen, M. C., Kommareddy, I., Kommareddy, A., Turubanova, S., Pickens, A., Adusei, B., Tyukavina, A., Ying, Q., 2020. Landsat analysis ready data for global land cover and land cover change mapping. *Remote Sens.* 12, p.16. <https://doi.org/10.3390/rs12030426>
- Frankenberg, E., Laurito, M. M., and Thomas, D., 2015. Demographic Impact of Disasters. Wright, James D. (Ed.) (pp. 101-108). Oxford, England: Elsevier.

E. Frankenberg, C. Sumantri, and D. Thomas, 2020. Effects of a natural disaster on mortality risks over the longer term. *Nat Sustain* 3, 614–619. <https://doi.org/10.1038/s41893-020-0536-3>

Frankenberg, E., Ingwersen, N., Iwo, R., Sumantri, C., and Thomas, D., 2023. Impacts of Disaster-Induced Death and Destruction on Health and Mortality over the Longer-Term. Chapter 1 in: Environmental Impacts on Families: Change, Challenge, and Adaptation. McHale, Susan, Valarie King, and Jennifer E. Glick (eds). National Symposium on Family Issues. Springer. <https://ipl.econ.duke.edu/dthomas/docs/ppr/23jan-FrankenbergIngerswenIwoSumantriThomas.pdf>

Lee, D. S., Batyra, E., Castro, A., and Wilde, J., 2023. Human fertility after a disaster: a systematic literature review. *Proc. R. Soc. B* 290: 20230211. <https://doi.org/10.1098/rspb.2023.0211>

Hossain S, Spurway K, Zwi A.B., Huq N.L., Mamun R, Islam R, Nowrin I, Ether S, Bonnitcha J, Dahal N and Adams A.M. (2017) What is the impact of urbanisation on risk of, and vulnerability to, natural disasters? What are the effective approaches for reducing exposure of urban population to disaster risks? London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Gruebner, O., Lowe, S.R., Sampson, L., and Galea, S., 2015. The geography of post-disaster mental health: spatial patterning of psychological vulnerability and resilience factors in New York City after Hurricane Sandy. *Int J Health Geogr* 14, 16. <https://doi.org/10.1186/s12942-015-0008-6>

Park, S., Yao, T. & Ukkusuri, S.V., 2024. Spatiotemporal heterogeneity reveals urban-rural differences in post-disaster recovery. *npj Urban Sustain* 4, 2. <https://doi.org/10.1038/s42949-023-00139-4>

CHAPTER 3

EVALUATING URBAN GROWTH STRUCTURES FOR RESETTLEMENT AREA DETECTION AFTER THE 2011 GREAT EAST JAPAN EARTHQUAKE

Martin Garcia-Fry and Osamu Murao

Abstract: Post-disaster population growth is one of the most essential requirements for sustainable recovery plans in a declining population context. Yet, studies that aim to explore the relationship between population growth and relocation sites do not consider geographic features to localize urban structures driving population growth, which is the purpose of this chapter. Evaluating these features in an era of big and open data is vital to secure sustainable resettlement sites. A ten-year interval before and after the 2011 Great East Japan Earthquake (GEJE) and Tsunami is assumed to be a strategic window for census-based population change data processing. A regression method is used to seek relationships between pre- and post-disaster data in rural and peri-urban areas located in two prefectures of the Tohoku region. Geographic features are extracted in a Geographic Information System (GIS) to characterize population growth areas using urban structures to find future settlement locations. Results demonstrate that pre-disaster population trends were not significantly changed and depict accessible built-up agglomerations surrounded by farmlands as key drivers of population growth in Sendai's peri-urban areas. Therefore, local governments can find sustainable resettlement areas through urban feature detection in the prospective areas of interest. This chapter renders a robust method for resettlement area detection in support of disaster recovery plans.

1. Introduction

The 2011 Great East Japan Earthquake (GEJE) and Tsunami caused massive destruction along the Japanese Pacific coastline from Ibaraki to Hokkaido prefectures instantly displacing more than 470,000 people from their homes. A smooth post-disaster transition from displacement to reconstruction and recovery is essential for the continuity of affected populations and businesses.

However, implementing these phases to meet the expectations of aid donors and beneficiaries requires several efficient decisions based on a series of operationally expensive tasks amid a plethora of interests and calamities. In addition, many past cases of resettlement allocated resources to housing stock and infrastructure needs, but settlement locations are rarely integrated to meet people's needs.

The reconstruction and recovery of housing and communities was at the core of recovery plans after the GEJE. However, government reconstruction support administered through collective relocation projects resulted in the formation of low-density areas and the separation of jobs and residences (Tanaka 2023). Only one third of the victims decided to join collective relocation projects causing population movement and community disintegration (Abe 2014). Therefore, how can we find resettlement areas that tighten communities and sustain population growth in a depopulating region?

The purpose of this chapter is to evaluate the relationship between post-disaster population growth and urban structures through a geographic lens to clarify whether pre-disaster data can be used to detect resettlement areas likely to undergo population growth. First, we collect 2000-2021 government-recorded panel data for 20 municipal areas and regress temporal (demographic) and static (geographic) features to find correlations between pre- and post-disaster population changes. We also compare rural and peri-urban areas where population growth is a stable trend. Then, through geographic feature detection, urban structures are extrapolated from population growth areas to inform local governments as to where future resettlement is likely to sustain urban growth.

Two contributions are reported in this study: First, several studies have monitored longitudinal change in terms of fertility (Lee et al. 2023), mortality (Nakasu 2021), and migration (Abe 2014; Yabe et al. 2020), but demographic studies in the post-disaster context have not combined population growth and geographic features to locate urban structures driving population growth in a depopulating region. Second, population growth was found associated to small municipal areas and distance away from train stations increasing after 2011. This is not a common measure of urban growth which can be used to detect resettlement areas by seeking vehicular accessibility and aggregate built-up land. It

also confirms that reconstruction projects exacerbated pre-disaster population movement to Sendai's peri-urban areas with most people favoring vehicle dependence to access jobs and amenities.

The next section introduces our study area followed by sub-section 3.3 with a detailed description of data and methods applied. Sub-section 3.4 integrates geographic features to detect relationships between rural and peri-urban population changes. Sub-section 3.5 analyzes urban structures with discusses results in sub-section 3.6 drawing concluding remarks.

3.2 Study Area

3.2.1 The 2011 Great East Japan Earthquake, Tohoku, Japan

On March 11, 2011, at 15:00 pm in the afternoon, a sudden and violent crustal deformation generated a 9.0 magnitude earthquake on the Richter scale setting off early warning systems throughout the Japanese nation (Hirose et al. 2011). The earthquake's epicenter shown in Figure 3.1, located off the Sanriku coast of Miyagi prefecture, Tohoku region, northeastern Japan, released a large-scale surge of energy triggering multiple tsunami tidal waves thirty minutes past 15:00 pm (Gusman et al. 2012). Given the rugged coastline in Iwate and Miyagi prefectures and the Tohoku-Fukushima-Ibaraki prefectoral coastal plains, a variety of wave heights were recorded.

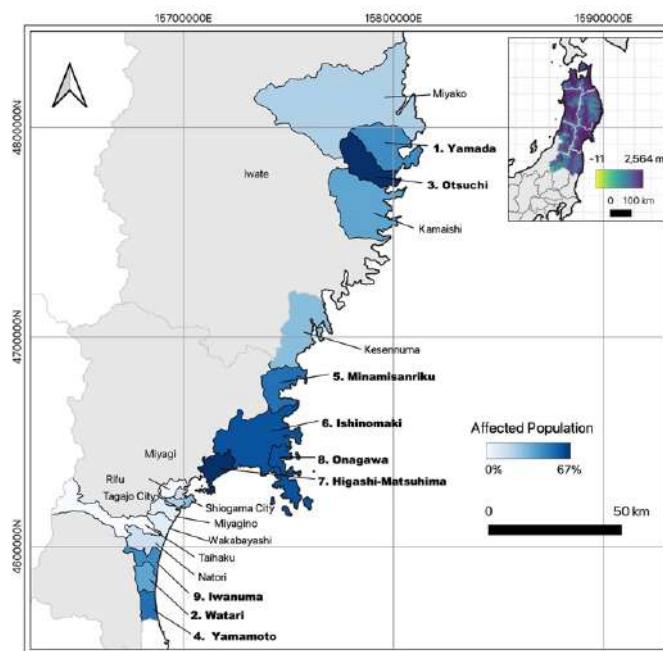


Figure 3.1. Study areas in the Tohoku region. The inset box illustrates elevation data. In addition, we show affected population shares as reported in a study by Miyazawa (2011).

3.3 Data and Methods

Datasets of this chapter can be found in Mendeley Data (Garcia-Fry 2024).

3.3.1 Data Collection

Data tools used in this study pertain to longitudinal tabular information derived from government-recorded census data sets compiled every 5 years and annual municipal population data sets offered in absolute counts. The selection of study areas concealed the following four criteria: First, study areas should be considered the smallest administrative unit under a local government's jurisdiction (Humanitarian Data Exchange, OCHA: <https://data.humdata.org/dataset/cod-ab-jpn>); second, coastal areas were targeted to retain locations damaged by the 2011 tsunami; third, heterogeneous geographies were sought to sustain a non-biased estimation of population dynamics; and fourth, data had to be available for the entire analysis period (2000-2021). The analysis period is divided into two main study periods: Before the disaster (2000-2010) and after the disaster (2012-2021), each of which contains 11 and 10 years, respectively. After defining these selection rules, we proceeded to extract the data from official government sources.

The smallest administrative unit by Japanese standards (shi/ku/machi/cho) is used to retrieve time-series annual demographic features from 20 coastal-affected municipal areas in the Tohoku region (see Table 3.1). The data was extracted from the statistical census bureau of Miyagi [Dataset 1] and Iwate prefectures [Dataset 2], Sendai city statistics [Dataset 3], and e-Stat (a government-led site for statistical data of Japan; [Dataset 4]). Population counts from the 2010 census and surface areas (in square-kilometers) were used to establish population densities. We also segmented locations based on three empirical categories generated to establish levels of damage after the disaster. Here,

affected populations represent the percentage of the 2010 census population count residing in tsunami inundation areas during the disaster event (Miyazawa 2011). All study areas were subsequently sorted in hierarchical order relative to affected populations categories, as either ‘severely affected’ (SA) (>30% affected populations), ‘affected’ (AF) (<30% affected populations), or ‘not affected’ (NA) (<1% affected populations). Categorical interpretations divided the study area into 9 SA, 9 AF, and 2 NA municipal areas. Furthermore, we labeled each municipal area with geography types and inundation rates to ensure the availability of heterogeneous locations for each category of the study area (International Research Institute of Disaster Science 2014).

Table 3.1. Selected municipalities sorted by levels of affected population.

Location	Geography	Inundation Rate (%)	Affected Population *1 (%)	Population in 2010	SA/AF/NA	Prefecture	Shi/Ku/Machi / Cho	Area (km ²)	Population Density (km ²)	Ocean Front	Mean Elevation	Mean Slope	Dist. to Train Stations (km)	
Higashimatsushima	Rias Coast	36%	28800	67%	42903	SA	Miyagi	Shi	101.4	423	36.9	12.1	1.6	1.42
Otsuchi	Rias Coast	2%	9300	61%	15276	SA	Iwate	Machi	200.6	76	17.2	410.0	14.8	10.85
Ishinomaki	Rias Coast	13%	92210	56%	163216	SA	Miyagi	Shi	554.4	294	205.9	82.8	6.7	5.42
Onagawa	Rias Coast	5%	5150	51%	10051	SA	Miyagi	Shi	64.1	157	43.4	133.2	10.3	0.73
Minamisanriku	Rias Coast	6%	8480	49%	17429	SA	Miyagi	Machi	164.0	106	52.5	127.5	7.2	0.19
Yamamoto	Flat Coast	38%	7460	45%	16735	SA	Miyagi	Machi	64.3	260	13.9	43.8	2.8	2.32
Iwanuma	Flat Coast	48%	6570	41%	15969	SA	Miyagi	Shi	61.4	260	12.4	32.6	2.1	0.95
Yamada	Rias Coast	2%	7050	38%	18617	SA	Iwate	Machi	263.0	71	55.3	274.5	12.9	4.43
Watari	Flat Coast	33%	10920	31%	34845	SA	Miyagi	Machi	73.3	475	7.7	17.3	1.5	1.12
Kamaishi	Rias Coast	2%	11390	29%	39574	AF	Iwate	Shi	444.7	89	82.8	410.6	13.7	3.47
Kesennuma	Rias Coast	5%	20880	28%	73489	AF	Miyagi	Shi	333.4	220	93.7	171.1	8.0	4.34
Shichigahama	Flat Coast	39%	5150	24%	21038	AF	Miyagi	Machi	13.2	1588	14.5	13.3	1.9	3.74
Tagajo	Flat Coast	30%	13160	21%	62870	AF	Miyagi	Shi	19.6	3214	0.8	8.6	0.9	0.73
Miyako	Rias Coast	8%	11740	20%	59430	AF	Iwate	Shi	1265.3	47	94.1	530.8	12.9	4.41
Natori	Flat Coast	27%	10430	14%	73134	AF	Miyagi	Shi	97.5	750	6.5	54.7	2.9	1.57
Shiogama	Rias Coast	33%	5270	9%	57469	AF	Miyagi	Shi	17.1	3351	26.6	25.3	2.2	0.68
Miyagino	Flat Coast	16%	12830	7%	183990	AF	Miyagi	Ku	58.2	3163	16.1	15.1	0.8	6.41
Wakabayashi	Flat Coast	16%	6750	5%	128267	AF	Miyagi	Ku	50.8	2525	8.7	6.3	0.3	6.32
Rifu	Rias Coast	1%	90	0%	34859	NA	Miyagi	Machi	45.2	771	3.2	60.0	3.8	1.45
Taihaku	Flat Coast	16%	20	0%	218012	NA	Miyagi	Ku	228.7	953	0.0	369.4	9.8	17.81

*Note: Population Census of 2010 used to determine affected population counts (Miyazawa, 2011).

3.3.2 Demographic Data

Demographic data was pre-processed to establish proportional relative percentages instead of absolute counts. This was done to ameliorate potential disparities caused by the difference between highly-populated areas and remote areas, surface area variations, and geographic features all of which cause demographic profile changes. Therefore, we proceeded to divide each annual demographic count by the total population of that area. This step rendered population changes in percentage (proportional to the annual population count) and cumulative change from 2000-2021 as shown in Figure 3.2. The features used in this study are: (1) Fertility, defined as the annual number of children born per unit area; (2) mortality, as the annual number of deceased people per unit area; (3) in-migration, defined as the annual number of floating populations migrating from a source unit area to a study unit area; and, (4) out-migration, as the annual number of floating populations migrating from a study unit area to a foreign unit area.

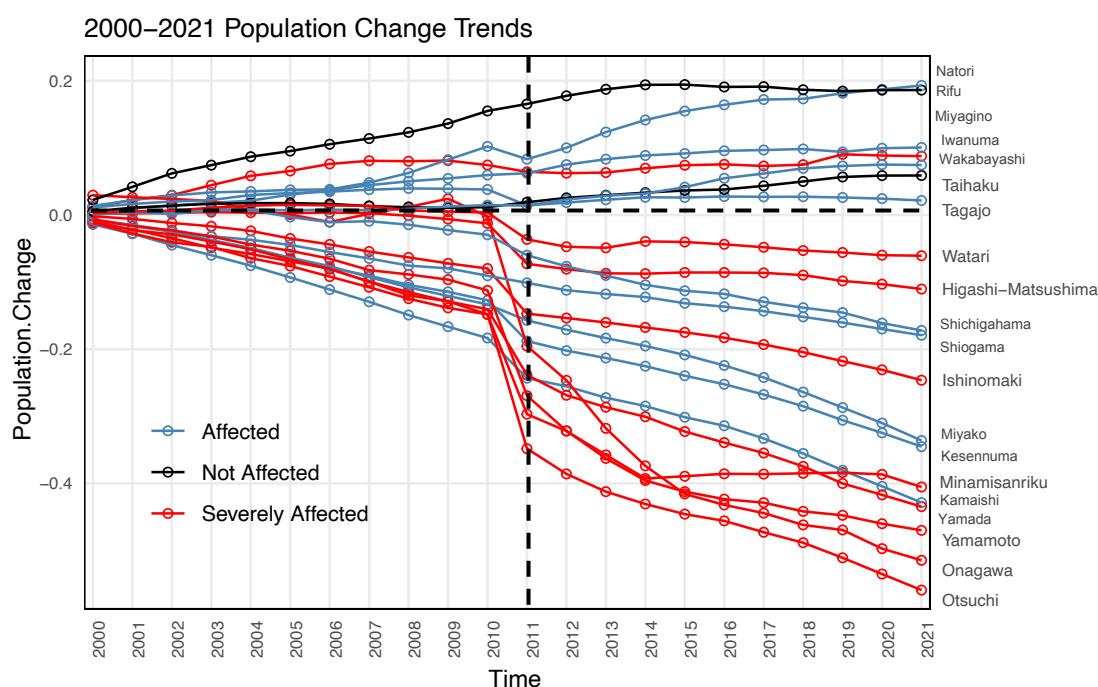


Figure 3.2. Population changes divided into two four analysis groups. The upper left quadrant (G1: 2000-2010) and the upper right quadrant (G2: 2012-2021) represent growing populations in peri-urban areas. The lower left quadrant (G3:

2000-2010) and the lower right quadrant (G4: 2012-2021) represent declining populations in rias coastal areas. All values are expressed in cumulative shares of total population counts.

(1) Fertility and Mortality

To analyze fertility trends before and after the 2011 earthquake, we plot the distribution of counts from 2000 to 2021 and estimate linear regressions to verify data trends. Figure 3.3 shows trends for fertility and mortality rates subdivided into before/after and rural/peri-urban groups for convenience.

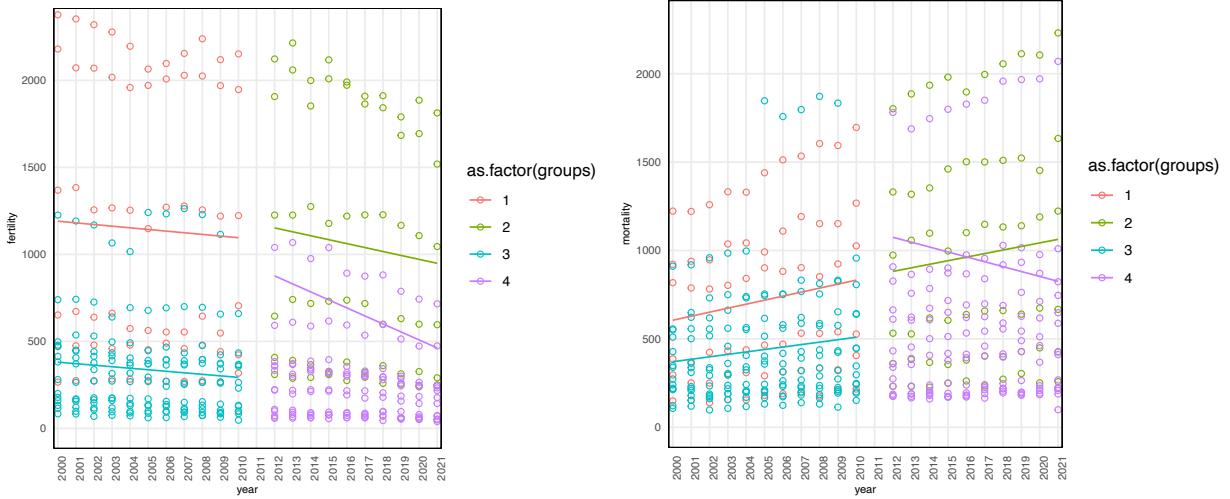


Figure 3.3. Mortality and fertility trends in Tohoku affected coastal areas. Group 1: Pre-disaster peri-urban areas. Group 2: Post-disaster peri-urban areas. Group 3: Pre-disaster rural areas. Group 4: Post-disaster rural areas. Values are expressed in feature counts.

The data exhibits a decreasing fertility rate exacerbated after 2011 with two outlying locations in G4 (rural areas) causing a significant change in slope. It is possible to detect that all other locations have continued to descend in births without significant change in trend caused by the 2011 GEJE. Although G1 and G2 exhibit higher fertility rates, they also saw exacerbated loss after the 2011 event.

Mortality rates denote increasing trends since before 2011 with a change in rural trends 5-6 years after the event, potentially indicating that people in rias coastal areas have gained resilience. Peri-urban areas (G1 and G2) show a stable increase with no apparent change in slope demonstrating that this disaster event did not cause a significant change in trends. Overall, the distribution of data indicates that both fertility and mortality rates have continued pre-disaster trends with fertility slightly affected and mortality moderately changed in rural towns of the Tohoku region. Post-disaster trends in rural areas exhibit higher disaster shocks than peri-urban areas. This can be a consequence of the difference in population density and more importantly, the percentage of affected people in rural areas as opposed to inland peri-urban areas.

(2) Migration

Migration in Figure 3.4 shows distributions in peri-urban areas growing after 2011 with in-migration slightly above out-migration values. In contrast to the pre-disaster in-migration and out-migration value similarities (G1), post-disaster trends show that people in-migrated to Sendai's peri-urban areas in a stable manner while out-migration dictates a steady increase since 2011. Although population is ultimately growing in this area, it is a very gentle growth which does not secure overall peri-urban growth as reported in similar case studies (King and Gurtner 2021).

Rural towns show a slight in-migration increase soon after 2011 in accordance with post-disaster populations temporarily migrating from devastated areas, but returning 1-3 months later as reported by Yabe et al. (2020). However, post-disaster out-migration and in-migration do show overall declining tendencies which means that population dynamics are less frequent in these areas.

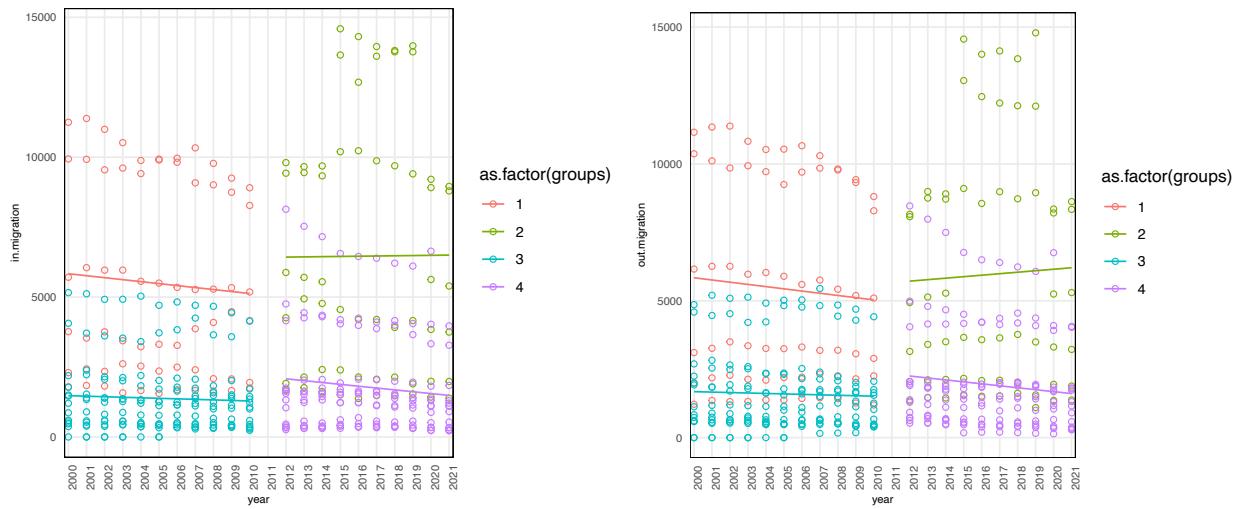


Figure 3.4. In-migration and out-migration trends in Tohoku affected coastal areas. Group 1: Pre-disaster peri-urban areas. Group 2: Post-disaster peri-urban areas. Group 3: Pre-disaster rural areas. Group 4: Post-disaster rural areas. Values expressed in feature counts.

3.3.3 Geographic Data

Geographic features were selected based on gridded population data studies reporting feature performances such as slope/elevation (Nieves et al. 2017) and population density (Hillson et al. 2014). Area (in square-kilometers; km²), population density (in people/km²), ocean front length (in kilometers length), and distance to train stations (municipal centroid to train station Euclidean distance in kilometers) were all sourced using GIS. Further, population density was estimated using 2010 official census data and mean slope and elevation were estimated for the surface area of each municipal boundary using the Multi-Error Removed Improved-Terrain Digital Elevation Model (MERIT-DEM) (Yamazaki et al. 2017). The following Figure 3.5 shows geospatial information applied in this study and the classification of SA, AF, and NA municipal areas.

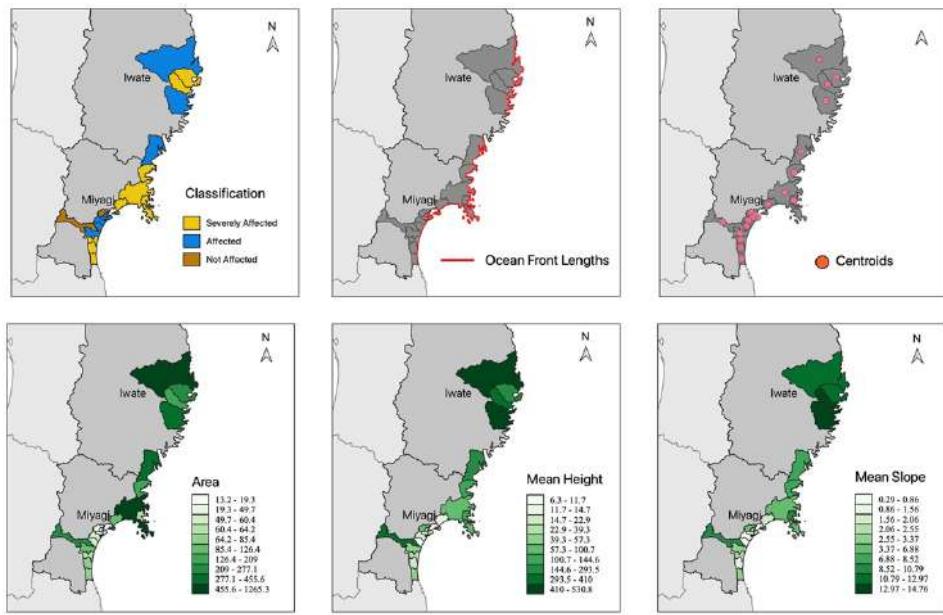


Figure 3.5. Geospatial data collected to characterize population changes in Tohoku affected coastal areas. Here mean elevation and slope data is estimated using basic statistical tools in GIS.

3.4 Geographic Feature Integration

3.4.1 Peri-Urban and Rural Areas

This sub-section utilizes traditional regression models to explore the relationship between population change and geographic features in rural and peri-urban areas. First, a preliminary regression was performed using all geographic features as the independent variables and population counts in rural and peri-urban areas as the dependent variable. The aim of this assessment was to identify correlated features (>0.80) introducing noise into the model. We found that elevation was found correlated with slope and inundation ratios were correlated with area, so the features were eliminated from the set increasing overall R^2 values (goodness-of-fit) from 0.61 to 0.93.

Each data group (G1-4) was separated into single file datasets to find population and geographic feature relationships. Figure 3.6 shows relationships found between G1 (peri-urban) and G3 (rural) in the pre-disaster context and G2 with G4 in the post-disaster context. The data shows

less difference between peri-urban and rural population changes after the disaster event. We can also note a significant difference in population change observing surface area rates. Peri-urban areas exhibit fewer people for every additional km^2 of land, while larger municipalities in rural areas have bigger populations. Population change for every km^2 is -95 (peri-urban) and 535 (rural) which explains the decrease in density for every km^2 in peri-urban areas (-1273) as opposed to rural areas (976). Therefore, population density increases moderately in smaller municipal areas near Sendai, while it grows in larger municipal rias coast areas. Another interesting observation is that peri-urban areas show significant increase of people distanced away from train stations after the 2011 GEJE in contrast to before, while rural areas do not express significant relationships. This may be a consequence of the restoration efforts carried out in Sendai's flat coast areas where newly designated disaster risk zones prompted collective relocation projects further inland (Ishihara and Tadono 2017). It also reflects the location choice of families reported to have moved closer to the city after the tsunami. In addition, mild population growth is found associated with fewer ocean front length in peri-urban areas and longer longer ocean front lengths in rural areas. This observation indicates that municipalities with shorter ocean fronts and deeper inland shapes constituting suburban and peri-urban neighborhoods are targeted and therefore exhibited dynamic growth after the GEJE.

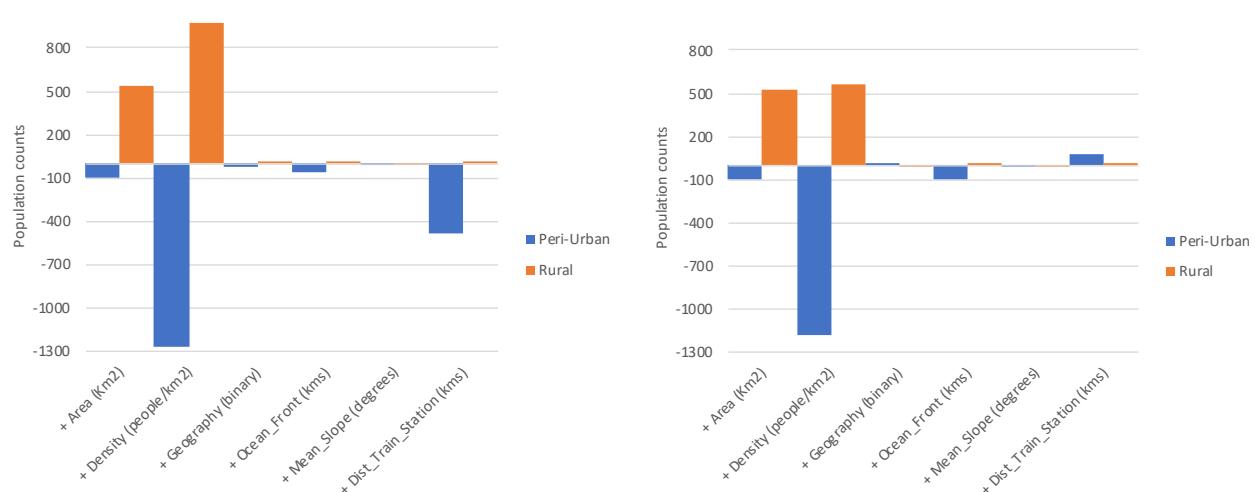


Figure 3.6. Population counts per feature unit increase in peri-urban (G1) and rural areas (G3). Full regressions are reported in Mendely Data (Garcia-Fry 2024). Peri-urban (left): $R^2 = 0.99$. Rural (left): $R^2 = 0.89$. Peri-urban (right): $R^2 = 0.96$. Rural (right): $R^2 = 0.94$.

3.4.2 Peri-Urban Municipal Area Analysis

The next step was to analyze differences among peri-urban municipal areas reflecting people's location choice in relation to geographic and urban features. Given that not all peri-urban municipalities currently experience population growth (e.g., Shiogama, Tagajo, and Shichigahama), we decided to evaluate the differences between growing municipal areas with a specific focus on municipalities with populations that underwent major tsunami inundation damage (i.e., Iwanuma).

As shown in Figure 3.2; Natori, Wakabayashi, and Iwanuma sustain trending growth. However, population and geographic feature relations have been identified only in terms of distance away from train stations (sparsely urbanized land) and short ocean front lengths (proximal to downtown Sendai). Therefore, we compared population changes for G1 and G2 to identify sudden change in urban structure to complement our analysis (Figure 3.7).

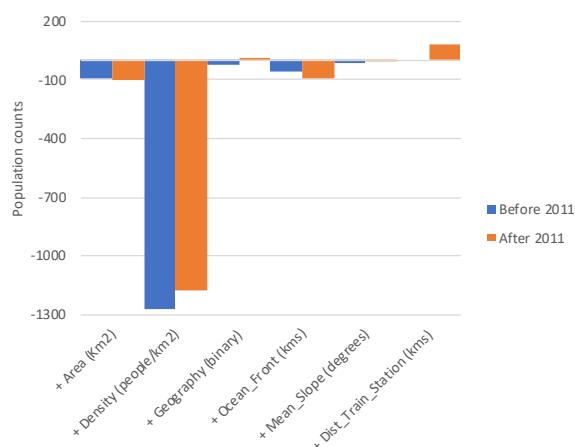


Figure 3.7. Population counts per feature unit increase in peri-urban areas before and after the 2011 GEJE.

The data shows two major differences: First, there is an 8% increase in population density among peri-urban areas and fewer people for every km length of coastline after the event. Slope is

also notable among the features with a 39% increase in location relevance for post-disaster population changes. As mentioned above, population growth increased in places distanced away from train stations. Therefore, we can infer that people's choice includes vehicular accessibility. Given that smaller municipal areas have larger densities, we can also note agglomeration effects in the urban fringe. However, further analysis is required to define which urban structures are intrinsically related.

3.5 Urban Structure Analysis

In order to analyze urban structures, we use the results found above to select sub-municipal areas likely experiencing sustained population growth. For example, vehicular accessibility is high near Sendai's coastal highway. Urban agglomerations in these areas are also hard to miss given the extensive low-density urban structure of suburban areas in Sendai. We also note that Iwanuma was severely damaged with 41% of the population living inside the tsunami inundation areas at the time of the event (Miyazawa 2011). Natori and Wakabayashi with superior growth trends were less affected with 14% and 5%, respectively. Therefore, Iwanuma with the aforementioned characteristics is selected for analysis. This area represents a resilient population growth case considered proximal to the urban center but located on the fringe with Natori, Taihaku, and Wakabayashi in between. To begin our analysis, we find that Iwanuma is a large peri-urban area with a central agglomeration of built-up land surrounded by large tracts of rural farmland and connected to two smaller urban conglomerates. To the east, the Tamaura relocation site is well positioned next to a larger residential set of neighborhoods. Together, they form a cluster surrounded by rich farmlands, rice paddies, and other coastal land cover areas. Intensive agriculture and sparse households to the north leads to an industrial built-up cluster conveniently located next to the coast. To the west of these two clusters is Iwanuma city with an approximate population of 43,964 (2021 population registration census data). Figure 3.8 illustrates the empirical analysis of urban structures over Iwanuma city.

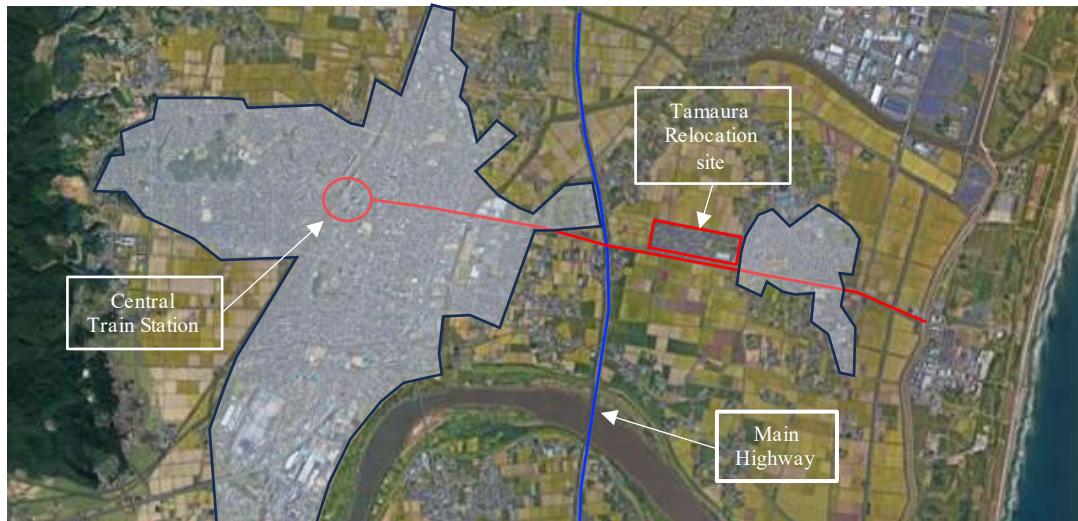


Figure 3.8. Urban structure analysis over Iwanuma municipal area.

Iwanuma has a central train station, surrounded by aggregate built-up areas and peripheral farmlands clearly connected to Tamaura relocation site, which is closely connected to the main highway and the surrounding farmlands. These relationships indicate that:

- Iwanuma is well connected.
- Is compact and distanced from the coast.
- Has a relocation site between two main urban clusters increasing people's access to transit.
- Has a central train station and is considered peri-urban.

3.6 Discussion and Concluding Remarks

3.6.1 Post-Disaster Population Changes

The effects of the 2011 GEJE and Tsunami were quantified in terms of population dynamics relative to ten years before and after the disaster event. While rural and peri-urban municipal areas have various differences, we also see commonalities between pre- and post-disaster trends. In particular, pre-disaster fertility trends were exacerbated by the event and continue to fall in coastal-affected areas. Long-term depopulation trends are accompanied by increased mortality rates after

2011 for 5-6 years. However, we note a trend change in rural areas while peri-urban areas see no change in trends. Peri-urban areas exhibit high in-migration rates, especially after the event, with an estimated 8% increase in population density.

Geographic feature relationships inform users as to where population growth is likely to thrive subject to the infusion of a planned investments in the area. We saw that small municipalities in peri-urban areas garnered higher densities as well as larger municipalities in rural areas. This may be a consequence of proximity to Sendai's central areas. However, small municipalities have manageable sizes leveraging their service, accessibility, and information dissemination. Here, the length of municipal coastal areas is directly associated with fewer populations after 2011. This is considered an interesting effect of new disaster risk designation zones and the subsequent collective relocation of populations to inland areas. Slope also gained importance which may be a product of perceived safety and policies that promoted movement of people to safer areas. All regressions saw agreement ($R^2 \geq 0.89$) with significant geographic feature relationships (99% p-values), except a binary feature (rias or flat coast), inundation area, and elevation correlated with area and slope, respectively.

3.6.2 Urban Structures

This chapter emphasized correlations between population changes in rural and peri-urban analysis areas with the purpose of finding geographic features that accurately guide users to population growth areas. Subsequent to this assessment is the analysis of urban structures as shown in Figure 3.9. This model helps local governments in peri-urban areas conduct a preliminary assessment of the region's internal land structures to extract geographic features that detect potential resettlement areas. However, analysis of population growth areas in contrast to depopulating areas is necessary to rely on urban features that are not disturbed by urbanization processes which may include several informal employment and real estate sub-processes causing spurious output data.

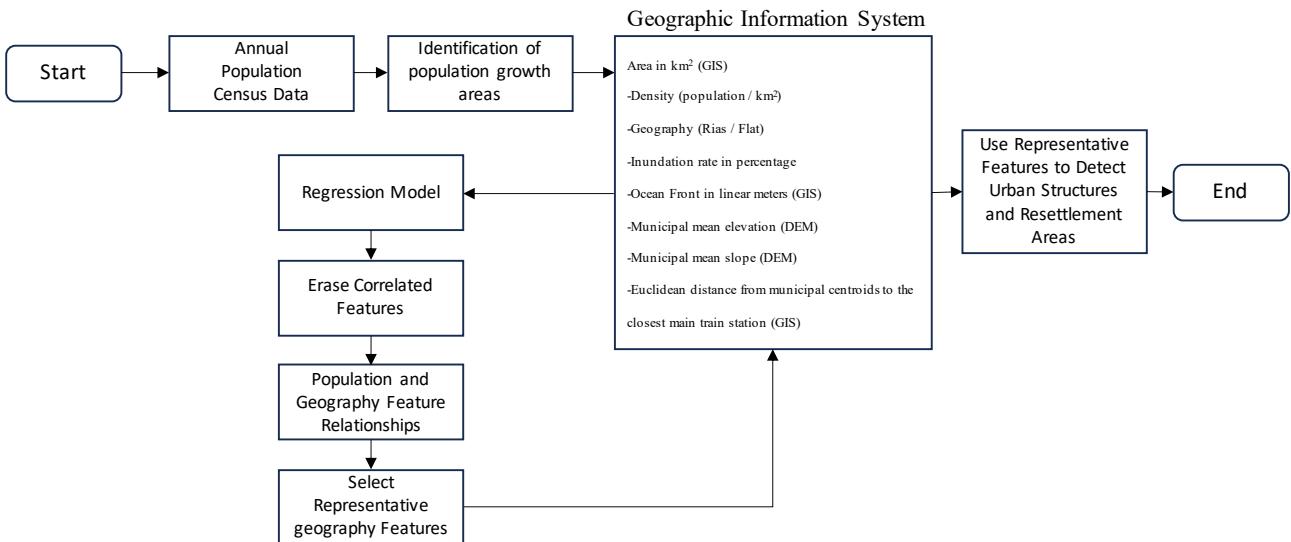


Figure 3.9. Proposed Method 1.

The geographic analysis rendered indications of conglomerate statistical population growth associated to vehicular accessibility, distanced from the ocean and train stations, as well as population damage the day of the event provides significant clues as to where urban growth is possible. Here, Iwanuma city was analyzed to represent a successful case of recovery as indicated in previous studies by Muroi (2022) and Tanaka (2023). Urban structures supplemented our findings with the notion of a central train station surrounded by compact built-up land and peripheral agricultural land including a well-connected coastal highway and more precisely; a relocation site in-between two major urban clusters with very similar land structures. As a result, a resettlement area can be detected in a different municipality where similar urban structures can be found.

3.6.3 Limitations

Limitations of this study are briefly noted below:

- Results are valid in the extent of Sendai's peri-urban area and are not representative of rural areas in the Tohoku region.

- Although the method is limited to regional cities with peripheral suburban and peri-urban areas, it may be applicable to sub-regional cities. However, integration of fine-scale population datasets may be required to enable deep feature analyses in very small urban areas.
- The method requires annual population census data for the past ten years.
- Additional geographic features may render better results and provide more plausible clues for future resettlement area detections.

3.6.4 Concluding Remarks

This chapter evaluated the relationship between post-disaster population growth and urban structures through a geographic lens to clarify whether pre-disaster data can be used to detect resettlement areas likely to undergo population growth. Pre-disaster population data trends were found to be exacerbated by the 2011 GEJE and Tsunami, but did not experience change. This chapter selected population growth areas in the Tohoku region because municipal peri-urban areas have overcome the depopulating trend. This poses a unique setting to find urban features that sustain population growth disassociated with peri-urbanization and informal processes often found in developing countries. Furthermore, the proposed method successfully extracted geographic features correlated with population growth and found that urban agglomerations in small municipal areas with access to main roads increase the probability of growth. The results proved that pre-disaster population data combined with geographic features can be used to locate urban structures likely to enable population growth after disaster events.

Local governments in Sendai's peripheral areas can now look for accessible built-up agglomerations surrounded by farmlands as key drivers of population growth. The method does not guarantee these results for sub-regional cities in the Tohoku region, but it may be useful to process these areas and improve the method. This method marks a pivotal change towards finding sustainable resettlement areas in preparation for future disaster events.

References

- Tanaka, S., 2023. The 2011 Great East Japan Earthquake and Tsunami: The highest casualties and largest reconstruction funds—Characteristics of major disasters and future challenges in developed countries. In Special Issue: Experience of Great Disasters: Revisioning 3/11 and beyond. Japanese Journal of Sociology 32, 7-24. <https://doi.org/10.1111/ijjs.12147>
- Lee, D. S., Batyra, E., Castro, A., and Wilde, J., 2023. Human fertility after a disaster: a systematic literature review. Proc. R. Soc. B 290: 20230211. <https://doi.org/10.1098/rspb.2023.0211>
- Nakasu, T., 2021. A fundamental vulnerability: contributions from population studies. J. Disaster Res. 16 (6), 936-941. <https://doi.org/10.20965/jdr.2021.p0936>
- Abe, T., 2014. Population movement in the Tohoku region after the Great East Japan Earthquake disaster. Science reports of the Tohoku University, 7th Series, Geography, 60, 83-95. <https://tohoku.repo.nii.ac.jp/records/11754>
- Yabe, T., Tsubouchi, K., Fujiwara, N., Sekimoto, Y., Ukkusuri, S.V., 2020. Understanding post-disaster population recovery patterns. J. R. Soc. Interface 17: 20190532. <http://dx.doi.org/10.1098/rsif.2019.0532>
- Hirose, F., Miyaoka, K., Hayashimoto, N., Yamazaki, T., and Nakamura, M., 2011. Outline of the 2011 off the Pacific coast of Tohoku Earthquake (Mw 9.0)—Seismicity: foreshocks, mainshock, aftershocks, and induced activity—. Earth Planet Sp 63, 1. <https://doi.org/10.5047/eps.2011.05.019>
- Gusman, A.R., Tanioka, Y., Sakai, S., Tsushima, H., 2012. Source model of the great 2011 Tohoku earthquake estimated from tsunami waveforms and crustal deformation data,
- Garcia Fry, M., 2024. Resettlement Area Detection Model. Mendeley Data, v1.
- [Dataset 1] Statistical census bureau of Miyagi Prefecture, 統計データ/推計人口(年報). Accessed: May 19, 2023. <https://www.pref.miyagi.jp/soshiki/toukei/suikei-nen.html>
- [Dataset 2] Statistical census bureau of Iwate Prefecture, 岩手の長期時系列データ(人口). Accessed: May 20, 2023. <https://www3.pref.iwate.jp/webdb/view/outside/s14Tokei/tyosaBtKekka.html/I139>
- [Dataset 3] Sendai city statistics, 仙台市の人口(人口動向). Accessed: May 20, 2023. <https://www.city.sendai.jp/chosatoke/shise/toke/jiho/jinko/index.html>
- [Dataset 4] e-Stat, Japan in terms of Statistics. Accessed: May 20, 2023. (<https://www.e-stat.go.jp/>)
- Miyazawa, H., 2011. Population in areas affected by the 2011 tsunami off the Pacific coast caused by the Tohoku Earthquake: From Sanriku coast to Sendai Bay area. The 2011 East Japan Earthquake Bulleting of the Tohoku Geographical Association. <http://tohokugeo.jp/articles/e-contents16.html>
- International Research Institute of Disaster Science, 2014. HFA IRIDEs Review Report: Focusing on 2011 Great East Japan Earthquake. Int. Res. Inst. Disaster. Sci., Tohoku University, Sendai, Japan, 1-62.
- King, D., and Gurtner, Y., 2021. Land Use Planning for Demographic Change After Disasters in New Orleans, Christchurch and Innisfail. In: Karácsonyi, D., Taylor, A., Bird, D. (eds) The Demography of Disasters. Springer, Cham. https://doi.org/10.1007/978-3-030-49920-4_6
- Nieves, J.J., Stevens, F.R., Gaughan A.E., Linard, C., Sorichetta, A., Hornby, G., Patel, N.N., and Tatem, A.J., 2017. Examining the correlates and drivers of human population distributions across low- and middle-income countries. J. R. Soc. Interface. 142017040120170401. <https://doi.org/10.1098/rsif.2017.0401>
- Hillson, R., Alejandre, J.D., Jacobsen, K.H., Ansumana, R., Bockarie, A.S., Bangura, U., Lamin, J.L., Malanowski, A.P., and Stenger, D.A., 2014. Methods for Determining the Uncertainty of Population Estimates Derived from Satellite Imagery and Limited Survey Data: A Case Study of Bo City, Sierra Leone. PLoS ONE 9, e112241. <https://doi.org/10.1371/journal.pone.0112241>

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C. C., Kanae, S., Bates, P.D., 2017. A high-accuracy map of global terrain elevations: Accurate Global Terrain Elevation map, Geophys. Res. Lett. 44, 5844–5853. <https://doi.org/10.1002/2017GL072874>

Ishihara, M., and Tadono, T., 2017. Land cover changes induced by the great east Japan earthquake in 2011. Sci Rep 7, 45769. <https://doi.org/10.1038/srep45769>

Muroi, K., 2022. Post-disaster reconstruction in the rural-urban fringe following the Great East Japan Earthquake. Proceedings of the 13th AIWEST-DR 2021, E3S Web of Conferences 340, 03001. <https://doi.org/10.1051/e3sconf/202234003001>

CHAPTER 4

Automatic Supervision of Land Cover Labels to Increment the Reliability of Training Datasets for Disaster Recovery Change Detection

Martin Garcia-Fry, Osamu Murao, Bruno Adriano, and Shunichi Koshimura

Abstract: Analyzing post-disaster resettlement in the periphery of urban areas is essential to avoid rural fragmentation. With the use of remote sensing satellite images, land change effects of resettlement can be quantified to inform future settlement locations. However, change detection is governed by the reliability of training samples and the efficiency of learning models. Therefore, we formulated an automatic label supervision model to secure reliable land cover samples from multisource remote sensing data and predict accurate land cover maps, aiming to evaluate land change interactions prompted by post-disaster resettlement sites. Pre-existing land cover samples were photo-inspected and then vetted against fused-canopy heights to alleviate interpretation biases. Then, a set of refined labels were held as reference samples while the prevailing quality samples derived statistical amplitude metrics to automatically supervise the feature space and structural height of candidate training samples. A classifier was then optimized to predict Sentinel-2 satellite images over Java, Indonesia. We observed 16.4% forest loss, 18.5% cropland growth, and 15.5% built-up growth as a direct consequence of urban resettlement after the 2010 Mt. Merapi eruptions. With 32% bias-reduced reference samples and 56% error-removed training data, 82%-78% overall accuracies were achieved on the first iteration. Codes available at: https://github.com/martingarciafry/GCC_model.

4.1 Introduction

Natural disaster events have surged during the past five decades increasing the vulnerability of populations at risk (MacManus et al. 2021). Ready-to-use, updated, and accurate land cover datasets are a necessary source of information to determine the extent of damage, response, and recovery after disaster events. The priority of a response stage is to distribute relief and rescue operations, while

recovery stages prioritize progress beyond the reconstruction of buildings and infrastructure (Contreras et al. 2016). A considerable amount of money from donors and governments is destined to finance the recovery process, but allocation and management are crucial to overcoming expectations of development endorsed by the Sendai Framework for Disaster Risk Reduction 2015-2030 (UNISDR 2015).

Monitoring the urban fringe of developing cities helps to understand the dynamics of risk associated with informal growth patterns (Fekete 2022), climate change (Gibson et al. 2016), and inconsistent planning schemes (Wijaya 2018). Multitemporal time-series analyses of remote sensing data have been increasingly applied in these areas (Huang and Zhang, 2022). Several studies have assessed post-disaster damage using optical imagery (Koshimura et al. 2020), SAR images (Moya et al. 2020), and multimodal/temporal datasets (Adriano et al. 2021). However, very few studies examine the effects of recovery on socioeconomic (land use) and environmental (land cover) change. Further, global targets such as biodiversity loss and food security have been linked to land use extent affecting carbon sources and partial land cover degradation (Winkler et al. 2021). This makes urban growth a necessary implication of recovery plans, and quantifying its dynamics, a critical task in supporting these debates.

In practical remote sensing applications, abrupt changes on the pixel level (e.g., change caused by human activity or natural disturbances) are difficult to estimate compared to synchronized change (e.g., vegetation growth and change in soil moisture). This is mainly due to the size of change in fractions of a pixel (Xu et al. 2022). The predictive uncertainty of algorithms can also stream inconsistent time-series data (Zhao et al. 2019). This can be attributed to machine and deep learning models in the computer vision field suffering from insufficient and inefficient labels (Huang et al. 2015). Manual labeling is hampered by the time-consuming process to supervise all samples, while domain expertise is limited to particular sites (Li et al. 2022). Another big challenge for remote

sensing data learners is label noise. Given the variety and complexity of applications, datasets with perfect annotations rarely exist.

Researchers have explored various ways to overcome these issues using unsupervised, semi-supervised, and self-supervised learning models. Transfer learning and domain adaptation networks are unsupervised methods where a model is pre-trained on a source image to be applied on a target image elsewhere (Saha et al. 2022). The accuracy of these networks depends on the difference between source and target data. Therefore, weakly-supervised models were developed entailing: (i) incomplete supervision, where a subset of training labels are available; (ii) inexact supervision, where coarse-grained labels are given; and, (iii) inaccurate supervision, where the given labels are not always true on the ground (Yue et al. 2022). Active learning is another method that requires a small number of labeled samples to draw instances according to some measures (Desai and Ghose, 2022). Likewise, self-supervised models use unlabeled data to identify features and labeled samples to adapt models for downstream tasks (Wang et al. 2022). However, the reliability of labels can be contested due to interpretation biases. Therefore, how can we obtain accurately-labeled samples to increase the reliability of model predictions?

Motivated by the complexity, computing requirements, and labeling bias of existing models, I formulated a simple and automatic label supervision model to secure reliable land cover samples from multisource remote sensing data and predict accurate land cover maps, aiming to evaluate land change interactions prompted by post-disaster resettlement sites. The proposed model utilizes a representative lidar acquisition to query pre-existing land cover samples and rectifies those labels with very-high resolution (VHR) satellite imagery. Then, canopy heights are fused with forest heights to vet the structural height of interpreted classes. These bias-reduced samples are held as reference or metric units to draw static and multitemporal amplitude metrics per class. Metrics converge in a canopy height fusion system to automatically supervise the three-dimensional space of candidate training samples.

The contribution of this study is two-fold: First, amplitude metrics have been used to seek labels via trained models (e.g., Tong et al. 2020), or to classify images directly, but have not been retained from bias-reduced samples to automatically supervise the feature space and vertical height of training samples. Second, bias-reduced reference labels were assessed to measure the size of error inherent to interpretation bias. I found that 32% of the mean variance was removed, thus increasing the reliability of mapping outcomes (Foody 2010). These potential novelties overcome the difficulties in modeling accurate land cover maps.

In the following section, I introduce the study area with a brief methodological description. Section 4.3 summarizes the data employed and Section 4.4 details the proposed model. Section 4.5 evaluates target areas with results discussed in Section 4.6, drawing concluding remarks.

4.2. Study Area and Methodology

4.2.1 The Cangkringan district, Yogyakarta, Java, Indonesia

From October 26 to November 30, 2010, the Merapi volcano erupted thrice displacing thousands of people from their homes (World Bank, 2012). The Cangkringan district, located in Yogyakarta, Java, Indonesia, was severely affected driving the relocation of 2,608 households onto 23 urban resettlement sites (Garcia-Fry et al. 2022).

Indonesia is situated at the intersection of three tectonic plates (the Eurasian, Australian, and Pacific plates) with 129 active volcanos and a large population living in close range (30 km) to a volcano (Orynbaikyzy et al. 2023). Java itself holds 33 active volcanos with Mt. Merapi being the most active stratovolcano in Indonesia (Rani and Khotimah 2021). Much of this land is explicitly fertile due to volcanic ash and sand deposits (Utami et al. 2018). Therefore, local farmers in the Cangkringan district choose to capitalize on increased yield after eruptions, but they are exposed to Yogyakarta's spillover effects in unprepared, informal, and high-risk areas. These peri-urban areas

are transitional spaces where neither decent urban services nor rural resource provision exist (Hutchings et al. 2022).

Given the extent of vulnerable areas in Java, a 10 km grid over >13.4 Mha of vegetation ecotones was used to partition the territory in a WGS84 geodetic datum (Figure 4.1). The resulting 1,392 tiles (111,100 pixels per tile) subdivide the tropical and subtropical moist broadleaf ecosystems of Java to evaluate target areas in a high spatial resolution domain.

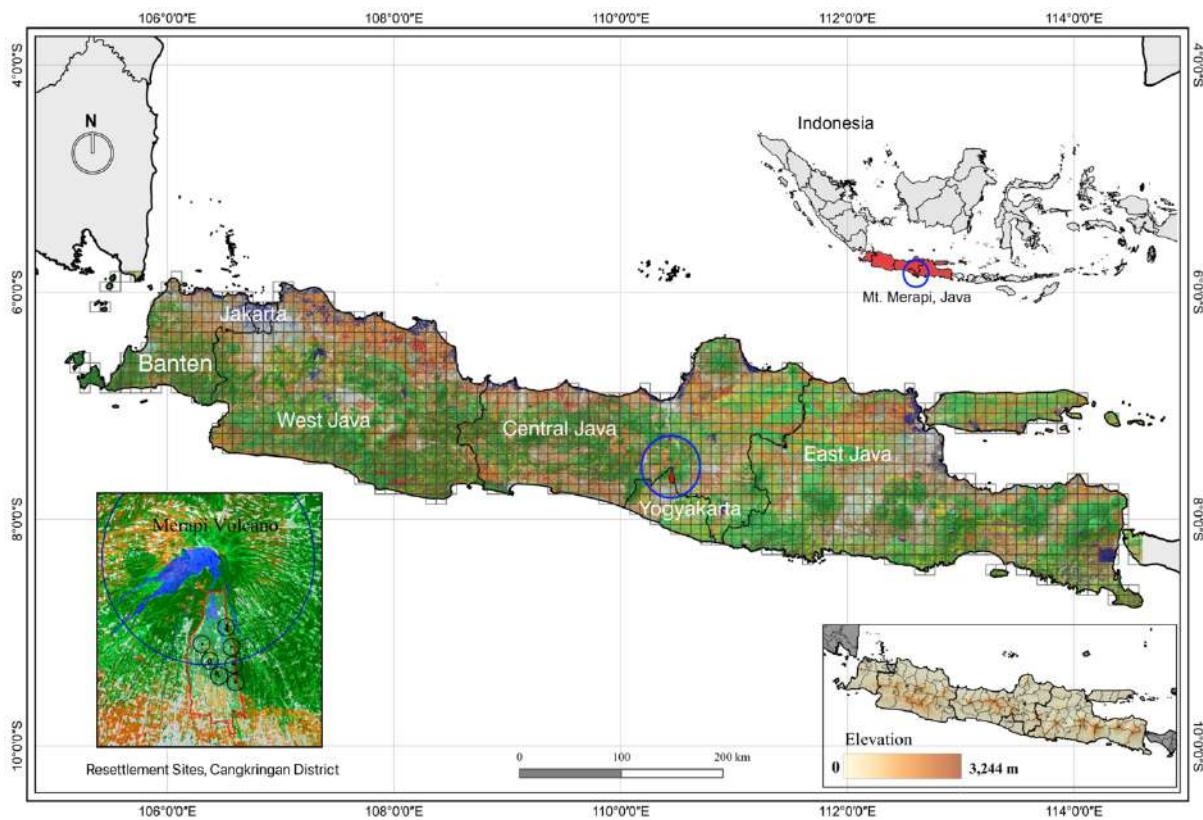


Figure 4.1. Land cover classification of Java. The left inset box shows target resettlement areas near the 10 km hazard-prone radius and the right inset box shows elevation data.

4.2.2 Methodology

The model, termed ground cover change (GCC), predicts satellite images with no limits to spatial resolutions, sensors, nor temporal acquisitions in regional extent. Here, I offer an overview of the

three data processing pipelines, illustrated in Figure 4.2 and detailed in sub-section 4.4.1 through sub-section 4.4.3. These steps do not require heavy computational resources.

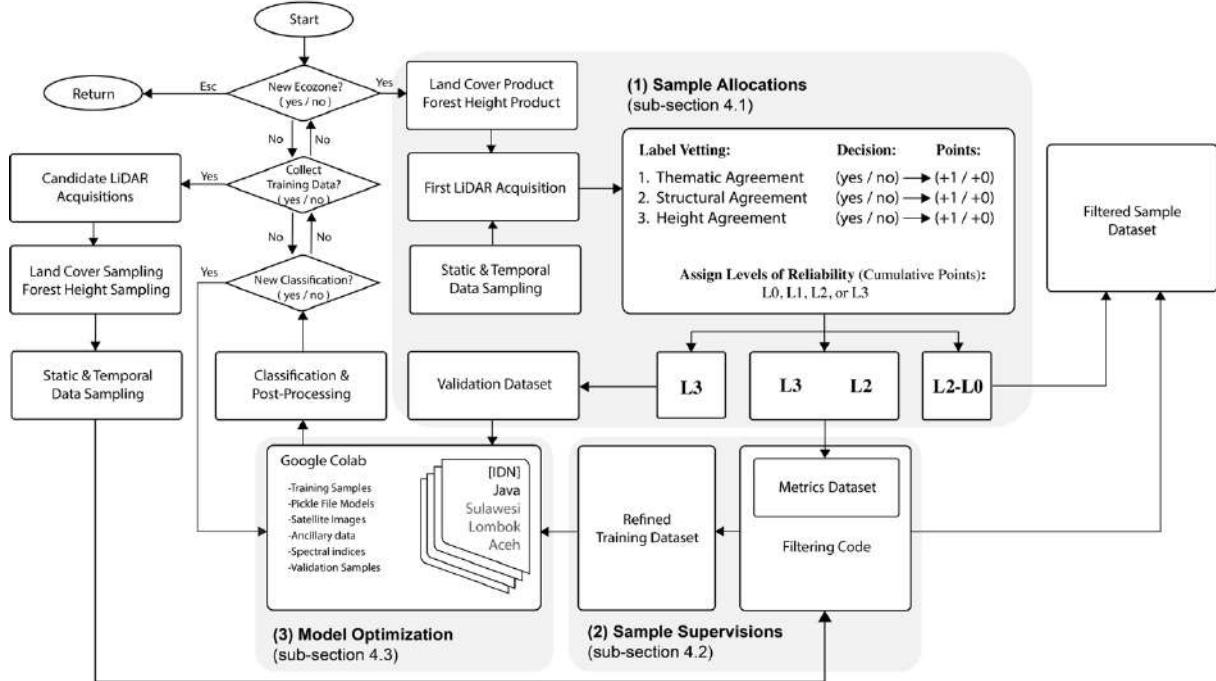


Figure 4.2. Proposed method 2.

(1) Sample allocations: First, a lidar acquisition sampled land cover (Hansen et al. 2022) and forest height (Potapov et al. 2021) pixel landings. Labeled samples were inspected and vetted against fused values of forest structure rendering three binary scores. Each cumulative score derived a level (L) of reliability to segment validation and metric samples with static and temporal features listed in Table 4.1.

(2) Sample supervisions: Multinomial regressions were subsequently performed over discrete metric samples to extrapolate statistically significant features. Their relative bounds were extracted and refined to supervise candidate training samples.

(3) Model optimization: Feature correlations and overfitting trials optimized SHAP (Shapley Additive exPlanations) values (Lundberg and Lee 2017) to detect land use features.

Table 4.1. Features used for image classifications.

Type	Resolution	Acronym/Indices	Dataset	Description/Formula
Lidar	11 m footprint	ICESat-2 (Canopy heights)	ATL08_20190904163339	First ground track acquisition
	11 m footprint	ICESat-2 (Canopy heights)	ATL08_20190904163339 ATL08_20190429105420 ATL08_20190523214429 ATL08_20190626080618 ATL08_20190713184723 ATL08_20190802180554 ATL08_20190621202031 ATL08_20190724184808	Training acquisitions
	11 m footprint	ICESat-2 (Canopy heights)	ATL08_20190605205350 ATL08_20190610183939	Post-informed acquisitions
	~30 m	GLCLU	Global Land Cover Land Use, 2019	Global land cover product
	~30 m	Forest Height	Global Forest Canopy Height, 2019	Forest canopy height product
	~90 m	MERIT-DEM	Multi-Error-Removed Improved-Terrain DEM	Digital Elevation Model
	~90 m	Slope	Slope	Ancillary Dataset
	~90 m	TSRI	Transformed Solar Radiation Index	Ancillary Dataset
	~90 m	TWI	Geomorpho90m-Topographic Wetness Index	Ancillary Dataset
	~10m, 20m	TCT	Tasseled Cap Transformation	Coefficients for TCG and TCW -0.3599*B -0.3533*G- 0.4734*R+0.6633*NIR-
Sentinel-2	~10m, 20m	TCG	Tasseled Cap Greenness	0.0087*SWIR1-0.2856*SWIR2 0.2578*B+0.2305*G+0.0883*R+0.1 071*NIR-0.7611*SWIR1- 0.5308*SWIR2
	~10m, 20m	TCW	Tasseled Cap Wetness	NIR-R/NIR+R NIR-G/NIR+G NIR-SWIR1/NIR+SWIR1
	~10m	NDVI	Normalized Difference Vegetation Index	G-NIR/G+NIR
	~10m	GNDVI	Green Normalized Difference Vegetation Index	R-SWIR2/R+SWIR2
	~10m	NDWI 1	Normalized Difference Water Index 1	NIR-SWIR2/NIR+SWIR2
	~20m	GN	Normalized Difference Water Index 2	(NIR-R/NIR+R+0.5)*1.5
	~10m, 20m	RS2	Normalized Difference Flooding Index	(B7-R / (B5 / B6))
	~20m	NS2	NIR-SWIR2	B-NIR/B+NIR
	~10m	SAVI	Soil-Adjusted Vegetation Index	SWIR1-SWIR2/SWIR1+SWIR2
	~10m, 20m	IRECI	Inverted Red-Edge Chlorophyll Index	
Wetness Index	~10m	BN	Vegetation index	
	~20m	S12	Vegetation index	

4.3 Data

Mapping products, shape files, and datasets supporting findings of this study can be found in Mendeley Data (Garcia-Fry 2023).

4.3.1 Lidar

Lidar is a unique source of comprehensive data for forest inventories and ice detection worldwide. Two spaceborne lidar instruments were launched in 2018, (1) the Global Ecosystem Dynamics Investigation (GEDI) on the International Space Station (ISS; Dubayah et al. 2020), and (2) the Advanced Topographic Laser Altimeter System (ATLAS) on-board the Ice, Cloud, and land Elevation Satellite-2 (ICESat-2; Neuenschwander and Pitts 2019).

A study by Liu et al. (2021) found that ICESat-2's Land and Vegetation Height (ATL08) data product presented fewer noise and almost same accuracy to GEDI data for nighttime and strong beam

acquisitions. However, the best estimate of vegetation structures is reasoned by the fusion of high-powered laser systems. Therefore, I selected a 2019 GEDI-derived Global Forest Canopy Height product (Potapov et al. 2021) and ICESat-2 (ATL08, RH98) in 20 m (`h_canopy_20`) and 100 m (`h_mean_canopy`) segments during evening and nighttime acquisitions between April 29 and September 04, 2019.

4.3.2 Land cover

The Global Land Cover and Land Use map for 2019 was selected to improve upon its land use accuracy (Hansen et al. 2022). This Landsat-derived mapping product includes maximum vegetation cover (percent per pixel), class heights, and ecozones in an overall climate domain classification system (FAO, 2012). Landsat images were classified using multitemporal statistical feature metrics. They retain information without regard to time or place, provided they represent a statistically valid representation of an area (Song et al. 2001).

4.3.3 Digital elevation model

The Multi-Error-Removed Improved-Terrain Digital Elevation Model (MERIT-DEM) is a general-purpose elevation product at ~90 m resolution over land areas between 90N-60S (Yamazaki et al. 2017). This product was selected given the multisource DEM's used to remove errors. MERIT-DEM offered topographical and hydrological features such as slope (in degrees), TSRI—a measure of solar radiation aspect (Roberts and Cooper 1989), and Geomorpho90m—a MERIT-DEM-derived wetness index (Amatulli et al. 2020).

4.3.4 Sentinel-2

Sentinel-2, Level-1C images provided in Top-of-Atmosphere (TOA) reflectance's were collected from the Google Earth Engine (GEE) cloud platform (Gorelick et al. 2017). My first concerns were

directed towards near-complete coverage from cloud cover as of 2016. The median value of image collections between April and August (i.e., the local vegetation cycle) were retrieved with <20% cloud cover over Java. Pre-processing from TOA to Bottom-of-Atmosphere (BOA) reflectance's followed with the SIAC module on GEE (Yin et al. 2022). 2016-2021 annual images were mosaicked into seamless composites west (10°S, 105°E–66,918,389 pixels) and east (10°S, 110°E–79,826,814 pixels) including intra-annual phenological stages of the Cangkringan district in 2017, 2018, 2020, and 2021. All images were projected to match the ATL08 product and wavelengths were resampled to a ~30 m resolution.

4.4. Ground Cover Change model

The GCC model adopts seven strata following FAO's Land Cover Classification System (FAO, 2012). Vegetated classes consist of dense short vegetation (≤ 5 m, stable clusters and sparse vegetation), open tree cover ($5 \text{ m} \leq h < 12$ m, stable open trees), dense tree cover (≥ 12 m, clustered trees), and rice paddies/wetlands (< 12 m, trees and water presence). Non-vegetated classes include built-up (< 8 m, built land) and water (0 m, inland bodies of water). Cropland is managed land (≤ 5 m) stratified into growing, senescence, and uncultivated classes.

4.4.1 Sample allocations

A first lidar acquisition was retrieved over a wide and heterogenous transect of Java. I sampled land cover pixel landings and applied a morphological erosion window to disregard points ≥ 45 m away from inter-class perimeters. This accounts for class variability while avoiding mixed pixels in areas with thematic homogeneity (Wulder et al. 2020). Prevailing samples were vetted against three quality control mechanisms to derive a pool of samples arranged by levels of reliability (Figure 4.3). The first mechanism queries each sample and seeks thematic agreement using VHR satellite images

in Google Earth. Suppose a lidar footprint lands over a fraction of a pixel p with label y_n and area A_p where two mixed classes, i and j , belong to a set of n classes:

$$T_p = \{(i, j) \mid j \in J_p, \forall A_i < A_j\}; \quad (1)$$

$$y_i = \{i \mid i \in I_p, \sum_j (A_p) < A_p/2\}; \quad (2)$$

$$y_j = \{j \mid j \in J_p, \sum_i (A_p) < A_p/2\}; \quad (3)$$

$$Agreement_p = \{y_n \mid y_n \in Y_p, \forall A_n \geq A_p/2\}. \quad (4)$$

Equations 1-4 define whether $\geq 50\%$ of a pixel is in agreement with the queried label. The second mechanism apportions a structural agreement to each sample k with forest heights, F_h^k , correlated with canopy heights, C_h^k , where $F_h^k = C_h^k \pm 6 m$. The third mechanism secures height agreements, h_p , for k in range of the forest height thresholds, d_k , in Section 4:

$$h_p = \{(F_h^k, y_n) \mid y_n \in Y_p, F_h^k \cap h_n, \text{if } h_n \text{ are 'true' estimates of forest height}\}; \quad (5)$$

$$h_n = \lim_k \left(\min_d h_n \rightarrow \max_d h_n \right) \vee F_h^k \gtrless d_k. \quad (6)$$

Equations 5-6 ratify the accuracy of label agreements using vegetation structures to circumvent interpretation biases (Stehman 1999). Then, each agreement is scored to define four levels of reliability (L0 being the least reliable and L3 the most reliable). I held L3 reference samples with $\geq 500 m$ distributions to avoid issues of spatial autocorrelation and allocated the remaining L3 and L2 samples with label agreements in a metrics dataset (Table 4.2).

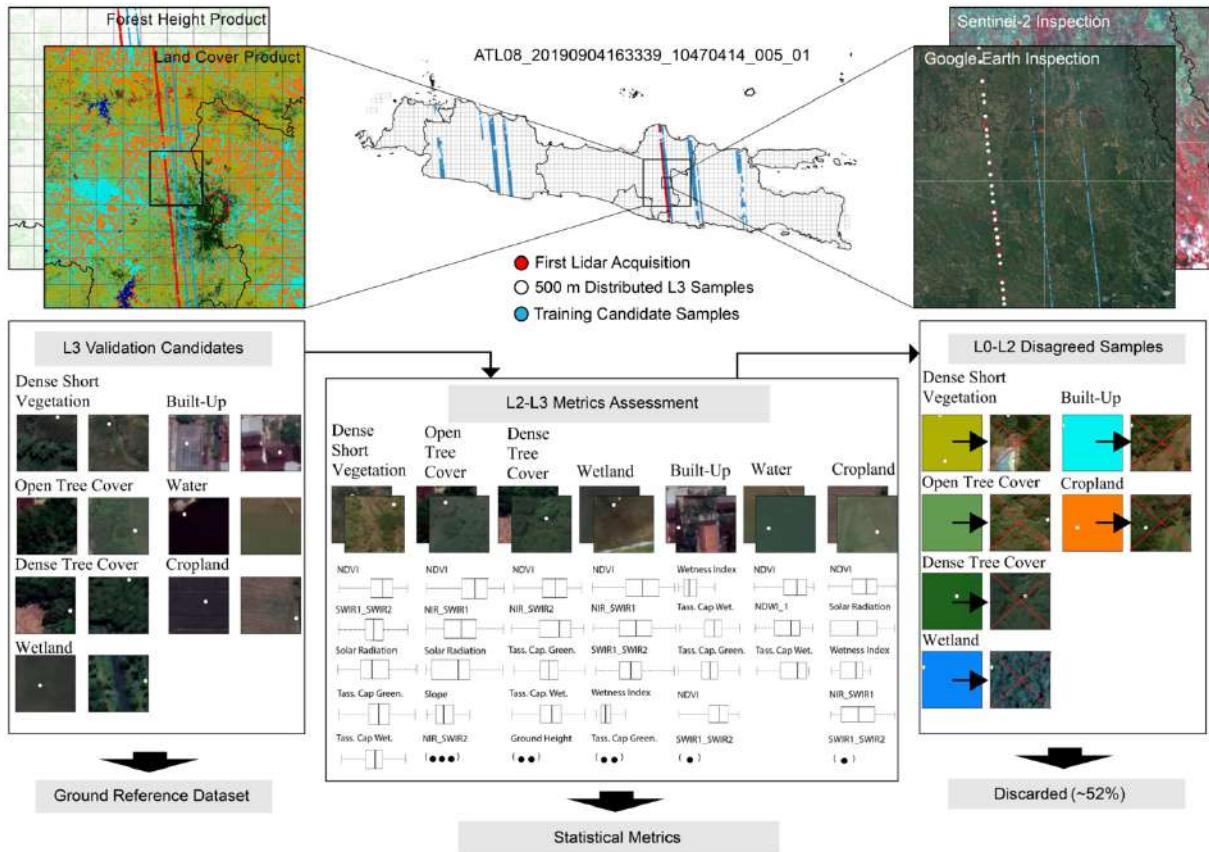


Figure 4.3. Sample distributions. Pixels correspond to the first lidar acquisition with metric features sorted by importance and unseen features represented with black dots.

Table 4.2. Sample allocations.

	Samples	L0	L1	L2	L3
First Lidar Acquisition	4,411 [2.67]	279 (1.186)	1,094 (1.055)	1,801 (0.426)	1,237 (0.004)
	100%	6%	25%	41%	28%
Metrics Dataset	1,888 [0.43]	-	-	896 (0.426)	992 (0.004)
	43%	-	-	47%	53%
Discarded	2,278 [2.67]	279 (1.186)	1,094 (1.055)	905 (0.426)	-
	52%	12%	48%	40%	-
Reference Dataset	245 [0.004]	-	-	-	245 (0.004)
	6%	-	-	-	100%

The predictive loss in error in parenthesis is estimated using native and vetted labels as in sub-section 4.4.4 Cumulative errors are in brackets (e.g., the metrics dataset has 83.9% of the cumulative error removed).

Metric samples were used to downstream 2019 Sentinel-2 signals from GEE. The extraction was defined by local periods of vegetation growth including the withdrawal of wet season (March-April),

peak growing period (May-June), and the senescence period (July-September). Then, multitemporal and static features forged a descriptive set of multidimensional samples.

4.4.2 Sample supervisions

(1) Metrics assessment

In this section, I extracted amplitude metrics to characterize regional land cover classes. First, I smoothed the spectral magnitude of raw signals with a moving-average window of 101 indices to retain noise-free temporal profiles. I also estimated the 25th, 50th, and 75th quartiles of cropland NDVI, blue, near-infrared, and infrared bands to stratify that class. Mean temporal signal estimations followed as shown in Supplementary Information (Figure S1). I performed multinomial regressions with the following dependent variables: forest heights (treed classes), SAVI (wetland and cropland), slope (water), and BN (built-up). Features with statistically significant p-values ($\geq 90\%$) were used to record the relative bounds (\pm standard error) of smoothed, mean, and raw amplitudes (Table S1).

(2) Detection of reliable training samples

Eighteen ICESat-2 ground-track acquisitions were retrieved to meet the training quota. I removed no data values and footprints landing ≥ 45 m away from class boundaries. The prevailing samples with forest heights, static, and multitemporal features were used as input.

A filtering (and vetting) module was parsed with raw metrics given the quantity of thematic agreements in Table S2. The module queried `h_mean_canopy` as the primary ICESat-2 value, or `h_canopy_20` as a secondary value to infill no data segments and return a densely-populated canopy height column. Then, each sample was vetted by 3-8 metric bounds (F_i) to find whether the signal of a sample k with class n belongs to the spectrum of that class: $\min_n F_i \leq k \leq \max_n F_i$. Forest and canopy height agreements followed as in sub-section 4.4.1. Lastly, forest (vegetated classes) or

ICESat-2 (non-vegetated classes) signatures were vetted against class thresholds using equations 5-6. The program ends after broadcasting reliability values in a dataset for user-based discriminations.

(3) Training data refinements

Refinements to a pool of 167,805 samples were carried out to alleviate label uncertainties. I applied the following steps to eliminate samples of dubious quality (see Table 4.3): First, L0-L1 samples failing to pass at least one filter were immediately erased. Samples with combined thematic and height disagreements were also eliminated. L2 samples with thematic disagreements were liable and therefore erased (except for rare classes).

The elimination of 64,702 samples rendered a set of 103,103 reliable samples for model calibration. As a result, L3 samples expanded to 60% of the training set with cropland stratified into growing (41%), senescence (37%), and cultivation (22%) classes detailed in Table S3.

Table 4.3. Data calibration summary.

	Samples	L0	L1	L2	L3
Pre-Filtering	167,805 [2.04]	4,004 (0.606)	34,053 (0.545)	68,365 (0.513)	61,383 (0.374)
	100%	2%	20%	41%	37%
Post-Filtering	103,103 [0.89]	- -	- (0.513)	41,720 (0.513)	61,383 (0.374)
	61%	-	-	25%	37%
Filtered	64,702 [1.66]	4,004 (0.606)	34,053 (0.545)	26,645 (0.513)	- -
	38%	2%	20%	16%	-

The predictive loss in error is in parenthesis. A model 10 classification in sub-section 4.4.3 provided the ‘y’ variable and native samples the ‘x’ variable. Brackets depict total errors. Refined samples saw 56.4% of the error removed. Replacing L2 with L3 samples would eliminate 81.7% of the error.

4.4.3 Model optimization

This section discriminates contending features based on the mean decrease in impurity (MDI). This method performs well with tree-based models, such as Random Forest classifiers, where the impurity is quantified by splitting the criterion of decision trees (Breiman 2001). However, MDI can convey up-ranked features when a model is overfitting. Therefore, permutation feature importance’s were tested over seen and unseen data. This method inspects non-linear estimators by signaling a

decrease in model score when a feature is shuffled. So, when two or more features are correlated, equal values are given and collinearity can be identified. I sought ten uncorrelated feature combinations for evaluation (Table S4).

Fast TreeSHAP (Yang 2021) estimations followed to identify features with the best relative contribution to land use classes (Figure 4.4). SHAP values explain predictions from black box learning models using ideas from game theory. They are sensitive to correlated features, but the trials above eliminated potential doubts.

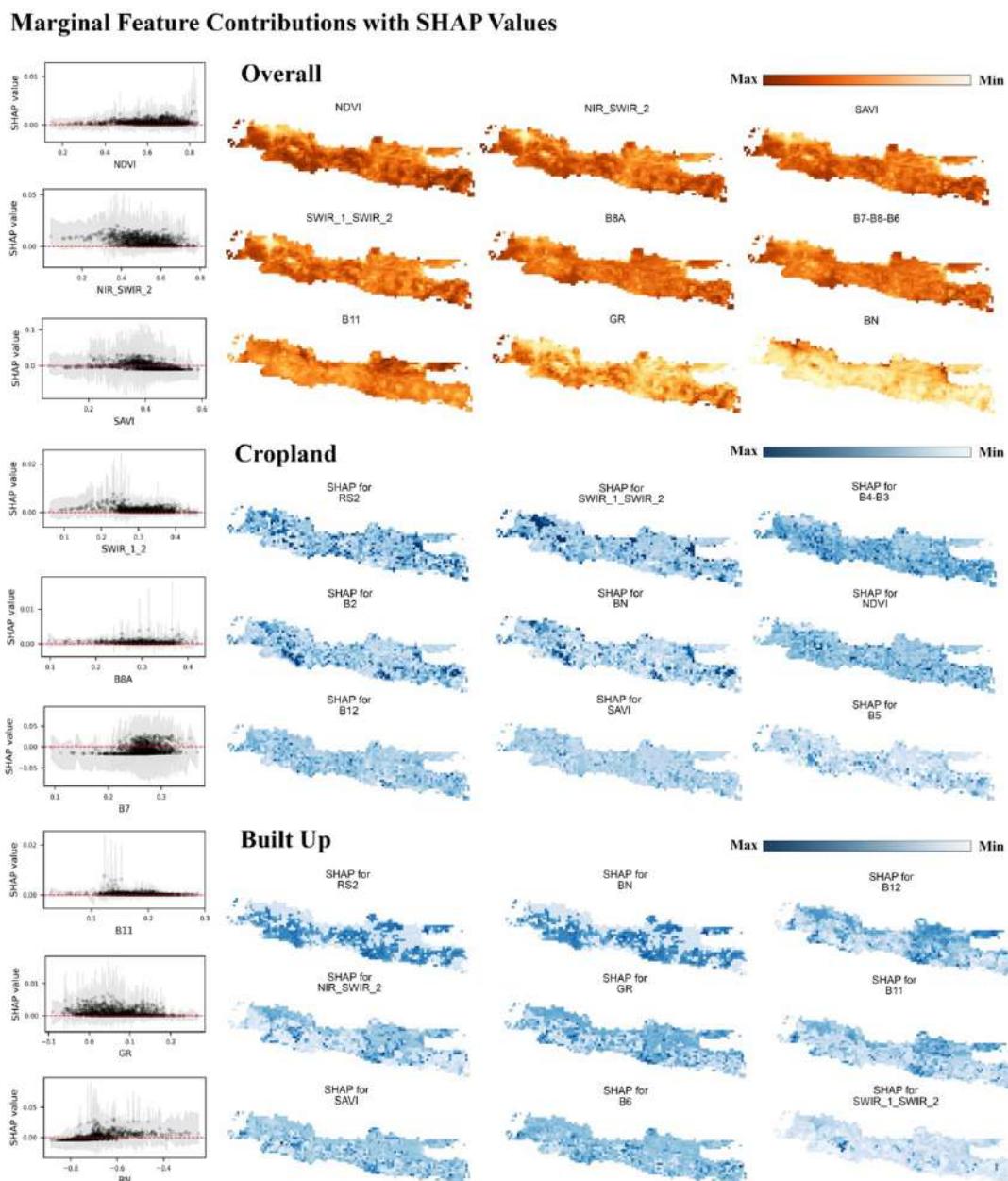


Figure 4.4. Feature contributions with SHAP values. Partial dependence scatter plots show overall SHAP values with shaded bootstrap 95% confidence intervals, sorted from highest to lowest.

Each model was optimized with red (RS2), green (GR), and blue (BN) band indices. However, learning models require unique parameters to apply their learnt capacity. Therefore, I used a cross-validated grid search approach with five K-fold splits per score to avoid error insertion. This was mediated through a stratified shuffle split algorithm of 100 bootstrapped sets of 1000 samples each. Final scores were based on the predictive mean accuracy of classifiers over unseen data. Table 4.4 shows model performances, features, and parameters with error terms.

Table 4.4. Model performance. Root mean square errors (RMSE) and 95th percentile standard errors in parenthesis.

No.	Features	mtry, depth, min.samp.split	#feat, max feat	max samples	RMSE	Test Data Scores	Bias- squared	Variance	Total error
Model 1	[N, N2, S12, TCG, RS2, GR, S]	391, 17, 23	18, log2	0.8	3.12	0.78 (0.005)	7.58	0.64	8.22
Model 2	[N, NDWI-2, NS2, S12, TCW, RS2, S]	391, 19, 15	18, log2	0.8	3.19	0.73 (0.005)	7.77	0.78	8.55
Model 3	[N, NS2, S12, TCW, RS2, S, BN]	391, 19, 16	18, log2	0.8	3.21	0.73 (0.005)	7.87	0.61	8.48
Model 4	[N, NDWI 2, NS2, S12, RS2, BN, I]	391, 19, 15	18, log2	0.8	3.19	0.73 (0.004)	7.97	0.56	8.53
Model 5	[N, NS2, S12, TCW, RS2, BN, I]	391, 17, 23	18, log2	0.8	3.07	0.78 (0.003)	7.94	0.75	8.69
Model 6	[N, NDWI 1, S12, TCW, RS2, GR, BN]	391, 17, 19	18, log2	0.8	3.31	0.78 (0.004)	7.68	0.56	8.24
Model 7	[N, NS2, S12, RS2, GR, SAVI, I]	391, 17, 22	18, log2	0.8	3.37	0.84 (0.002)	7.44	0.49	7.93
Model 8	[N, NDWI 1, S12, TCW, RS2, BN, I]	390, 17, 25	18, log2	0.8	3.37	0.79 (0.003)	7.81	0.63	8.44
Model 9	[N, NS2, S12, RS2, GR, BN, I]	391, 17, 20	18, log2	0.8	3.35	0.78 (0.003)	8.07	0.66	8.73
Model 10	[N, NS2, S12, RS2, GR, S, BN]	391, 17, 22	18, log2	0.8	3.08	0.78 (0.005)	8.11	0.61	8.72

Sentinel-2 bands B1, B2, B3, B4, B5, B6, B8A, B11 and three full-zero dummy variables were included to boost high ranking features. N= NDVI, S= SAVI, and I= IRECI.

4.4.4 Classifications and Accuracy Assessments

(1) Preliminary Classifications

Here, pixel-based accuracy assessments over the 2019 east Java composite were compared using supervised mean-weighed (SMW) and Stratified Random Sampling (SRS) designs (Table S5). SRS includes 269 stratified samples in area weighed proportion with ≥ 500 m buffered areas. Pixel landings were photo-interpreted using VHR imagery (2019-2021) in Google Earth. The reliability of SMW, x_i , and SRS, y_j , was evaluated with the Mean Absolute Difference (MAD) of producer's (10.55%) and overall (9.07%) accuracies in Table S6 as follows:

$$MAD = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |x_i - y_j| \quad (7)$$

Equation 7 describes a sample size n for a distributed population of absolute sample values, j_i , where $i = 1$ to n estimated as the average of all possible absolute differences.

The MAD led us to estimate the 0-1 loss decomposition with an expectation (E) over the training set subdivided into bias and variance (Dietterich and Kong 1995). The loss function states that models not predicting true labels have a bias value of 1, or otherwise 0.

$$\text{Bias} = \begin{cases} 1 & \text{if } y \neq E[\hat{y}], \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

$$\text{Var} = P(\hat{y} \neq E[\hat{y}]) \quad (9)$$

Equation 8 shows that predictions not always agree with the true label, y , while the variance in equation 9 defines the probability, P , of a predicted label not matching the main prediction. However, if the bias is 1, the loss is entirely defined by the variance; $\{1 - P(\hat{y} \neq E[\hat{y}])\}$. Table 4.5 shows accuracy and total errors using SRS and SMW reference labels.

Table 4.5. Error quantification with two reference sample designs.

	Accuracy		Loss		Bias		Variance		Total Error	
	SRS	SMW	SRS	SMW	SRS	SMW	SRS	SMW	SRS	SMW
Model 1	0.649	0.739	0.335	0.278	0.333	0.278	0.064	0.009	0.397	0.287
Model 2	0.638	0.746	0.343	0.229	0.333	0.222	0.073	0.026	0.406	0.248
Model 3	0.657	0.746	0.306	0.213	0.315	0.204	0.109	0.033	0.424	0.237
Model 4	0.627	0.705	0.366	0.263	0.352	0.259	0.108	0.024	0.46	0.283
Model 5	0.66	0.728	0.305	0.21	0.296	0.204	0.048	0.03	0.344	0.234
Model 6	0.657	0.75	0.361	0.241	0.37	0.222	0.028	0.069	0.398	0.291
Model 7	0.66	0.75	0.323	0.259	0.315	0.259	0.055	0.02	0.37	0.279
Model 8	0.657	0.754	0.315	0.223	0.315	0.222	0.036	0.051	0.351	0.273
Model 9	0.657	0.743	0.335	0.223	0.333	0.222	0.066	0.024	0.399	0.246
Model 10	0.66	0.769	0.325	0.219	0.315	0.204	0.054	0.039	0.369	0.243

The mean total error for SRS (0.39) and SMW (0.26) labels revealed that L3 samples were 33% more reliable (Table S7). Then, model 10 was selected to estimate map accuracy in area-proportion and compare user accuracies for SMW (Table S8) and SRS (Table S9) designs. This step confirmed the results above for classes with large samples sizes.

(2) Informed classifications

Final-stage post-processing steps are reported here with parameters, sequences, and random seeds to improve the temporal consistency of preliminary maps. I applied a morphological processing sequence starting with a dilation operation over rare classes (water and wetlands) with 1-pixel size to avoid pixel intrusions. Pixels smaller than a two-pixel threshold with 0-pixel connections and a random seed of 0 were removed. A second dilation was performed over target classes (open tree cover, dense tree cover, built-up, and cropland) with 1-pixel size to reshape their domain. Temporal co-registrations with the 2019 template concluded the sequence and increased the overall accuracy from 78% to 82% (Table S10).

(3) Post-informed classifications

Final accuracy was performed over the 2017-2019 east composite of Java. I added five transition classes shown in Table 4.6. Additional reference samples were sought with power allocations over rare and transition classes to ensure lower variance in the error. All labels were then reviewed using the time-stamp tool in Google Earth. I applied a 3-pixel kernel window to ensure that 50% of the neighboring pixels went through the same change as the central pixel. Also, trends before and beyond target years were considered to secure transitions unrelated to weather change. The reference set reflects ‘from-to’ labels for transition classes.

Table 4.6. Confusion matrix of post-informed land cover change with error-adjusted areas expressed in percentages of area proportion with 95% confidence intervals.

Reference	Error-Adjusted Area (ha)														
	Dense Short Open Tree Cover	Dense Tree Cover	Wetland	Built Up	Water	Cropland	Built Up Loss	Built Up Gain	Cropland Loss	Cropland Gain	Cropland > Built Up	Built Up > Cropland	User's Accuracy		
1	1.49%	0.45%	0.15%	0.00%	0.00%	0.00%	0.00%	0.00%	0.15%	0.00%	0.00%	0.00%	66.67% ± 23.9%	318,916 ± 21.2%	
2	0.00%	8.41%	0.49%	0.00%	0.32%	0.00%	0.81%	0.00%	0.00%	0.00%	0.00%	0.00%	83.87% ± 9.2%	1,146,244 ± 9.1%	
3	0.22%	0.67%	12.34%	0.00%	0.22%	0.00%	0.90%	0.00%	0.00%	0.00%	0.00%	0.00%	85.94% ± 8.5%	1,687,818 ± 9.8%	
4	0.00%	0.00%	0.00%	1.02%	0.00%	0.51%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	66.67% ± 53.3%	359,167 ± 48.2%	
5	0.00%	0.00%	0.00%	0.00%	8.08%	0.21%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	97.44% ± 5.0%	971,793 ± 7.5%	
6	0.00%	0.00%	1.26%	2.53%	0.00%	17.70%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	82.35% ± 18.1%	1,752,631 ± 11.5%	
7	1.35%	1.35%	0.68%	0.23%	0.45%	0.00%	23.22%	0.00%	0.00%	0.00%	0.00%	0.00%	85.12% ± 6.3%	2,710,664 ± 4.6%	
8	0.00%	0.00%	0.00%	0.00%	0.37%	0.00%	0.37%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00% ± 0.0%	0 ± 0%
9	0.00%	0.00%	0.00%	0.00%	0.16%	0.00%	0.16%	0.00%	0.47%	0.00%	0.00%	0.00%	60.00% ± 42.9%	45,069 ± 40.8%	
10	0.00%	1.01%	2.52%	0.00%	0.00%	0.00%	0.00%	0.00%	1.51%	0.00%	0.00%	0.00%	30.00% ± 28.4%	158,101 ± 47.2%	
11	0.00%	0.00%	0.30%	0.00%	0.61%	0.00%	0.61%	0.00%	0.00%	2.43%	0.00%	0.00%	61.54% ± 26.4%	231,001 ± 22.8%	
12	0.00%	0.16%	0.00%	0.00%	0.00%	0.00%	1.29%	0.00%	0.00%	0.00%	0.81%	0.00%	35.71% ± 25.1%	76,830 ± 37.2%	
13	0.29%	0.00%	0.00%	0.00%	0.00%	0.00%	1.14%	0.00%	0.00%	0.00%	0.00%	0.57%	28.57% ± 33.5%	54,195 ± 64.5%	

Producer's Accuracy	44.46%	69.77%	69.57%	27.06%	79.09%	96.07%	81.49%	0.00%	100.00%	91.03%	100.00%	100.00%	100.00%	OA = 78.06 ± 5.06%
	± 6.8%	± 6.5%	± 7.0%	± 11.6%	± 7.2%	± 17.3%	± 5.8%	± 0.0	± 12.2%	± 24.4%	± 13.2%	± 14.5%	± 27.3%	

4.5. Evaluation

4.5.1 Mapping outcomes

Bi-temporal land change maps (2016-2021) were produced for the geographical unit of Java. The reporting scheme analyzes open and dense tree cover as a merged forest class. Further, stable strata were used as a base layer pertaining to the latter year of change instead of the former to contrast a target year's ecosystem. These results are reported in percentages rather than absolute values and map names use the ISO 3166 country code, a three-letter code (alpha-3) followed by classification periods and unit coordinates (e.g., [IDN]2017201910S110E.tif).

Resettlement sites were subsequently selected including two non-affected baseline areas as illustrated in Figure 4.5. I considered settlement size and overlapping influence areas. Areas of influence were drawn from the centroid of each settlement to represent assumed daily median walking distances, equivalent to 800 m (Yang and Diez-Roux 2012).

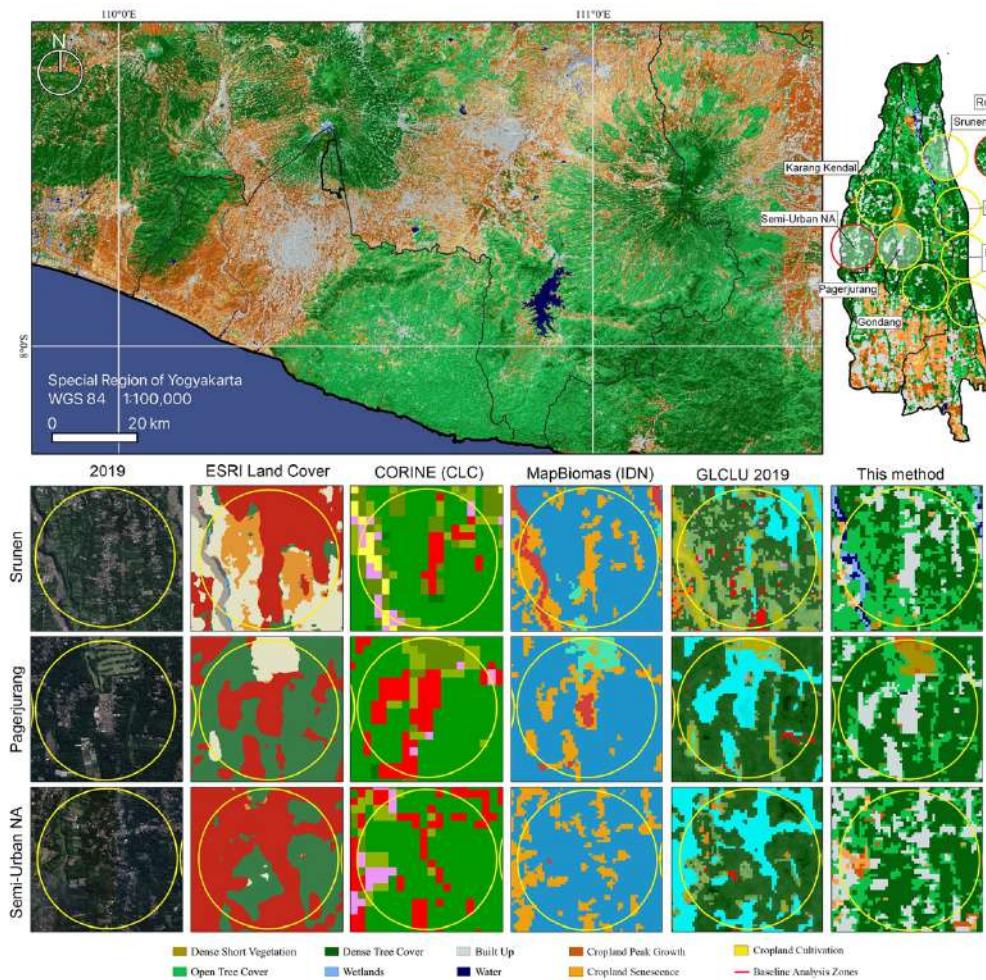


Figure 4.5. The Special Region of Yogyakarta in 2019. The top right is the Cangkringan district with seven affected (AF) and two non-affected (NA) areas. Google Earth images are shown for comparative purposes and global products are sorted by relative precision.

4.5.2 Land interactions

Land change interactions compare sub-district and local trends to explain local distributions, intensities, and livelihoods in terms of crop area change. The first interaction was identified in Figure 4.6, where rural areas of the Cangkringan district show built growth caused by forest loss.

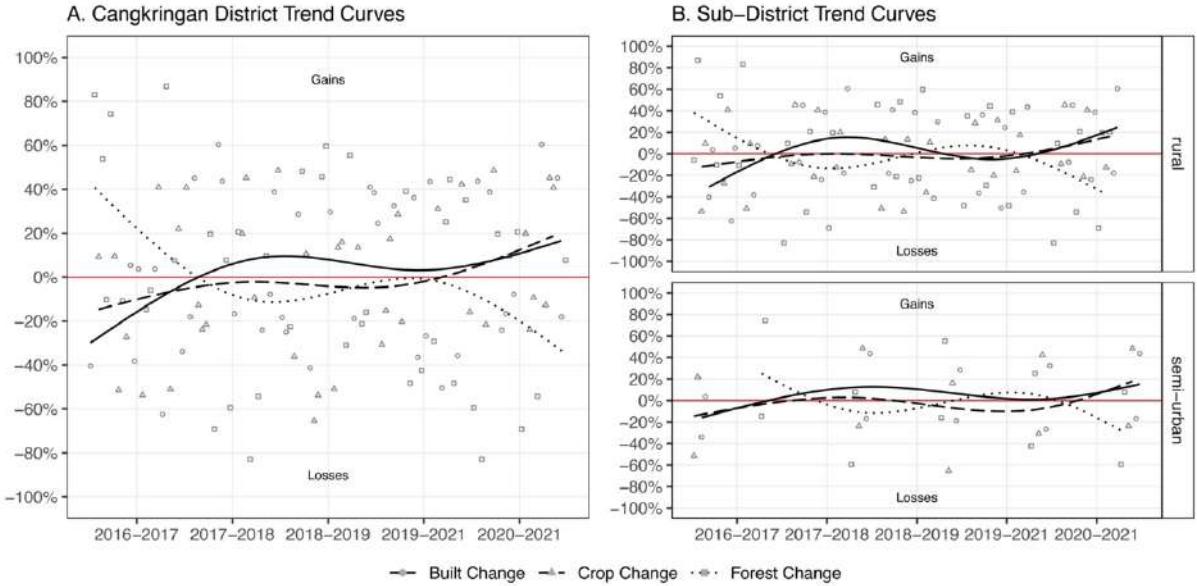


Figure 4.6. Land change in rural and semi-urban sub-districts of the Cangkringan district. Smoothed trends use Gaussian regressions with four-degree polynomials.

Inter-annual seasons s of cropland area change were then compared in sub-districts i through pixel-wise gain g and loss l detections (Table 4.7). Annual changes, Δy_i , between 2017, 2018, 2020, and 2021 were estimated with the following logic:

$$f_g = \{(g, s) \mid g \in s \forall y_i, \sum_{i>1}^y (\Delta y_i)\}; \quad (10)$$

$$f_l = \{(l, s) \mid l \in s \forall y_i, \sum_{i>1}^y (-\Delta y_i)\}. \quad (11)$$

Equations 10-11 describe annual gross changes with net changes subject to seasons x_s per annum;

$\sum_{y>1}^{\Delta y_i} (f_g - f_l)/x_s$. Seasonal image pairs were analyzed to confirm results as follows:

$$k_g = \{(g, s) \mid g \in \Delta s_i \forall s_i, \sum_{i>1}^y (\Delta s_i)/x_s\}; \quad (12)$$

$$k_l = \{(l, s) \mid l \in \Delta s_i \forall s_i, \sum_{i>1}^y (-\Delta s_i)/x_s\}; \quad (13)$$

where equations 12-13 describe mean intra-annual change with net estimates; $\sum_{y>1}^{\Delta y_i} (k_g - k_l)$.

Table 4.7. Cropland area change (Ha)

	Inter-annual change				Intra-annual change			
	2017	2018	2020	Net	Growth	Peak	Seasons	Net
	2018	2020	2021	Change	Change	Peak	Closure	Change
Glagaharjo	-23.0	-17.6	11.7	-28.9	-49.0	-7.90	-29.7	-28.9
Kepuharjo	-41.8	-14.6	21.9	-34.5	-54.4	72.3	-121.3	-34.5
Umbulharjo	-51.8	-8.1	45.3	-44.6	-23.0	115.0	-136.0	-14.6

Total Affected	-116.6	-40.3	78.9	-78.0	-126.1	179.2	-287.0	-78.0
Total Non-Affected	27.8	-80.9	72.4	19.3	114.8	181.8	-238.4	19.3

I also analyzed 2016-2021 mean land use transitions to illustrate the overall distribution of post-disaster changes in Figure 4.7. However, several topological and topographical variations exist. Therefore, a contrastive analysis was performed in resettlement areas.

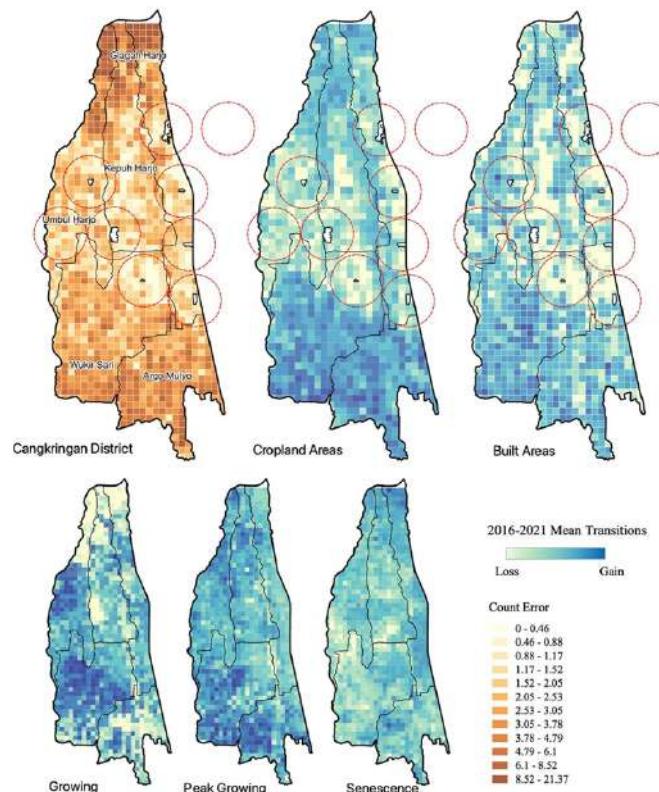


Figure 4.7. Mean transitions in 200-m tiles. Rural-affected areas include Umbulharjo, Kepuharjo, and Glaga Harjo, while semi-urban non-affected areas encompass Wukirsari and Argo Mulyo. Count errors are shown with class intervals by standard deviation from the mean error, which is zero.

4.5.3 Resettlement areas of influence

Land change interactions of the analysis period now focus on homogenous areas. Table S11 provides land change measures in affected and non-affected areas. I saw significantly higher cumulative changes in non-affected areas. This can be explained by the intensity of peri-urban growth

increasing since before 2010. Another observation is the intensity of forest to built-up change growing from rural to semi-urban areas. Figure 4.8 shows a steep increment of built land in 2018 increasing until 2021. This is assumed to be a long-term effect of resettlement (2014-2018), further exacerbated by Yogyakarta's spillover effects.

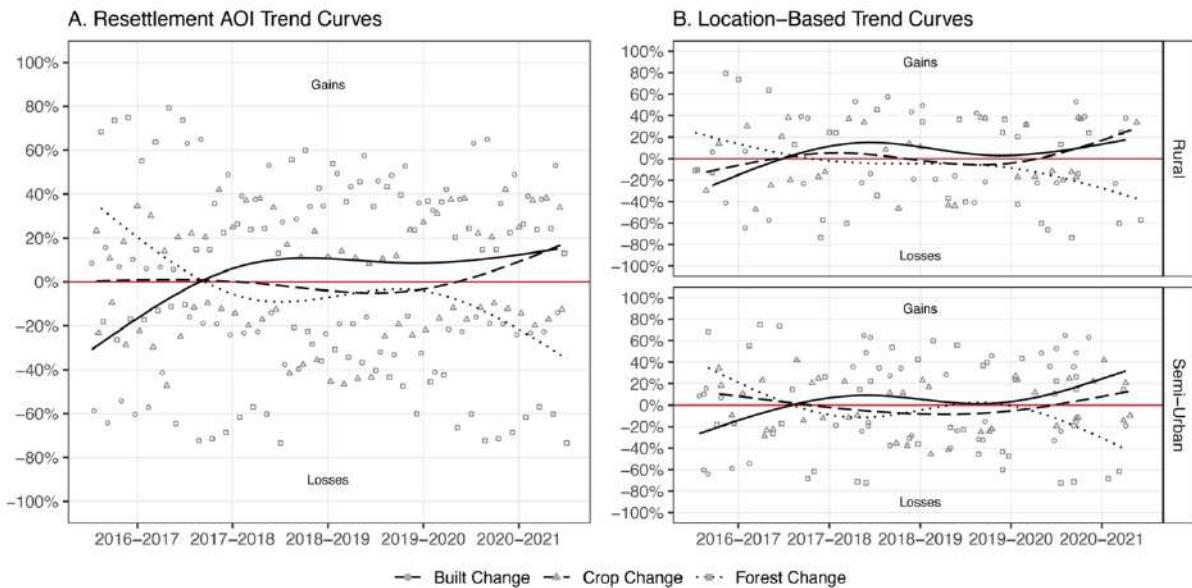


Figure 4.8. Resettlement areas of influence. Smoothed trends use Gaussian regressions defined by four-degree polynomials.

I proceeded to evaluate mean change in affected sub-districts and resettlement areas (Table 4.8). Here, relative gross and net change intensities are sorted from rural to semi-urban areas. Forests have the highest net change values in sub-districts (4.5%) and resettlement areas (-6.9%). Conversely, built-up growth denotes 0.3% on the aggregate level and 4.7% in localized areas. Particularly, cropland growth in resettlement areas shows that spillover effects were successfully undermined given the overall declining tendency in the region.

Table 4.8. 2016-2021 Mean intensity of change over total change.

		Crop Gain	Crop Loss	Net Change	Built Gain	Built Loss	Net Change	Forest Gain	Forest Loss	Net Change
Sub-District	Glagaharjo	14.3%	30.5%	-16.2%	34.4%	31.4%	3.0%	51.3%	38.1%	13.2%
Settlement	Srunen	26.6%	22.0%	4.6%	38.1%	29.8%	8.3%	35.3%	48.2%	-12.9%
Settlement	Gading	29.4%	27.4%	2.0%	31.7%	34.2%	-2.5%	38.9%	38.4%	0.5%

	Settlement	Banjarsari	28.9%	26.4%	2.5%	33.6%	25.4%	8.2%	37.5%	48.3%	-10.8%
	Settlement	Jetis Sumur	31.4%	21.1%	10.3%	31.8%	27.5%	4.3%	36.8%	51.3%	-14.5%
Sub-District	Kepuharjo	33.0%	29.9%	3.1%	28.5%	37.7%	-9.1%	38.5%	32.5%	6.0%	
	Settlement	Pagerjurang	20.0%	22.3%	-2.3%	37.8%	35.0%	2.8%	42.3%	42.7%	-0.4%
Sub-District	Umbulharjo	28.0%	29.4%	-1.4%	30.0%	22.8%	7.1%	42.0%	47.8%	-5.8%	
	Settlement	Karang Kendal	15.4%	17.2%	-1.8%	45.5%	33.9%	11.6%	39.1%	48.9%	-9.8%
	Settlement	Gondang	24.6%	24.4%	0.3%	34.2%	34.2%	0.0%	41.1%	41.4%	-0.3%
Mean Total	Sub-Districts		25.1%	29.9%	-4.8%	31.0%	30.6%	0.3%	43.9%	39.5%	4.5%
	Settlements		25.2%	23.0%	2.2%	36.1%	31.4%	4.7%	38.7%	45.6%	-6.9%

4.6. Discussion and Concluding Remarks

4.6.1 Automatic Sample Calibration and Model Optimization

The model funneled multisource remote sensing data to avail regionally-consistent multitemporal metrics. These metrics support—via rule-based combination of attributes—multiple classification schemes (Gomez et al. 2016). Smoothed metrics delivered 66,917 reliable samples and a 56% overall accuracy. I used mean temporal profiles to increase the number of samples rendering a 74% overall accuracy. However, raw metrics defined the spectral boundary of classes securing a 78% accuracy.

The Random Forest classifier also showed considerable sensitivity to the number of training samples per class. Therefore, I maintained pixel landing proportions by filtering entire reliability levels instead of class-specific samples. Then, SHAP values highlighted the utility of NIR-SWIR2 binding the feature set and generating predictive power. NIR-SWIR2 was not identified using the MDI algorithm. Also, SWIR1-SWIR2 was particularly useful in detecting land use classes. However, Red-SWIR2 and Blue-NIR contributed the most (Figure S3).

The loss function subsequently revealed the size of error inherent to simple photo-interpretation. Bias-reduced samples narrowed the mean variance from 17.35% (Table S9) to 11.83% (Table 6) eliminating 32% of the error. Consequently, the net worth of producer and user accuracies grew by 8.4% over the land cover product enabling change detection analyses.

4.6.2 Land change effects of resettlement

The Cangkringan district saw forest to cropland (Pearson's $r=68-87$), forest to built-up (Pearson's $r=45-62$), and cropland to built-up (Pearson's $r=26-52$) change correlations (Figure 4.9). This sequence translates to forests loss and cropland growth as a medium to built-up growth.

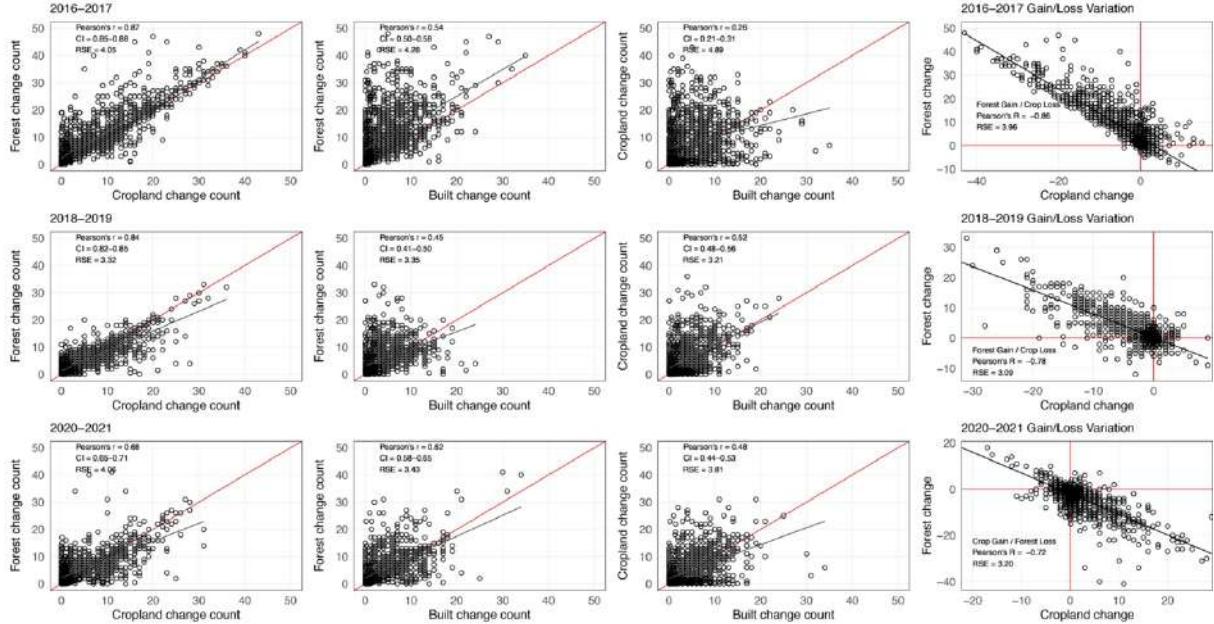


Figure 4.9. Correlation of transition counts with best-fit linear regressions. The full series is available in Figure S4. The right column illustrates gain/loss variations for the highest correlation. Pearson's r , 95% confidence intervals (CI), and residual standard errors (RSE) complement the analysis.

Then, a localized analysis identified transitions caused by resettlement sites. If peri-urban dynamics (d_i) extend homogeneously over 'not affected', x_i , and 'affected', y_i , areas, then total changes of class i in not affected (n_a) and affected (a) areas can be defined as:

$$Total_{ch_i} = \sum_i^{n_a} \sum_i^a \left| \left[\frac{x_i}{n_a} \right] + \left[\frac{y_i}{a} \right] \right|; \quad (14)$$

$$d_i = \left\{ \left(\frac{x_i}{Total_{ch_i}} \right) \times 100, \text{ where 'not affected' represents urban dynamics} \right\}; \quad (15)$$

$$l_e = \left\{ 1 - \left[\left(\frac{d_i}{100} \right) \times j_i \right] \right\}, \text{ and 'affected' includes resettlement effects, } j_i \}. \quad (16)$$

Equations 14-16 estimate localized effects of investments made, l_e . Therefore, applying surface areas in Table S12, I found that 16.4% forest loss, 18.5% cropland growth, and 15.5% built-up growth were a product of resettlement.

4.6.3 Implications and limitations

The following points summarize the implications of this study:

- The GCC model renders a two-step label supervision model. First, bias-reduced labels are obtained to characterize an eco-zone of interest. Second, amplitude metrics are converged in a filtering module to automate high-quality label retrievals.
- Bias-reduced labels eliminated 32% of the mean variance from the reference set. The filtering module also removed 56% of the error from the training set.
- Peri-urban resettlement availed 18.5% cropland growth and 15.5% built-up growth at the expense of 16.4% forest loss. In addition, the resiliency of recovered livelihoods was quantified with a 2.2% mean net intensity.

Limitations of this study encompass the time-consuming steps to attain statistically significant metrics and reliance on hand-labeled sample agreements to achieve automatic label supervisions.

4.6.4 Concluding remarks

This chapter formulates an automatic label supervision model that predicts accurate land cover maps to evaluate land change interactions prompted by post-disaster resettlement sites. Mapping accuracies of 78%-82% enabled site-specific analyses after the 2010 Merapi volcano eruptions. Experimental results indicate that resettlement sites caused land use growth at the expense of forest land fueled by Yogyakarta's spillover effects. Local governments can use these results to find candidate resettlement sites next to roads that sustain land use growth. The estimated land use growth values are reliable in the extent of Yogyakarta's region, but can be useful to future disaster resettlement cases within the Indo-Malayan ecozone. Further, developed countries can apply these results as a safety margin near shrinking cities to avoid land use dispersion prompted by post-disaster resettlement sites. For these cases, we suggest using use cumulative land use change values only.

This method opens the accessibility of remotely sensed image classification tasks to people with no extensive domain knowledge by vetting pre-existing land cover labels against very-high resolution satellite images and using canopy height datasets to alleviate interpretation biases. This allows users to create and maintain regionally consistent and highly reliable training datasets for land cover classifications on-demand. It also offers possibilities for continued disaster recovery monitoring in peri-urban areas where deforestation is likely and preservation is necessary. This study marks a major step towards disaster recovery planning as technological improvements continue to rise with big data models at the forefront of Earth observation.

References

- Adriano, B., Yokoya, N., Xia, J., Miura, H., Liu, W., Matsuoka, M., Koshimura, S., 2021. Learning from multimodal and multitemporal earth observation data for building damage mapping. *ISPRS J. Photogramm. Remote Sens.* 175, 132-143. <https://doi.org/10.1016/j.isprsjprs.2021.02.016>.
- Amatulli, G., McInerney, D., Sethi, T., Strobl, P., Domisch, S., 2020. Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. *Sci. Data* 7, 162. <https://doi.org/10.1038/s41597-020-0479-6>.
- Breiman, L., 2001. Random Forests. *Machine Learn.* 45, 5-32. <https://doi.org/10.1023/A:1010933404324>.
- Contreras, D., Blaschke, T., Tiede, D., Jilge, M., 2016. Monitoring recovery after earthquakes through the integration of remote sensing, GIS, and ground observations: the case of L'aquila (Italy). *Cartogr. Geog. Inf. Sci.* 43, 115-133. <https://doi.org/10.1080/15230406.2015.1029520>.
- Desai, S. and Ghose, D., 2022. Active learning for improved semi-supervised semantic segmentation in satellite images: in 2022 IEEE/CVF Winter Conf. Appl. Computer Vis. (WACV), Waikoloa, HI, USA, 1485-1495. <http://www.doi.org/10.1109/WACV51458.2022.00155>.
- Dietterich, T.G. and Kong, E.B., 1995. Machine learning bias, statistical bias, and statistical variance of decision tree algorithms. Tech. Rep. Dept. Computer Sci., Oregon State University, OR. <https://hunch.net/~coms-4771/tr-bias.pdf>.
- Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's Forests and topography. *Sci. Remote Sens.* 1, 100002. <https://doi.org/10.1016/j.srs.2020.100002>.
- Gibson, T.D., Pelling, M., Ghosh, A., Matyas, D., Siddiqi, A., Solecki, W., Johnson, L., Kenney, C., Johnston, D., Du Plessis, R., 2016. Pathways for transformation: disaster risk management to enhance resilience to extreme events. *J. Extreme Events* 3, 1671002. <https://doi.org/10.1142/S2345737616710020>.
- FAO (United Nations Food and Agricultural Organization), 2012. Global Ecological Zones for FAO forest reporting: 2010 Update (Rome: FAO). <https://www.fao.org/forest-resources-assessment/remote-sensing/global-ecological-zones-gez-mapping/en/> (accessed 14 September 2022).
- Fekete, A., 2022. Peri-urban growth into natural hazard-prone areas: mapping exposure transformation of the built environment in Nairobi and Nyeri, Kenya, from 1948 to today. *Nat Hazards*, 1-24. <https://doi.org/10.1007/s11069-022-05515-4>.
- Foody, G.M., 2010. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sens. Environ.* 114, 2271-2285. <https://doi.org/10.1016/j.rse.2010.05.003>.
- Garcia Fry, M., 2023. Ground Cover Change Model. Mendeley Data, v5. <http://www.doi.org/10.17632/mzp3k6fmtz.5>
- Garcia-Fry, M., Murao, O., Bachri, S., Moya, L.A., 2022. Land-Use microsimulation model for livelihood diversification after the 2010 Merapi volcano eruptions. *Transp. Res. D: Transp. Environ.* 104, 103189. <https://doi.org/10.1016/j.trd.2022.103189>.
- Gomez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* 116, 55-72. <http://dx.doi.org/10.1016/j.isprsjprs.2016.03.008>.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18-27. <http://dx.doi.org/10.1016/j.rse.2017.06.031>.

- Hansen, M.C., Potapov, P.V., Pickens, A.H., Tyukavina, A., Hernandez-Serna, A., Zalles, V., Turubanova, S., Kommareddy, I., Stehman, S.V., Song, X.-P., Kommareddy, A., 2022. Global land use extent and dispersion within natural land cover using Landsat data. Environ. Res. Lett. 17, 034050. <https://doi.org/10.1088/1748-9326/ac46ec>.
- Huang, X., Weng, C., Lu, Q., Feng, T., Zhang, L., 2015. Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas. Remote Sens. (Basel) 7, 16024-16044. <https://doi.org/10.3390/rs71215819>.
- Huang, C. and Zhang, C., 2022. Characterizing urban growth in Vientiane from 2000 to 2019 using time-series optical and SAR-based estimates of urban land. Int. J. Appl. Earth Obs. Geoinf. 109, 102798. <https://doi.org/10.1016/j.jag.2022.102798>.
- Hutchings, P., Willcock, S., Lynch, K., Bundhoo, D., Brewer, T., Cooper, S., Keech, D., Mekala, S., Mishra, P.P., Parker, A., Shackleton, C.M., Venkatesh, K., Vicario, D.R., Welivita, I., 2022. Understanding rural–urban transitions in the Global South through peri-urban turbulence. Nat. Sustain. 5, 924-930. <https://doi.org/10.1038/s41893-022-00920-w>.
- Koshimura, S., Moya, L., Mas, E., Bai, Y., 2020. Tsunami Damage Detection with Remote Sensing: A Review. Geosci. (Basel) 10, 177. <https://doi.org/10.3390/geosciences10050177>.
- Li, Y., Zhou, Y., Zhang, Y., Zhong, L., Wang, J., Chen, J., 2022. DKDFN: domain knowledge-guided deep collaborative fusion network for multimodal unitemporal remote sensing land cover classification. ISPRS J. Photogramm. Remote Sens. 186, 170-189. <https://doi.org/10.1016/j.isprsjprs.2022.02.013>.
- Liu, A., Cheng, X., Chen, Z., 2021. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. Remote Sens. Environ. 264, 112571. <https://doi.org/10.1016/j.rse.2021.112571>.
- Lundberg, S.M. and Lee, S.-I., 2017. A unified approach to interpreting model predictions. Proc. 31st Conf. Neural Inf. Process. Syst., Long Beach, CA, USA, 4768–4777. https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.
- MacManus, K., Balk, D., Engin, H., McGranahan, G., Inman, R., 2021. Estimating population and urban areas at risk of coastal hazards, 1990-2015: how data choices matter. Earth Syst. Sci. Data 13, 5747-5801. <https://doi.org/10.5194/essd-13-5747-2021>.
- Moya, L., Mas, E., Koshimura, S., 2020. Learning from the 2018 Western Japan Heavy Rains to Detect Floods during the 2019 Hagibis Typhoon. Remote Sens. (Basel) 12, 2244. <https://doi.org/10.3390/rs12142244>.
- Neuenschwander, A. and Pitts, K., 2019. The ATL08 land and vegetation product for the ICESat-2 mission. Remote Sens. Environ. 221, 247-259. <https://doi.org/10.1016/j.rse.2018.11.005>.
- Orynbaikyzy, A., Plank, S., Vetrita, Y., Martinis, S., Santoso, I., Ismanto, R.D., Chusnayah, F., Tjahjaningsih, A., Suwarsono, Genzano, N., Marchese, F., Khomarudin, M.R., Strunz, G., 2023. Joint use of Sentinel-2 and Sentinel-1 data for rapid mapping of volcanic eruption deposits in Southeast Asia. Int. J. Appl. Earth Obs. Geoinf. 116, 103166. <https://doi.org/10.1016/j.jag.2022.103166>.
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J.B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. Remote Sens. Environ. 112165. <https://doi.org/10.1016/j.rse.2020.112165>.
- Rani, M. and Khotimah, N., 2021. Disaster risk analysis of Merapi volcano eruption in Cangkringan district Sleman regency. IOP Conf. Ser.: Earth Environ. Sci. 884, 012051. <https://doi.org/10.1088/1755-1315/884/1/012051>.
- Roberts, D.W. and Cooper, S.V., 1989. Concepts and techniques of vegetation mapping. In: Ferguson, D., Morgan, P., Johnson, F.D., (eds.). Land classifications based on vegetation: applications for resource management. USDA For. Serv. Gen. Tech. Rep. INT-257, Ogden, UT, 90-96.

- Saha, S., Ebel, P., Zhu, X.X., 2022. Self-Supervised multisensor change detection. *IEEE Transac. Geosci. Remote Sens.* 60, 4405710. <http://www.doi.org/10.1109/TGRS.2021.3109957>.
- Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., Macomber, S.A., 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sens. Environ.* 75, 230–244. [https://doi.org/10.1016/S0034-4257\(00\)00169-3](https://doi.org/10.1016/S0034-4257(00)00169-3).
- Stehman, S.V., 1999. Basic probability sampling designs for thematic map accuracy assessment. *Int. J. Remote Sens.* 20, 2423-2441. <https://doi.org/10.1080/014311699212100>.
- The World Bank, 2012. REKOMPAK; Rebuilding Indonesia's Communities After Disasters [WWW Document]. Financial Literacy and Education Russia Trust Fund, pp. 43-44. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/128361468267347740/rekompak-rebuilding-indonesias-communities-after-disasters>. (accessed 10 January 2019).
- Tong, X.-Y., Xia, G.-S., Lu, Q., Shen, H., Li, S., You, S., Zhang, L., 2020. Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sens. Environ.* 237, 1-20. <https://doi.org/10.1016/j.rse.2019.111322>.
- UNISDR (United Nations Office for Disaster Risk Reduction), 2015. Sendai framework for disaster risk reduction 2015-2030. <https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030> (accessed 14 September 2022).
- Utami, S.N.H., Purwanto, B.H., Marwasta, D., 2018. Land management for agriculture after the 2010 Merapi eruption. *Planta Tropika* 6, 32-38. <https://doi.org/10.18196/pt.2018.078.32-38>.
- Wang, Y., Albrecht, C.M., Braham, N.A.A., Mou, L., Zhu, X.X., 2022. Self-supervised learning in remote sensing: A review. *IEEE Geosci. Remote Sens. Mag.* 10, 213-247. <http://www.doi.org/10.1109/MGRS.2022.3198244>.
- Wijaya, N., 2018. Disaster risk reduction and climate change adaptation integration into peri-urban development planning: a case study of Bandung metropolitan area, Indonesia. *IOP Conf. Ser.: Earth Environ. Sci.* 145, 012079. <http://www.doi.org/10.1088/1755-1315/145/1/012079>.
- Winkler, K., Fuchs, R., Rounsevell, M., Herold, M., 2021. Global land use changes are four times greater than previously estimated. *Nat. Commun.* 12, 2501. <https://doi.org/10.1038/s41467-021-22702-2>.
- Wulder, M.A., Hermosilla, T., Stinson, G., Gougeon, F.A., White, J.C., Hill, D.A., Smiley, B.P., 2020. Satellite-based time series land cover and change information to map forest area consistent with national and international reporting requirements. *For.: An Int. J. For. Res.* 93, 331–343. <https://doi.org/10.1093/forestry/cpaa006>.
- Xu, C., Du, X., Yan, Z., Zhu, J., Xu, S., Fan, X., 2022. VSDF: A variation-based spatiotemporal data fusion method. *Remote Sens. Environ.* 283, 113309. <https://doi.org/10.1016/j.rse.2022.113309>.
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., Bates, P.D., 2017. A high-accuracy map of global terrain elevations. *Geophys. Res. Lett.* 44, 5844-5853. <https://doi.org/10.1002/2017GL072874>.
- Yang, Y. and Diez-Roux, A.V., 2012. Walking distance by trip purpose and population subgroups. *Am. J. of Prev. Med.* 43, 11-19. <https://doi.org/10.1016/j.amepre.2012.03.015>.
- Yang, J., 2021. Fast TreeSHAP: Accelerating SHAP value computation for Trees. <https://doi.org/10.48550/arXiv.2109.09847>.
- Yin, F., Lewis, P.E., Gomez-Dans, J.L., 2022. Bayesian atmospheric correction over land: Sentinel-2/MSI and Landsat 8/OLI. *Geosci. Model Dev.* 15, 7933-7976. <https://doi.org/10.5194/gmd-15-7933-2022>.

Yue, J., Fang, L., Ghamisi, P., Xie, W., Li, J., Chanussot, J., Plaza, A., 2022. Optical remote sensing image understanding with weak supervision: Concepts, methods, and perspectives. *IEEE Geosci. Remote Sens. Mag.* 10, 250-269. <http://doi.org/10.1109/MGRS.2022.3161377>.

Zhao, K., Wulder, M.A., Hu, T., Bright, R., Wu, Q., Qin, H., Li, Y., Toman, E., Mallick, B., Zhang, X., Brown, M., 2019. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote Sens. Environ.* 232, 111181. <https://doi.org/10.1016/j.rse.2019.04.034>.

Supplementary Information

This file provides supplementary information complementing Chapter 4 results.

Figure S1. Cropland NDVI time-series from March (growth) through October (senescence). Smoothed, mean, and raw amplitudes are compared. Raw metrics outperformed smoothed and mean metrics given their ability to characterize signal variability while avoiding a model to overfit. Further, low levels of sensor-derived noise are seen at the ~80, ~163, ~462, and ~736 markers.

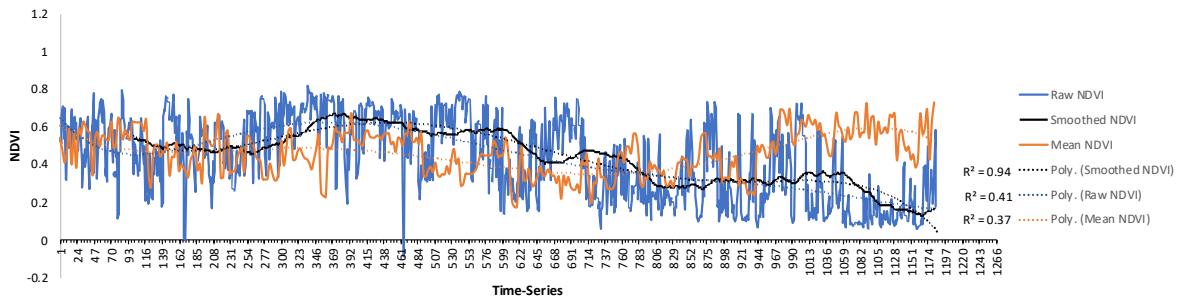


Table S1. Multitemporal metrics for the Indo-Malayan tropical ecozone. The enumerated ranges represent smoothed, mean, and multitemporal raw metrics. Water and wetland samples were constrained to the number of available units and may not cast a significant representation of the ecozone.

01 Dense Short Vegetation	Samples	Min	Max	Standard Error	Range
<i>Smoothed (NIR-Red / NIR+Red)</i>	1,192	0.254	0.742	0.004	0.258 0.738
<i>Mean</i>	323	0.255	0.755	0.006	0.261 0.749
<i>Temporal</i>	1,292	0.085	0.858	0.005	0.090 0.853
<i>Smoothed (Wetness Index)</i>	1292	-3.866	8.466	0.056	-3.81 8.41
<i>Mean</i>	323	-3.866	8.466	0.112	-3.753 8.353
<i>Temporal</i>	1,292	-3.866	8.466	0.056	-3.810 8.410
<i>Smoothed (Solar Radiation)</i>	1292	4633.647	7884.382	13.617	4647.26 7870.77
<i>Mean</i>	323	4633.65	7884.38	27.27	4660.91 7857.12
<i>Temporal</i>	1,292	4633.65	7884.38	13.62	4647.26 7870.77
<i>Smoothed (NIR-SWIR2/NIR+SWIR2)</i>	1192	0.226	0.715	0.004	0.230 0.711
<i>Mean</i>	323	0.163	0.685	0.005	0.168 0.680
<i>Temporal</i>	1,292	-0.037	0.761	0.005	-0.033 0.756
<i>Smoothed (Tasseled Cap Greenness)</i>	1292	1.49E-06	2.01E-01	1.45E-03	0.001 0.200
<i>Mean</i>	323	0.004	0.225	0.002	0.006 0.223
<i>Temporal</i>	1,292	-0.066	0.321	0.002	-0.064 0.319
02 Open Tree Cover	Samples	Min	Max	Standard Error	Range
<i>Smoothed (NIR-Red / NIR+Red)</i>	2,144	0.437	0.749	0.002	0.439 0.747
<i>Mean</i>	560	0.305	0.795	0.004	0.309 0.791
<i>Temporal</i>	2,244	0.140	0.858	0.003	0.143 0.855
<i>Smoothed NDWI_I (NIR-SWIR1 / NIR+SWIR1)</i>	2,144	0.115	0.377	0.002	0.117 0.376
<i>Mean</i>	560	0.007	0.478	0.003	0.011 0.475
<i>Temporal</i>	2,244	-0.119	0.505	0.002	-0.116 0.502
<i>Smoothed (Tasseled Cap Greenness)</i>	2,144	0.056	0.171	0.001	0.057 0.171
<i>Mean</i>	560	0.015	0.206	0.001	0.016 0.204
<i>Temporal</i>	2,244	-0.053	0.278	0.001	-0.052 0.277
<i>Smoothed (Solar Radiation)</i>	2244	3804.301	7598.074	14.280	3818.58 7583.79
<i>Mean</i>	560	3804.30	7598.07	25.61	3829.91 7572.46
<i>Temporal</i>	2,244	3804.30	7598.07	14.28	3818.58 7583.79
<i>Smoothed GNDVI (NIR-Green / NIR+Green)</i>	2144	0.445	0.656	0.001	0.446 0.655
<i>Mean</i>	560	0.336	0.701	0.003	0.339 0.698
<i>Temporal</i>	2,244	0.218	0.745	0.002	0.220 0.743
<i>Smoothed (Ground Height)</i>	2244	20.85	979.44	4.28	25.13 975.16
<i>Mean</i>	560	20.848	979.445	8.587	29.43 970.86
<i>Static</i>	2,244	20.848	979.445	4.283	25.13 975.16
<i>Smoothed (Slope)</i>	2244	0.05	33.03	0.17	0.22 32.86
<i>Mean</i>	560	0.049	33.033	0.337	0.386 32.695
<i>Static</i>	2,244	0.049	33.033	0.168	0.217 32.864
<i>Smoothed (NIR-SWIR2 / NIR+SWIR2)</i>	2144	0.385	0.685	0.002	0.386 0.683
<i>Mean</i>	560	0.195	0.743	0.004	0.199 0.739
<i>Temporal</i>	2,244	0.083	0.776	0.003	0.086 0.773
03 Dense Tree Cover	Samples	Min	Max	Standard Error	Range
<i>Smoothed (NIR-Red / NIR+Red)</i>	1060	0.558	0.731	0.002	0.560 0.730

<i>Mean</i>	290	0.314	0.809	0.005	0.320	0.804
<i>Temporal</i>	1,160	0.140	0.852	0.003	0.143	0.848
<i>Smoothed (NIR-SWIR2 / NIR+SWIR2)</i>	1060	0.506	0.697	0.002	0.508	0.695
<i>Mean</i>	290	0.250	0.748	0.005	0.255	0.742
<i>Temporal</i>	1,160	0.083	0.781	0.003	0.086	0.778
<i>Smoothed NDWI_I (NIR-SWIR1/NIR+SWIR1)</i>	1060	0.217	0.392	0.001	0.219	0.391
<i>Mean</i>	290	0.065	0.457	0.005	0.070	0.453
<i>Temporal</i>	1,160	-0.055	0.497	0.003	-0.052	0.494
<i>Smoothed (Tasseled Cap Greenness)</i>	1060	0.090	0.169	0.001	0.090	0.169
<i>Mean</i>	290	0.025	0.270	0.003	0.027	0.267
<i>Temporal</i>	1,160	-0.053	0.304	0.001	-0.052	0.302
<i>Smoothed (Tasseled Cap Wetness)</i>	1060	-0.129	-0.088	0.000	-0.129	-0.089
<i>Mean</i>	290	-0.205	-0.032	0.002	-0.204	-0.034
<i>Temporal</i>	1,160	-0.326	-0.020	0.001	-0.325	-0.021
<i>Smoothed (Ground Height)</i>	1160	26.20	940.67	6.58	32.79	934.09
<i>Mean</i>	290	26.20	940.67	13.19	39.39	927.48
<i>Static</i>	1,160	26.20	940.67	6.58	32.79	934.09
<i>Smoothed (Solar Radiation)</i>	1160	4254.61	7709.27	20.44	4275.05	7688.83
<i>Mean</i>	290	4254.61	7709.27	40.93	4295.55	7668.34
<i>Temporal</i>	1,160	4254.61	7709.27	20.44	4275.05	7688.83
04 Wetland	Samples	Min	Max	Standard Error	Range	
<i>Smoothed (NDVI)</i>	90	0.361	0.591	0.006	0.367	0.585
<i>Mean</i>	26	-0.056	0.643	0.029	-0.026	0.614
<i>Temporal</i>	104	-0.109	0.717	0.018	-0.092	0.699
<i>Smoothed (SWIR1-SWIR2/SWIR1+SWIR2)</i>	90	0.268	0.384	0.003	0.271	0.381
<i>Mean</i>	26	0.176	0.382	0.008	0.185	0.373
<i>Static</i>	104	0.166	0.460	0.006	0.172	0.455
<i>Smoothed (Tasseled Cap Greenness)</i>	90	0.020	0.104	0.002	0.022	0.102
<i>Mean</i>	26	-0.043	0.142	0.009	-0.033	0.132
<i>Temporal</i>	104	-0.062	0.175	0.006	-0.056	0.169
<i>Smoothed (Tasseled Cap Wetness)</i>	90	-0.150	-0.035	0.003	-0.146	-0.038
<i>Mean</i>	26	-0.205	0.015	0.013	-0.193	0.002
<i>Temporal</i>	104	-0.225	0.017	0.007	-0.218	0.010
<i>Smoothed (Wetness Index)</i>	104	-0.366	7.620	0.289	-0.078	7.332
<i>Mean</i>	26	-0.366	7.620	0.586	0.220	7.034
<i>Temporal</i>	104	-0.366	7.620	0.289	-0.078	7.332
05 Built Up	Samples	Min	Max	Standard Error	Range	
<i>Smoothed (Slope)</i>	1220	0.03	14.30	0.08	0.11	14.22
<i>Mean</i>	305	0.028	14.30	0.156	0.183	14.14
<i>Static</i>	1,220	0.028	14.30	0.078	0.105	14.220
<i>Smoothed NDVI (NIR-R / NIR+R)</i>	1120	0.253	0.557	0.002	0.256	0.555
<i>Mean</i>	305	0.080	0.730	0.007	0.087	0.722
<i>Temporal</i>	1,220	0.038	0.798	0.004	0.042	0.794
<i>Smoothed (NIR-SWIR2/NIR+SWIR2)</i>	1120	0.168	0.469	0.002	0.170	0.467
<i>Mean</i>	305	0.007	0.646	0.008	0.015	0.638
<i>Temporal</i>	1,220	-0.037	0.726	0.004	-0.033	0.722
<i>Smoothed (SWIR1-SWIR2/SWIR1+SWIR2)</i>	1120	0.153	0.270	0.001	0.154	0.269
<i>Mean</i>	305	0.038	0.391	0.004	0.042	0.387
<i>Temporal</i>	1,220	0.018	0.447	0.002	0.020	0.445
<i>Smoothed (Tasseled Cap Greenness)</i>	1120	-0.014	0.088	0.001	-0.014	0.087
<i>Mean</i>	305	-0.149	0.175	0.003	-0.147	0.173
<i>Temporal</i>	1,220	-0.169	0.207	0.001	-0.167	0.206
<i>Smoothed NDWI_I (NIR-SWIR1/NIR+SWIR1)</i>	1120	0.015	0.228	0.002	0.016	0.226
<i>Mean</i>	305	-0.140	0.371	0.005	-0.134	0.366
<i>Temporal</i>	1,220	-0.163	0.423	0.003	-0.160	0.420
06 Water	Samples	Min	Max	Standard Error	Range	
<i>Mean NDWI_I (NIR-SWIR1/NIR+SWIR1)</i>	8	0.233	0.465	0.027	0.260	0.438
<i>Temporal</i>	32	-0.024	0.523	0.029	0.005	0.494
<i>Mean (Tasselled Cap Greenness)</i>	8	-0.016	0.018	0.005	-0.012	0.013
<i>Temporal</i>	32	-0.046	0.055	0.003	-0.043	0.051
<i>Mean GNDVI (NIR-Green/NIR+Green)</i>	8	0.045	0.309	0.035	0.080	0.274
<i>Temporal</i>	32	-0.318	0.519	0.040	-0.277	0.479
07 Cropland	Samples	Min	Max	Standard Error	Range	
<i>Smoothed NDVI (NIR-Red / NIR+Red)</i>	1404	0.254	0.744	0.003	0.257	0.740
<i>Mean</i>	376	0.176	0.735	0.006	0.182	0.729
<i>Temporal</i>	1,504	-0.078	0.833	0.005	-0.073	0.828
<i>Smoothed (BN)</i>	1404	-0.789	-0.456	0.002	-0.787	-0.458
<i>Mean</i>	376	-0.791	-0.283	0.004	-0.787	-0.287
<i>Temporal</i>	1,504	-0.845	-0.014	0.003	-0.841	-0.017
<i>Smoothed (Tasseled Cap Greenness)</i>	1404	-0.004	0.153	0.001	-0.003	0.152
<i>Mean</i>	376	-0.032	0.195	0.002	-0.030	0.193
<i>Temporal</i>	1,504	-0.067	0.249	0.002	-0.066	0.247
<i>Smoothed (Tasseled Cap Wetness)</i>	1404	-0.224	-0.078	0.001	-0.223	-0.079

Mean	376	-0.306	-0.040	0.003	-0.303	-0.043
Temporal	1,504	-0.406	0.033	0.002	-0.405	0.031
Smoothed (Ground Height)	1504	6.07	700.61	4.31	10.38	696.30
Mean	376	6.07	700.61	8.63	14.70	691.97
Temporal	1,504	6.070	700.61	4.31	10.38	696.30
Smoothed (Slope)	1504	0.088	21.536	0.094	0.182	21.443
Mean	376	0.088	21.536	0.187	0.276	21.349
Temporal	1,504	0.088	21.536	0.094	0.182	21.443
NDVI_Cultivation/Harvest_First Quartile	1,081			R2 = 0.93	0.133	0.321
B2 (Blue)	1,081			R2 = 0.84	0.042	0.058
B8 (NIR)	1,081			R2 = 0.76	0.222	0.240
B11 (SWIR1)	1,081			R2 = 0.86	0.147	0.178
NDVI_Senescence_Third Quartile	1,081			R2 = 0.93	0.321	0.563
B2 (Blue)	1,081			R2 = 0.84	0.058	0.074
B8 (NIR)	1,081			R2 = 0.76	0.240	0.273
B11 (SWIR1)	1,081			R2 = 0.86	0.178	0.252
NDVI_Peak growing_Last Quartile	1,081			R2 = 0.93	0.563	0.780
B2 (Blue)	1,081			R2 = 0.84	0.074	0.109
B8 (NIR)	1,081			R2 = 0.76	0.273	0.294
B11 (SWIR1)	1,081			R2 = 0.86	0.252	0.337

Table S2. Smoothed, mean, and raw metric performances. Thematic agreement variations can be seen for each metric set.

Metric Performances								
01_Dense Short Vegetation	Samples	Level 0	Level 1	Level 2	Level 3	Thematic Agreement	Height Agreement	Structural Agreement
<i>Smoothed Metrics</i>	36,728	1,561	8,981	16,991	9,195	25,600	19,364	25,584
		4.3%	24.5%	46.3%	25.0%	69.7%	52.7%	69.7%
<i>Mean Metrics</i>	36,728	1,505	8,737	17,001	9,485	26,246	19,364	25,584
		4.1%	23.8%	46.3%	25.8%	71.5%	52.7%	69.7%
<i>Temporal Raw Metrics</i>	36,728	110	5,715	17,572	13,331	35,904	19,364	25,584
		0.3%	15.6%	47.8%	36.3%	97.8%	52.7%	69.7%
02_Open Tree Cover	Samples	Level 0	Level 1	Level 2	Level 3	Thematic Agreement	Height Agreement	Structural Agreement
<i>Smoothed Metrics</i>	36,788	3,459	11,078	15,974	6,277	15,333	22,582	23,942
		9.4%	30.1%	43.4%	17.1%	41.7%	61.4%	65.1%
<i>Mean Metrics</i>	36,788	1,967	9,332	15,891	9,598	25,987	19,979	23,942
		5.3%	25.4%	43.2%	26.1%	70.6%	54.3%	65.1%
<i>Temporal Raw Metrics</i>	36,788	912	7,802	16,442	11,632	31,661	19,979	23,942
		2.5%	21.2%	44.7%	31.6%	86.1%	54.3%	65.1%
03_Dense Tree Cover	Samples	Level 0	Level 1	Level 2	Level 3	Thematic Agreement	Height Agreement	Structural Agreement
<i>Smoothed Metrics</i>	28,397	2,005	14,733	10,301	1,358	3,669	21,221	14,519
		7.1%	51.9%	36.3%	4.8%	12.9%	74.7%	51.1%
<i>Mean Metrics</i>	28,397	751	8,659	12,815	6,172	17,065	21,221	14,519
		2.6%	30.5%	45.1%	21.7%	60.1%	74.7%	51.1%
<i>Temporal Raw Metrics</i>	28,397	450	7,090	13,503	7,354	20,418	21,221	14,519
		1.6%	25.0%	47.6%	25.9%	71.9%	74.7%	51.1%
04_Wetland Cover	Samples	Level 0	Level 1	Level 2	Level 3	Thematic Agreement	Height Agreement	Structural Agreement
<i>Smoothed Metrics</i>	1,245	162	252	739	92	134	1,000	872
		13.0%	20.2%	59.4%	7.4%	10.8%	80.3%	70.0%
<i>Mean Metrics</i>	1,245	151	220	611	263	359	1,000	872
		12.1%	17.7%	49.1%	21.1%	28.8%	80.3%	70.0%
<i>Temporal Raw Metrics</i>	1,245	103	212	492	438	638	1,000	872
		8.3%	17.0%	39.5%	35.2%	51.2%	80.3%	70.0%
05_Built Up	Samples	Level 0	Level 1	Level 2	Level 3	Thematic Agreement	Height Agreement	Structural Agreement
<i>Smoothed Metrics</i>	24,602	2,722	5,316	12,506	4,058	6,742	16,618	19,142
		11.1%	21.6%	50.8%	16.5%	27.4%	67.5%	77.8%
<i>Mean Metrics</i>	24,602	986	4,006	8,430	11,180	18,646	16,618	19,142
		4.0%	16.3%	34.3%	45.4%	75.8%	67.5%	77.8%
<i>Temporal Raw Metrics</i>	24,602	424	3,627	7,450	13,101	22,070	16,618	19,142
		1.7%	14.7%	30.3%	53.3%	89.7%	67.5%	77.8%
06_Water	Samples	Level 0	Level 1	Level 2	Level 3	Thematic Agreement	Height Agreement	Structural Agreement
<i>Smoothed Metrics</i>	1,864	-	49	1,132	683	754	1,860	1,748
		0.0%	2.6%	60.7%	36.6%	40.5%	99.8%	93.8%
<i>Mean Metrics</i>	1,864	2	110	1,732	20	26	1,860	1,748
		0.1%	5.9%	92.9%	1.1%	1.4%	99.8%	93.8%
<i>Temporal Raw Metrics</i>	1,864	0	49	1,132	683	754	1,860	1,748
		0.0%	2.6%	60.7%	36.6%	40.5%	99.8%	93.8%

07_Cropland	Samples	Level 0	Level 1	Level 2	Level 3	Thematic Agreement	Height Agreement	Structural Agreement
<i>Smoothed Metrics</i>	38,181	4,747 12.4%	8,725 22.9%	17,219 45.1%	7,490 19.6%	13,620 35.7%	23,741 62.2%	28,272 74.0%
<i>Cultivation/Harvest</i>	8,216	1,199 7.6%	807 5.1%	5,576 35.4%	634 4.0%	952 6.0%	6,149 39.0%	6,760 42.9%
<i>Senescence</i>	14,192	1,202 8.5%	3,443 24.3%	5,302 37.4%	4,245 29.9%	7,539 53.1%	8,781 61.9%	10,462 73.7%
<i>Peak Growing</i>	15,773	2,346 28.6%	4,475 54.5%	6,341 77.2%	2,611 31.8%	5,129 62.4%	8,811 107.2%	11,050 134.5%
<i>Mean Metrics</i>	38,181	3,258 8.5%	9,209 24.1%	14,403 37.7%	11,311 29.6%	19,935 52.2%	23,741 62.2%	28,272 74.0%
<i>Cultivation/Harvest</i>	8,216	633 4.0%	1,129 7.2%	4,103 26.0%	2,351 14.9%	3,479 22.1%	6,149 39.0%	6,760 42.9%
<i>Senescence</i>	14,192	923 6.5%	3,483 24.5%	4,742 33.4%	5,044 35.5%	8,856 62.4%	8,781 61.9%	10,462 73.7%
<i>Peak Growing</i>	15,773	1,702 20.7%	4,597 56.0%	5,558 67.6%	3,916 47.7%	7,600 92.5%	8,811 107.2%	11,050 134.5%
<i>Raw Temporal Metrics</i>	38,181	2,005 5.3%	9,558 25.0%	11,774 30.8%	14,844 38.9%	25,625 67.1%	23,741 62.2%	28,272 74.0%
<i>Cultivation/Harvest</i>	8,216	248 1.6%	1,355 8.6%	2,501 15.9%	4,112 26.1%	5,784 36.7%	6,149 39.0%	6,760 42.9%
<i>Senescence</i>	14,192	688 4.8%	3,590 25.3%	4,221 29.7%	5,693 40.1%	9,868 69.5%	8,781 61.9%	10,462 73.7%
<i>Peak Growing</i>	15,773	1,069 13.0%	4,613 56.1%	5,052 61.5%	5,039 61.3%	9,973 121.4%	8,811 107.2%	11,050 134.5%

Table S3. Training sample classification, refinement, and filtering.

Classes	1. Classification					2. Refinement				3. Filtering			
	Sample (-in)	Level-0	Level-1	Level-2	Level-3	Samples (-out)	Level-2	Level-3	Level-0	Level-1	Level-2		
1. Dense Short Vegetation	36,728	110	5,715	17,572	13,331	36,728	17,172	13,331	110	5,715	400		
2. Open Tree Cover	36,788	912	7,802	16,442	11,632	36,788	5,413	11,632	912	7,802	11,029		
3. Dense Tree Cover	28,397	450	7,090	13,503	7,354	28,397	7,279	7,354	450	7,090	6,224		
4. Wetland	1,245	103	212	492	438	1,245	492	438	103	212	0		
5. Built Up	24,602	424	3,627	7,450	13,101	24,602	5,887	13,101	424	3,627	1,563		
6. Water	1,864	0	49	1,132	683	1,864	1,132	683	0	-	0		
7. Cropland	38,181	2,005	9,558	11,774	14,844	0	-	-	0	49	0		
7. Peak Growing	0	1,069	4,613	5,052	5,039	15,773	2,203	5,039	1,069	4,613	2,849		
8. Senescence	0	688	3,590	4,221	5,693	14,192	1,606	5,693	688	3,590	2,615		
9. Cultivation	0	248	1,355	2,501	4,112	8,216	536	4,112	248	1,355	1,965		
Total	167,805	4,004	34,053	68,365	61,383	167,805	41,720	61,383	4,004	34,053	26,645		
	100%	2%	20%	41%	37%	100%	25%	37%	2%	20%	16%		

Table S4. Permutation-based feature importance's. First, all features were tested. Then, 10 feature combinations were identified. Overfitting features are shaded and correlated features are in bold. Note that band correlations were not considered and default parameters were used for 20% unseen test data.

Overfitting and Correlation Evaluations		Features																									
Training Dataset	Features	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	Total
		B25	SWIR_1_2	NDWI_I	B11	RS2	MCARI	Tass Wet	B3	B2	B1	BG	NIR SWIR_2	NDWI_II	NDVI	B12	BG	BN	B4	B5	Tass GN	IRECI	B6	B8A	B7	B8	SAVI
Test Data	Overfitting	0.11	0.035	0.032	0.032	0.031	0.03	0.023	0.022	0.02	0.02	0.019	0.018	0.018	0.017	0.015	0.015	0.012	0.012	0.012	0.01	0.008	0.007	0.006	0.005	0.550	
	2	+/- 0.001	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.001	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.002	
Test Data	Correlations	0.108	0.035	0.033	0.033	0.032	0.031	0.025	0.024	0.022	0.02	0.02	0.019	0.019	0.018	0.017	0.016	0.015	0.014	0.013	0.013	0.012	0.009	0.009	0.008	0.008	0.879
	0	+/- 0.001	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.001	+/- 0.000	+/- 0.001	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	+/- 0.000	0.004	

Feature Permutations		Features																								
Model 1	Features	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total									
		15	SWIR_1_2	B11	RS2	NDWI_II	B1	B3	BG	NDVI	B2	B5	B4	TassGN	B6	SAVI	B8A									
Correlations	0.135	0.097	0.027	0.023	0.021	0.021	0.018	0.018	0.017	0.015	0.01	0.008	0.008	0.006	0.006	0.43										
1	+/- 0.002	+/- 0.003	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001									
Model 2	14	SWIR_1_2	B11	TassWet	B3	RS2	NDWI_II	B2	NDVI	B1	NIR_SWIR_2	B5	B4	B6	B8A											

	<i>Correlations</i>	0.126	0.029	0.027	0.025	0.018	0.016	0.015	0.015	0.013	0.012	0.011	0.08	0.07	0.002	0.459	
	0	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.016	
Model 3	14	SWIR_1_2	B3	B11	TassWet	NDVI	RS2	NIR_SWIR_2	B2	B5	B1	BN	B4	B6	B8A		
	<i>Correlations</i>	0.125	0.029	0.026	0.023	0.019	0.017	0.012	0.011	0.011	0.009	0.006	0.005	0.002	0.002	0.297	
	0	+/- 0.002	+/- 0.001	+/- 0.002	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.018	
Model 4	15	SWIR_1_2	B11	RS2	B3	NIR_SWIR_2	NDWI_II	IRECI	NDVI	B1	B2	BN	B5	B4	B6	B8A	
	<i>Correlations</i>	0.127	0.046	0.027	0.021	0.019	0.015	0.015	0.014	0.012	0.012	0.01	0.009	0.007	0.006	0.003	0.343
	1	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.015	
Model 5	15	SWIR_1_2	B11	TassWet	B3	RS2	NDVI	NIR_SWIR_2	B2	IRECI	B1	BN	B4	B5	B6	B8A	
	<i>Correlations</i>	0.118	0.025	0.024	0.024	0.019	0.017	0.014	0.01	0.008	0.008	0.008	0.006	0.005	0.003	0.003	0.292
	1	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.017	
Model 6	15	SWIR_1_2	NDWI_I	NDVI	B3	TassWet	B11	RS2	GB	B2	B5	B1	BN	B4	B6	B8A	
	<i>Correlations</i>	0.12	0.03	0.022	0.021	0.019	0.019	0.018	0.015	0.01	0.09	0.08	0.06	0.06	0.03	0.004	0.598
	0	+/- 0.002	+/- 0.002	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.000	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.017	
Model 7	15	SWIR_1_2	B11	RS2	B3	NIR_SWIR_2	GB	IRECI	B2	B1	B5	B4	NDVI	B6	SAVI	B8A	
	<i>Correlations</i>	0.121	0.036	0.024	0.021	0.018	0.017	0.01	0.01	0.08	0.07	0.05	0.04	0.03	0.01	0.01	0.547
	0	+/- 0.002	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.018	
Model 8	15	SWIR_1_2	NDWI_I	B11	B3	RS2	B2	BN	B1	NDVI	IRECI	B5	TassWT	B4	B6	B8A	
	<i>Correlations</i>	0.113	0.04	0.029	0.026	0.025	0.012	0.011	0.01	0.01	0.009	0.008	0.007	0.007	0.005	0.004	0.316
	0	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.017	
Model 9	15	SWIR_1_2	B11	RS2	B3	NIR_SWIR_2	BG	IRECI	B2	NDVI	BN	B5	B1	B4	B6	B8A	
	<i>Correlations</i>	0.12	0.037	0.025	0.02	0.019	0.017	0.01	0.009	0.007	0.007	0.006	0.006	0.004	0.004	0.003	0.294
	1	+/- 0.002	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.018	
Model 10	14	SWIR_1_2	B11	RS2	B3	NIR_SWIR_2	BG	NDVI	B2	B5	B1	B4	BN	B6	B8A		
	<i>Correlations</i>	0.121	0.036	0.023	0.021	0.02	0.018	0.011	0.009	0.009	0.007	0.005	0.005	0.004	0.003	0.292	
	0	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.002	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.001	+/- 0.015	

Table S5. Supervised mean-weighted sample design. The upper table provides class weights from a preliminary classification using model 10. The lower table denotes mean-weighted allocations.

Classes	PixelSum	Percentage %	Area [metre^2]	Class Weight
1	7392461	6.788441372	6677624451	0.06788
2	12937193	11.88012709	11686191690	0.1188
3	17522660	16.09092697	15828252982	0.16091
4	1314928	1.207488499	1187777029	0.01207
5	11404567	10.47272815	10301767632	0.10473
6	25923388	23.80525235	23416646982	0.23805
7	13605844	12.49414428	12290185443	0.12494
8	15131535	13.89517486	13668345102	0.13895
9	3665190	3.365716428	3310773281	0.03366

ICESat-2 Ground Track ATL0820190904

Classes	Weighed	Equal	Mean-Weighted	L3 reference samples available
1	20	33	27	19
2	35	33	34	53
3	47	33	40	42
4	4	33	18	2
5	31	33	32	39
6	70	33	52	10
7	37	33	35	35
8	41	33	37	57
9	10	33	21	12
Total	295	295	295	269

Table S6. Mean absolute difference estimations. Values represent per-class absolute differences to estimate the mean difference for each model and the net average of model estimates.

Mean Absolute Difference									
	Dense Short Vegetation	Open Tree Cover	Dense Tree Cover	Wetland	Built Up	Water	Cropland Strata	Mean Difference	OA
Model 1	SMW	37.5%	66.7%	75.0%	-	77.3%	80.0%	80.8%	73.9%
	SRS	36.7%	42.6%	66.7%	-	84.4%	84.6%	78.9%	64.9%
	Absolute Difference	0.8%	24.0%	8.3%	-	7.1%	4.6%	1.9%	7.80% 9.0%
Model 2	SMW	75.0%	61.4%	82.5%	-	90.9%	80.0%	74.0%	74.6%
	SRS	43.3%	37.7%	71.8%	-	87.5%	84.6%	73.3%	63.8%
	Absolute Difference	31.7%	23.7%	10.7%	-	3.4%	4.6%	0.7%	12.47% 10.8%
Model 3	SMW	75.0%	63.2%	80.0%	-	86.4%	80.0%	76.9%	74.6%
	SRS	50.0%	41.0%	71.8%	-	84.4%	84.6%	75.6%	65.7%

	Absolute Difference	25.0%	22.2%	8.2%	-	2.0%	4.6%	1.4%	10.56%	9.0%
Model 4	SMW	75.0%	59.7%	75.0%	-	81.8%	80.0%	72.1%	70.5%	
	SRS	50.0%	36.1%	64.1%	-	81.3%	84.6%	74.4%	62.7%	
	Absolute Difference	25.0%	23.6%	10.9%	-	0.6%	4.6%	2.3%	11.17%	7.8%
Model 5	SMW	50.0%	64.9%	80.0%	-	77.3%	80.0%	76.9%	72.8%	
	SRS	36.7%	44.3%	71.8%	-	84.4%	84.6%	78.9%	66.0%	
	Absolute Difference	13.3%	20.7%	8.2%	-	7.1%	4.6%	2.0%	9.31%	6.7%
Model 6	SMW	75.0%	63.2%	80.0%	-	81.8%	80.0%	77.9%	75.0%	
	SRS	33.3%	45.9%	71.8%	-	84.4%	84.6%	77.8%	65.7%	
	Absolute Difference	41.7%	17.3%	8.2%	-	2.6%	4.6%	0.1%	12.40%	9.3%
Model 7	SMW	50.0%	64.9%	77.5%	-	84.1%	80.0%	79.8%	75.0%	
	SRS	40.0%	44.3%	69.2%	-	84.4%	84.6%	78.9%	66.0%	
	Absolute Difference	10.0%	20.7%	8.3%	-	0.3%	4.6%	0.9%	7.46%	9.0%
Model 8	SMW	75.0%	64.9%	82.5%	-	79.5%	80.0%	79.8%	75.4%	
	SRS	40.0%	42.6%	69.2%	-	81.3%	84.6%	80.0%	65.7%	
	Absolute Difference	35.0%	22.3%	13.3%	-	1.7%	4.6%	0.2%	12.85%	9.7%
Model 9	SMW	50.0%	64.9%	77.5%	-	77.3%	80.0%	80.8%	74.3%	
	SRS	36.7%	42.6%	69.2%	-	87.5%	84.6%	78.9%	65.7%	
	Absolute Difference	13.3%	22.3%	8.3%	-	10.2%	4.6%	1.9%	10.11%	8.6%
Model 10	SMW	50.0%	75.4%	77.5%	-	77.3%	80.0%	80.8%	76.9%	
	SRS	30.0%	49.2%	69.2%	-	84.4%	84.6%	78.9%	66.0%	
	Absolute Difference	20.0%	26.3%	8.3%	-	7.1%	4.6%	1.9%	11.35%	10.8%
MAD										

Note: Pixel-based producer's and overall accuracies (OA) are estimated for the Mean Absolute Difference (MAD). Wetland areas were not included given the lack of unit samples.

Table S7. Mean total error estimations. Filters are evaluated to determine which performed best. Structural agreements consider height correlations and thematic agreements. Height agreement consider the height thresholds and thematic agreements. Lastly, SRS considers photo-inspected samples and SMW considers photo-inspected samples with structural and height agreements. Note the 33% difference between SRS and SMW.

	L2: Structural Agreement	L2: Height Agreement	SRS	SMW
Model 1	0.743	0.556	0.397	0.287
Model 2	0.575	0.515	0.406	0.248
Model 3	0.53	0.509	0.424	0.237
Model 4	0.721	0.501	0.46	0.283
Model 5	0.696	0.559	0.344	0.234
Model 6	0.553	0.484	0.398	0.291
Model 7	0.659	0.591	0.37	0.279
Model 8	0.601	0.562	0.351	0.273
Model 9	0.672	0.543	0.399	0.246
Model 10	0.671	0.571	0.369	0.243
Total Mean	0.64	0.54	0.39 (1.00)	0.26 (0.67)

Table S8. Confusion matrix of preliminary land cover classification with a mean-weighted sampling design estimated with area proportions and shown in percentages with 95th percentile confidence intervals. The overall accuracy (OA) and user's accuracy variances were compared with Table S9.

Reference	Dense Short Vegetation 1	Open Tree Cover 2	Dense Tree Cover 3	Wetland 4	Built Up 5	Water 6	Cropland Peak Growth 7	Cropland Senescence 8	Cropland Cultivation 9	User's Accuracy 10
1	1.07%	2.86%	0.72%	0.00%	0.36%	0.00%	0.72%	1.07%	0.00%	15.79% ± 16.4%
2	0.00%	9.64%	0.67%	0.00%	0.45%	0.00%	0.22%	0.90%	0.00%	81.13% ± 10.5%
3	0.38%	1.15%	12.26%	0.00%	0.00%	0.00%	1.53%	0.77%	0.00%	76.19% ± 12.9%
4	0.00%	0.00%	0.00%	1.21%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00% ± 0%
5	0.00%	0.00%	0.27%	0.00%	9.13%	0.00%	0.00%	0.81%	0.27%	87.18% ± 10.5%
6	0.00%	0.00%	0.00%	2.65%	0.00%	21.16%	0.00%	0.00%	0.00%	88.89% ± 20.5%
7	0.36%	1.07%	1.07%	0.36%	0.00%	0.00%	9.64%	0.00%	0.00%	77.14% ± 13.9%
8	0.49%	0.24%	0.00%	0.00%	1.71%	0.00%	0.24%	11.21%	0.00%	80.70% ± 10.2%
9	0.00%	0.00%	0.00%	0.28%	0.00%	0.28%	0.00%	0.00%	2.81%	83.33% ± 21.1%
Producer's Accuracy	46.61%	64.43%	81.81%	26.89%	78.42%	98.69%	78.02%	76.00%	91.26%	OA = 78.12% ± 6.1%
	± 23.9%	± 6.9%	± 11.1%	± 0.0%	± 7.6%	± 21.5%	± 11.3%	± 8.0%	± 21.9%	

Table S9. Confusion matrix of preliminary land cover classification with a stratified random sampling design estimated with area proportions and shown in percentages with 95th percentile confidence intervals.

Reference

	Dense Short Vegetation	Open Tree Cover	Dense Tree Cover	Wetland	Built Up	Water	Cropland Peak Growth	Cropland Senescence	Cropland Cultivation	Users's Accuracy
	1	2	3	4	5	6	7	8	9	10
1	2.78%	2.47%	0.31%	0.00%	0.31%	0.00%	0.00%	0.62%	0.31%	40.91% ± 20.5%
2	2.16%	8.10%	0.54%	0.00%	0.00%	0.00%	0.27%	0.81%	0.00%	68.18% ± 13.8%
3	1.46%	4.39%	9.87%	0.00%	0.00%	0.00%	0.37%	0.00%	0.00%	61.36% ± 14.4%
4	0.13%	0.00%	0.00%	0.27%	0.27%	0.13%	0.13%	0.13%	0.13%	22.22% ± 27.2%
5	0.47%	0.93%	0.93%	0.00%	6.28%	0.00%	0.47%	0.70%	0.70%	60.00% ± 14.3%
6	0.00%	0.00%	0.00%	0.00%	0.00%	23.81%	0.00%	0.00%	0.00%	100.00% ± 0.0%
7	0.89%	1.79%	1.79%	0.00%	0.00%	0.00%	8.03%	0.00%	0.00%	64.29% ± 17.7%
8	1.11%	0.83%	0.28%	0.28%	0.56%	0.00%	0.00%	10.84%	0.00%	78.00% ± 11.5%
9	0.00%	0.00%	0.00%	0.00%	0.00%	0.22%	0.00%	0.00%	3.14%	93.33% ± 12.6%
Producer's Accuracy	30.84%	43.77%	71.99%	49.12%	84.73%	98.52%	86.67%	82.75%	73.36%	OA = 73.12% ± 4.5%
	± 8.4%	± 6.9%	± 10.3%	± 57.3%	± 10.1%	± 18.3%	± 12.7%	± 8.7%	± 8.0%	

Table S10. Confusion matrix of informed land cover classification with a mean-weighted sampling design estimated with area proportions and shown in percentages with 95th percentile confidence intervals. The overall accuracy (OA) and the 95% confidence interval margin are compared with Table S8.

Reference	Dense Short Vegetation	Open Tree Cover	Dense Tree Cover	Wetland	Built Up	Water	Cropland Peak Growth	Cropland Senescence	Cropland Cultivation	User's Accuracy
	1	2	3	4	5	6	7	8	9	10
1	1.20%	0.72%	0.24%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	55.56% ± 32.5%
2	0.00%	10.42%	0.93%	0.00%	0.23%	0.00%	0.46%	0.46%	0.00%	83.33% ± 9.9%
3	0.40%	0.79%	13.11%	0.00%	0.40%	0.00%	1.99%	0.00%	0.00%	78.57% ± 12.4%
4	0.00%	0.00%	0.00%	1.48%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00% ± 0
5	0.00%	0.25%	0.00%	0.00%	9.48%	0.25%	0.00%	0.75%	0.25%	86.36% ± 10.1%
6	0.00%	0.00%	0.00%	4.82%	0.00%	19.26%	0.00%	0.00%	0.00%	80.00% ± 24.8%
7	0.00%	1.52%	1.14%	0.38%	0.00%	0.00%	10.24%	0.00%	0.00%	77.14% ± 13.9%
8	0.50%	0.50%	0.00%	0.00%	0.99%	0.00%	0.00%	13.61%	0.00%	87.30% ± 8.2%
9	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.26%	100.00% ± 0.0%
Producer's Accuracy	57.39%	73.40%	85.04%	22.20%	85.41%	98.72%	80.71%	91.83%	92.90%	OA = 82.05% ± 6.9%
	± 22.7%	± 7.3%	± 11.1%	± 0.0%	± 8.9%	± 28.7%	± 11.9%	± 8.0%	± 0.8%	

Table S11. 2016-2021 land transitions expressed in percentage of area over total change. Cumulative hectares of change are in parenthesis sorted from rural to semi-urban to depict differences if any.

	Srunen (rural)	Gading (rural)	Banjarsari (rural)	Gondang (semi-urban)	Jetis Sumur (semi-urban)	Pagerjurang (semi-urban)	Karang Kendal (semi-urban)	Not Affected (rural)	Not Affected (semi-urban)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Forest > Crop	18% (16.4)	17% (12.3)	19% (14.5)	13% (8.1)	23% (23.4)	11% (8.5)	10% (10.9)	18% (19.5)	14% (20.9)
Forest > Built	34% (30.8)	24% (17.3)	28% (20.7)	30% (18.7)	27% (25.6)	35% (24.7)	38% (42.0)	30% (30.5)	34% (45.5)
Crop > Built	7% (6.6)	9% (7.1)	5% (3.8)	6% (3.5)	6% (4.7)	7% (4.7)	7% (7.2)	5% (5.0)	5% (5.8)
Crop > Forest	15% (15.0)	17% (14.1)	22% (15.4)	18% (12.3)	17% (15.4)	16% (12.7)	11% (11.4)	18% (19.5)	16% (22.4)
Built > Crop	9% (9.6)	13% (10.4)	9% (6.6)	11% (7.8)	7% (6.6)	8% (7.0)	5% (5.8)	8% (9.0)	6% (8.6)
Built > Forest	17% (19.6)	20% (16.0)	17% (12.0)	22% (14.3)	20% (19.7)	24% (20.2)	30% (30.9)	20% (22.1)	25% (33.4)
Total (Ha)	97.9	77.2	73	64.8	95.5	77.8	108.1	105.7	136.5

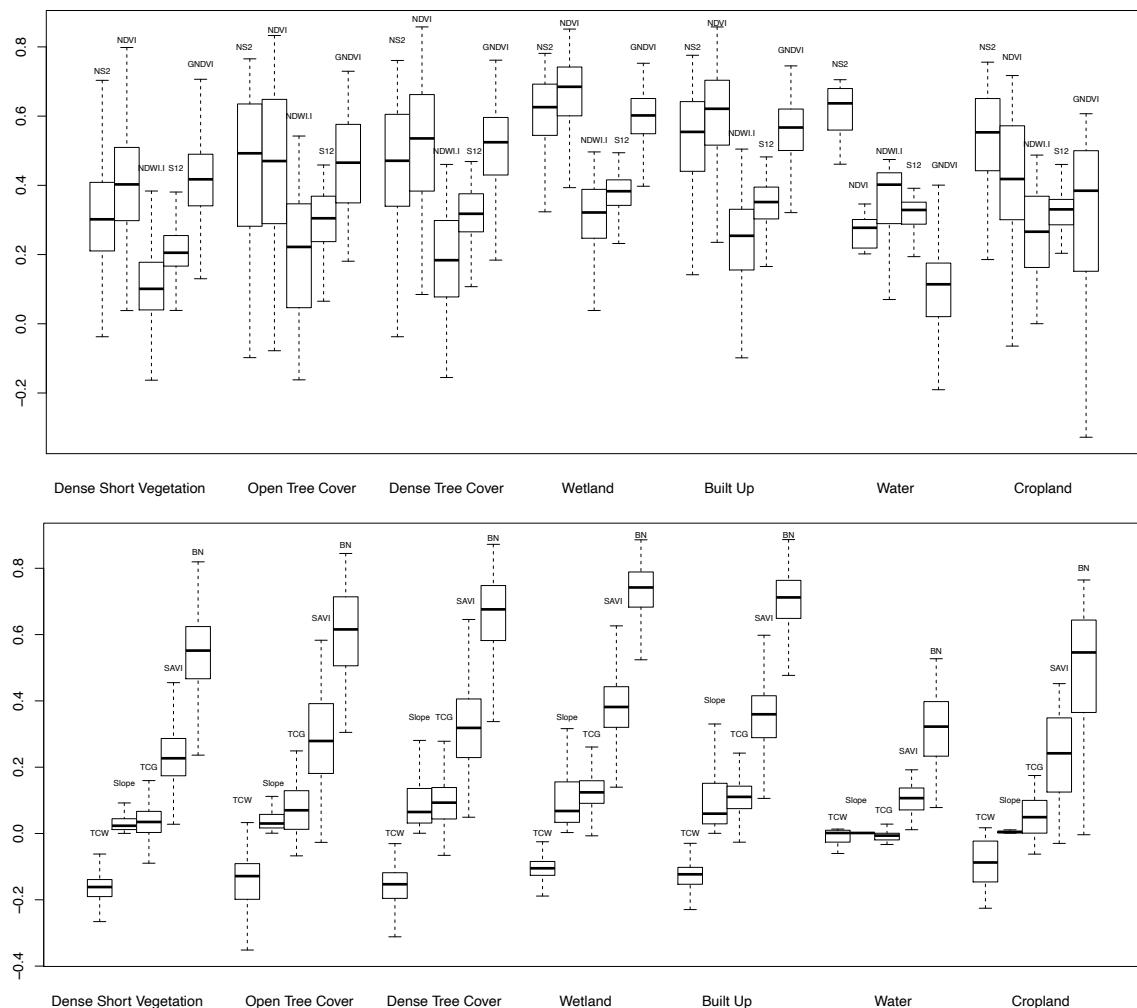


Figure S2. Box plot distributions of spectral amplitude for static and multitemporal metric units. Whiskers are located at 25th, median, and 75th quartiles. Inter-class comparisons highlight meaningful differences between predictors (e.g., BN for water and slope for built-up).

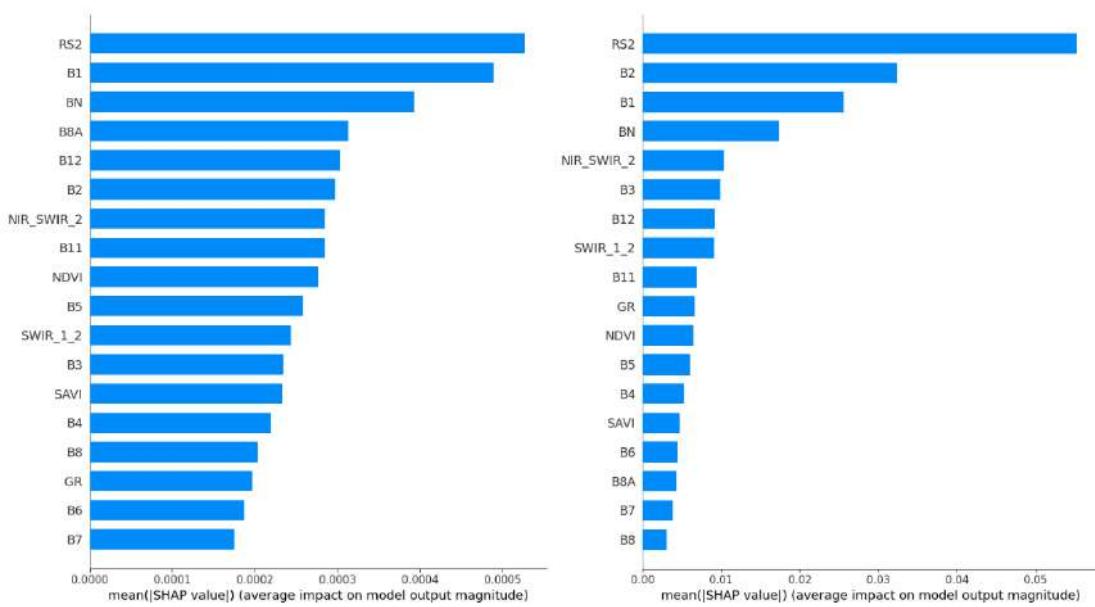


Figure S3. Mean SHAP feature importance's for cropland (left) and built-up (right) classes, sorted from highest to lowest. Both classes show RS2 as the highest predictor and BN follows with very high distinction.

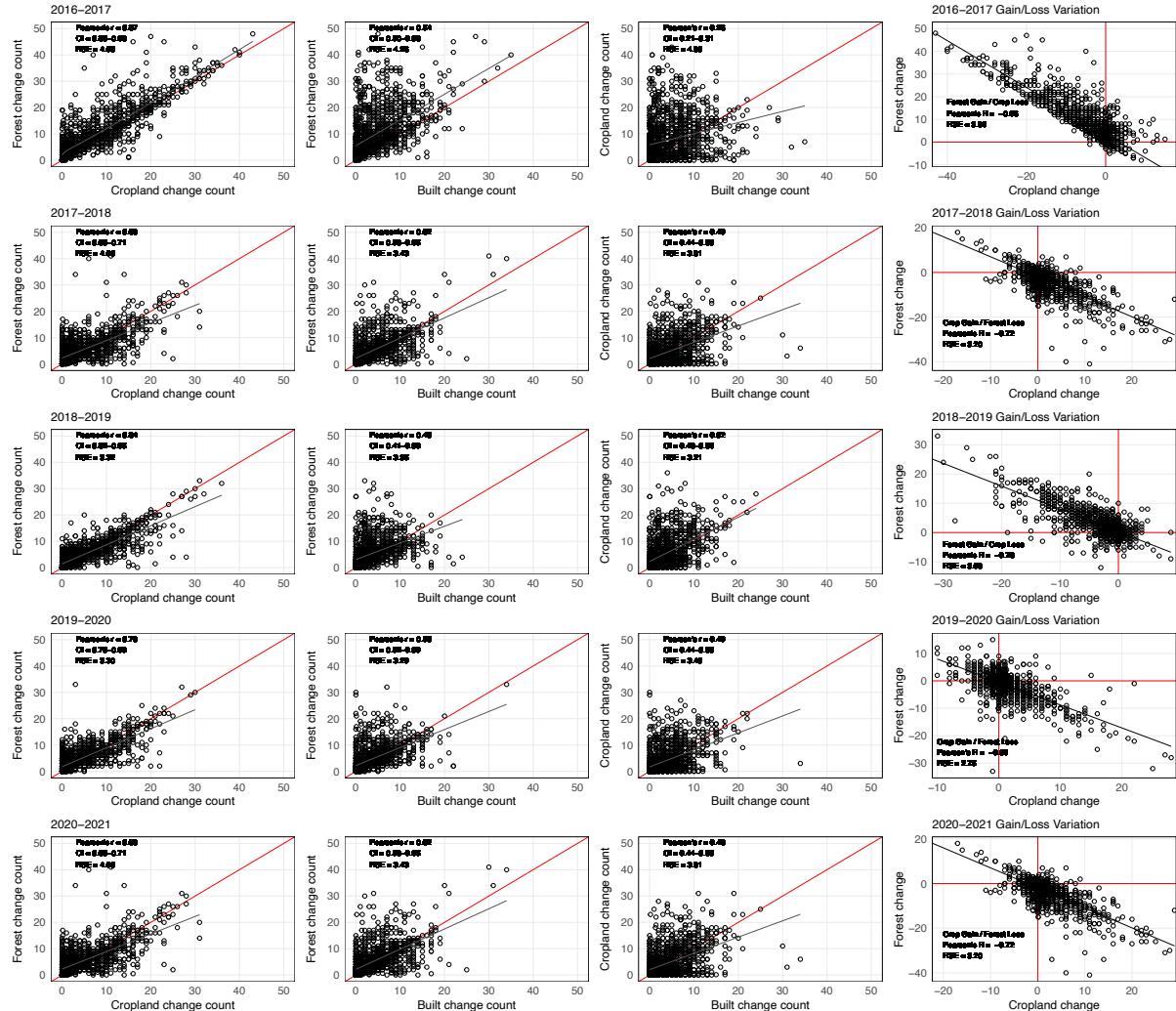


Figure S4. 2016-2021 Time-series of land change correlations sorted by correlation intensity.

Table S12. 2016-2021 Land change with area-reference estimators and error-adjusted standard errors expressed in 95% confidence intervals.

		Gross Cropland Change (ha)			Gross Built Up Change (ha)			Gross Forest Change (ha)		
		Gain	Loss	Net Change	Gain	Loss	Net Change	Gain	Loss	Net change
1. Rural	Srunen	26.01 ± 1.19	21.54 ± 0.98	4.47 ± 5.46	37.33 ± 2.78	29.21 ± 2.18	8.12 ± 5.2	34.59 ± 3.27	47.19 ± 4.46	-12.60 ± 4.93
2. Rural	Gading	22.73 ± 1.04	21.18 ± 0.97	1.55 ± 5.48	24.46 ± 1.82	26.38 ± 1.97	-1.92 ± 5.31	30.03 ± 2.84	29.66 ± 2.80	0.37 ± 5.13
3. Rural	Banjarsari	21.08 ± 0.96	19.26 ± 0.88	1.83 ± 5.49	24.55 ± 1.83	18.53 ± 1.38	6.02 ± 5.37	27.38 ± 2.59	35.23 ± 3.33	-7.85 ± 5.11
4. Urban	Gondang	15.97 ± 0.73	15.79 ± 0.72	0.18 ± 5.53	22.18 ± 1.65	22.18 ± 1.65	0.00 ± 5.36	26.65 ± 2.52	26.84 ± 2.54	-0.18 ± 5.19
5. Urban	Jetis Sumur	29.94 ± 1.37	20.17 ± 0.92	9.77 ± 5.45	30.39 ± 2.27	26.29 ± 1.96	4.11 ± 5.27	35.14 ± 3.32	49.02 ± 4.63	-13.87 ± 4.9
6. Urban	Pagerjurang	15.52 ± 0.71	17.34 ± 0.79	-1.83 ± 5.52	29.39 ± 2.19	27.20 ± 2.03	2.19 ± 5.27	32.86 ± 3.11	33.22 ± 3.14	-0.37 ± 5.07
7. Urban	Karang K.	16.61 ± 0.76	18.62 ± 0.85	-2.01 ± 5.51	49.20 ± 3.67	36.60 ± 2.73	12.60 ± 5.06	42.26 ± 4.00	52.85 ± 5.00	-10.59 ± 4.8
8. Rural	Not Affected	28.57 ± 1.30	24.55 ± 1.12	4.02 ± 5.44	35.51 ± 2.65	31.13 ± 2.32	4.38 ± 5.2	41.62 ± 3.93	50.02 ± 4.73	-8.40 ± 4.83
9. Urban	Not Affected	29.48 ± 1.35	28.20 ± 1.29	1.28 ± 8.18	51.30 ± 3.82	41.99 ± 3.13	9.31 ± 7.56	55.77 ± 5.27	66.36 ± 6.27	-10.59 ± 6.83
Total Affected		147.87 ± 6.75	133.90 ± 6.11	13.97 ± 8.84	217.51 ± 16.22	186.39 ± 13.90	31.13 ± 11.09	228.92 ± 21.64	274.01 ± 25.90	-45.09 ± 13.93

Note: 95% Confidence intervals for relative net change values consider the sum of squares for gain/loss estimates.

Sub-section 4.6.2

Not Affected (NA)

Affected (AF)

1. Cropland

Mean AF = 2.00 ha (43%)

Mean NA = 2.65 ha (57%)

Total = 4.65 ha (100%)

$$0.57 * (0.43) = 0.2451$$

Urban Dynamics = 24.5%

$$0.43 - 0.2451 = 0.185$$

Effects = 18.5%

2. Forest

Mean AF = -6.44 ha (40.4%)

Mean NA = -9.50 ha (59.6%)

Total = 15.94 ha (100%)

$$0.596 \times 0.404 = 0.24$$

Urban Dynamics = 24%

$$0.404 - 0.24 = 0.164$$

Effects = 16.4%

3. Built-up

Mean AF = 4.45 ha (39.4%)

Mean NA = 6.85 ha (60.6%)

Total = 11.30 ha (100%)

$$0.606 \times 0.394 = 0.239$$

Urban Dynamics = 23.9%

$$0.394 - 0.239 = 0.155$$

Effects = 15.5%

CHAPTER 5

LAND USE MICROSIMULATION MODEL FOR LIVELIHOOD DIVERSIFICATION AFTER THE 2010 MERAPI VOLCANO ERUPTIONS

Martin Garcia-Fry, Osamu Murao, Syamsul Bachri, and Luis Moya

This Chapter is based on: Garcia-Fry, M., Murao, O., Bachri, S., and Moya, L., 2022. Land use microsimulation model for livelihood diversification after the 2010 Merapi volcano eruptions.

Transportation Research Part D: Transport and Environment Vol. 104, No. 103189.

<https://doi.org/10.1016/j.trd.2022.103189>

Abstract: The 2010 Merapi Volcano eruptions caused significant damage in the Sleman Regency of Yogyakarta, Indonesia. Several households (HHs) were relocated to safely designated areas, but to sustain their livelihoods they have to travel back to their farmlands or change their identities as farmers. This study evaluates post-disaster mobility, aiming to clarify whether a land-use network with distributed livelihood options can complement rural labor in a recovery scenario. We collected sociodemographic and travel routine data in the affected areas. Then, the operational land use and transport microsimulation (OLUTM) model was performed to forecast a future scenario and gauge livelihood changes. Middle-aged farmers with upper median incomes and risk-taking behaviors were found to venture home businesses that contribute to rural labor. A 63% share of livelihoods diversified by farmers and 30% travel utility savings rendered a robust microsimulation system.

5.1. Introduction

On Tuesday, October 26, 2010, Indonesia's most active stratovolcano erupted thrice, forcing the evacuation of 400,000 people and destroying 2,847 homes (Maly et al. 2015; World Bank 2012). In response, the government relocated 2,608 households (HHs) on 23 resettlement sites in the Cangkringan district, Sleman Regency, Yogyakarta, Indonesia (World Bank 2016). The affected livelihoods were centered on dairy farming and agriculture, while tourism was the secondary employment with the standard minimum wages (Rindrasih 2018). As a result, people's recovery relied on daily trips to farmlands, dairy cooperatives, and tourist destinations.

HHs of the largest resettlement site, Pagerjurang, collectively believe that resettling has granted them safety from the volcano. However, resilience did not translate into secured livelihoods: (1) even though tourism became a main source of livelihood diversification, it failed to promote choices of livelihood in ample spheres of life; (2) middle-aged farmers and the elderly (48% of the active economic population) ([Dataset 2] Indonesia 2020) are “too tired from the daily grind of commuting to the dairy cooperative and their land on Mt. Merapi” (Miller M.A. 2018, p.202); and, (3) Yogyakarta’s rural urbanization is shifting land-use structures that generate place-attachment. As a result, identity loss is at stake, risking tourist-attractive value (Umaya et al. 2020). Therefore, how can we attract livelihood options closer to HHs while sustaining agricultural areas in the rural periphery of urban centers?

This paper describes our recent efforts to develop a sequential land use/transport model to evaluate post-disaster mobility and clarify whether a land-use network with distributed livelihood options can complement rural labor in a recovery scenario. The operational land use and transport microsimulation (OLUTM) model delivers a predictive network of land-use change. The method establishes key driving factors of a city’s mobility system and applies them to rural areas. Then, through behavioral and urban form analysis, the outcome is evaluated through land use and travel simulations to explain that mobility with livelihood changes.

Specific contributions are: first, a significant probability of residential to mixed land-use change was found for middle-aged farmers with risk-taking lifestyles and upper median incomes venturing home-business entrepreneurship (HBEs). HBEs diversify livelihoods with non-farm activities that capitalize on local products and complement rural labor. Second, a land-use treatment in nodes incited land-use conversion between the nodes. This is not common in land-use modeling research because it targets HBEs in an agronomic-centric network.

In the following chapter, transportation planning is shown to be a channel of development consistent with the needs of disaster-affected areas. Chapter 3 introduces the study area and our

method. Chapters 4 and Chapter 5 describe and perform sub-models of the OLUTM model, leading to an evaluation in Chapter 6 with findings and concluding remarks.

5.2. The Influence of Urban Form on Travel

Transportation and employment are important constituents of the urban form. Together, they can provide access to an equitable range of opportunities. Several meanings of accessibility exist (see Litman 2021), but this study follows the importance of reducing vehicle dependence as a factor of neighborhood accessibility (Krizek 2003). This means shorter distances in daily routines and having mobility choices for a variety of travel purposes. For instance, after the 2009 earthquake in L'Aquila, Italy, the isolation of houses made people dependent on private transport to access services, employment, or amenities (Contreras et al. 2017). In Yogyakarta, travel expenditures are not considered a burden because people adapted to the travel costs, lack of alternatives, and the perception that public transport is expensive (Herwangi et al. 2017).

While transport satisfies the need for mobility owing to the spatial separation of activities, urban form research examines physical characteristics that compose urban environments (size, shape, and the configuration of its parts) (Wegener and Fuerst 2004). For example, residents and firms want to be close to each other, saving on travel utilities and time. This causes economies of scale driven by agglomeration effects (Levinson and Wu 2020). These effects commonly expand in nodes of a transportation network, where local and inflowing resources meet profitably, creating building densities, activities, and people (Hynynen 2006). Nodes are defined by their scale of influence and employment shares in a traffic analysis zone (TAZ). TAZ's help to assign the expected growth and monitor trips in the network (Clifton et al. 2008).

In this paper, we evaluate development in a TAZ with three urban form measures: (i) density, defined by Handy et al. (2002) as the amount of population, employment, or building square-meters (m^2) per unit of area; (ii) land use, as the proximity of activities measured by the area of neighborhood

shopping in a 400-meter radius, equivalent to a walking distance of 800 m, assumed by Yang and Diez-Roux (2012) to be a median distance for walking; and, (iii) accessibility, quantified by connectivity, road intersections, and block lengths (Krizek 2003).

5.2.1 Model of Microsimulation

Land use and transport microsimulation models are modeling systems in which transportation and land use co-evolve over time (Miller 2018). The underlying concept is that development of the urban form and location of microdata (households, firms, agents, etc.) rely on accessibility to reach valued locations (Litman 2021). The OLUTM model follows the prospect theory, developed by Kahneman and Tversky (1979), to simulate random decisions based on the potential value of loss or gains in a two-stage decision-making process. The models here maximize the utility behavior of agents (people and vehicles) and objects (jobs, buildings, land uses, etc.) with sociodemographic data found by Kitamura (2009) to affect travel behavior.

While several influential studies have significantly advanced the computational barriers inherent to microscopic simulation (Miller and Salvini 2001; Waddell et al. 2003; Wagner and Wegener 2007; Zhu et al. 2018; Zondag et al. 2015), they mainly derive findings from urban areas and do not address post-disaster mobility in rural areas. Moreover, similar realities exist (see Yokohari 2006 and Bray 1994), but they were not meant to address urban growth.

5.3. Objective Area and Methodology

5.3.1 The Cangkringan Sub-District

Yogyakarta is the capital city of the Special Region of Yogyakarta (DIY), with a recorded population of 3,842,932 in 2019 ([Dataset 2] Indonesia 2020). The region is separated into four regency districts: Bantul, Kulon Progo, Sleman, Gunung Kidul, and Yogyakarta. Sleman is north of

Yogyakarta, with Mt. Merapi at its northern apex surrounded by fertile agrarian land and the Cangkringan sub-district shown in Figure 5.1.

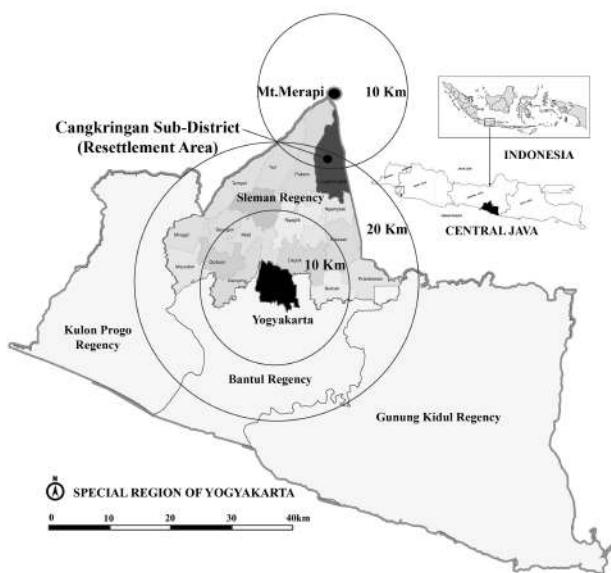


Figure 5.1. Study area in the Special Region of Yogyakarta. Suburban areas of the city extend through all four regencies reaching 20 km distances from the center. The Cangkringan sub-district is home to 75% of the relocated HHs, previously living within 10 kms from Mt. Merapi's summit.

The Sleman Regency has been facing rapid urbanization since before 2010, and is in need of better public service provision (Giyarsih 2010). A new spatial plan was adopted in 2012 called Rencana Tata Ruang Wilayah (RTRW) Kabupaten Sleman Tahun 2011-2031 as a top-down planning approach clarifying the distinction between urban and rural areas and regulating the scale and intensity of activities. Also, a law concerning sustainable agriculture food crops was issued by the Special Region of Yogyakarta in 2011, but urban expansion has continued to threaten food security in the region (Widowaty and Oktavian Artanto 2018).

There are three types of urban growth management initiatives imposed by governments worldwide: greenbelts (permanently limit an urban area with open spaces), urban growth boundaries (a line that separates urban and rural areas), and urban service boundaries (limiting the provision of public utilities) (Pramana 2016). Given that the Sleman Regency has partly become urban and partly

rural, we consider it imperative to implement a fourth type of urban growth management strategy as a mid-range option between greenbelts and urban growth boundary lines. An agronomic urban boundary (AUB) network is proposed herein as large tracts of rural land, interconnected with point-specific development nodes, with assigned influence areas and mixed-use economic corridors to form a revitalization network that promotes preservation and creation of new agricultural areas.

5.3.2 Methodology

First, travel routines were simulated to set the baseline results. Then, the OLUTM model was processed to evaluate future land use-transport interactions with output data seen in Figure 5.2.

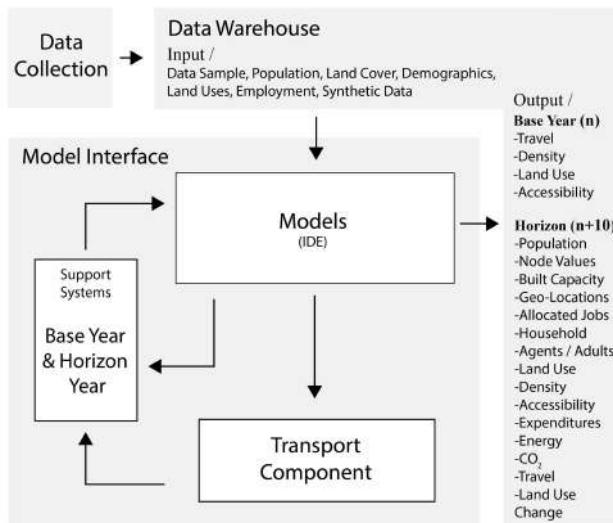


Figure 5.2. Model architecture.

(1) Transport Component

A field survey was conducted in September 2019 in the affected areas of the Cangkringan sub-district. The data collected was processed and simulated with sociodemographic data in the transport component of the model to estimate travel utilities. Urban form measures were then incorporated to enable a comprehensive diagnosis of livelihoods in 2019.

(2) OLUTM Model (2019-2030)

The models in Figure 5.3 reflect sequential steps for AUB network consolidation. It took seven data inputs from an exogenous data warehouse and ten sub-models sorted in two integration processes. The models are not expected to bring land use and travel to equilibrium because reality dictates that land use responds to the intensity of change. Therefore, the transport component evaluates travel performance on demand and the models are intended to reflect the observed behavior of agents. However, simplifications and abstractions of real behavior do exist.

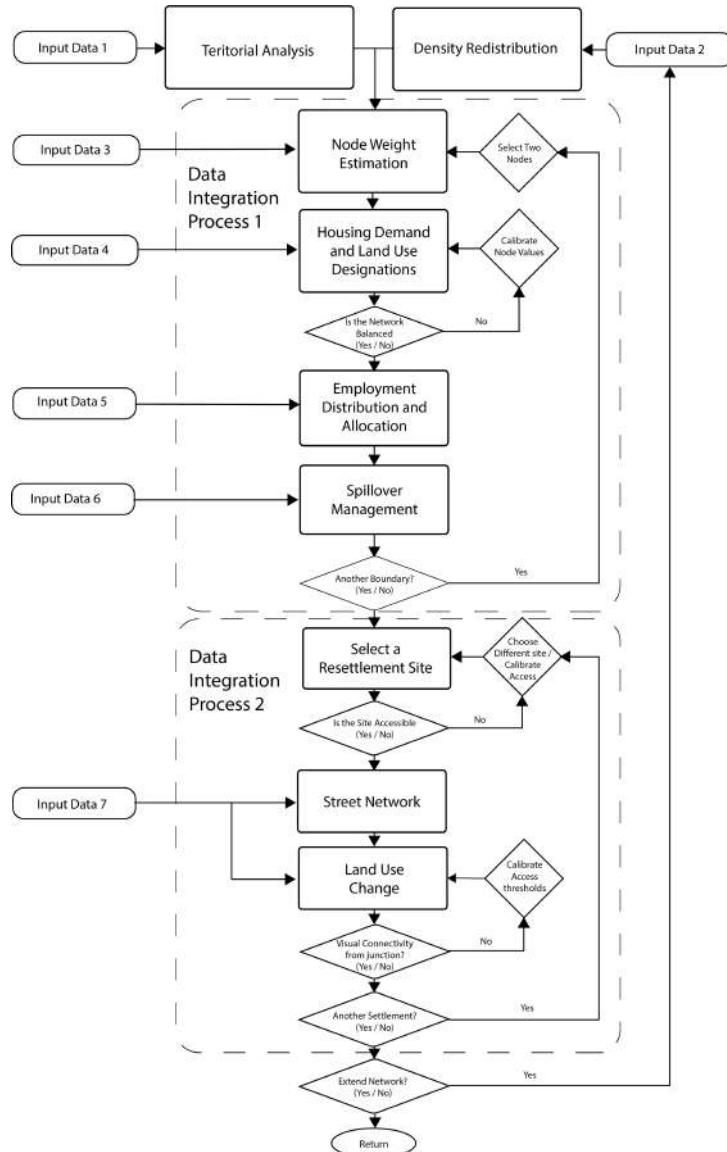


Figure 5.3. Layout of the OLUTM model (proposed method 3). Each iteration corresponds to one edge of the network. A future site-selection model is specified to locate resettlement sites in the second integration process.

5.4. Transport Component

Data of this study including our data sample, a synthetic dataset, model instruments, and travel simulations are available in Mendeley Data (Garcia-Fry 2021).

5.4.1 Data

First, field observations and survey questionnaires were conducted in the Pagerjurang resettlement site to gather sociodemographic data and travel routines. Subsequently, online open municipal datasets (Indonesia 2020; Indonesia 2018; Indonesia 2019) and OpenStreetMap polygon data (OSM 2021) were sought for model input. Since high-quality spatial data is required, management with a Google Earth Engine (GEE) in a geographic information system (GIS) was essential for data integration processes.

The settlement was first partitioned into zones and HHs were randomly selected to represent a statistically accurate data sample. In-depth interviews took place with 41 Javanese citizens over three days, for 1–2 hours each. The questionnaire (see Appendix 1), consisted of a census survey, activity-based schedules, Rolando Arellano's Lifestyle questionnaire (selected for its adaptability to developing countries¹), and HH location preferences with appropriate classifications for peri-urban areas (Lauf et al. 2012). The interviewed HHs accounted for 8.3% of the settlement and the refusal rate (the ratio of denied interviews over a total share), 10% of the settlement. Additionally, absent HH members contributed at a later date and all questionnaires were revised with a mapping activity to catalog travel schedules (Supplementary Fig. S11). Lastly, road networks were selectively sampled to extract the basic values of street attributes listed in Table 5.1. In accordance with Han et al. (2020), the appearance of cities may differ greatly, but the main pattern differences lie in intersection form and road density, as evidenced in (Supplementary Fig. S9).

Table 5.1. Road network accessibility measures.

Categories		Connectivity		Intersections		Permeability	
Zone	Building Density	Boundary Roads	Block Size (m)	Semi-Urban	Urban	Road Crossing	Level
Residential	High	Local / Collector / Sub-Arterial	80 / 150 / 200	30 – 40	40 – 50	0	Zero
	Low	Local / Collector	80 / 100	40 – 50	50 – 60	2	Low
Commercial	High	Collector / Sub-Arterial	100 / 150	20 – 30	30 – 40	3	High
	Low	Local / Collector	80 / 100	30 – 35	30 – 40	2	Limited

¹ Publications are available only in Spanish. The English version translated by the authors is available in (Supplementary Fig. S13). Segment terms and definitions are available in (Garcia-Fry and Murao 2020).

Office	High	Sub-Arterial / Arterial	200 / 150	10 – 25	20 – 35	2	High
Small Industry	High	Local / Collector / Sub-Arterial	80 / 100 / 300	10 – 20.	15 – 25	2	Low
Large Industry	High	Local / Collector / Sub-Arterial	80 / 100 / 300	15 – 20	20 – 30	3	High

Values represent topological characteristics of sampled networks (0.25 km^2) in the urban periphery of Lima. As a classification survey, only measures of the reference point or network were necessary.

5.4.2 Activity-Based Schedules

Bowman and Ben-Akiva (2001) used the “primary destination” approach, assuming that one activity and destination are identifiable as a travel tour generator, representing a start and end time for the tour. We used this method to identify work destinations as the motivation for trip generation from Monday to Saturday, from 5:00 am to 4:00 pm (Mohammadi and Taylor 2019). The tour aggregation method, attributed to geographer Hagerstraand (1970), Chapin (1976), and Jones (1977) was used to retrieve a “home-work-other-home” schedule of activities reflecting daily travel routines. To model the individual activities of a population, a microscopic dataset is required. Therefore, we generated a synthetic population in proportion to our data sample, where the 820 agents created have the same number of statistical features as the real population and only exist as files in the computer (more in Moeckel et al. 2003).

5.4.3 Travel Simulation

To begin a travel microsimulation, first, identification of economic polar growth areas was reconciled with primary activities at the provincial level. Kaliurang, a small tourist town, and Yogyakarta represent two economic poles connected by a north and south axis road with “strategic” value (Pramana 2016). This road provides service through the Sleman Regency connecting the city to a growth destination (Ministry of Public Works in the RTRW Sleman Regency 2011-2031). The local road network was surveyed to identify road hierarchies (local, collector, sub-arterial, and arterial), road widths (in width-meter lengths), and destinations structuring travel tours in Figure 5.4. Travel schedules from our data sample, reflecting origins and stops made on the parcel (raster) level,

were then used to identify locations using GIS. Sociodemographic data established travel route choices based on the shortest path with the least number of directional changes known to have a detectable impact in a study by Peponis et al. (2008). Sectoral speed identifiers, as either high (H) or low (L), were subsequently incorporated every 100–200 m, depending on roads (either paved or non-paved) and slope.

The consistent mode choice for single-occupant vehicles in Asia is the motorcycle, which was the only choice for farmers in Pagerjurang. The operation cost in suburban areas of Yogyakarta was calculated with fuel consumption in free-flow transit and potential traffic jams, with monthly maintenance and repairs of 3.00 USD and 11.00 USD, respectively (Herwangi et al. 2017). Here, 29.20 km/L of petrol in free-flow transit for Yogyakarta, with a speed of 42.42 km/h, represents a low performance *Lp* speed in the Cangkringan district (Sugiyanto 2016). While 45 km/L represents a high performance *Hp* speed, applied to a velocity of more than 55 km/h on straight, uninterrupted road segments. Fuel consumption is a function of tour distance *D*, factored by the sum of segments *i*, between an origin *P_i* and a destination *P_j*, under two mean performance levels.

$$F = \sum_{i=1}^n [(D_i(P_i - P_j) / Lp) + (D_i(P_i - P_j) / Hp)] P_l \quad (1)$$

Equation 1 estimates fuel consumption *F* subject to the local price per liter of fuel *P_l* for *n* number of road segments. Topography and weight were not considered due to a high probability (84%) of individual ridership for HHs with less than three adults, each possessing more than one vehicle. However, shared ridership was contemplated whenever a shared travel agenda was detected. The fuel price was set at 0.49 USD/L during the survey, and CO₂ emissions were estimated by multiplying the emission rate of a four-stroke motorcycle engine (55 g/km) by the total vehicle kilometers traveled (VKT) (Meszler 2007). Input data is shown in Table 5.2, defining access to the settlement as travel origins to estimate intra-settlement travel.

Table 5.2. Travel tours in segments with distances, road widths, and route choices.

N.	Job Type	Round Trip		Travel Route	Mean Width (m)	Velocity Segments					Travel Distance (km)		
		Travel Origins	Destination			1. High / R. Width	2. Low / R. Width	3. High / R. Width	4. Low / R. Width	5. High / R. Width	Low	High	Total
1	Farmer	Access North	Kaliadem	Jl. Kaliadem Raya	7.31	-	1,212m / 7.75m	1,286m / 7.85m	680m / 7.29m	3,053.7m / 6.35m	3.784	8.678	12.46
2	Farmer	Access South	Ngudi Makmur 1	Jl. Wukirsari	4.80	968m / 3.62m	1,438m / 4.42m	1,564m / 7.10m	693m / 4.09m	-	4.262	5.064	9.33
3	Tourism	Access North	Desa Petung	Jl. Petung Merapi	7.49	-	1,212m / 7.75m	702m / 7.85m	1,522m / 6.88m	-	5.468	1.404	6.87
4	Admin.	Access South	Hyatt Hotel	Jl. Kaliurang	-	-	-	-	-	-	-	-	17.90
5	Teacher	Access North	Negeri Cangkringan	Jl. Petung Merapi	7.51	-	1,212m / 7.75m	400m / 7.28m	-	-	2.424	0.800	3.22
6	Tourism	Access NW	Tlogo putri Jeep	Jl. Kaliurang	8.07	-	2,716m / 6.87m	3,057m / 9.53m	862m / 7.82m	-	7.156	6.114	13.27
7	Farmer	Access NW	UPP Kaliurang	Jl. Kaliurang	8.32	-	2,716m / 6.87m	2,672m / 9.78m	-	-	5.432	5.344	10.77
8	Tourism	Access NW	Stonehenge	Jl. Raya Merapi Golf	6.34	-	558.8m / 6.36m	1,950m / 6.52m	1,075m / 6.15m	-	3.268	3.900	7.17
9	Tourism	Access NW	Lava Tour	Jl. Kaliurang	8.35	-	2,716m / 6.87m	2,059m / 9.83m	-	-	5.432	4.118	9.55
10	Teacher	Access North	SMP Cangkringan	Unnamed road	8.00	-	950m / 8.00m	-	-	-	1.900	-	1.90

Note: Farming (66%), Tourism (24%), Administration (2.4%), Teachers (7.3%). The Hyatt Hotel is in the city.

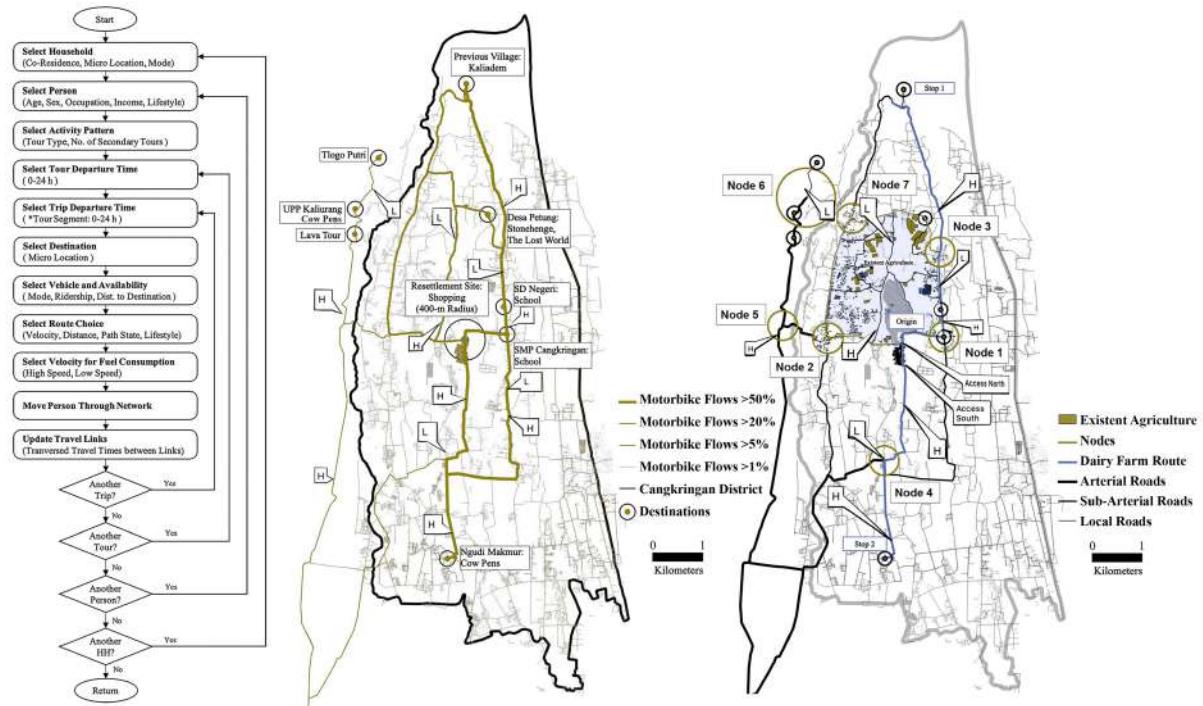


Figure 5.4. The transport component (left). Travel flow classified with travel intensities, velocities, and road quality in the Cangkringan district (middle). The roads were structured with three types of hierarchies in edge-of-urban areas (right).

The following output data corresponds to the baseline results of this study. The daily fuel consumption totaled 5.22 USD and the mean estimation for monthly travel expenditures per person was 20.39 USD, equal to 26% of the average monthly income (78.15 USD). We also registered 22,519 g of CO₂ emissions from the 421 VKT daily with 8% (33.56 km) originating inside the settlement.

Farmers, tourist vendors, administrators, and teachers travel 35%, 40%, 19%, and 6% of the total distance, respectively. They spend in average 3.91 USD more than people in other disadvantaged areas of Yogyakarta DIY (see Herwangi et al. 2017, p.6). These results revealed a restrained potential to increase HH profits after utilizing 26% of the average monthly income for travel. This is exacerbated by the low-income margins and the limited access to livelihood options for middle-aged people and the elderly.

5.5. OLUTM Model (2019-2030)

The model system in Figure 5.5 divides compact modules into three environments: The geographic environment, the socioeconomic environment, and the settlement environment.

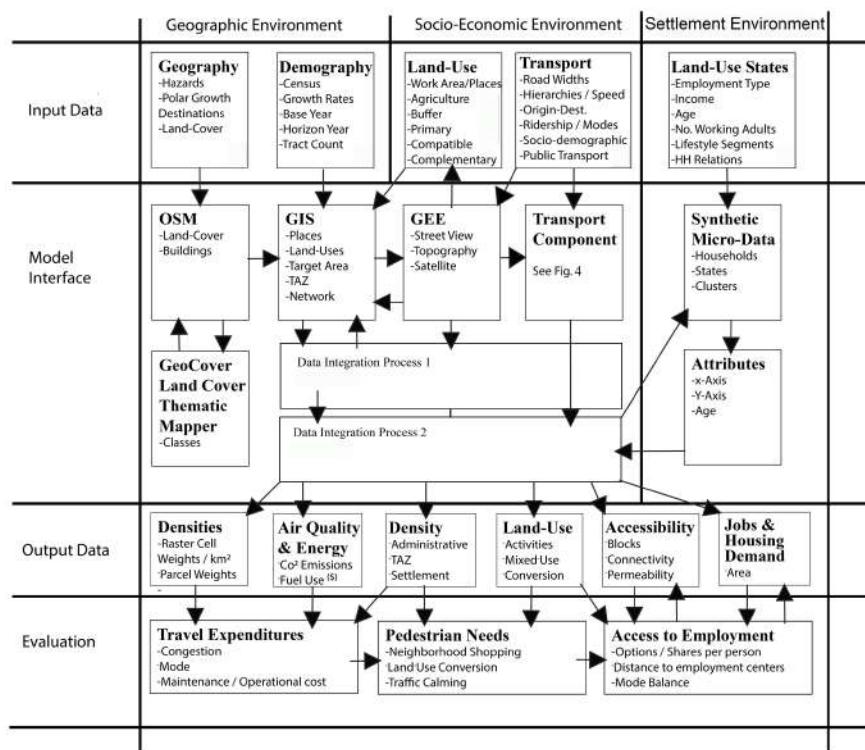


Figure 5.5. Model integration flow.

5.5.1 Geographic Environment

(1) Territorial Analysis

The first module uses future population projections estimated with conglomerate statistical trends (fertility, birth, mortality, migration, etc.), subject to past and current data, and assumed to continue. The local government's statistical census bureau (Statistik Badan Pusat Kabupaten Sleman) offers open datasets for the district's population and its growth rate from 2010 to 2017 (Indonesia 2018). The projections for 2019 and 2030 using the district's growth rate (0.5%) resulted in 29,740 and 31,302 people with densities of 619 and 652 people/km², respectively (Supplementary Eq. S1 and S2). These values were validated after comparing the 2018 population count with the government-recorded count in (Indonesia 2019).

(2) Density Redistribution

The Cangkringan sub-district has a 48 km² area and is expected to have an increase of 33 people/km² by 2030. Open source 2019 building footprint polygon data (OSM 2021) were compiled and used as ancillary data for a dasymetric redistribution of the census-based district population count. We used this method to distribute a total count (population census) across an area (administrative/census unit). Instead of distributing the count equally by areas, land cover data were used at a finer spatial resolution to unequally distribute the total count in weights reflecting a more accurate distribution of people over the study area.

Selective sampling of the ancillary data in a km² raster cell grid was compiled and GeoCover Landsat Thematic Mapper (TM) classifications from MDA Federal Inc. (1995–1997) were normalized to accurately classify the samples in our target district. The first redistribution of density weights was possible using these samples. Subsequently, a 4.17 km² TAZ was traced around the Pagerjurang resettlement site shown in Figure 5.6. We used building footprint areas from a 200 m spatial resolution raster grid and interpolated them with each km² “parent” raster cell's density weight to redistribute densities to a higher resolution. Land cover normalization was carried out following the methods of Mennis and Hultgren (2006), Gaughan et al. (2013), Stevens et al. (2015), and Linard

et al. (2011). The projected population count in the TAZ was 3,409 (2019) and 3,590 (2030) with a total of 181 new persons by 2030 and a 92.4% accuracy with deviation range of ± 17 persons (Supplementary Table S3). Ages below 18 (29%) (Indonesia 2020) were then excluded from the count, leaving 128.5 adults.

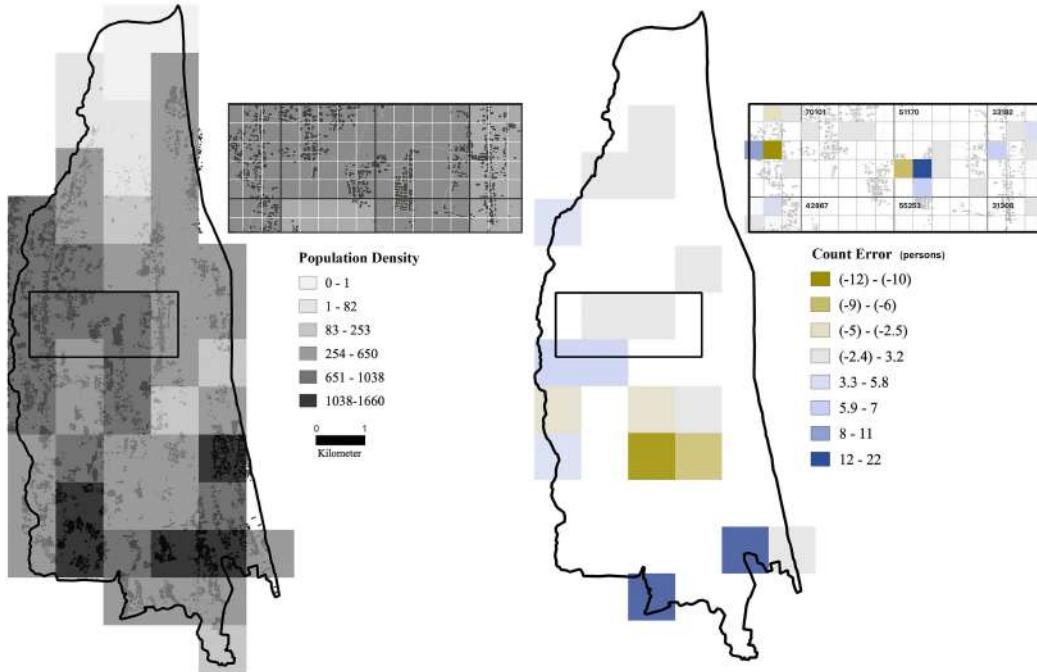


Figure 5.6. Density map of the Cangkringan district in 2019. The inset box defines the TAZ's boundary with ancillary data in a 200-meter spatial resolution and the Pagerjurang resettlement site located in the 51,170 (m^2) “parent” raster cell. A map of the count error by block for the dasymetric map to the right is shown with class intervals by standard deviation from the mean error, which is zero. Green areas indicate underestimation while blue areas indicate overestimation.

(3) Node Weight Estimation

The projected population reflects young adults with higher education living in the target area and having similar profiles, values, and attributes of those surveyed in 2019 (compatible with testimonies in Miller M.A. 2018, p.202). Residential preferences for social neighborhoods (71%), areas surrounded with amenities (12%), and in central areas (10%) were used to estimate the number of adults choosing to live in serviced areas (119.5 adults). Subsequently, nodes identified in Section 5.4.3 were assigned a typology from Table 5.3 by evaluating road volume capacity at the junction

level (in width-meters length). Each road adopted a street enclosure ratio (height:width) to set a built capacity and an influence radius (r) to limit that growth. Enclosure ratios were estimated with theoretical constructs found in Alkhresheh (2007) and Rudlin and Falk (1999) to target urban environments with rural streetscapes. The model aggregated road volume capacities per node and used those weights to distribute adults among nodes with the same scale: 61 in node 1 and 58.5 in node 2 (Supplementary Fig. S4).

Table 5.3. Road network types.

Width (m)	Road Type	Type 1	Type 2	Type 3
		Local Node $r = 300$ (m)	District Node $r = 600$ (m)	Inter-District Node $r = 900$ (m)
3 - 4	Local Road	1:1		
4 - 5	Local Road	1.5:1		
5 - 8	Collector Road		1:1	
8 - 10	Collector Road		1:1.5	
15 - 18	Sub-Arterial Road		1:2	
18 - 21	Sub-Arterial Road		2:1	
24 - 30	Arterial Road			2:1.5
28 - 30	Arterial Road			2:1
16 - 32	Highway Road			0.5:1
				3:1

Note: Areas of the Sleman Regency were sampled to measure road widths, network hierarchies, and enclosure ratios. Roads are classified with 4 types found to establish commonalities among sampled networks (Han et al. 2020). Highways were also sampled to fit the model. Enclosure ratios follow a height-to-width ratio rationale of 3/4 for urbanized areas (Alkhresheh 2007, p.139) and 17/20 as the local equivalent for nodes of the TAZ.

(4) Housing Demand

This module estimated the demand for new adults using HH compositions sourced from our data sample to retrieve the percentage of adults per household (APH): 1-adult (52%), 2-adults (32%), and 3-adults (or more) (16%). These ratios were then applied to the 61 adults in node 1, totaling 45 HHs partitioned into 32 (1-adult), 10 (2-adult), and 3 (3-adult) HHs, followed by node 2 with 43 HHs, partitioned into 31 (1-adult), 9 (2-adult), and 3 (3-adult) HHs.

The model proceeded using residential surface areas from resettlement studies in Sleman and an HBE study in Malang to estimate spatial requirements. In Sleman, HHs received 36 m^2 houses on 100 m^2 lots. However, most of them expanded their housing coverage to the maximum footprint available (Maly et al. 2015). Also, lot sizes were deemed insufficient for large families (Singghi and Asano 2019, p.120). Therefore, we assigned lots of 70 m^2 (1-adult HH), 100 m^2 (2-adult HH), and

120 m^2 (3-adult HH) considering the size of families with HBEs (Tutuko 2014, pp.51-63), thus resulting in $3,600 \text{ m}^2$ (node 1) and $3,430 \text{ m}^2$ (node 2) coverages.

Using a Google Street View built-in environment, we recorded the current (2019) building heights in node influence areas and estimated the potential building capacity of each node with the assigned enclosure ratios as in Figure 5.7. A user-based input was required to establish the mean inter-story height from our survey (Supplementary Table S4). Then, the model queued lots in asphalted roads having 25.4% higher value according to Pramana (2016) and estimated the built capacity in node 1 ($20,555 \text{ m}^2$) and node 2 ($10,589 \text{ m}^2$).

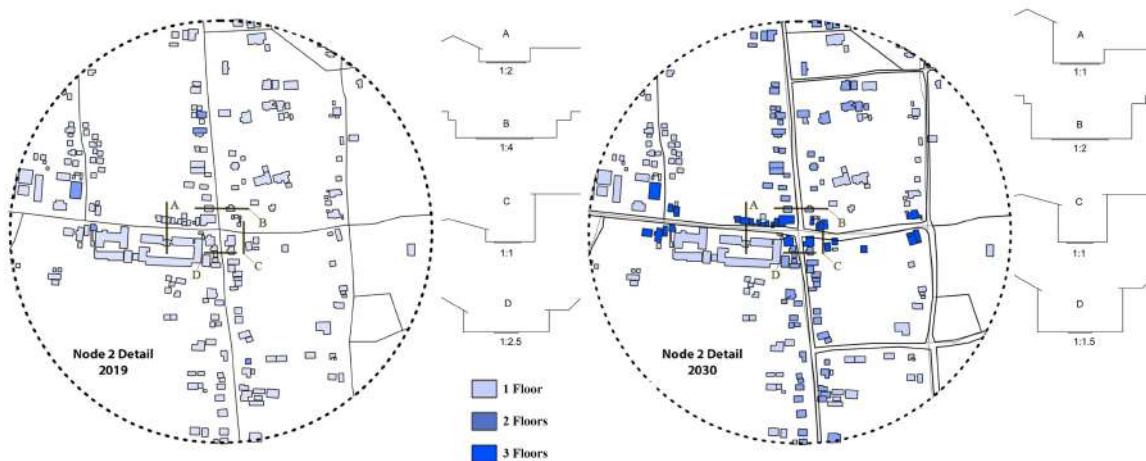


Figure 5.7. Built capacity of buildings sharing a side with intersecting roads.

5.5.2 Socioeconomic Environment

(1) Land-Use Designations

In this module, primary activities were classified and new activities were assigned as user-based inputs to trigger local employment in the TAZ. First, land uses were surveyed with Google Street View to classify nodes with a land-use category (residential, commercial, light industry, or service). A category was assigned by identifying primary activities factored with intersection proximity. Primary activities maximize business utilities attracting people from various scales of a city, depending on the degree of specialization.

We classified node 1 as “local-institutional; service” and node 2 as “local-educational; service” (node type-primary activity; land use). Both were identified as service nodes because the primary activity belongs to this land use and attracts people from the local district. We then assigned a technical research institute compatible with node 1 ($4,060 \text{ m}^2$), a cultural language institute ($4,027 \text{ m}^2$), and a medical veterinary center in node 2 ($2,300 \text{ m}^2$). The model distributed them unevenly to reinforce the node with less potential for development. A total of 164 job positions were estimated with density ratios (floor-space requirement per employee), built corridors (10%), and a 20% gross external to net internal space transition (Dancer 2015).

(2) Employment Distribution

This module simulates the multiplier effect, classified into sectors based on the Ministry of Manpower and Transmigration which separates labor within the local economic structure. Employment values were sourced from the Secretariat Governance Bureau of Population for Yogyakarta DIY (a translated version is available in Garcia-Fry 2021).

The model factors local multipliers to capture the demand for local services following an empirical study by Moretti (2010), to estimate long-term employment of a local area as a result of attraction to a successful firm or service. There is a shock in employment for tradable jobs under agglomeration effects and for non-tradable sectors. Tradable jobs follow a logic of domestic manufacturing with production consumed in a different geographic entity. Informality was also considered, contributing to the size of multipliers related to informal jobs often created in developing countries (Charpe 2019). The model proceeds as follows:

$$\Delta N_{c,t}^{NT} = \beta \Delta N_{c,t}^T + \eta X_{m,t-1} + \varepsilon_{c,t} \quad (2)$$

Equation 2 captures the relation between the change in the tradable sector, $\Delta N_{c,t}^T$, and the change in the non-tradable sector, $\Delta N_{c,t}^{NT}$, measured in an administrative entity c with an error term $\varepsilon_{c,t}$, and regressor coefficients β and β' . Population size of the last term is denoted by $X_{m,t-1}$ with an elasticity

coefficient η . We used six observances for all entities in the sub-district's census (2015-2020) to estimate local multipliers.

$$\Delta N_{c,t}^{T1} = \beta' \Delta N_{c,t}^{T2} + \eta X_{m,t-1} + \varepsilon_{c,t} \quad (3)$$

Equation 3 was used to find the changes in the tradable sector $\Delta N_{c,t}^{T1}$ for sub-sector $T1$, having a shock effect on the rest of the tradable sector $\Delta N_{c,t}^{T2}$, in administrative entity c . Equations were computed with a standard ordinary least square estimation. This method has two limitations: (1) reverse causality and (2) omitted variable. An instrumental variable estimation followed to verify the results based on the shift-share approach coined by Bartik (1991) for tradable, non-tradable, and agricultural sectors. The instrument assumes that a change in national employment is unrelated to the local labor market. However, the net change of administrative entity c replaced the national employment in $t - s$ as a weight for variance.

$$\Delta \tilde{N}_{m,t}^T = \sum_{j \in T} N_{j,c,t-s} (\log(N_{j,t} - N_{j,c,t}) - \log(N_{j,c,t-1} - N_{j,c,t-s})) \quad (4)$$

Equation 4 estimates the inner product change of industry-location shares for tradable employment $\Delta \tilde{N}_{m,t}^T$ and the local industry growth rate by subtracting the log change in national employment $(N_{j,c,t-1} - N_{j,c,t-s})$ from the log variance in local sectoral shares $(N_{j,t} - N_{j,c,t})$. Multipliers for the Cangkringan sub-district can be seen in Table 5.4, with 2.95 non-tradable jobs for each tradable job added, and 2.09 agricultural jobs removed. One tradable sub-sector job created 1.96 jobs in the tradable sector, confirming the existence of agglomeration economies.

Table 5.4. Local multipliers for tradable, non-tradable, and agriculture jobs in the Cangkringan district.

Tradable (T _i)			Non-Tradable (NT)			Agriculture (A)			
	OLS	IV	Additional jobs for each new job	OLS	IV	Additional jobs for each new job	OLS	IV	Additional jobs for each new job
	1	2	3	4	5	6	7	8	9
T				0.703** (0.45) [0.08]	1.207*** (0.249) [0.644]	2.95 (0.021)	-1.26** (0.81) [1.419]	-2.88*** (0.508) [0.713]	-2.09 (0.002)
NT	0.113** (0.072) [0.08]	0.533*** (0.11) [0.64]	1.51 (9.78E-5)				-1.05* (0.277) [0.341]	-2.14*** (0.211) [0.887]	-0.58 (0.000)
A	-0.062*** (0.040) [0.08]	-0.246*** (0.043) [0.712]	0.68 (0.000)	0.322*** (0.084) [0.341]	0.414*** (0.0409) [0.887]	0.37 (0.278)			

T ₂	0.379*** (0.124) [0.25]	1.01*** (0.015) [0.996]	1.96 (0.005)
----------------	-------------------------------	-------------------------------	--------------

Note: The estimation results are depicted for combinations between tradable, non-tradable, and agriculture employment change. The table includes both the ols and the instrumental variable estimation. Standard errors are in parenthesis and r-square values in brackets. P-values are significant at ***99%, **95%, *90%.

Multiplicity values were applied to the pool of jobs in 2020, adding 840 jobs (2020-2025) and 1,417 jobs (2025-2030). For job removals, the vacancy relative to that space became available to other jobs in that sector. Removals added to 88 tradable jobs secured by HHs with higher education (from subsection 5.1.3), and 36 entrepreneurship jobs for commercial vendors of Pagerjurang. Negative coefficients deducted 183 agricultural jobs and 6 entrepreneur jobs confirming the rural labor trend and totaling 1,944 new jobs (Supplementary Table S6).

(3) Employment Allocation

The model for job allocation applies variables that follow traditional bid-rent literature in urban economics. Access to people is expected to increase bids on commerce and service venues closer to intersections where the density is higher. In this module, we allocated jobs with a multinomial logit model to estimate the probability of a company choosing a location among the alternative vacancy areas in proportion to the built capacity of a node.

Jobs were located in non-residential lots using floor-space density ratios (square-meters per employee) with the entrepreneur sub-sector in residential lots sharing a side with roads. Vacancy in node 1 totaled 20,555 m² with 1,200 job vacancies, 5,687 m² of residential area, and 5,224 m² of residential space subject to land-use conversion. In node 2, a total of 18,255 m² were recorded, with 830 job vacancies, 7,817 m² of residential area, and 2,379 m² of residential space subject to conversion. Vacant lots in node 2 (26.5% of the recorded total) were included in the pool of HH vacancies for families that migrate in search of employment.

First, the allocation choice model used variables found by Pramana (2016) to explain land pricing with the shortest path to services and with job density ratios to determine a land value

coefficient. A Monte Carlo sampling process followed to generate a decision as to whether a job would be placed, subject to the accommodation of that job. Sectors were ranked to compete for vacancies using a discrete random sampling distribution constrained by the log-likelihood coefficient of the job's density ratio. A final decision was based on whether the mean log-likelihood and standard deviation of the real estate coefficients in Table 5.5 were higher than 1. This followed a variant of the location choice model in the Urban Sim with the logic that all jobs j from a given sector s are paired to a location l with a probability p :

$$P_s = \{(j, l) \mid j \in J_s, l_s \in L_s^j, \text{ sectoral set of job pairs and locations}\}, \quad (5)$$

$$T_s = \{j \mid j \in J_s, \forall l \in L_s^j (j, l) \notin P_s\}, \quad (6)$$

$$J_s = \{l \mid l \in L_s^j, \forall j \in J_s (j, l) \notin P_s\}, \quad (7)$$

$$F_s = \{(l, p) \mid l \in J_s, \text{ denoting the probability } p \text{ of a job } j \text{ locating in } l\}. \quad (8)$$

Equations 5-8 trial sectoral jobs competing for geo-coded sectoral vacancies with land value coefficients. Once allocated pairs were removed, the 160 non-tradable jobs left were assumed to hold vacant residential lots in node 2, totaling 9,523 m² of HBES. The housing demand then surged (137,520 m²), with a comparatively low offer (26,016 m²) causing nodal spillovers.

Table 5.5. Sensitivity of job location choice variables.

<i>Variables</i>	Regression Scores		Maximized Model				
	<i>Coefficients</i>	<i>Coefficients</i>	<i>Min</i>	<i>Max</i>	<i>Probability</i>	<i>Likelihood Ratio Test</i>	
Intercept	1.364*** (0.226)	1.895*** (0.210)	1.34	2.45	2.30	-13.51	
Vacancy	-0.032* (0.024)	-0.042*** (0.022)	-0.10	0.02	1.00	-11.84	
Built Density Ratio	0.994*** (0.202)	0.838*** (0.187)	0.34	1.33	1.26	-12.30	
Building Capacity	0.002 (0.002)	0.003*** (0.002)	0.00	0.01	1.00	-11.84	
Land Value	-0.057** (0.021)	-0.051*** (0.020)	-0.10	0.00	1.00	-11.85	

Note: ***significant at 99%, **significant at 95%, *significant at 90%. Standard errors are in parenthesis. Building capacity did not have significant effect on location choice. Land value accounts for job density ratios making them significant predictors. The regression and a Lorenz-curve distribution for location choice is reported in (Supplementary Fig. S5). F-test is significant at 99% with 7.78. Chi-square =4. Predictive efficiency = 91%.

(4) Spillover Management

In-migration was addressed with a land-use treatment on “edges” of the network (i.e., the connection between nodes) (Nystuen and Dacey 2005; Arlinghaus et al. 2002). This follows regional accessibility, urban economics, and social urban theory with the general assumption that HHs place a higher bid on parcels with access to services, bus stops, and markets (Rhondi et al. 2018; Alonso 2013). This is common for people seeking employment in urban agglomerations where changes in land price account for vacancy rates and intensity of use.

This module used a multinomial logit model to find the probability of location choices for census-based and in-migrating HHs. As in the employment allocation model, they were sectored and ranked with a discrete random sampling distribution using income ratios from our data sample. This was done to pair competing HHs with vacancies in nodes using the Monte Carlo instrument. Constraints included vacancies below 70 m², 100 m², or 120 m². A final decision was based on the mean and standard deviation of the log-likelihood for variables listed in Table 5.6. A total of 301 pairs were allocated and removed from the pool of HHs and locations.

Table 5.6. Sensitivity of HH location choice variables.

Variables	Regression Scores		Maximized Model			
	Coefficients	Coefficients	Min	Max	Probability	Likelihood Ratio Test
Intercept	0.903*** (0.269)	0.707*** (0.175)	0.24	1.17	0.71	-24.85
Vacancy	-2.500** (1.177)	-1.287* (0.77)	-3.32	0.74	0.08	-20.37
Built Density	1.413*** (0.241)	0.788*** (0.157)	0.37	1.20	0.80	-22.27
Building Capacity	0.026** (0.012)	0.013* (0.007)	-0.01	0.03	0.51	-24.17
Land Value	-0.859** (0.363)	-1.021*** (0.237)	-1.65	-0.40	0.30	-23.10

Note: ***significant at 99%, **significant at 95%, *significant at 90%. Standard errors are in parenthesis. Built density is a good location choice predictor with land value negatively correlated, hence the preference for lower prices in nodal areas. The regression with a Lorenz-curve distribution for location choices is available in (Supplementary Fig. S6). F-test is significant at 99% with 23.23. Chi-square = 4. Predictive efficiency = 81%.

To accommodate the overspill, the edge was assigned with an enclosure ratio from Table 3 expanding the pool of vacancies by 5,658 m². The resettlement site boundary road also delivered 2,620 m². The model then used a value for non-working adults (16%) in the sub-district (Indonesia 2020) to provide an estimate for housing demand in node 1 (75,079 m²) and node 2 (36,425 m²).

These were divided by the number of intersecting roads in each node to distribute that demand along segments of an edge and consolidate the network.

The model proceeded to pair 75 HHs with new vacancies, leaving a deficit of 19,598 m². Residential demand weights \mathbf{RDW} were then used to supply that deficit in a vector of n terms with segments \mathbf{W} of 400-800 m each from the midpoint of the edge to each node;

$$\mathbf{RDW} = \mathbf{aW}_1 + \sum_{i=1}^n \mathbf{a}(W_i - i). \quad (9)$$

Equation 9 distributes the remaining demand to segment weights with vacant lots added as user input to the data warehouse. Lots were classified as bare areas and developed as residential pockets with mixed uses shown in Figure 5.8. By our definition, these are HHs around a public space with access to the road network. The allotment left a 12,202 m² (44%) housing deficit.

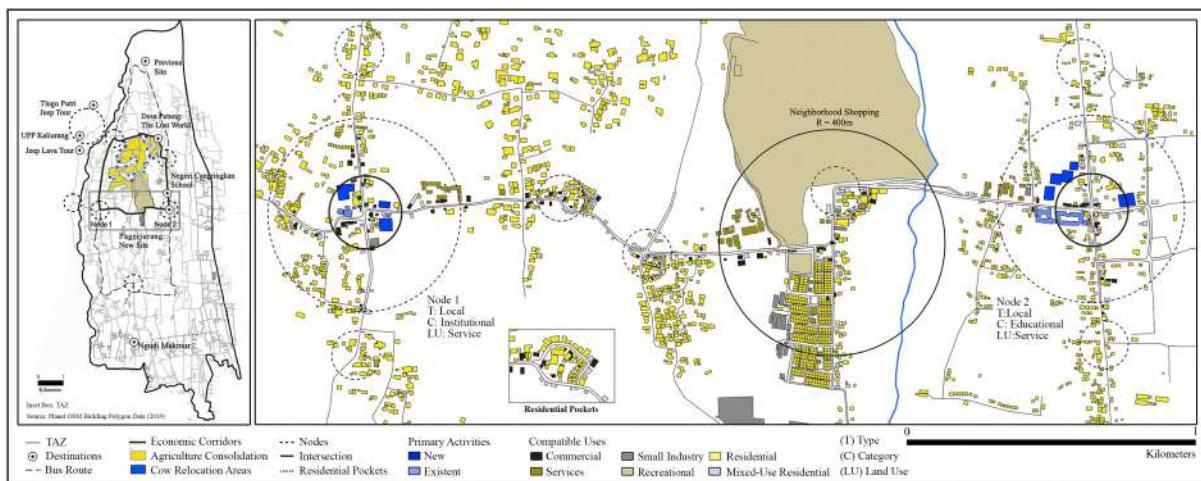


Figure 5.8. Land uses by 2030. New primary activities caused mixed-use development between the nodes. The inset box shows the AUB network with consolidated agricultural areas and cow relocation sites. The aggregation of dairy farming inside the network would contribute to further reduce VKT. A tentative public bus route is established to simulate its feasibility and promote directive extensions of the network.

5.5.3 Settlement Environment

(1) Street Network

Development of the edge was expected to have a substantial effect on livelihood diversification. Therefore, a clustering model grouped displaced HHs with travel attributes to locate them in a simulated settlement and evaluate micro land-use change. First, the model clustered part of our synthetic dataset into 12 neighborhoods of 26 HHs each. Note that this number considers the average of the neighborhood system in Indonesia (20-40 HHs) and the average of manageable decision-making neighborhoods (20-25 HHs) according to McCamant et al. (1994). Destinations and demographic data were used to weigh the polarity of travel attributes on-site. A majority of farmers travel to a dairy cooperative south of Pagerjurang. Therefore, the variables selected were: south destinations for the y-axis and age-above-50 for the x-axis. The latter established a control variable for the labor force currently affected by travel, testing our hypothesis that under equal access to customers, HBEs are driven by income, HH composition, and risk (changing residential land-use structures to mixed-use development).

The settlement module followed to locate HH clusters in a road network. A residential category was selected from Table 5.1 to define the permeability level. Then, blocks of 80 m and 100 m lengths were populated with 120 m² lots to accommodate HHs of all sizes. User inputs included recreational areas (2.5 m²/person), agricultural areas (35 m²/lot), and religious buildings as constituents of the local culture. Then, each block was sub-divided into 26 m, 33 m, or 65 m lengths to stimulate walking in slow travel environments (Krizek 2003). Given that landmarks were found to influence navigational decisions by Millonig and Schechtner (2007), a market and a large open space were included to guide potential customers through the road network. Then, road widths were selected from Table 5.3 to model the settlement in Figure 5.9.

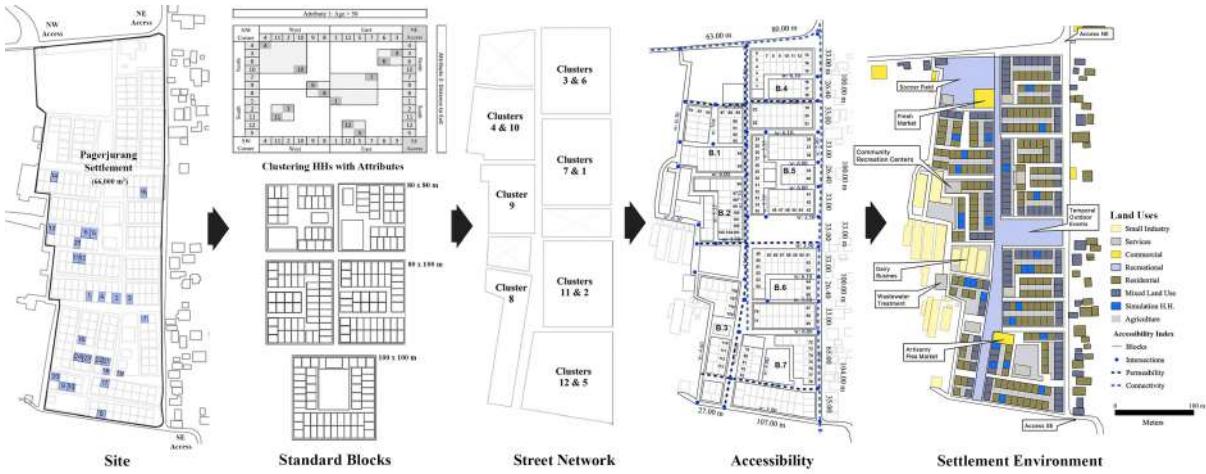


Figure 5.9. Settlement simulation process. A total of 113 mixed-use vacancies became available for HBEs. Note that road hierarchies are a product of sampled networks in Tables 5.1 and 5.3.

(2) Land-Use Change

This last module evaluates the relationship between accessibility and sociodemographic data to explain land-use change. Access to people, income, and HH compositions are considered to facilitate economic activity (Tutuko 2014). Also, livelihood diversification associated to services in the region is related to the percentage of motorcycle ownership (70%) (Rijanta 2009). Therefore, this model focused on HBEs with farming as the primary occupation.

A logit model was used to identify the intensity of demographic effects on land-use change with our data sample (e.g., see Sun and Robinson 2018). We assumed that livelihoods diversified by 2019 reflect a choice of land-use change, subject to the increment of people (Rijanta 2008). A regression followed to evaluate farming with a nominal value between commerce and services and compare the efficiency of variables in (Supplementary Table S8).

Using the HH allocation model, agents were extracted from our synthetic data to pair HHs with mixed-use vacancies relative to their block. Land-use change was constrained by the mean income of segmentations found to take higher risks in (Garcia-Fry and Murao 2020). These thresholds were used to train our data and maximize the model's predictive accuracy shown in Table 5.7. The outcome

revealed that 94 HBEs (84% of the lots) diversified their livelihoods with 61 farmers in 59 HBEs (63% of HBEs) complementing rural labor.

Table 5.7. Predictive capacity of the model.

Variables	Regression Scores			Maximized Model			
	Coefficients	Min	Max	Coefficients	Probability	Log-Likelihood	Likelihood Ratio Test
Intercept	0.227 (0.161)	-0.19	0.65	0.465			
Risk-Taking Levels	0.484*** (0.042)	0.37	0.59	33.682	1.00	-1.75	-6.75
Occupation Classes	0.226*** (0.070)	0.04	0.41	1.585	0.83	-0.37	-5.37
Education	0.075*** (0.024)	0.01	0.14	-0.380	0.41	-0.37	-5.37
HH Composition	-0.291*** (0.045)	-0.41	-0.17	-13.602	0.00	-0.37	-5.37
Income	0.010*** (0.002)	0.01	0.01	0.414	0.60	-0.37	-5.37
Mean				3.69	0.57	-0.65	-5.65
Standard Deviation				15.75	0.39	0.62	0.62

All P-values are significant predictors excluding the y-intercept. Standard errors are in parenthesis. Farming was constrained as a median value to train our data. Results indicate that HH composition (with negative correlation to land-use change) does not predict ventures. However, income and risk-taking segments have the largest predictive outcome. This is because size is not the only factor to affect business ventures. In reality, education may seem more appropriate to measure livelihood diversification, but it does not stimulate changes in the region yet. It is, however, well understood that education is a major driver of specialization, but does not necessarily imply occupational multiplicity.

5.5.4 Transport Component

A random selection of 25 HHs and 41 agents was performed to initiate the transport component.

Intra-settlement travel was simulated to evaluate allocation with travel attributes. Routines were then established and simulated to account for travel utilities. Diversified HHs were not considered for scheduling adjustment because their primary occupations were unchanged.

(1) Travel within the Settlement

To begin the simulation, all origins and destinations were reconfigured. The shortest paths to destinations were measured, enabling agents with destinations within a 400 m radius from their lot to presumably walk. Pedestrians sell their farm produce in the settlement's market or work in the nodes nearby. Consequently, the vehicle occupancy rate decreased by 50% (compared to the base year). This simulation was performed with motorcycles moving at low speeds.

Intra-settlement travel denotes a 55% reduction of cumulative travel distance. This indicates that HH allocation with individual attributes had a substantial effect on travel efficiency. Consequently, 49% less fuel consumption was observed—88.73 USD savings for the entire settlement. Additionally, walking modes detracted 45% energy consumption and CO₂ emissions as a result of nine agents (22%) living closer to their workplaces.

(2) Travel in Total

New travel agendas were processed for those who secured a non-tradable job in nodes 1 and 2. Walking shares increased to 13 persons (32%) and congestion in nodal areas decreased travel speed adding expense to each tour as shown in Figure 5.10.

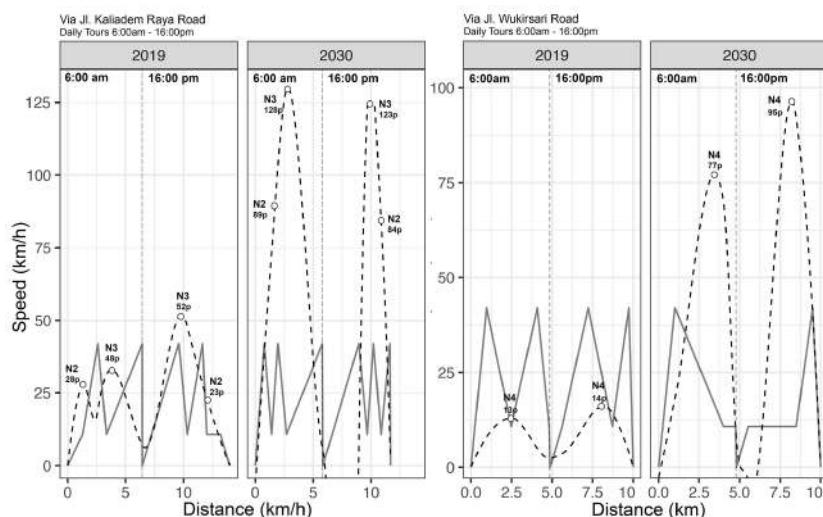


Figure 5.10. Vehicular congestion at peak hour transit (nodes 2,3 and 4). The curve shows congestion levels, while lines follow velocity change in the dairy farming route (Figure 5.4). Most trips are made through Kaliadem Raya road to farmlands. Then, they travel south through Pagerjurang to Wukirsari, where dairy cooperatives were relocated totaling 21.8 km round trips. Congestion was estimated for 15-minute intervals.

Travel through congested areas increased fuel consumption, but nearby workplace areas conveyed higher rates of modal shift. The effect was reflected in 20% less vehicle ownership for

agents with new jobs near their lot. VKT were reduced by 23%, consequently dropping fuel consumption by 28%, thus improving on travel performance as depicted in Figure 5.11.

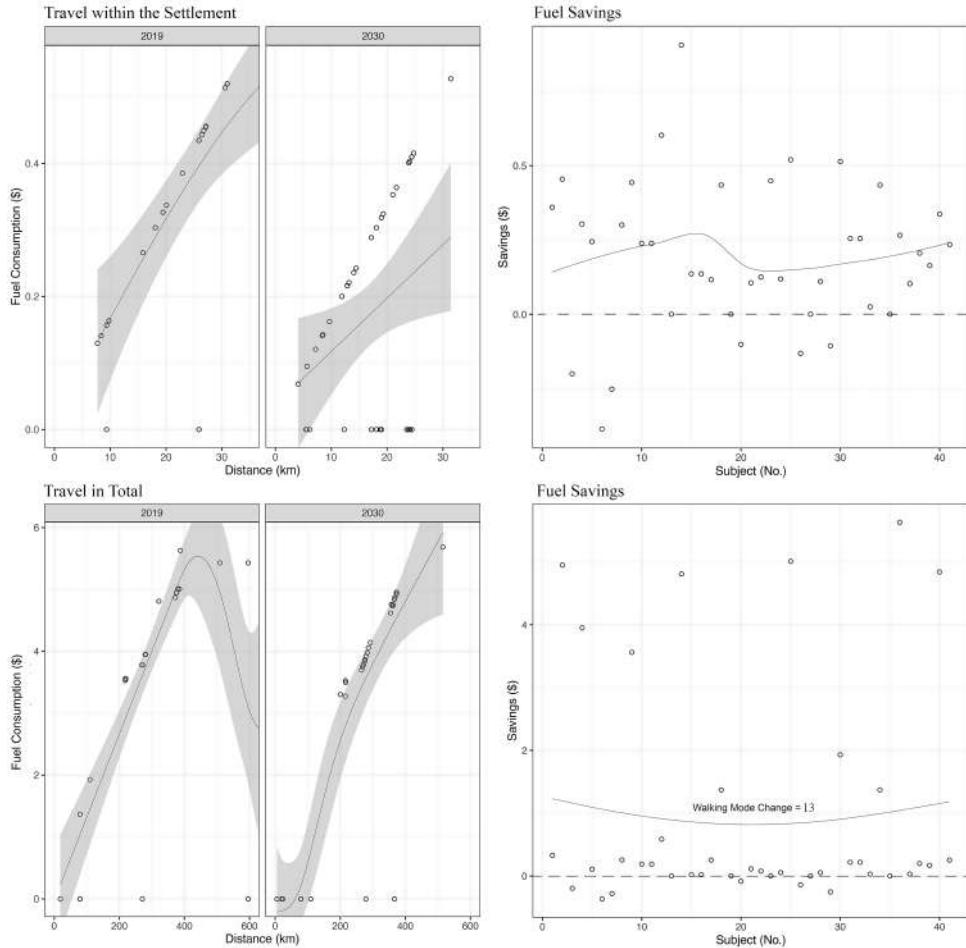


Figure 5.11. Comparative assessment of fuel consumption. Smoothed trend lines use Gaussian regression with a confidence interval affected by multi-modal travel, thus resulting in utility savings (right). Density conveys to low amount of savings when correlated with congestion and compared to modal shift, causing the majority of travel savings. The latter provides an indication of shorter intra-tour distances.

5.6. Discussion and Concluding Remarks

Travel utilities were measured in savings, totaling 98.8 USD/day for the entire settlement with 28% less CO₂ emissions by 2030. Considering the time frame (2019-2030), if the model were applied after the 2010 Mt. Merapi eruptions, with an operational expense of 18.36 USD per motorcycle

(Herwangi et al. 2017), then 2.39 VKT/person/day would save 105.6 USD from daily travel expenses incurred by all HHs in 2019. Mobility profiles show an approximate 2:1 ratio for motorcycles (68%) and walking modes (32%). Consequently, travel expenditures fell to 18.4% of the average monthly income with 6 USD savings per person as seen in Table 5.8.

Table 5.8. Monthly travel expenditures.

		2019	2030		Total	
	Agents	Dist. (km)	Fuel (USD)	Dist. (km)	Fuel (USD)	Savings (USD)
1	Travel in Settlement	41	1007.01 100%	452.80 45%	7.60 51%	7.37 49%
2	Travel in Total	41	12,632 100%	9,687 77%	112.2 72%	44.32 28%
3	Travel Utilities	41	836 100%	590 70%	246 30%	
4	Expenditure per Person (income %)	1	20.39 (26%)	14.38 (18%)	6.01 (30%)	
5	Expenditure per Household	1	33.43	23.58	9.85	
6	Before Intervention (2019) (2030)	528 / 605	10,766	12,336		
7	After Intervention (2019) (2030)	528 / 605	7,598	8,700		
8	Savings after Implementation	528 / 605	3,168	3,636		

All values are in USD and pertain to 41 travel diaries sought in the Pagerjurang settlement. Hypothetical statements exemplify the results to illustrate the effects of implementation after resettlement in 2010.

The interpolation of results reveals their extent (urban form data in Supplementary Table S10):

- Access to 3.5 times more transit volume and a 32% pedestrian mode shift rendered a suitable environment for livelihood diversification.
- People over 50-years-old with upper median incomes and risk-taking behaviors had the highest probability of venturing a home business.
- Land-use treatment in nodes had a direct effect on land-use change between the nodes. The model satisfied 56% of the housing demand, leading to mixed-use development.
- The residential allocation model revealed that HHs sought locations close to valued destinations in low-density areas with a negative correlation to vacancy.
- Allocation of HHs with individual attributes reduced intra-settlement travel distance by half. This had a strong effect on agents shifting from vehicle to walking modes (22%).

- The OLUTM model demonstrates that people utilized 21%–26% of their monthly income before the intervention and 11%–18.4% after with a modest bus service (Supplementary Figure S10). Monthly travel savings amount to 3,172 USD (2019) and 3,636 USD (2030).

5.6.1 Limitations and Assumptions

A series of constraints were met. For example, clustering with occupational profiles was insufficient due to the scale of the settlement having a small number of HHs. Simplifications follow the prospect theory for utility-based decisions, but the assumptions in Table 5.9 entail further research. The data collected represents a portion of the ideal amount for microscopic simulation. Furthermore, large datasets are not easily found for rural areas.

Table 5.9. Model assumptions.

Section	Module	Description	Theoretical Logic	Reference
5.1.1	Territorial Analysis	The government-recorded population growth rate is assumed to continue (the change in population over unit time period).	Population Growth Rates Cohort models	Statistical Census Bureau
5.2.2	Employment Distribution	The shift-share approach assumes that a change in national employment is unrelated to change in the local labor market.	Shift-Share Approach (National employment change replaced with district variance)	Charpe 2019
5.2.3	Employment Allocation	160 Non-tradable jobs were assumed to take hold of vacant lots in node 2, thus becoming HBEs.	Bid-Rent Theory of urban accessibility.	Alonso 2013
5.2.4	Spillover Management	Relocating HHs place higher bids on lots with access to serviced areas.	Accessibility, Urban Economics, and Social Urban Theory.	Rhondi et al. 2018; Alonso 2013
5.3.1	Street Network	HBEs are driven by income, HH composition, and risk.	Prospect Theory	Kahneman and Tversky 1979
5.3.2	Land-Use Change	Livelihoods currently diversified reflect a choice of land-use change subject to a rise in people.	Relative Location in Bid-Rent Theory.	Rijanta 2008

5.6.2 Concluding Remarks

This study provides a detailed description of the OLUTM model to evaluate post-disaster mobility, aiming to clarify whether a land-use network with distributed livelihood options can complement rural labor in a resettlement site. The outcome revealed reduction of travel expenses with a land use network. Livelihood diversification is predicted based on sociodemographic data and local employment structures influencing a broad spectrum of job opportunities. As a result, 94 livelihood

diversifications were recorded in the simulation with 63% accounting for farmers. Findings suggest that they are risk-oriented individuals with incomes above the average. Together with travel utility savings, the settlement's location was improved with farming as the primary employment and HBEs as the secondary employment.

While the expected development of a land use network provides travel efficiency and access to opportunities, resettlement offers a chance to locate HHs that will potentially venture micro-industries, known to be compatible with agriculture in the region. AUB networks are envisioned to accommodate public transport on existent roads to avoid further rural fragmentation. However, incentives are required to promote intensive agriculture in contiguous parcels and deaccelerate extensive agricultural land-use conversion. These policies should reinforce dairy cooperatives with business training to avoid large real estate transactions. There is also a need for land-use planning and hazard mitigation infrastructure. Together with forest protection policies, tree farming could become a source of labor for paper, wood, and biopolymer industries that could help absorb future hazard shocks. AUB networks target resource depletion with rural labor, land cover, energy-efficient travel, and semi-urban development to reduce urban intensities and enclose that growth in city-wide areas. This is especially important for Yogyakarta and its environmental stability.

The OLUTM model represents a continued effort to address urban growth and develop land use-transportation policy. Although the OLUTM model is not automated, the results reported here document a major step in that direction.

References

- Alkhresheh, M.M., 2007. Enclosure as a Function of Height-to-Width Ratio and Scale: Its Influence on User's Sense of Comfort and Safety in Urban Street Space, University of Florida, Gainesville, FL.
<https://ufdc.ufl.edu/UFE0019676/00001>
- Alonso, W., 2013. Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press, Cambridge, MA and London. Accessed: August, 2019. <http://dx.doi.org/10.4159/harvard.9780674730854>
- Arlinghaus, S.L., Arlinghaus, W.C., Harary, F., 2002. Graph Theory and Geography: An Interactive View (Ebook). Wiley and Sons, New York. <http://hdl.handle.net/2027.42/58623>
- Bartik, T.J., 1991. Who Benefits from State and Local Economic Development Policies?, W.E. Upjohn Institute Press, Michigan, 354 pp. <https://www.jstor.org/stable/j.ctvh4zh1q>
- Bowman, J.L., Ben-Akiva, M.E., 2001. Activity-based disaggregate travel demand model system with activity schedules. *Transp. Res. Part A: Policy and Pract.* 35(1), 1-28. [https://doi.org/10.1016/S0965-8564\(99\)00043-9](https://doi.org/10.1016/S0965-8564(99)00043-9)
- Bray, F., 1994. Agriculture for Developing Nations. *Scientific America*.
https://www.academia.edu/35944993/Agriculture_for_Developing_Nations
- Chapin, F.S., 1974. Human activity patterns in the city: Things people do in time and in space, Wiley- Interscience, New York, 507 pp. <https://www.abebooks.com/9780471145639/Human-activity-patterns-city-Things-0471145637/plp>
- Charpe, M., 2019. Sectoral Employment Multipliers in Rwanda: Comparing local multipliers and input-output analysis. *Int. Labor Org.-Geneva*, 1-26. https://www.ilo.org/wcms5/groups/public/-/-ed_emp/-/-ifp_skills/documents/publication/wcms_723283.pdf
- Clifton, K., Ewing, R., Knaap, G., Song, Y., 2008. Quantitative analysis of urban form: A multidisciplinary review. *J. of Urbanism: Int. Res. on Placemaking and Urban Sust.* 1(1), 17-45. <https://doi.org/10.1080/17549170801903496>
- Contreras, D., Blaschke, T., Hodgson, M.E., 2017. Lack of spatial resilience in a recovery process: Case L'Aquila, Italy. *Tech. Forecast. and Soc. Change* 121, 76-88. <https://doi.org/10.1016/j.techfore.2016.12.010>
- Dancer, S., 2015. Employment Densities Guide, 3rd. ed. Government UK: Homes and Communities Agency, 1-40. https://www.kirklees.gov.uk/beta/planning-policy/pdf/examination/national-evidence/NE48_employment_density_guide_3rd_edition.pdf
- Garcia-Fry, M., Murao, O., 2020. A Post-Disaster Forecast Analysis for Relocation Settlements using a Combined Model Sequence, 17th World Conf. on Earthq. Eng: N.C000108, Sendai, 1-12.
https://www.researchgate.net/publication/344756575_A_POSTDISASTER_FORECAST_ANALYSIS_FOR_RELLOCATION_SETTLEMENTS_USING_A_COMBINED_MODEL_SEQUENCE
- Garcia-Fry, M., 2021. OLUTM Model Dataset. Mendeley Data, v3. <http://dx.doi.org/10.17632/t9p23k3pyn.3>
- Gaughan, A.E., Stevens, F.R., Linard, C., Jia, P., Tatem, A.J., 2013. High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015. *PLoS ONE* 8(2), e55882. <https://doi.org/10.1371/journal.pone.0055882>
- Giyarsih, S.R., 2010. Urban Sprawl of the City of Yogyakarta, Special Reference to the Stage of Spatial Transformation (Case Study at Maguwoharjo Village, Sleman District). *Indonesian J. of Geog.* 42(1), 47–58.
<https://jurnal.ugm.ac.id/ijg/article/view/1576>
- Hägerstrand, T., 1970. What about People in Regional Science? *Papers in Reg. Sci.* 24(1), 7-24.
<https://doi.org/10.1111/j.1435-5597.1970.tb01464.x>
- Han, B., Sun, D., Yu, X., Song, W., Ding, L., 2020. Classification of Urban Street Networks Based on Tree-Like Network Features. *Sustainability* 12(2), 628-641. <https://doi.org/10.3390/su12020628>

- Handy, S.L., Boarnet, M.G., Ewing, R., Killingsworth, R.E., 2002. How the built environment affects physical activity. *Am. J. of Prev. Med.* 23(2), 64-73. [https://doi.org/10.1016/S0749-3797\(02\)00475-0](https://doi.org/10.1016/S0749-3797(02)00475-0)
- Herwangi, Y., Pradono, P., Syabri, I., Kustiwan, I., 2017. Transport Affordability and Motorcycle Ownership in Low-income Households: Case of Yogyakarta Urbanised Area, Indonesia, Proceedings of the Eastern Asia Society for Transportation Studies 12th conference. Ho Chi Minh, Vietnam. https://pdfs.semanticscholar.org/cc76/3d8152d8b58434732d7b6ca9df8c4a559580.pdf?_ga=2.180259811.67623511.1589852803-1018558182.1589195121
- Hynynen, A., 2006. Node-Place-Model-A strategic tool for regional land use planning. *Nordisk arkitekturforskning - Nordic J. of Arch. Res.* 18(4), 21-29. <http://arkitekturforskning.net/na/article/view/164>
- Indonesia, Statistics of D.I. Yogyakarta Province, 2020. Delivering Data to Inform Development Planning. Division of Integrated Data Processing and Statistical Dissemination, Catalog: 1102002.34, pp. xxiv–552. Accessed: April, 2020. <https://yogyakarta.bps.go.id/publication/2020/02/28/b734b0efb96248fa39801969/provinsi-di-yogyakarta-dalam-angka-2020--penyediaan-data-untuk-perencanaan-pembangunan.html>
- Indonesia, Statistics Sleman Regency, 2018. Population and Population Growth Rate by Subdistrict in Sleman Regency, 2010, 2016, and 2017. Badan Pusat Statistik Kabupaten Sleman. Accessed: May 19, 2020. <https://slemankab.bps.go.id/statictable/2018/10/12/363/jumlah-penduduk-dan-laju-pertumbuhan-penduduk-menurut-kecamatan-di-kabupaten-sleman-2010-2016-dan-2017.html>
- Indonesia, Statistics of D.I. Yogyakarta Province, 2019. Population Distribution and Density by Subdistrict in Sleman Regency, 2018”, Badan Pusat Statistik Kabupaten Sleman. Accessed: Oct. 25, 2021. <https://slemankab.bps.go.id/statictable/2019/07/09/520/distribusi-dan-kepadatan-penduduk-menurut-kecamatan-di-kabupaten-sleman-2018.html>
- Jones, P.M., 1977. New Approaches to Understanding Travel Behaviour: The Human Activity Approach, Working Paper 28, Transp. Stud. Unit, University of Oxford, Oxford, England.
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk, *Econometrica* 47 (2), 263–292. <https://doi.org/10.2307/1914185>
- Kitamura, R., 2009. Life-style and travel demand. *Transp.* 36, 679-710. <https://doi.org/10.1007/s11116-009-9244-6>
- Krizek, K.J., 2003. Operationalizing Neighborhood Accessibility for Land Use-Travel Behavior Research and Regional Modeling. *J. of Plan. Edu. and Res.* 22(3), 270-287. <https://doi.org/10.1177/0739456X02250315>
- Lauf, S., Haase, D., Seppelt, R., Schwarz, N., 2012. Simulating Demography and Housing Demand in an Urban Region under Scenarios of Growth and Shrinkage. *Environ. and Plan. B: Plan. and Design* 39(2), 229-246. <https://doi.org/10.1068/b36046t>
- Levinson, D.M., Wu, H., 2020. Towards a general theory of access. *J. of Transp. and Land Use* 13(1), 129-158. <https://doi.org/10.5198/jtlu.2020.1660>
- Linard, C., Gilbert, M., Tatem, A.J., 2011. Assessing the use of global land cover data for guiding large area population distribution modelling. *GeoJournal* 76, 525-538. <https://doi.org/10.1007/s10708-010-9364-8>
- Litman, T., 2021. Evaluating Accessibility for Transport Planning, Victoria Transp. Policy Inst., pp. 1-62. <https://www.vtpi.org/access.pdf>
- Maly, E., Iuchi, K., Nareswari, A., 2015. Community-Based Housing Reconstruction and Relocation: Rekompak Program after the 2010 Eruption of Mt. Merapi, Indonesia. *J. of Soc. Safety Sci.* 27, 205–214. <https://doi.org/10.11314/jisss.27.205>
- McCamant, K., Durret, C., Hertzman, E., 1994. Cohousing: A Contemporary Approach to Housing Ourselves, 2nd Edition. ed. Ten Speed Press, Berkeley.

- MDA Federal Inc. (1995–2007) EarthSat GeoCover LC Overview. Rockville, MD: MacDonald, Dettwiler, and Associates Ltd. Available: <http://www.mdafederal.com/geocover/geocoverlc/gclcoverview>
- Mennis, J., Hultgren, T., 2006. Intelligent Dasymetric Mapping and Its Application to Areal Interpolation. *Cartogr. and Geog. Inf. Sci.* 33(3), 179-194. <https://doi.org/10.1559/152304006779077309>
- Meszler, D., 2007. Emissions Issues Related to Two-and Three-Wheeled Motor Vehicles. Meszler Engineering Services, pp. 20–24. Accessed: April 2020. https://theicct.org/sites/default/files/publications/twothree_wheelers_2007.pdf
- Miller, M.A., 2018. Reconfiguring rural aspirations through urban resettlement: Navigating futurity after the 2010 eruption of Mount Merapi, Indonesia. In: Bunnell, T., and Goh, D., (Ed.), *Urban Asias: Essays on Futurity Past and Present*, Jovis, Berlin, pp. 193-202. <https://urbanaspirationsasia.com/urban-asias/>
- Miller, E.J., 2018. The case for microsimulation frameworks for integrated urban models. *J. of Transp. and Land Use* 11(1), 1025-1037. <https://doi.org/10.5198/jtlu.2018.1257>
- Miller, E.J., and P.A., Salvini., (2005). ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems. *Netw. Spat. Econ.* 5, 217-234. <https://doi.org/10.1007/s11067-005-2630-5>
- Millonig, A., Schechtner, K., 2007. Developing Landmark-Based Pedestrian-Navigation Systems. *IEEE Transactions on Intelligent Transportation Systems* 8(1), 43-49. <https://doi.org/10.1109/TITS.2006.889439>
- Moeckel, R., Wegener, M., Spiekermann, K., 2003. Creating a Synthetic Population. In: 8th International Conference on Computers in Urban Planning and Urban Management (CUPUM), Sendai, pp. 1–18. https://www.bgu.tum.de/fileadmin/w00bj/msm/publications/2003_moeckel_etal_sympop_cupum.pdf
- Mohammadi, N., Taylor, J.E., 2019. Recurrent Mobility: Urban Conduits for Diffusion of Energy Efficiency. *Sci. Rep.* 9, 20247, pp. 1-12. <https://doi.org/10.1038/s41598-019-56372-4>
- Moretti, E., 2010. Local Multipliers. *Amer. Econ. Rev: Papers and Proceedings* 100(2), pp. 1-7. <https://eml.berkeley.edu/~moretti/multipliers.pdf>
- Nystuen, J.D., Dacey, M.F., 2005. A Graph Theory Interpretation of Nodal Regions. *Papers in Reg. Sci.* 7, 29-42. <https://doi.org/10.1111/j.1435-5597.1961.tb01769.x>
- OSM-OpenStreetMap, 2021. The Free Wiki World Map. Open-Street Map contributors, CC BY-SA. Accessed April 15, 2021. Available: <http://www.openstreetmap.org/> [Internet]
- Peponis, J., Bafna, S. and Zhang, Z. (2008). The Connectivity of Streets: Reach and Directional Distance. *Environ. and Plan. B: Plan. and Design* 35(5), 881–901. <https://doi.org/10.1068/b33088>
- Pramana, A.Y.E., 2016. Impact of Land Use Regulation on Land Values. Erasmus University, Rotterdam, 61 pp. <http://hdl.handle.net/2105/42298>
- Rondhi, M., Pratiwi, P.A., Handini, V.T., Sunartomo, A.F., Budiman, S.A., 2018. Agricultural Land Conversion, Land Economic Value, and Sustainable Agriculture: A Case Study in East Java, Indonesia. *Land* 7(4), 148-167. <https://doi.org/10.3390/land7040148>
- Rijanta, R., 2008. Livelihood Strategies, Responses to the Crisis, and the Role of Non-Agricultural Activities in Five Villages in the Special Region of the Yogyakarta. In: Titus, M. J. and Burgers, P.P.M. (Eds.). *Rural Livelihoods, Resources and Coping with Crisis in Indonesia: A Comparative Study*. ICAS Publication Series. Amsterdam University Press, Amsterdam, p. 155. <http://www.jstor.org/stable/j.ctt46mz9x.11>
- Rijanta, R., 2009. Determinants to Rural Diversification in Java, Indonesia: The Case of Yogyakarta Special Province (DIY). *Reg. Views* 22, 1-15. https://www.researchgate.net/publication/272941622_Determinants_to_Rural_Diversification_in_Java_Indonesia_the_Case_of_Yogyakarta_Special_Province_DIY

- Rindrasih, E., 2018. Under the Volcano: Responses of a Community-Based Tourism Village to the 2010 Eruption of Mount Merapi, Indonesia. *Sustainability* 10(5), 1-17. <https://doi.org/10.3390/su10051620>
- Rudlin, D., Falk, N., 1999. Building the 21st Century Home: The Sustainable Urban Neighborhood. Architectural Press, Oxford, pp. 215-217.
- Singgih, B.A., Asano, J., 2019. Study on Impact of Relocating Settlement as a Post-Disaster Mitigation. *Urb. and Reg. Plan. Rev.* 6, 111-124. <https://doi.org/10.14398/urpr.6.111>
- Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J., 2015. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLoS ONE* 10(2), e0107042. <https://doi.org/10.1371/journal.pone.0107042>
- Sugiyanto, G., 2016. The Effect of Congestion Pricing Scheme on the Generalized Cost and Speed of a Motorcycle. *Walailak J. of Sci. and Tech. (WJST)* 15(1), 95-106. <https://doi.org/10.48048/wjst.2018.2347>
- Sun, B., Robinson, D.T., 2018. Comparison of Statistical Approaches for Modelling Land-Use Change. *Land* 7(4), 144-177. <https://doi.org/10.3390/land7040144>
- The World Bank, 2016. Community-based Settlement Rehabilitation and Reconstruction Project for Central and West Java and Yogyakarta Special Region. Report No: ICR00003662, pp. 5-25. <https://documents1.worldbank.org/curated/en/655251467992043861/text/ICR3662-P103457-OOU-9.txt>
- The World Bank, 2012. REKOMPAK; Rebuilding Indonesia's Communities After Disasters [WWW Document]. Financial Literacy and Education Russia Trust Fund, pp. 43-44. Accessed January, 2019. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/128361468267347740/rekompak-rebuilding-indonesias-communities-after-disasters>
- Tutuko, P., 2014. Residential Land Use and Housing Development in Indonesia from the Perspective of Sustainable Urban Form, Thesis Abstract, Graduate School of Natural Science and Technology, Division of Environmental Science and Engineering, Kanazawa University. <http://hdl.handle.net/2297/40520>
- Umaya, R., Hardjanto, Soekmadi, R., Sunito, S., 2020. Livelihood adaptation patterns of sub-villages communities in the slope of Merapi Volcano. *IOP Conf. Ser: Earth and Environ. Sci.* 528, p. 1. <https://doi.org/10.1088/1755-1315/528/1/012020>
- Waddell, P., Borning, A., Noth., M., Freier, N., Becke, M., Ulfarsson, G., 2003. Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim. *Netw. Spat. Econ.* 3, 43-67. <https://doi.org/10.1023/A:1022049000877>
- Wagner, P., Wegener, M., 2007. Urban Land Use, Transport and Environment Models. *disP - The Plan. Rev.* 43(170), 45-56. <https://doi.org/10.1080/02513625.2007.10556988>
- Wegener, M., Fuerst, F., 2004. Land-Use Transport Interaction: State of the Art. *Integration of Transport and Land Use Planning (TRANSLAND) 2A*. SSRN Electronic Journal, pp. 1-119. <https://doi.org/10.2139/ssrn.1434678>
- Widowaty, Y., Oktavian Artanto, I., 2018. Protection of Agricultural Land Sustainable Food for the Realization of Food Security in The Special Region of Yogyakarta. In: Proceedings of the International Conference on Food, Agriculture and Natural Resources (FANRes 2018), Atlantis Press, Paris, pp. 35-40. <https://doi.org/10.2991/fanres-18.2018.8>
- Yang, Y., Diez-Roux, A. v., 2012. Walking Distance by Trip Purpose and Population Subgroups. *Am. J. of Prev. Med.* 43(1), 11-19. <https://doi.org/10.1016/j.amepre.2012.03.015>
- Yokohari, M., 2006. Agro-activities in the Fringe of Asian Mega-Cities. *Landsc. Ecology and Manag.* 10(2), 75-79. <https://doi.org/10.5738/jale.10.75>
- Zhu, Y., Diao, M., Ferreira, J., Zegras, C., 2018. An integrated microsimulation approach to land-use and mobility modeling. *J. of Transp. and Land Use* 11(1), 663-659. <https://doi.org/10.5198/jtlu.2018.1186>

Zondag, B., de Bok, M., Geurs, K.T., Molenwijk, E., 2015. Accessibility modeling and evaluation: The TIGRIS XL land-use and transport interaction model for the Netherlands, Comp., Environ. and Urb. Syst. (49), 115-125.
<https://doi.org/10.1016/j.compenvurbsys.2014.06.001>

Supplementary Information

Data sample and models in Mendeley Data, v.3: <http://dx.doi.org/10.17632/t9p23k3pyn.3>

This file operates each step of the Operational Land Use and Transport Microsimulation (OLUTM) model corresponding to Chapter 5 of the manuscript. The compact modules described herein use seven data inputs to process data which is then used in the next modules, thus integrating steps of the model within the process. The model interface has three main environments and three data support systems to output datasets in visual representations. The transport component of the model simulates travel accounting for activity-based travel scheduling. The model's first integration of data holds the Geographic and the Socioeconomic Environments. While a different level of integration occurs in the Settlement Environment. Finally, the resulting scenario is packed with numerical data and mapping.

5.5.1 Geographic Environment

(1) Territorial Analysis

The last administrative census tract available for the Cangkringan sub-district was in 2017; 2010-2017 (Growth Rate: 0,5%) found in [Dataset 1].

Data Input 01

Area: 48 sq/m (square-meter)

Growth Rate (2010-2017) census: 0,5%

Starting Population Count (2016): 29,321 persons

Population for Evaluation (2017): 29,456 persons

Growth Projection for 2018:

$$29,456 - 29,321 = 135 \text{ (0,46\%)}$$

$$0.0046 * (29,456) = 135.62$$

$$29,456 + 135.62 = 29,591.62$$

Total: 29,592 persons*

*Note: This value was confirmed by projections made by the Indonesian Statistical Census Bureau (2010-2035) [Dataset 2]. <https://slemankab.bps.go.id/statictable/2019/07/09/520/distribusi-dan-kepadatan-penduduk-menurut-kecamatan-di-kabupaten-sleman-2018.html>

Growth Projection for 2019:

$$(2018) 29,592 - (2017) 29,456 = 136 \text{ (0,46\%)}$$

$$0.0046 * (29,592) = 147.96$$

$$29,592 + 148 = 29,740$$

Total: 29,740 persons

Density: 619 / sq. km (square-kilometer)

The latter procedure is summarized with the following equations (i) and (ii) for n number of years;

$$n = (y - y_s) - 1 \quad (1)$$

$$f(n) = \left[\left(\frac{(P_l - P_s)}{P_s} \right) P_l \right] + P_l \quad (2)$$

Where:

y =Target year (desired future population count);

y_s =Starting year (the last administrative census count for a region, district, or sub-district)

P_l =Last population count;

P_s =Starting population count;

Growth Projection (2030):

$$n = (2030-2019) + 1$$

$$n=10$$

$$f(2030)^{10}: [(\frac{(29,740 - 29,592)}{29,592}) \times 29,740] + 29,740 = 29,888.7, \text{ iterate this function 10 times.}$$

$$f(2030)^{10}=31,302 \text{ persons}$$

(2) Density Redistribution

Cangkringan Sub-District

-Growth Rate: 0,5% (see growth estimations sub-section 5.1):

-Area: 48 km²

-Population 2019: 29,740

-Density in 2019: 619 / km²

-Population 2030: 31,302 persons

-Density in 2030: 652 / km²

Ancillary Data

Emergency planning for disaster-affected areas requires population distribution counts for density that is hard to come by in remote locations, especially after a disaster or human displacement event. We used planet-OSM building polygon data, an open-source product that provides world-wide geographic datasets (<https://planet.openstreetmap.org/>) [Dataset 3] as ancillary data for a dasymetric redistribution of the smallest census-based enumeration regarded as a trustworthy source published annually by the Cangkringan sub-district.

Dasymetric Redistribution

Density weights were calculated for a sq./km spatial resolution of the administrative area using QGIS 3.4, Google Earth Engine, and EarthSat GeoCover Land Cover Thematic Mapper from MDA Federal Inc. adapted from [1, 2] to provide us with a spectrum of global land cover designations within our target district.

A Target Density Weight (TDW) can be dasymetrically calculated by gathering building footprint areas for all raster cells comprising a census area using a population count less than 10 years old (step 5.1.2; 2010-2017 count), by using equation (iii):

$$TDW_1 = \frac{P_1 (ADV_1)}{Vp_1} \quad (\text{iii})$$

Where the TDW (first weight) of a sq./km spatial resolution (raster cell) is a product of the administrative-unit level population count P_1 (for the year in question) and the Ancillary Data Value (ADV) of that cell. This is sub-divided by the output value Vp_1 corresponding to the total built footprint area in the administrative unit representing the population count. This calculation is applied to all raster cells.

A selective sampling approach is herein described and adapted from [3] to normalize GeoCover land cover designations with a local disaggregated value. Samples provide a mean density and a target mean estimation, with a standard deviation of that mean used to classify a range of values to validate accuracy of each target density weight for each normalized class, as shown in Table S1 and associated to Fig. S1(A) and S1(B).

Table S1: Land Cover Selective Sampling for Weight 1 (W_1)

Land Cover (1 sq. / km res.) WGS 84 / Pseudo-Mercator	Sample Number	Sample Mean (SM)	Sample SD	Estimation Method	Target Mean (TM)	Target SD	Density Target Range
Terrestrial Vegetation - Tree Cover				Preset	0	0	0 - 1
Bare Areas	3	4,069	5,191	Sample	55	70,6	2 - 82
Cultivated and Managed Lands	4	8,058	3,766	Sample	110	51,3	83 - 253
Rural Settlement	5	29,238	2,974	Sample	398	40,5	254 - 650

Semi-Urban Settlement	4	66,382	6,368	Sample	903	86,7	651 - 1038
Urban Settlement	6	86,285	27,640	Sample	1174	376,2	1,039+

*Note: "Sample Number" is the number of sampled source zones. "Sample Mean" is the mean data density (for built area per sq. / km) of the sampled source zones. "Sample SD" is the standard deviation of the samples collected. "Estimation Method" is the method used to estimate the target mean density. "Target Mean" is the mean data density of target zones associated to Sample Means (the dasymetric map density weight). "Target SD" is the Standard deviation of the Target Mean data density providing a comprehensive "Density Target Range" of values for each classification.

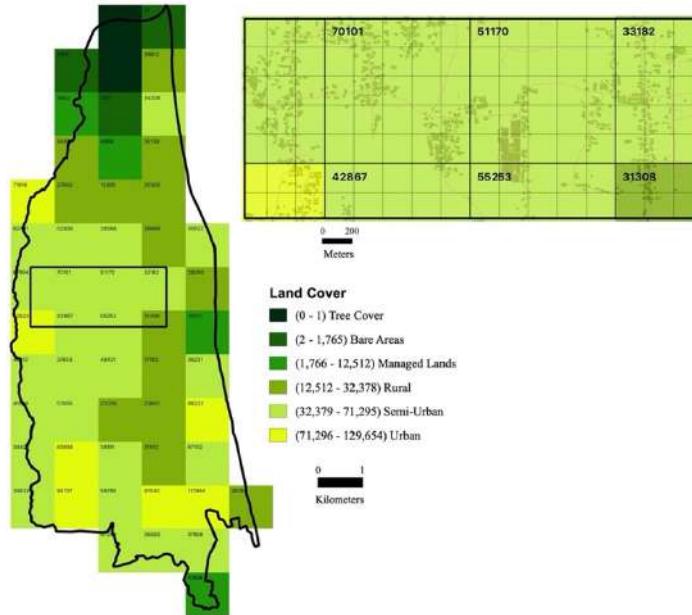


Fig. S1A: Land cover classifications normalized to building footprint areas (in sq./m) after performing a selective sampling approach. Classifications are adapted from [1], except “urban” settlements, applied using a Google Earth Engine designation and planet-OSM ancillary data. The inset box defines the interest area where resettlement is intended and illustrates ancillary data used for TDWs.

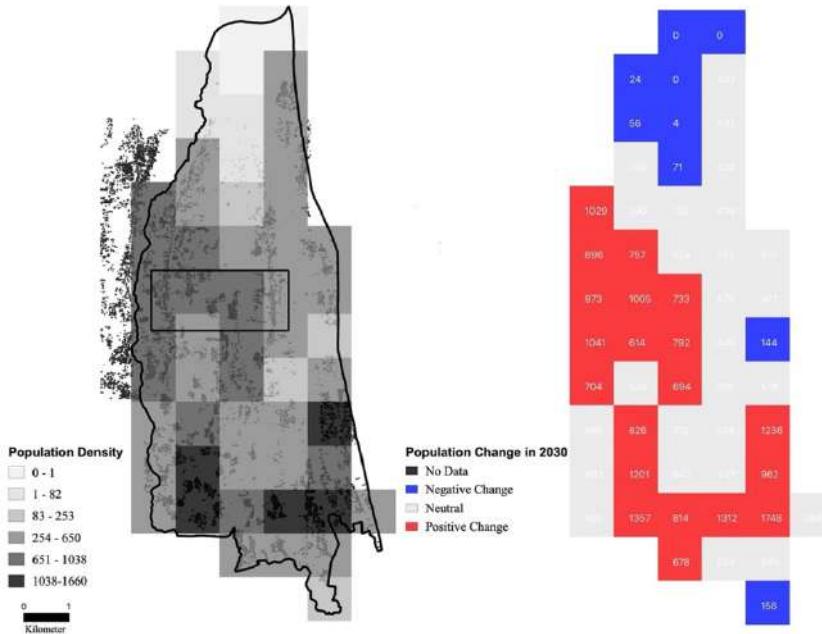


Fig. S1B: Dasymetric mapping of density weights for the Cangkringan sub-district on a square-kilometer spatial resolution with regions and presets of zero density for “tree cover” areas. Note that class intervals are the same as for Fig. 1(A) and color schemes are inversely proportional. Population Change in 2030 illustrates the number of people with a “negative change” equal to less than one person per year growth rate while “Neutral” refers to a 1-3 person per year growth rate. Inscribed data is the population count for each cell.

Target Area Density Weights (2019) and (2030)

-Area: 4.17 sq. km
-Population Growth: 0,5%

To find the population count for the inset box in Fig. S1(A) and S1(B), density weights obtained in section 5.2.2 reveal a population density count for each of the eight raster cells that make up the target area. To estimate a density weight that is accurately bounded by our target area, a fine-level (200 sq./m) resolution is obtained from sub-dividing each “parent” cell into 25 raster cells (Fig. S1A). A Target Density Weight (TDW_2) is determined with areal interpolation of land cover population density weights (TDW_1) and fine-resolution building footprint areas to estimate land cover class entities normalized to a higher spatial resolution using equation (iv):

$$TDW_2 = \frac{TDW_1(Tw_2)}{VW_1} \quad (iv)$$

Where, TDW_2 is a product of TDW_1 and the value for ancillary building footprint data in a higher spatial resolution Tw_2 , is divided by the total building footprint value VW_1 of its “parent cell”.

Once target densities are obtained for TDW_2 , the classification is normalized to determine land cover classifications adequate to this resolution and corresponding to the target area in the district. The Target Mean (TM) from each “parent cell’s” class in Table 1 becomes the Sample Mean (SM) for a land cover class in Table S2. This states that each class in Table S2 is normalized with TDW_2 , computed using a TDW_1 , corresponding to the SM and the SD in Table S1. The new set of land cover density weights are used to estimate a Density Target Range of values for a TM value and the subsequent value of that column. The process is illustrated in Fig. S2.

Table S2: Land Cover Selective Sampling for Target Density Weight 2 (W_2)

Land Cover (200-m res.) WGS 84 / Pseudo-Mercator	Sample Number	Sample Mean (SM)	Sample SD	Estimation Method	Target Mean (TM)	Target SD	Density Target Range
Terrestrial Vegetation - Tree Cover				Preset	0	0	0 - 1
Bare Areas	4	55	70,6	Sample	7	1,1	2 - 16
Cultivated and Managed Lands	4	110	51,3	Sample	26	6,3	17 - 34
Rural Settlement	4	398	40,5	Sample	42	6,9	35 - 54
Semi-Urban Settlement	4	903	86,7	Sample	67	13,8	55 - 127
Urban Settlement	5	1174	376,2	Sample	188	95,8	128+

*Note: Classes are sampled using density weights (from Table 1) and built area per 0.2 sq./km as ancillary data input for refined spatial resolution density weights.
“Sample Number” is the number of sampled source zones. “Sample Mean” is the mean population density (for a single type of land cover) of the sampled source zones.
“Sample SD” is the standard deviation of the samples collected in (Table 1). “Estimation Method” is the method used to estimate the target mean density weight.
“Target Mean” is the mean data density of target zones in a finer-spatial resolution associated to the Sample Mean (the dasymetric map density weight). “Target SD” is the Standard deviation of the Target Mean (TM) data density providing a comprehensive “Target Range” of data values for each class.

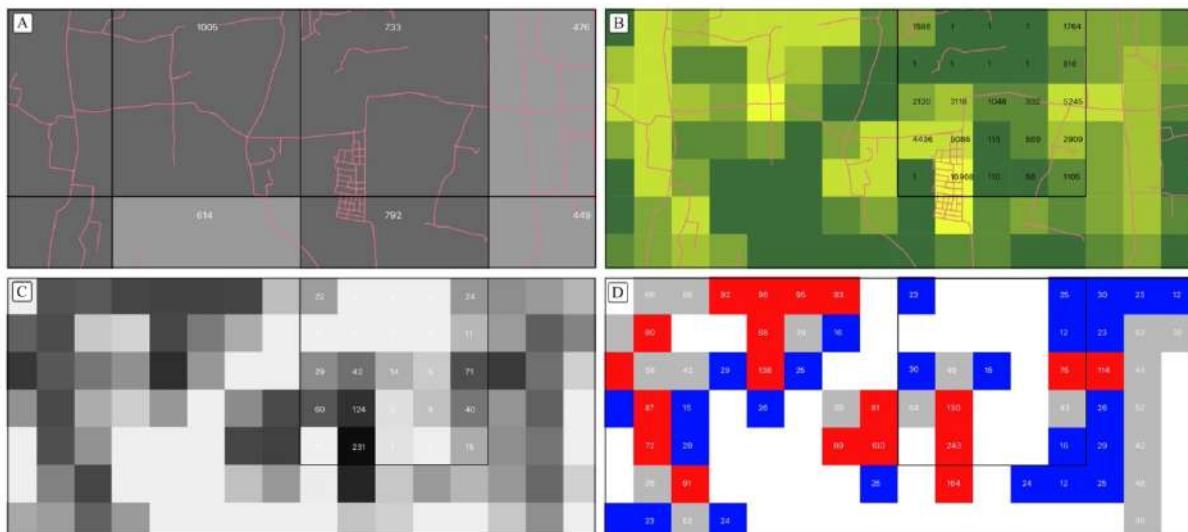


Fig. S2: Maps of the inset box shown in Fig.1 (A) and (B). Population density weight (1) by raster tract (A); normalized land cover with ancillary building data (B); higher-resolution population density weight (2) by dasymetric redistribution (C); population change between 2019 and 2030 (D). Class intervals and color codes are as shown in previous figures.

Target Area Population Count (2019): 3,409 persons (sum of all TDW_2)

Target Area Population Count (2030): 3,590 (with a 0,5% growth rate)

Target Area Population Growth Forecast (2019-2030): 181 persons

Target Area Density (2019): 817.5 persons / km²
 Target Area Density (2030): 860.9 persons / km²

Target Area Population Density Range

A population range can be estimated to have a more accurate value of the population count for urban, semi-urban, or rural settlement land cover areas that compile two or more raster cells, visualized in Fig. S3. The values for TDW_2 represent the median value of a population density range using the following equation (v);

$$D_{min} = TDW_2 - \left[\frac{(TDM)(TDW_2)}{SMW_2} \right] \quad (v)$$

Where, the minimum value for a density range D_{min} is the product between the mean of all TDW_2 in a “parent cell” (TDM) and the TDW_2 in question, sub-divided by the Sample Mean (SMW_2) of its corresponding land cover type using Table S2. The resulting value is subtracted from the target density weight TDW_2 providing the minimum value (or added to the TDW to obtain the maximum value of that range).

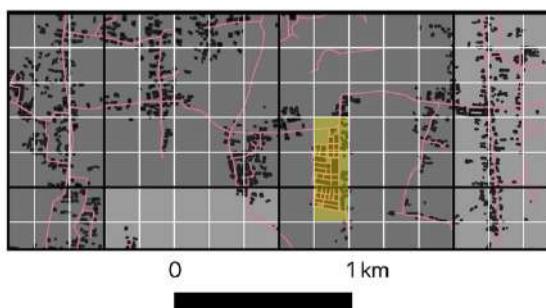


Fig. S3: Map of inset box in Fig. S1(B) corresponding to the census tract dasymetric mapping. Note that Pagerjurang settlement is located in the extent of 3 high-resolution (200-meter) raster cells within two semi-urban “parent” class. Color intervals are as shown in Fig. S1(B).

Accuracy Assessment

The following calculation is performed to validate the aforementioned dasymetric distribution sequence with an official high-spatial resolution census count in [Dataset 4] and [Dataset 5] for Pagerjurang in Table S3.

1. “Parent” cell id: 731

$TDW_2 = 124$ (id. 804)

TDM= 28 (the mean value for all density weights inside a “parent” raster cell)

$SMW_2 = 920$ (Sample Mean density for Semi-Urban settlements in Table 2)

$$D_{min} = 124 - \left[\frac{(124)x(28)}{920} \right]$$

$$= 120$$

$$D_{max} = 124 + \left[\frac{(124)x(28)}{920} \right]$$

$$= 128$$

2. “Parent” cell id: 731

$TDW_2 = 231$ (id. 805)

$D_{min} = 224$

$D_{max} = 238$

3. “Parent” cell id: 736

$TDW_2 = 156$ (id. 806)

TDM= 21

$SMW_2 = 920$

$$D_{min} = 152$$

$$D_{max} = 160$$

4. Pagerjurang Population Estimation (± 17 persons)

$$\sum D_{min} = 496; \quad \sum TDW_2 = 511; \quad \sum D_{max} = 526$$

Population Count for 2019 in Table S3 is: 528.3 persons

Table S3: Pagerjurang Population Count for 2014 and 2017 with a growth rate of 0,3%					
2014	2015	2016	2017	2018	2019
495	501.66	508.32	515	521.64	528.3

Note: The value for the year 2019 is calculated by applying the 0,3% growth rate from the 2014 and 2017 official census counts sourced from: Monografi Desa Kepuharjo 2017 [Dataset 4], and Data Sekunder Laporan Profil Desa Kepuharjo 2015 [Dataset 5].

Therefore, the upper value of the threshold (526) is 2 persons away from the actual value (528) in 2019, illustrating robustness of the model with an accuracy of $96.78\% \pm 3.2\%$.

Population by Age Group (Adults: 18 years or older)

Target Area Population Growth Forecast (2019-2030): 181 persons

Population according to Age Group (2019) [Dataset 6]:

0 – 19 years: 29%

20+ years: 71%

Adults = 128.5

(3) Node Weight Estimation

Input Data 2: Household Preferences from housing-preference-driven demand in typical urban structures.

Social Neighborhoods: 71%

Surroundings: 12%

Centrality Near Job: 10%

Education Facility: 2%

Crime Rate: 0%

Green Areas Supply: 0%

NA: 5%

$$(71\% + 12\% + 10\%) = 93\%$$

$$0.93 \times 128.5 = 119.5$$

Total New Adults: 119.5

Node 1

Category: Local Node

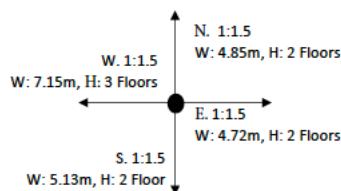
Influence Radius: 300 meters

Road Type:

A. Vertical axis: Local Road (width (m): 4.85 north, 5.13 south) 2 Lanes / Ratio 1.5:1

B. Horizontal axis (left side): Collector Road (width (m): 7.15 west) 2 Lanes / Ratio 1.5:1

C. Horizontal axis (right side): Local Road (width (m): 4.72 east) 1 Lane / Ratio 1.5:1



Node 2

Category: Local

Influence Radius: 300 meters

Road Type:

A. Vertical axis: Collector Road (width (m): 4.62 north, 5.10 south) 2 Lanes / Ratio 1:1.5

B. Horizontal axis: Local Road (width (m): 5.60 west, 6.00 east) 1 Lane / Ratio 1:1.5

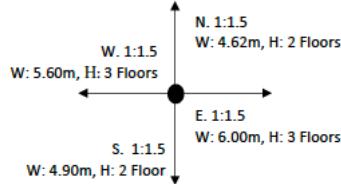


Fig. S4: Road width and enclosure ratios for comparison between typologies of equal categories. Note that each road segment is treated independently to achieve the highest building density output.

Weights

Node 1: 21.85 meters (51%)

Node 2: 21.12 meters (49%)

Total sum: 42.97 (100%)

Total Adults: 119.5

$119.5 \times 0.51 = 61$ adults in Node 1

$119.5 \times 0.49 = 58.5$ adults in Node 2

(4) Housing Demand

Input Data 3: Household composition of Adults per Household (APH):

1-Adult HHs: 52%

2-Adult HHs: 32%

3-Adult (or more) HHs: 16%

Node 1

$61 \times 0.52 = 31.72 \rightarrow 32$ HHs

$61 \times 0.32 = 19.52 \rightarrow 10$ HHs

$61 \times 0.16 = 9.76 \rightarrow 3$ HHs

Total: 45 HHs

Node 1 – 51% (0.51)

$0.51 \times 88 = 45$ HHs

Node 2

$58.5 \times 0.52 = 30.42 \rightarrow 31$ HHs

$58.5 \times 0.32 = 18.72 \rightarrow 9$ HHs

$58.5 \times 0.16 = 9.36 \rightarrow 3$ HHs

Total: 43 HHs

Node 2 – 49% (0.49)

$0.49 \times 88 = 43$ HH

Node 1

$32 (1 \text{ adult}) \times 70\text{m}^2 = 2,240 \text{ m}^2$

$10 (2 \text{ adult}) \times 100\text{m}^2 = 1000 \text{ m}^2$

$3 (+3 \text{ adult}) \times 120\text{m}^2 = 360 \text{ m}^2$

Total Area: $3,600\text{m}^2$

Node 2

31 (1 adult) x 70m² = 2,170 m²

9 (2 adult) x 100m² = 900 m²

3 (+3 adult) x 120m² = 360 m²

Total: 3,430 m² Housing Demand Area: 7,030 m²

The following Table S4 corresponds to the data collected using Google Street View for the current (2019) building heights. Road widths together with enclosure ratios were used to increment space for each lot with access to roads pertaining to a node. Finally, the net-built capacity was calculated in surface area (m²).

Table S4: Built Density Thresholds. Enclosure ratios provide a maximum height parameter for building growth in roads leading to the junction of each node. This table shows the net-built capacity of buildings with access to these roads by 2030.

Road	Hierarchy	Node 1			Enclosure Ratio (H:W)		Lots with direct street access (m ²) 2019	Building Levels (m)	Growth / Level (m ²) 2030	Net Growth / Level (m ²) 2030	Max. Built Area (m ²) 2030
		Direction	Width (m)	Lanes (unit)	2019	2030					
1	Local	North	4.85	2	0.6:1	1.5:1	4,292 356 -	L1: +0.00 L2: +2.85 L3: +5.70	3885 763	3885 1170	7,770 2,289
2	Local	South	5.13	2	0.5:1	1.5:1	3,798 -	L1: +0.00 L2: +2.85 L3: +5.70	3,723 75	3,723 150	7,446 225
3	Collector	West	7.15	2	0.4:1	1.5:1	3,835 296 -	L1: +0.00 L2: +2.85 L3: +5.70	681 3450	681 6604	1,362 10,054
4	Local	East	4.72	1	0.6:1	1.5:1	2,428 269 -	L1: +0.00 L2: +2.85 L3: +5.70	783 1,914	783 3,559	1,028 6,011
							15,274			20,555	36,185
Road	Hierarchy	Node 2			Enclosure Ratio (H:W)		HHs with direct street access (m ²) 2019	Building Levels (m)	Growth / Level (m ²) 2030	Net Growth / Level (m ²) 2030	Max. Built Area (m ²) 2030
		Direction	Width (m)	Lanes (unit)	2019	2030					
1	Collector	North	4.62	2	0.6:1	1.5:1	1,934 261 -	L1: +0.00 L2: +2.85 L3: +5.70	2,195	1,934	4,390
2	Collector	South	5.1	2	0.6:1	1.5:1	3,351 90 -	L1: +0.00 L2: +2.85 L3: +5.70	2,877 564	2,877 1,038	5,754 1,692
3	Local	West	5.6	1	0.5:1	1.5:1	1,511 452 -	L1: +0.00 L2: +2.85 L3: +5.70	255 1,705	2,958	4,985
4	Local	East	6	1	0.5:1	1.5:1	933 264 -	L1: +0.00 L2: +2.85 L3: +5.70	177 1,023	1,782	3,063
							8,796			8,796	10,589
											19,884

Note: the mean inter-story height is calculated as 2.85 m. This value corresponds to the mean first level height of buildings surveyed using the ruler tool and standard finishing materials in Google Street View.

5.5.2 Socioeconomic Environment

(1) Land Use Designation Module

Data Input 3: User-defined activities compatible with the structuring activities of a node.

Growth Areas by 2030

Category:

Node (1): Local-Institutional Service Node

Node (2): Local-Educational Service Node

Proposed Primary Activities:

(1) Technical Research Institute (primary “Institutional” node compatible with “educational” land use) [4].

*Area: 4,060m²

(2) Cultural Language Institute (primary “Educational” node compatible with “educational” land use) [5].

*Area: 4,027 m²

(3) Local Medical/Veterinary Center (primary “Educational” node compatible with “service” land use) [6, 7].

*Area: 2,300 m²

Total: 10,387 m²

(2) Employment Distribution

Table S5: Data collected for the OLS regression based on sectoral areas of the local employment structure in the Cangkringan district (see Mendeley Repository). Data source is openly available in Bahasa language. (<https://kependudukan.jogjaprov.go.id/statistik/penduduk/statusperkawinan/9/0/17/04/34.clear>).

Year	T1	NT	T	A	Pop	NT*T	NT*A	T*A	(xi-xmean)^2	(yi-beta(xi))	(yi-ymean)^2	yhat	(yhat-ymean)^2	Error	Error^2
2015	18	64	20	-106	7979	1280	-6784	-2120	21.78	49.94	1021.87	35.31	107.76	28.69	822.88
2015	10	93	24	-238	10694	2232	-22134	-5712	75.11	76.13	3716.93	38.13	37.12	54.87	3011.14
2015	0	22	30	-78	4026	660	-1716	-2340	215.11	0.91	100.67	42.34	106.32	-20.34	413.89
2015	1	25	-1	-36	3334	-25	-900	36	266.78	25.70	49.47	20.55	131.85	4.45	19.80
2015	9	36	24	-55	4950	864	-1980	-1320	75.11	19.13	15.73	38.13	37.12	-2.13	4.52
2016	21	17	8	-62	7864	136	-1054	-496	53.78	11.38	226.00	26.88	26.58	-9.88	97.57
2016	14	20	12	-81	10726	240	-1620	-972	11.11	11.56	144.80	29.69	5.49	-9.69	93.89
2016	0	38	1	-90	4018	38	-3420	-90	205.44	37.30	35.60	21.96	101.54	16.04	257.39
2016	4	1	5	-9	3379	5	-9	-45	106.78	-2.52	963.07	24.77	52.77	-23.77	564.96
2016	9	34	16	-23	5041	544	-782	-368	0.44	22.75	3.87	32.50	0.22	1.50	2.24
2017	18	20	27	-123	7777	540	-2460	-3321	136.11	1.02	144.80	40.24	67.27	-20.24	409.47
2017	15	10	23	-90	10607	230	-900	-2070	58.78	-6.17	485.47	37.42	29.05	-27.42	752.03
2017	0	3	2	-36	3948	6	-108	-72	177.78	1.59	842.93	22.66	87.86	-19.66	386.50
2017	4	21	-1	-23	3383	-21	-483	23	266.78	21.70	121.73	20.55	131.85	0.45	0.20
2017	9	-5	17	-20	5058	-85	100	-340	2.78	-16.95	1371.47	33.21	1.37	-38.21	1459.62
2018	23	44	36	-50	7811	1584	-2200	-1800	427.11	18.69	143.20	46.56	211.09	-2.56	6.57
2018	17	98	34	-82	10713	3332	-8036	-2788	348.44	74.10	4351.60	45.16	172.21	52.84	2792.45
2018	1	22	13	-11	4000	286	-242	-143	5.44	12.86	100.67	30.39	2.69	-8.39	70.44
2018	3	11	18	-18	3396	198	-198	-324	7.11	-1.65	442.40	33.91	3.51	-22.91	524.78
2018	10	52	30	-12	5186	1560	-624	-360	215.11	30.91	398.67	42.34	106.32	9.66	93.23
2019	25	59	20	-52	7829	1180	-3068	-1040	21.78	44.94	727.20	35.31	10.76	23.69	561.02
2019	17	33	20	-69	10755	660	-2277	-1380	21.75	18.94	0.93	35.31	107.76	-2.31	5.36
2019	1	50	-3	-19	4045	-150	-950	57	336.11	52.11	322.80	19.14	166.12	30.86	952.05
2019	3	17	9	-18	3453	153	-306	-162	40.11	10.67	226.00	27.58	19.82	-10.58	111.95
2019	11	10	23	-35	5227	230	-350	-805	58.78	-6.17	485.47	37.42	29.05	-27.42	752.03
2020	27	18	21	-41	7898	378	-738	-861	32.11	3.24	196.93	36.02	15.87	-18.02	324.62
2020	21	97	9	-108	10842	873	-10476	-972	40.11	90.67	4220.67	27.58	19.82	69.42	4819.02
2020	2	26	3	-52	4080	78	-1352	-156	152.11	23.89	36.40	23.36	75.18	2.64	6.96
2020	3	8	16	-34	3489	128	-272	-544	0.44	-3.25	577.60	32.50	0.22	-24.50	600.35
2020	11	17	4	-33	5332	68	-561	-132	128.44	14.19	226.00	24.07	63.48	-7.07	49.92

Source: Semester II Employment Data, Secretariat of Governance Bureau DIY [Dataset 7]. Tradable sub-sector (T1), Non-Tradable (NT), Tradable sector (T), and Agriculture (A) jobs are listed.

Table S6: Jobs Placed and Removed

Total	-0.2	30.2	0.0	30		67	120	-17	-	653	-53	-	1417	-243		2257	-313
Total	-51	-30.2	0	-51													

Note: *The Employment Density Guide 3rd Edition, 2015, **Area Guidelines for Mainstream Schools. 2014, ***International Health Facility Guidelines Part C: Version 4, 2015. Administrative entities correspond to the Cangkringan district's sub-divisions Argomulyo Village (A); Wukirsari (W); Glagaharjo (G); Kepuharjo (K); and, Umbulharjo (U). Local Multipliers (LM) are applied to two observances (2020-2025) and (2025-2030).

Density Guide 3rd Edition, 2015

Commerce: 15 m²/job

Small Service: 10m²/job

Industry: 47 m²/job

Manufacturing: 36 m²/job

Small retail (HBE): 12 m²/job

General teaching 55 m²/job (30-person classroom)

Medical center 9 nurses/1000 persons

Medical center 3 nurses/1 doctor

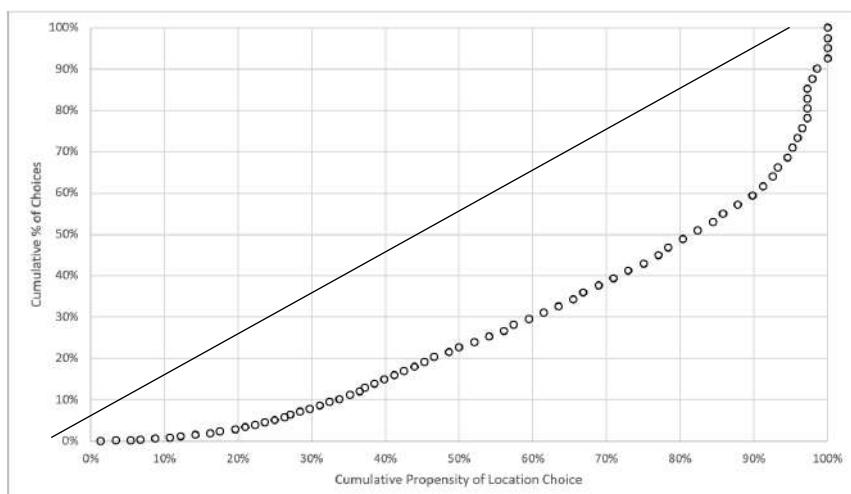
5.2.3 Employment Allocation

Regression Statistics	
Multiple R	0.56425996
R Square	0.318389302
Adjusted R Square	0.277696126
Standard Error	0.766674846
Observations	72

F-Test of Overall Significance					
	df	SS	MS	F	Significance F
Regression	4	15.79022172	3.947555429	7.781854432	3.266E-05
Residual	67	33.98755606	0.507276956		
Total	71	49.77777778			

Residual square-sums are too high and the f-test result indicates a low value with high significance. However, in Fig. S5, the efficiency dictates prediction of job location choices with 91% efficiency.

Fig S5: Lorenz-curve Predictive Efficiency.



*Note: 80% of location choices were made with 50% of the total set of choices.

<i>Likelihood Ratio Test</i>	
Vacancy	-11.84
Built Density	-12.30
Building Capacity	-11.84
Land Value coef.	-11.85
OBS	72
Probability Log-Likelihood	0.79
df	4

Note: An overall 79% probability of prediction. Predictors are all above the differential chi-square value 4. Therefore, all are significant predictors for job location choice.

Pool of Vacancy Areas by Land Use

Node 1	Built Capacity (m ²)	Jobs
Structuring Activity	4027	60
Residential Capacity	5687	
HBE Capacity	5224	435
Services	3912	391
Commerce	5524	368
Light Industry	208	6

Node 2	
Structuring Activity	6327
Residential	7817
HBE	2379
Vacant Lots	4909
Services	309
Commerce	2720
Light Industry	370

Aggregate Built Capacity

Total built capacity node 1: 24,582 m²

Total built capacity node 2 (Structuring Activity excluded): 24,831 m²

Total: 49,414 m²

Aggregate Employment Demand / Jobs

Tradable: 1,022 jobs

Non-Tradable: 692 jobs

Entrepreneurs: 230 jobs

Total: 1,944 jobs

Tradable demand: 1,022 – (164 Services + 549 commerce + 422 services) = 113 jobs vacancy available.

Entrepreneur demand: 230 – (633 HBEs) = 403 job vacancies available*

*Note: become rental space for non-tradable jobs below.

Non-tradable demand: 692 – (113+403+16 Light Industry) = 160 jobs left without space for allocation.
91% allocated

Non-Allocated (non-tradable): 160 jobs

*These were assumed to take hold of vacant lots in node 2 with residential use, thus becoming HBE vacancies.

Vacant Residential Lots node 2: 4,909 m²

Vacant Residential Area used for 160 HBE jobs: 1,920 m²

$$160 \times 12 \text{ m}^2 = 1,920 \text{ m}^2$$

Vacant Area left: 2,989 m²

Land use change (residential → mixed-use):

$$7,603 \text{ m}^2 (\text{HBE Capacity}) + 1,920 \text{ m}^2 (\text{Vacant Lots used for HBES}) = 9,523 \text{ m}^2$$

(4) Spillover Management

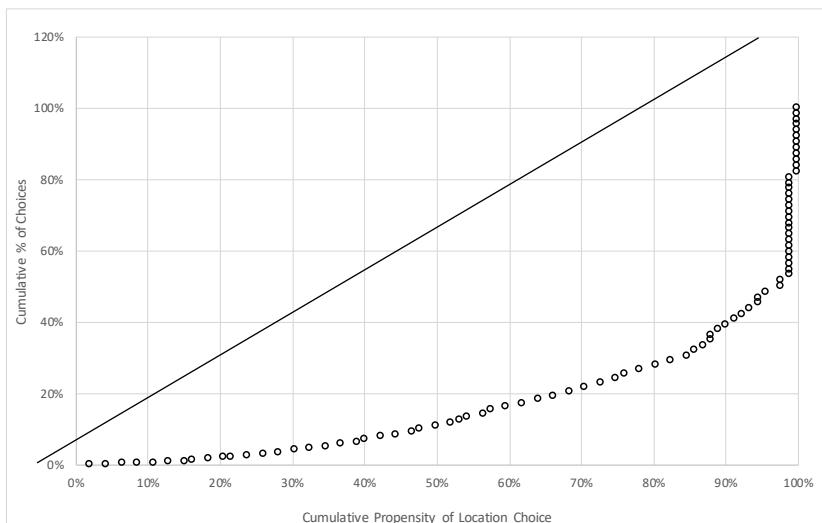
Regression to predict the probability of a site being selected. The R² value indicates that 50-53% of the choices can be predicted with the model. However, the f-test with a significance higher than 99% provides an overall fitted measure of robustness.

<i>Regression Statistics</i>	
Multiple R	0.728863266
R Square	0.53124166
Adjusted R Square	0.5083754
Standard Error	0.902066226

<i>F-Test of Overall Significance</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	4	75.61950252	18.9048756	23.2325553	7.35327E-13
Residual	82	66.72532507	0.81372348		
Total	86	142.3448276			

Note: High f-value with very significant rating (99%). However, residual values were high.

Fig. S6: Lorenz-curve distribution for predictive efficiency.



Note: 90% of location choices were predicted with 40% of total choices. Predictive efficiency = 81%.

Likelihood Ratio Test

Intercept	-24.85
Vacancy	-20.37
Built Density	-22.27
Building Capacity	-24.17
Land Value coef.	-23.10
OBS	87
Probability log-Likelihood	0.77
df	4

Note: resulting values are very high with an overall prediction probability of 77%. Positive or negative signs are not considered.

Floating Population - Residential Demand

Node 1

In-Migration for the 1,260 jobs allocated

$$1,260 \times 0.16 = 202 \text{ jobs} \rightarrow 202 \text{ HHs of 3-adults} = 606 \text{ adults}$$

$$1,260 \times 0.32 = 403 \text{ jobs} \rightarrow 403 \text{ HHs of 2-adults} = 806 \text{ adults}$$

$$1,260 \times 0.52 = 655 \text{ jobs} \rightarrow 655 \text{ HHs of 1-adult} = 655 \text{ adults}$$

Estimation of Adults that do not work (16% of Kepuharjo sub-district [Dataset 7])

3-adult HHs + 2-adult HHs

$$(933 \text{ adults} + 1244 \text{ adults}) \times 0.16 = 348 \text{ adults in total do not work in the TAZ.}$$

16% → X

$$(16+32=48) \rightarrow 100\%$$

X=33%, Y=77%

$$348 \times 0.33 = 115 \text{ (3-adult HHs)}$$

$$348 \times 0.77 = 233 \text{ (2-adult HHs)}$$

3-Adult HHs Node 1

$$115 \text{ adults} \times 0.51 \text{ (Node Weight Estimations in sub-section 5.4.1)} = 58.65 \text{ adults}$$

$$59 \text{ adults that do not work} \rightarrow 59 \text{ HHs} \times 2 \text{ adults} = 118 \text{ jobs} + (59 \text{ adults not working})$$

Therefore, 59 HHs have 2 working adults and 59 adults that do not work.

$$202 \text{ jobs} - 118 \text{ jobs} = 84 \text{ jobs} / 3 \text{ adults} = 28 \text{ HHs with 3-adults working.}$$

Total: 28 HHs (3 working adults) + 59 HHs (2 working adults) = 87 HHs, 202 jobs, 261 adults.

2-Adult HHs Node 1

$$233 \times 0.51 \text{ (Node Weight Estimations in sub-section 5.4.1)} = 118.83 \text{ adults}$$

$$119 \text{ adults that do not work} \rightarrow 119 \text{ HHs} \times 1 \text{ adult} = 119 \text{ jobs} + (119 \text{ adults not working})$$

Therefore, 119 HHs have 1 working adult and 119 adults do not work.

$$403 \text{ jobs} - 119 \text{ jobs} = 284 \text{ jobs} / 2 \text{ adults} = 142 \text{ HHs with 2-adults working.}$$

Total: 119 HHs (1 working adult) + 142 HHs (2 working adults) = 261 HHs, 403 jobs, 522 adults.

1-Adult HHs in Node 1

$$1,260 \times 0.52 = 655 \text{ jobs} \rightarrow 655 \text{ HHs of 1-adult} = 655 \text{ adults (all work)}$$

Total Residential Demand for Node 1 (1,012 HHs, 1,456 adults)

$$87 \text{ (3-adult HHs)} \times 120 \text{ m}^2 = 10,440 \text{ m}^2$$

$$261 \text{ (2-adult HHs)} \times 100 \text{ m}^2 = 26,100 \text{ m}^2$$

$$655 \text{ (1-adult HH)} \times 70 \text{ m}^2 = 45,850 \text{ m}^2$$

Sub-Total: 82,390 m²

Additional Housing Demand (sub-section 5.1.4): 3,600 m²

Total Demand: 85,990 m²

Residential vacancy: 10,911 m²

Deficit: 75,079 m²

Node 2

In-Migration for 684 jobs allocated (+160 HBE jobs)
 $684 \times 0.16 = 109$ jobs → 109 HHs of 3-adults = 327 adults
 $684 \times 0.32 = 219$ jobs → 219 HHs of 2-adults = 438 adults
 $684 \times 0.52 = 356$ jobs → 356 HHs of 1-adult = 356 adults

3-Adult HHs Node 2

115×0.49 (Node Weight Estimations in sub-section 5.4.1) = 56.35 HHs
 56 adults that do not work → 56 HHs x 2 adults = 112 jobs + (56 adults not working)
 Therefore, 54 HHs have 2 working adults and 54 adults that do not work.
 109 jobs – 108 jobs = 1 jobs / 3 adults = 0 HHs with 3-adults working.

Total: 1 HHs (3 working adults) + 53 HHs (2 working adults) = 54 HHs, 109 jobs, 162 adults.

2-Adult HHs Node 2

233×0.49 (Node Weight Estimations in sub-section 5.4.1) = 114.17 HHs
 114 adults that do not work → 114 HHs x 1 adult = 114 jobs + (114 adults not working)
 Therefore, 114 HHs have 1 working adult and 114 adults do not work.
 219 jobs – 114 jobs = 105 jobs / 2 adults = 52.5 HHs with 2-adults working.

Total: 115 HHs (1 working adult) + 52 HHs (2 working adults) = 167 HHs, 219 jobs, 334 adults.

1-Adult HHs in Node 2

Total: $684 \times 0.52 = 356$ jobs → 356 HHs of 1-adult = 356 adults (all work)

Total Residential Demand for Node 2 (573 HHs, 844 adults)

54 (3-adult HHs) x $120 \text{ m}^2 = 6,480 \text{ m}^2$
 167 (2-adult HHs) x $100 \text{ m}^2 = 16,700 \text{ m}^2$
 356 (1-adult HH) x $70 \text{ m}^2 = 24,920 \text{ m}^2$
 Sub-Total: $48,100 \text{ m}^2$
 Additional Housing Demand (sub-section 5.1.4): $3,430 \text{ m}^2$
 Total Demand: $51,530 \text{ m}^2$
 Residential vacancy: $15,105 \text{ m}^2$
Deficit: 36,425 m²

Residential Demand in TAZ (aggregate values)

$87 + 54 = 141$ (3-adult HHs) x $120 \text{ m}^2 = 16,920 \text{ m}^2$
 $261 + 167 = 428$ (2-adult HHs) x $100 \text{ m}^2 = 42,800 \text{ m}^2$
 $655 + 356 = 1,011$ (1-adult HH) x $70 \text{ m}^2 = 70,770 \text{ m}^2$
 Sub-Total: $130,490 \text{ m}^2$
 Additional Housing Demand (sub-section 5.1.4): $7,030 \text{ m}^2$
 Total Demand: $137,520 \text{ m}^2$
 Residential vacancy: $26,016 \text{ m}^2$
Deficit: 111,504 m²

Household Allocations in Nodes

From the estimations above: 301 HHs

The Horizontal Boundary road between node 1 and node 2 is applied with a (H:W) enclosure ratio applicable to its width.

Service: 943 m^2
 Commerce: $2,544 \text{ m}^2$
 Industry: 207 m^2
 Residential: $2,088 \text{ m}^2$
 Mixed-Use Residential: $3,570 \text{ m}^2$
Total residential: 5,658 m²
 Total non-residential: $3,694 \text{ m}^2$
 Total: $9,352 \text{ m}^2$

The vertical boundary road is applied with a (H:W) enclosure ratio and delivers the following growth areas.

Residential: 1,232 m²

Mixed-Use Residential: 1,388 m²

Total residential: 2,620 m²

Non-Allocated in nodal Areas (HH Type, Income USD: 100>HH income>100, HH APH ratios) = Max. of 1,284 HHs using the Household Allocation Model (in Mendeley Data).

*Note: Max. = Equilibrium. 100 USD is an approximation to the mean income in our data sample (78 USD).

(3-Adult, >100 USD, 0.16) = 141 HHs (549 p), >100 → 100% allocated

(3-Adult, < 100USD, 0.16) = 0% allocated

(2-Adult, >100 USD, 0.32) = 377 HHs (644 p) > 100 → 100% allocated

(2-Adult, < 100USD, 0.32) = 51 HHs (88 p) < 100 → 90 HHs → 100% allocated (+39 vacancies left)

(1-Adult, >100 USD, 0.52) = 607 HHs, (356 p) >100 → 100% allocated

(1-Adult, < 100USD, 0.52) = 404 HHs (238 p) < 100 → 146 HHs → (+/-45 vacancies) → 219 HHs (80%)

Non-Allocated in Nodes = 1,284 HHs

Non-Allocated at all = 254 HHs / 86% of 1-adult HHs and 14% of 3-adult HHs.

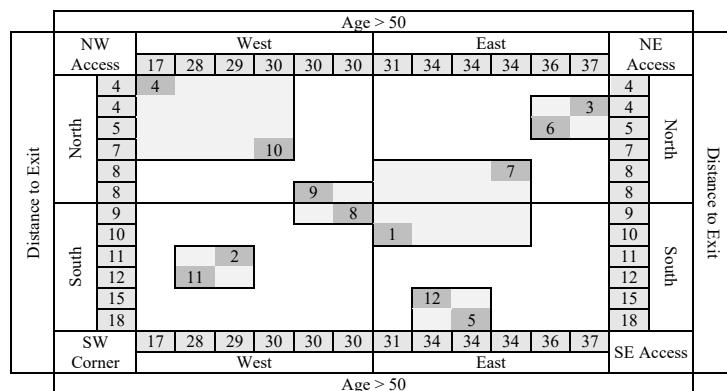
5.5.3 Settlement Environment

(1) Household Clustering

Table S5: Geo-referenced Attribute Values by Cluster. (X, Y) = (Age>50, Dist. To Exit)

X-Axis Cluster	4	11	2	10	9	1	12	5	7	6	3
Land Use	Education	Commerce	Industry	Services	Services	Commerce	Commerce	Commerce	Industry	Industry	Services
Age>50	17	28	29	30	30	31	34	34	34	36	37

Y-Axis Cluster	Land-Use	Distance to Exit
4	Education	4
3	Services	4
6	Small Industry	5
10	Small Industry	7
7	Small Industry	8
9	Services	8
8	Commerce	9
1	Commerce	10
2	Commerce	11
11	Commerce	12
12	Commerce	15
5	Commerce	18



(2) Street Network

Block Types

Three block types are analyzed to satisfy 26 HHs (1 cluster) or 52 HHs (2 clusters). Lots are equal to 120 m² in order to satisfy household sizes of more than 2 adults. However, 30% is established to be the minimum footprint area for gardens and 250 m² recreational areas are integrated to the three standard block types in Table S6.

Table S6: Standard Block Types

Neighborhood Cell Estimation			
Unit	1 Small Cell	1 Big Cell	Combined Cell
Unit Area	80 x 80	100 x 100	80 x 100
Land Use	-	1 cluster	2 cluster
Persons	50 persons	100 persons	100 persons
Households	26 HH	52 HH	52 HH
Plots	1 Plot	26 Plots	52 Plots
Plot Area	120	3,120	6,240
Unbuilt Area	35	910	1,820
Built Area	85	2,460	4,420
Total Unbuilt Area	3,940	5,330	3,580
Public Space	3,030	3,760	1,760
Density m ² / person	78 / Ha	100 / Ha	125 / Ha
Recreational Areas	2.5 / person	250	200
Agro-Community	35 / HH	910	1820
Roads (10%)	454	564	234

*Note: values are in (m; m²; m³) unless indicated differently. Interior lot open space was designated to small agricultural plots, which in combination, are very useful to the community. Roads take 10% of each block area.

Selective Sampling of the Accessibility Index

The sampling method is performed to quantify the accessibility index for settlement simulation using Table S7. The following was evaluated: (i) Intersection density in a 0.25 (km²) target area; (ii) road hierarchies and connectivity between opposite sides of a network; (iii) number of access points to the network; and, a permeability designation for the target area based on the results as either (1) High, (2) Limited, (3) Low, and (4) none.

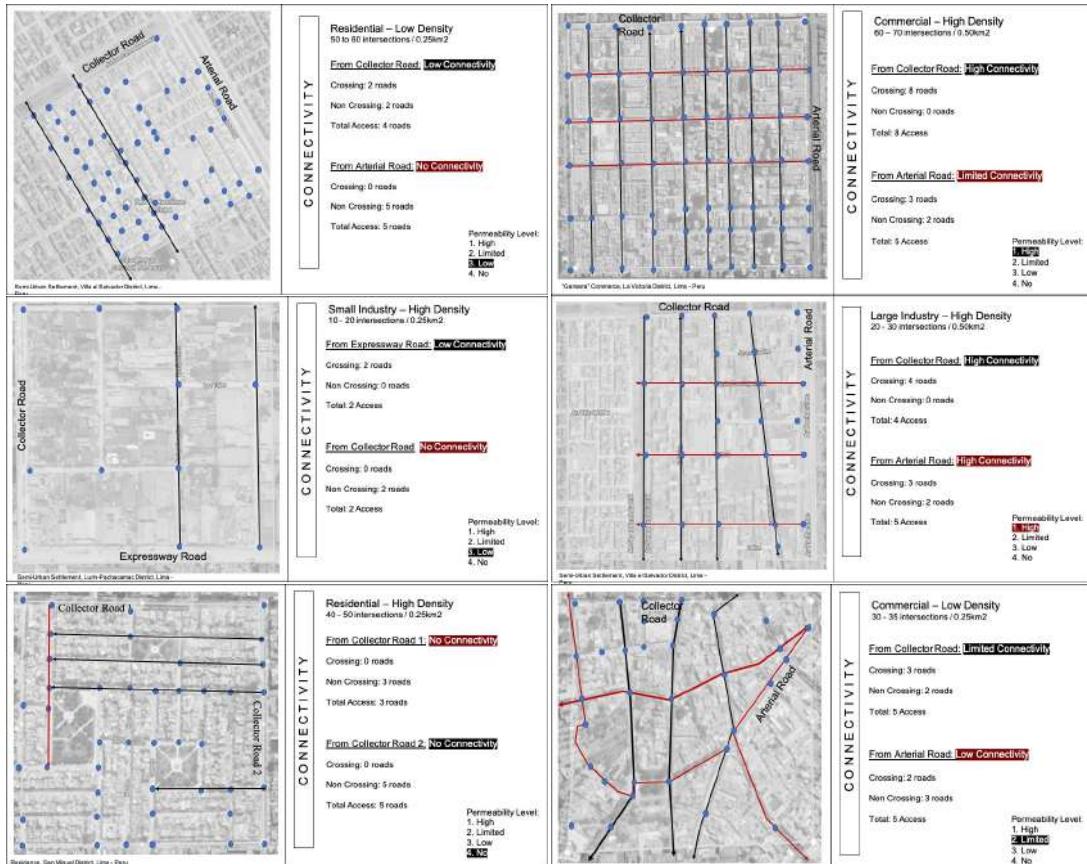


Fig. S9: Selective Sampling method for six different settlement areas in Peru, evaluated using Google Earth to quantify accessibility measures and density indices in [12] for: high- and low-density residential and commercial areas, office areas, and small- and large-industry densities, classified with two urban designations (i.e., semi-urban and urban). The number of intersections, block sizes, block proportions, and connectivity are variables sampled and they convey standard topological attributes of street networks for this study. Further sampling might be required to replicate this data for other locations around the world.

Table S7: Results from selective sampling of spatial characteristics in Lima city

Categories		Connectivity		Intersections		Permeability	
Zone	Density	Boundary Roads	Block Size (m)	Semi Urban	Urban	Road Crossing	Level
Residential	High	Local / Collector / Sub-Arterial	80 / 150 / 200	30 – 40	40 – 50	0	Zero
	Low	Local / Collector	80 / 100	40 – 50	50 – 60	2	Low
Commercial	High	Collector / Sub-Arterial	100 / 150	20 – 30	30 – 40	3	High
	Low	Local / Collector	80 / 100	30 – 35	30 – 40	2	Limited
Office	High	Sub-Arterial / Arterial	200 / 150	10 – 25	20 – 35	2	High
Small Industry	High	Local / Collector / Sub-Arterial	80 / 100 / 300	10 – 20.	15 – 25	2	Low
Large Industry	High	Local / Collector / Sub-Arterial	80 / 100 / 300	15 – 20	20 – 30	3	High

(3) Land Use Change

Here, we report the full statistic results for the first and second iteration of the predictive variables with our data sample. Then, the land use multinomial logit model is provided with full regression scores for the maximized log-likelihood of conversion choices. Our trained data and optimized logit scores are available online in our Mendeley data repository (Find the link at the top of this document).

First Iteration (Only those with primary occupations other than farming, services, or education.)

Regression Statistics

Multiple R	0.96
R Square	0.93
Adjusted R Square	0.92
Standard Error	0.16
Observations	41

F-Test of Overall Significance

	df	SS	MS	F	Significance F
Regression	6	12.21687472	2.036145786	76.48961888	2.70502E-18
Residual	34	0.905076502	0.026619897		
Total	40	13.12195122			

Regression Scores

Variables	Coefficients	Standard Error	t Stat	P-value
Intercept	-0.569	0.24	-2.407	0.022
HH comp	0.029	0.04	0.675	0.505
Occupation	0.789	0.05	17.476	0.000
Risk	0.031	0.04	0.840	0.407
income	0.001	0.00	0.994	0.327
age	-0.001	0.00	-0.618	0.541
education	-0.013	0.02	-0.798	0.430

*Note: P-value significance is low due to the number of observations.

Second Iteration (Only those with commercial and farming as their primary occupation. Assigned values to Commercial = 2 and Farming = 1).

<i>Regression Statistics</i>	
Multiple R	0.79
R Square	0.63
Adjusted R Square	0.56
Standard Error	0.20
Observations	41

<i>F-Test of Overall Significance</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	6	2.269614294	0.378269049	9.596855821	3.48989E-06
Residual	34	1.340141803	0.039415935		
Total	40	3.609756098			

<i>Regression Scores</i>					
<i>Variables</i>	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	
Intercept	0.844	0.267	3.161	0.003	
HH comp	-0.101	0.052	-1.941	0.061	
Occupation	0.290	0.066	4.38	0.000	
Risk	-0.034	0.046	-0.745	0.461	
income	-0.002	0.001	-1.943	0.060	
age	0.003	0.003	1.174	0.249	
education	-0.017	0.020	-0.823	0.416	

*Note: Higher P-values correspond to a trained dataset using the abovementioned values.

Table S8. Analysis of Predictive Efficiency

<i>Variables</i>	<i>First Iteration</i> 1	<i>Min</i> 2	<i>Max</i> 3	<i>Second Iteration</i> 4	<i>Min</i> 5	<i>Max</i> 6
Intercept	-0.569** (0.24)	-1.21	0.08	0.844*** (0.267)	0.12	1.57
HH composition	0.029 (0.04)	-0.09	0.15	-0.101* (0.05)	-0.24	0.04
Occupation	0.789*** (0.05)	0.67	0.91	0.29*** (0.06)	0.11	0.47
Segmentations	0.031 (0.04)	-0.07	0.13	-0.034 (0.04)	-0.16	0.09
Income	0.001 (0.00)	0.00	0.00	-0.002* (0.001)	0.00	0.00
Age	-0.001 (0.00)	-0.01	0.01	0.003 (0.003)	0.00	0.01
Education	-0.013 (0.02)	-0.06	0.03	-0.017 (0.02)	-0.07	0.04

***significant at 99%, **significant at 95%, *significant at 90%. Note that education with negative coefficient values indicate a negative correlation, while income and HH composition represent accurate predictors of land use change. F-significance values for the first and second iteration account for 76.49 and 9.59 with R² values of 93% and 63%, respectively. Given that the first iteration is statistically significant, occupation (constant variable), income, and risk were considered.

Multinomial Logistic Regression with Training Data

<i>Regression Statistics</i>	

Multiple R	0.893
R Square	0.797
Adjusted R Square	0.789
Standard Error	0.281
Observations	139

F-Test of Overall Significance

	df	SS	MS	F	Significance F
Regression	5	41.26	8.25	104.3	2.83989E-44
Residual	133	10.52	0.079		
Total	138	51.78			

Note the high F value with upper 99% significance for the overall regression model. The sum of squares residual accounts for approximately 20% of the total, which is adequately low.

Regression Scores						
Variables	Coefficients	Standard Error	t Stat	P-value	Min	Max
Intercept	0.227	0.161	1.406	0.162	-0.19	0.65
Segmentation	0.484	0.042	11.464	0.000	0.37	0.59
Occupation	0.226	0.070	3.223	0.002	0.04	0.41
Education	0.075	0.024	3.134	0.002	0.01	0.14
HH Composition	-0.291	0.045	-6.408	0.000	-0.41	-0.17
Income	0.010	0.002	5.801	0.000	0.01	0.01

All P-values are significant predictors excluding the y-intercept with standard errors in the low 5 – 10%. Additionally, the instrument reveals that HH composition has a negative correlation indicating that land use change may not account for only large families. Income and Risk-taking lifestyle segmentations have the largest predictive outcome. This is followed by Education and then Occupation. All variables are relevant, indicating the possibility of various indicators of land use change among sociodemographic attributes.

Multinomial logistic regression model with trained data

Variables	Maximized Coef.	Probability	Log-Likelihood	Likelihood Ratio Test
Segmentation	33.682	1.00	-1.75	-6.75
Occupation	1.585	0.83	-0.37	-5.37
Education	-0.380	0.41	-0.37	-5.37
HH Composition	-13.602	0.00	-0.37	-5.37
Income	0.414	0.60	-0.37	-5.37
Mean	3.69	0.57	-0.65	-5.65
Standard Deviation	15.75	0.39	0.62	0.62

Risk is a potentially very prominent predictor of land use change followed by occupation and Income. Together, the mean probability is low, but the standard deviation of the log-likelihood is also low. Meaning that education and HH composition were not able to suggest a predictive potential in the change of land use. This is because HHs in our target area mainly small, and do not represent a major indication of livelihood change, while education may seem appropriate to measure livelihood diversification, in reality, it does not stimulate livelihood changes in the region, yet. It is however, well understood that education is a major driver of specialization, but not necessarily an occupational multiplicity in rural areas.

5.5.4 Transport Component

(1) Travel within the settlement

25 HHs are randomly selected and origins and destinations are calibrated in the model. Travel schedules are revised for individuals with destinations in a 400-meter radius from their lots. Route directness is calculated between each selected lot and their corresponding entrance/exit point.

Change of Job into Node 2

4 vendors from the 36 adults who earn a job in node 2 are found within the 25 randomly selected HHs and their destination is changed to node 2.

Change of Job into Market

4 vendors are found within the selected 25 HHs and their destination is changed to the local market. This is based on couples with a same occupational land-use, which sum up to 30% of settlement population. The decision is also based on their individual itineraries (if they represent a part of the 30% of the population who have a side job as tourist vendors).

Pedestrians

Walking only to very close destinations (Market and Schools near the site)

Non-essential tours in these itineraries reflect stops in-between tour origin and final destination found to provide a measure for non-essential travel in neighborhood areas.

(2) Travel in Total

Congestion

Density and transit volume calculations can be seen in Table S9, where congestion levels substantially rise along the Dairy Farming Route by 2030.

Volume of vehicles per minute = (Density + Through Transit) / (Time Threshold for peak hour estimation) / (mean road width)

*Note: Time threshold deviation at 6:00±30min and 16:00±30min.

**Assumption: subjects with a same routine depart within a 30-minute interval.

Node 1 (Via “Jl. Raya Merapi” to “Jl..Kaliurang”)

Density

2019 → 319 (OSM land use “residential” data) = 319 persons

2030 → 1,084 persons (Floating Population) + 66 (growth rate) + 319 = 1,469 persons

Through Transit

2019 → 6:00am: 27.42 persons (survey data x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].

2019 → 16:00pm: 27.42 persons

2030 → 6:00am: 200 persons (survey data x 10.5 → factor of new employment locations statistically applied to Pagerjurang Population (605 persons) in 2030 [Dataset 4 growth rate].

2030 → 16:00pm: 200 persons

Congestion Node 1

Road mean width: 7.15m + 4.72m / 2 = 5.93m (Section 5.3)

(1) V (2019) = [Density (319) + Through Transit (27.42) / 30 min] / 5.93m = 1.95 persons / meter-wide / min. at 6:00 and 16:00

(2) V (2030) = [(1,084+200) / 30 min] / 5.93m = 7.22 persons / meter-wide / min. at 6:00 and 16:00

Node 2 (Via “Jl. Petung Merapi”)

Density

2019 → 140 (OSM land use “residential” data) = 140 persons

2030 → 1,073 persons (Floating Population) + 65 (growth rate) + 140 (OSM polygon “residential” data) = 1,278 persons

Through Transit

2019 → 6:00: 283 persons (survey data sample x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].

2019 → 16:00: 210 persons (survey data sample)

2030 → 6:00: 273 persons (survey data sample x 10.5 → factor of New Employment Locations statistically applied to Pagerjurang Population (605 persons) in 2030 [Dataset 4 growth rate].

2030 → 16:00: 189 persons

Congestion Node 2

Road mean width: $5.60m + 4.62m / 2 = 5.11m$ (section 5.3)

$$V(2019) = [(140+283) / 30 \text{ min}] / 5.11m = 2.76 \text{ persons / meter-wide / min. at 6:00}$$

$$V(2019) = [(140+210) / 30 \text{ min}] / 5.11m = 2.28 \text{ persons / meter-wide / min. at 16:00}$$

$$V(2030) = [(1,073+273) / 30 \text{ min}] / 5.11 = 8.78 \text{ persons / meter-wide / min. at 6:00}$$

$$V(2030) = [(1,073+189) / 30 \text{ min}] / 5.11 = 8.23 \text{ persons / meter-wide / min. at 16:00}$$

Node 3 (*Via “Jl. Petung Merapi”*)

Density

2019 → 501 (OSM polygon “residential” area) = 501 persons

2030 → 1000~ persons (Floating Population) + 250 (growth rate) + 501 (OSM polygon “residential” data) = 1,751 persons

Through Transit (same as Node 2 – 95 people staying at Node 2)

2019 → 6:00 : 283 persons (survey data sample x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].

2019 → 16:00 : 210 persons (survey data sample)

2030 → 6:00 : 178 persons (survey data sample x 10.5 → factor of New Employment Locations statistically applied to Pagerjurang Population (605 persons) in 2030 [Dataset 4 growth rate].

2030 → 16:00 : 94 persons

Congestion Node 3

Road Width = 7.49m (Table 1-Manuscript)

$$V(2019) = [(501+283) / 30 \text{ min}] / 7.49m = 3.49 \text{ persons / meter-wide / min. at 6:00}$$

$$V(2019) = [(501+210) / 30 \text{ min}] / 7.49m = 3.16 \text{ persons / meter-wide / min. at 16:00}$$

$$V(2030) = [(1,751+178) / 30 \text{ min}] / 7.49 = 8.58 \text{ persons / meter-wide / min. at 6:00}$$

$$V(2030) = [(1,751+94) / 30 \text{ min}] / 7.49 = 8.21 \text{ persons / meter-wide / min. at 16:00}$$

Node 4 (*Via “Jl. Wukirsari”*)

Density

2019 → 159 (Dasymetric Redistribution Data) = 159 persons

2030 → 1000~ persons (Floating Population) + 79 (growth rate) + 159 = 1,238 persons

Through Transit

2019 → 6:00am: 36.56 persons (survey data sample x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].

2019 → 16:00pm: 54.84 persons (survey data sample)

2030 → 6:00am: 164.5 persons (survey data sample x 10.5 →(Factor of new employment locations statistically applied to Pagerjurang's population (605 persons) in 2030 [Dataset 4].
 (128 (census count) + 36.56 (survey data sample))
 2030 → 16:00pm: 191 persons (54.84+128)

Congestion Node 4
 Road Width = 4.80m

$$V(2019) = [(159+36.56) / 30 \text{ min}] / 4.80\text{m} = 1.36 \text{ persons / meter-wide / min. at 6:00am}$$

$$V(2019) = [(159+54.84) / 30 \text{ min}] / 4.80\text{m} = 1.48 \text{ persons / meter-wide / min. at 16:00pm}$$

$$V(2030) = [(1,238+164.5) / 30 \text{ min}] / 4.80 = 9.74 \text{ persons / meter-wide / min. at 6:00am}$$

$$V(2030) = [(1,238+191) / 30 \text{ min}] / 4.80 = 9.92 \text{ persons / meter-wide / min. at 16:00pm}$$

Table S9: Results for Congestion along the Dairy Farming Route by 2030.

Travel Route Via Kaliadem Raya (persons / 15 min)					
Year		Dist (km)	Speed (km/h)	Congestion (6:00)	Congestion (17:00)
2019	2019	1.2	10.77	41	34.2
	2019	1.28	42	0	0
	2019	0.68	10.77	52	47.4
	2019	3.05	42	0	0
	2030	0.72	42	0	0
2030	2030	0.6	10.77	132	123.45
	2030	0.49	42	0	0
	2030	0.68	10.77	129	123.15
	2030	3.05	42	0	0
Travel Route Via Wukirsari (persons / 15 min)					
2019	2019	0.96	42	0	0
	2019	1.44	10.77	26.4	22.2
	2019	1.56	42	0	0
	2019	0.69	10.77	0	0
	2030	0.96	42	0	0
2030	2030	2.95	10.77	146	149
	2030	0.69	10.77	0	0

Table S10: Comparative Urban Form Measures for 2019 and 2030

		2019	2030	Total
	Density Index (person / sq.km)		818	1,589
1	Census Population		3,409	3,590
2	Density (person / sq./km)	818	861	43
3	Census Growth HHs (0.5%)	1,774	1,868	94
4	Population (census growth rate)	3,409	3,590	105%
5	Population (with intervention)	3,409	6,193	181%
6	Transit Volume Node 2 (6:00 am)	2.76	8.78	169%
7	Transit Volume Node 3 (6:00 am)	3.49	8.58	159%
8	Transit Volume Node 4 (6:00 am)	1.36	9.74	186%
	Land Use Index (neighborhood shopping in a 400 m radius from main entrance in sq./m)	1,994	6,424	169%
9	Employment Demand		1,944	154%
10	Employment Offer		1,260	100%
11	Entrepreneurs (HBE land use change)		1,023	100%
12	Census Housing Demand (sq./m)		7,030	5.2%
13	Total Housing Demand (sq./m)		130,490	94.8%
14	Edge Housing Demand (Spillover)		27,876	20.3%
15	Housing Supply (sq./m)		15,674	56%
16	Housing Deficit (sq./m)		12,202	44%

17	Total Development (sq./m)		61,386	83.4%
	Accessibility Index		"No"	"Low"
18	Block Length (m)	30, 55, 85	26, 33, 65	Lower
19	Access Points	14	13	-7%
20	Intersections	57	45	-17%
21	Connectivity	0	1	Higher

Public Transportation Assessment

Estimation

A convenient route is traced on GIS using node hierarchies and sub-centers visualized in Fig. S10. A Hyundai H-1, 2.4 Wagon GLS with a consumption level of 10.2 L /100 km and 241 g/km emission rate of CO₂ is selected for semi-urban terrain. The vehicle has a capacity of nine passengers but only six are considered for real-time operation between 5:00am and 5:00pm, Monday through Saturdays. Subjects going to Ngudi Makmur, Desa Petung, or Kaliadem take this mode from node 2. They are selected by age, distance to bus stop, and HH expenditure comparisons without exceeding 30% of all subject candidates as the minimum number considered for sustainable enterprises. The latter assertion results in 13 subjects in total (32%). The choice modelled considers all previous requirements if their daily fuel consumption rate is equal or more than \$0.17 daily. It is worth mentioning that the price of a complete tour to Kaliadem (the farthest destination) is \$0.12 using public transportation as opposed to \$0.17 using a motorcycle. Operation costs are \$57.65 and net income adds to \$75.00 / month working at 8-route cycles daily with 19.84 VKT per cycle. The results account for simulated HHs (8.3% of the settlement) and gathers a 57% total reduction of VKT, 33% reduction for travel within the settlement, and 34% in fuel savings equal to \$1.75 per day as well as 32% in energy savings and 41% in CO₂ emissions.

Vehicle Capacity: 9 Passengers

Motor: 10.2L = 100 km,

1L = 9.80 km,

Petrol: 1L = \$0.49,

1 km = \$0.05 + 30% Income = \$0.065

Cost/km: \$0.065

****Assumption:** 6 Persons/km (average)

Table S11: Calculation and Estimation for Travel Fares and Fuel Consumption

H1 2.4 Wagon GLS (Petrol \$0.49/L, 9 Passenger Capacity)					Calculation with 6 Persons per/Km	
Origin	Destination	Total VKT (one way)	Total VKT (Km)	Consumption	Total Consumption	Travel Cost (+30% Income)
Node 2	Kaliadem	5.28	10.56	0.05 /Km	\$ 0.69	\$ 0.120
Node 2	Desa Petung	2.48	4.96	0.05 /Km	\$ 0.32	\$ 0.053
Node 2	Ngudi Makmur	4.525	9.05	0.05 /Km	\$ 0.59	\$ 0.098
10.2L = 100 Km, 1L = 9.80Km, \$0.05 = 1Km					https://www.group1hyundai.co.za/hyundai-h1-bus/	

Destination Kaliadem: 10.56 km x \$0.065 = 0.686/6 = \$0.12 (Total Fare)

Destination Desa Petung: 4.96 km x \$0.065 = 0.322/6 = \$0.053 (Total Fare)

Destination Ngudi Makmur: 9.05 km x \$0.065 = 0.588/6 = \$0.098 (Total Fare)



Fig. S10: Tentative Public Transportation Route and Convenient Bus Stops

Travel Origins and Household Mapping



Fig. S11: Household numbers in order of interviews used for the first travel simulation (Left). Households randomly selected for the second travel simulation (Right).

References

- [1] MDA Federal Inc. (1995–2007) EarthSat GeoCover LC Overview. Rockville, MD: MacDonald, Dettwiler and Associates Ltd. <http://www.mdafederal.com/geocover/geocoverlc/gclcoverview> In: Gaughan AE, Stevens FR, Linard C, Jia P, Tatem AJ (2013), “High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015”. PLoS ONE 8 (2), p. 4. Apr. 19, 2021. <http://doi:10.1371/journal.pone.0055882>
- [2] Nirav N. P., et al., (2015), “Multitemporal settlement and population mapping from Landsat using Google Earth Engine, International Journal of Applied Earth Observation and Geoinformation, Vol. 35, Part B, pp. 199-208. <https://doi.org/10.1016/j.jag.2014.09.005>
- [3] Mennis, J., and Hultgren, T., (2006), "Intelligent dasymetric mapping and its application to areal interpolation." Cartography and Geographic Information Science, vol. 33, no. 3, p. 186. Accessed 24 Apr. 2021. <https://www.tandfonline.com/doi/abs/10.1559/152304006779077309>
- [4] Gaughan AE, et al., (2013), “High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015”, PLOS ONE, vol. 8, Issue 2, pp.1-11.<https://doi.org/10.1371/journal.pone.0055882>
- [5] Stevens FR, et al., (2015), “Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data”, PLOS ONE, vol. 10, Issue 2, pp.1-22. <https://doi.org/10.1371/journal.pone.0107042>
- [6] Linard C, et al., (2011), “Assessing the use of global land cover data for guiding large area population distribution modelling”, GeoJournal, vol.76, Issue 5, pp.525-538. <https://doi.org/10.1007/s10708-010-9364-8>
- [7] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). “Google Earth Engine: Planetary-scale geospatial analysis for everyone”, Remote Sensing of Environment. Accessed: May 05, 2021.
- [8] Garcia-Fry, M., Murao, O., (2020), “A Post-disaster Forecast Analysis for Relocation Settlements Using a Combined Model Sequence”, 17th World Conference on Earthquake Engineering, N.C000108.
- [9] Alonso, W. (1964), “Location and Land Use. Toward a General Theory of Land Rent”, Cambridge: Harvard University Press. Accessed: August, 2019. <http://dx.doi.org/10.4159/harvard.9780674730854>
- [10] Krizek, K. J. (2003), “Operationalizing Neighborhood Accessibility for Land Use–Travel Behavior Research and Regional Modeling” *Journal of Planning Education and Research*, Vol. 22, No. 3, pp. 270–287. doi.org/10.1177/0739456X02250315
- [11] “Hyundai H1 Bus – Price, Specs, Fuel Consumption”, Hyundai Vehicles at Group 1 Motors, Accessed: May, 2020. <https://www.group1hyundai.co.za/hyundai-h1-bus/>

Dataset Citations

1. Indonesia, Statistics Sleman Regency, “Population and Population Growth Rate by Subdistrict in Sleman Regency, 2010, 2016, and 2017”, Badan Pusat Statistik Kabupaten Sleman, 12, Oct. 2018. Accessed: May 19, 2020. <https://sleman.kab.bps.go.id/statictable/2018/10/12/363/jumlah-penduduk-dan-laju-pertumbuhan-penduduk-menurut-kecamatan-di-kabupaten-sleman-2010-2016-dan-2017.html>
2. Indonesia, Statistics Sleman Regency, “Population Distribution and Density by Subdistrict in Sleman Regency, 2018”, Badan Pusat Statistik Kabupaten Sleman, 09, Jul. 2019. Accessed: Apr. 19, 2021. <https://sleman.kab.bps.go.id/statictable/2019/07/09/520/distribusi-dan-kepadatan-penduduk-menurut-kecamatan-di-kabupaten-sleman-2018.html>
3. OpenStreetMap (2021) OpenStreetMap The Free Wiki World Map. Open-Street Map contributors, CC BY-SA. Accessed April 15, 2021. <http://www.openstreetmap.org/>

4. Monografi Desa Kepuharjo, 2017 in: Ruswantoro, R., 2018. Evaluation of the Impact of Government Policies on Relocation Post Eruption of Merapi 2010, [Doctoral Dissertation, Sekolah Tinggi Pengembunan Masyarakat Desa "APMD"] APMD Repository, p.44 – Table 5. Access on Apr. 20, 2021.
<http://repo.apmd.ac.id/618/1/repo%20Tesis%20Ruwantoro.pdf>
5. Data Sekunder Laporan Profil Desa Kepuharjo, 2015 in: Wibowo, P., D., 2016. Kebutuhan Pelayanan Dasar Desa Di Daerah Rawan Bencana. PT Sulaksana Watinsa Indonesia, p. 46 – Table 8. Access Apr. 20, 2021.
https://www.academia.edu/40312242/Kebutuhan_Pelayanan_Dasar_Desa_Di_Daerah_Rawan_Bencana
6. Indonesia, Yogyakarta DIY, (2019), “Total Population of Cangkringan district, Sleman district, DI Yogyakarta, Semester II-2019 According to Age Group”, *Data on the Consolidation and Clearance of Population Databases by the Directorate General of Population for Civil Registration of the Ministry of Home Affairs, processed by the DIY Secretariat's Administration Bureau*. Accessed Apr. 26, 2021.
<https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/13/5/17/04/34.ez>
7. Garcia-Fry, Martin (2021), “OLUTM Model Dataset”, Mendeley Data, V3, doi:
<https://data.mendeley.com/datasets/t9p23k3pyn/3>

CHAPTER 6

CONCLUSIONS

6.1 Main Summary

This study clarified relationships between preliminary population (Chapter 3), land cover data (Chapter 4), and resettlement location, as well as mobility and livelihoods (Chapter 5) after the 2011 Great East Japan Earthquake and the 2010 Merapi volcano eruptions to develop a model that will contribute to the formulation of future disaster recovery plans. A resettlement area detection method (method 1) was proposed to help municipal local governments in Sendai's peri-urban areas, Tohoku region, Japan, evaluate the relationship between population growth and urban structures through a geographic lens and detect future resettlement areas likely to undergo population growth. This method combined population and geographic data to identify local drivers of growth in a depopulating region.

In particular, Chapter 3 evaluated settlement location using post-disaster population data for Great East Japan Earthquake affected areas and found no significant change in population trends. It was useful to classify municipal areas by affected population shares to secure reliable urban structures driving population growth. However, geography features were tied to the internal land structures of the region. Therefore, a comparative analysis of peri-urban municipal areas found population growth associated to small municipal areas and distance away from train stations. This is not a common measure of urban growth because it can be used to detect resettlement areas by seeking vehicular accessibility and aggregate built-up structures. As a result, Iwanuma city with 41% of the population severely affected during the tsunami leveraged accessible and interconnected built-up agglomerations surrounded by farmlands are key drivers of population growth. These results can help peri-urban municipal governments of the Tohoku region to identify similar urban structures and detect sustainable resettlement areas in preparation for future disaster events.

Following resettlement area detections in a depopulating context, Chapter 4 evaluated settlement location using land cover data for Merapi volcano affected areas and saw small causal effects on land cover change. This chapter assumed that resettlement sites and new infrastructure triggers land use growth. However, the study found that resettlement sites caused minor land use growth effects (15-18% of the district's total) with no significant trend alterations. Chapter 4 selects Merapi affected areas because urbanizing regions experience a higher intensity of land use growth than depopulating regions. Therefore, a ground cover change method (method 2) was applied in a developing country to improve our understanding of investment effects on land cover change and to effectively locate future resettlement sites next to roads that sustain that growth. In contrast, results can be used as a surplus measure to help local governments in shrinking peri-urban areas locate resettlement sites next to roads that help generate growth. Further, this method serves to monitor recovery and anticipate where land use changes need to be made. Practical applications of this model bridge multiple areas of study in monitoring and evaluating spatially-explicit phenomena.

After finding a resettlement area and estimating the effects of resettlement on land cover change, Chapter 5 surveyed the local road network and selected a candidate resettlement site in Merapi affected areas to evaluate post-disaster mobility data. This chapter found that settlement locations could be improved if livelihood options were located nearby. Therefore, it proposed a land use microsimulation method (method 3) to deliver a predictive network of land use change and verify whether livelihoods were improved. Merapi affected areas were selected to evaluate the method's performance in a dynamic land use change setting. Based on the model's result, it found that middle-aged farmers with upper median incomes and risk-taking behaviors launched home businesses that contribute to rural labor. A 63% share of livelihoods diversified by farmers and 30% travel utility savings improved livelihoods and mobility estimating the settlement's location.

6.2 Key Take Aways

Natural disaster events in the periphery of urban areas poses an increasingly unique opportunity to manage urban growth. While developed countries experience population decline and shrinking cities, developing countries experience unprecedented population growth and rapid urbanization. These trends can only hope to balance or sustain the present state of resource consumption, emissions, or waste. Therefore, this study targets peri-urban areas to incentivize post-disaster land use change as a medium to effectively regenerate population, land cover, and mobility in the aftermath of disaster events.

For this reason, pre-disaster recovery plans that contemplate post-disaster resettlement and recovery needs amid comprehensive urban development plans can successfully manage urban growth. In terms of population, 2011 Great East Japan Earthquake affected areas provided insightful results which contribute to the utility of census data and the necessity of data surveys worldwide. The latter may not be a consequential issue in developed countries currently undergoing issues of population decline, but using these data to the advantage of displaced populations in need of fast, efficient, and long-term recoveries es less understood. Therefore, Chapter 3 is addressed to local governments near Sendai, but may promote finding candidate resettlement areas as a preventive measure in other municipalities at risk. The following take-aways summarize these findings:

- (1)** Pre-disaster population data saw no significant change in trends. Therefore, annual census population data and geographic features can be used to identify resettlement areas that will potentially undergo population growth subject to a well-planned investment.
- (2)** Population growth was observed in interconnected and accessible built-up agglomerations sustained by agricultural land, industrial facilities, and a central train station.
- (3)** Dispersed urban structures exacerbated population decline after the disaster event.

Chapter 4 estimated land cover change prompted by post-disaster resettlement sites after the 2010 Merapi volcano eruptions and saw land use growth at the expense of forest land. In particular,

this chapter emphasizes the need to secure resettlement sites taking the environment into account. For this reason, results indicate that local governments should find candidate resettlement sites next to roads that sustain land use growth. The estimated land use growth values are reliable in the extent of Yogyakarta's region, but can be useful to future disaster resettlement cases near shrinking cities in the Indo-Malayan ecozone. Additionally, other past cases of disaster resettlement can be examined to source new estimates for new parts of the world. The following take-aways were learnt:

- (1)** Resettlement sites caused small effects on land cover change with no significant trend change.
- (2)** Specifically, urban resettlement in rural urbanizing areas availed 18.5% cropland growth and 15.5% built-up growth at the expense of 16.4% forest loss.
- (3)** Increasing land prices have substantial effects on land change, especially cropland to built-up land.
- (4)** Lastly, 33% of photo-interpreted labels may contain interpretation biases.

Chapter 5 evaluated post-disaster mobility input data and livelihood changes in an overall development plan. Socioeconomic improvement was rendered as a result of livelihood options, land use changes, and increased access to transit. In particular, the location of a resettlement site was evaluated to determine whether it was well located and whether people had access to choices of livelihood in ample spheres of life. Results from this chapter are summarized as follows:

- (1)** Short intra-network distances generate mixed-use patterns of development.
- (2)** Risk-oriented people with upper-median incomes and ages-above-50 are likely to entrepreneurial business.
- (3)** A household's primary occupation defines a parcel's tentative land use change subject to a rise in transit.
- (4)** Transit caused by the demand for specialized goods, services, and employment can supersede transit generated through accessibility and urban density.

6.3 Limitations

Methods applied in this study encompass time-consuming steps to attain significant results. Therefore, future applications should consider using methods before disasters. The population method is not meant to resettle populations in rural towns of Miyagi and Iwate prefecture. However, to avoid promoting rural to urban population movement in the future, and to reinforce sub-regional cities in a nationally centralized population trend, targeting rural areas can be useful to identify fine-level urban structures driving growth. Further, these applications may require fine-grained population data instead of small area municipal census data to successfully identify population growth.

This study also assumes that resettlement in depopulating regions will not have a significant impact on land cover change (based on the results). Even though previous research (Ishihara and Tadono 2017; Harada 2015) has shown that disasters do impact land cover class transitions (e.g., from cropland to grassland), here we focus on agriculture and built infrastructure effects on the environment. Another important land cover limitation is that moderate-resolution remote sensing imagery (~30 m) may be available via open data websites (e.g., ESA, NASA), but the temporal resolution of image collections is limited dating back to 2016 (Sentinel-2) and 2013 (Landsat 8/OLI). In addition, this study only considers the extent of Merapi and East Japan affected areas, but it may be useful to other disaster cases in peri-urban areas near cities of relative size to Sendai and Yogyakarta. Lastly, the third method in Chapter 5 makes six simulation assumptions based on theoretical knowledge (see Table 5.9) such as bid-rent theory (Alonso 2013), urban economics (Rhondi et al. 2018), prospect theory (Kahneman and Tversky 1979), and shift-share (Moretti 2011).

6.4 Policy Implications

This study proposes a potentially new urban growth management scheme, termed Agronomic Urban Boundary networks (pp. 103-104), proven to provide travel efficiency and access to employment opportunities while sustaining rural labor as the primary employment and home-

business entrepreneurship as the secondary employment. These networks are envisioned to accommodate public transport on existing roads and agricultural parcels of land in central areas of the network thereby claiming a balance of urban and rural systems. The following policies are considered fundamental to reduce urban intensities and enclose that growth in city-wide areas:

- (1) Promoting intensive agriculture in central areas of the network: Land use should not be limited by interventions that condone imposition of laws and regulations, but rather in a freedom-oriented society, land use subject to a disperse city growth. However, I believe local governments can use land use readjustment tools to incentivize the disaggregation of large real estate assets into small and more productive intensive agricultural land. These subdivisions can naturally occur between household members in the form of premeditated inheritance, neighborhood cooperative associations, or time-sharing-based proposals between local farmers and institutions interested in the dissemination of agronomic education and widespread development.
- (2) To help transition the urban intensity from urban to rural landscapes: The rural periphery requires a new administrative conception and a constitution to accompany that concept which addresses the importance of land management and rural preservation through its legal tools. Further, the whole idea behind a transitional “band” is the mitigation of informal land use dispersion and uncontrolled resource consumption in a single geographical area.
- (3) Food security: In addition to the obvious need of food systems in the periphery of cities, healthy soil is a major issue nowadays and constitutes a paramount health problem for society. This too must be regulated to the extent of its use and lifecycle. Crop cycles is another requirement which should never end in use of toxic chemicals or synthetic fertilizers and should rather swap crops for trees and vice versa in a continued effort to sustain soil nutrients.
- (4) Decarbonization: Large agronomic urban boundary networks might also pose opportunities for forest conservation initiatives with policy implications and reforestation programs that aim to restore disaster-affected land, absorb future hazard shocks, and expand over unused land to

expand the carbon sinks in the periphery of cities. These programs not only serve primary objectives, but also constitutes an ample source of foods, medicines, and building material.

(5) Tree Farming: This activity should be regulated to provide victims and local communities a new source of rural labor for paper, wood, and biopolymer industries. In fact, small and medium sized light industries should be located next to these managed lands in order to develop sustainable chains on-site and avoid transportation issues. Further, these industries would need to abide strict regulations focused on ecological footprint, emissions, and waste management as well as strict adherence to local culture, religion, and other less quantitative motions at play. The system's performance should envision rural identity as the mission.

6.5 Concluding remarks

This study evaluated post-disaster data and found no significant change in pre-disaster trends. Therefore, it can be applied in areas that have not experienced past disaster resettlement. However, the application of methods rely on annual population census data for the past ten years and lessons learned from past cases of resettlement in the ecozone of interest to apply land cover change results. Another important source of data are mobility profiles, which can be collected in the target area using a survey questionnaire attached to this study (Appendix 1). Local governments should conduct the questionnaire survey after permanent displacement to register all households that will relocate in the government's jurisdiction area. They need to survey 100% of the people that will relocate, including those who lost their homes and choose to reconstruct by themselves near resettlement sites. If the survey is conducted for research purposes, then at least 8-10% of the target households should answer the questionnaire to develop a synthetic dataset for mobility assessments.

Past cases of resettlement can bring uncertainty, but the results obtained state that only 15-18% of the total land changes were a product of resettlement. To handle this uncertainty, this study suggests that only local roads able to sustain more than the expected land use growth are surveyed as

a safety measure to avoid land use dispersion. The model proposed is for municipal local governments in need of identifying where future resettlement is likely to enable a successful recovery of the population and the affected area. It is also addressed to urban planners in private and public institutions who may act as consultants to the recovery planning efforts. A user's guide is attached (Appendix 2) to facilitate steps taken to conduct this study. This dissertation marks a major step towards a sustainable and resilient future.

References

- Ishihara, M., Tadono, T., 2017. Land cover changes induced by the great east Japan earthquake in 2011. *Sci Rep* 7, 45769. <https://doi.org/10.1038/srep45769>
- Harada, I., Hara, K., Park, J., Asanuma, I., Tomita, M., Hasegawa, D., Short, K., Fujihara, M., 2015. Monitoring of rapid land cover changes in eastern Japan using Terra/Modis data. Proceedings of the 36th Int. Symp. Remote Sens. Environ., 11–15, Berlin, Germany, ISRSE 36-541-6. <https://meetingorganizer.copernicus.org/ISRSE36/ISRSE36-541-6.pdf>
- Alonso, W., 2013. Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press, Cambridge, MA and London. Accessed: August, 2019. <http://dx.doi.org/10.4159/harvard.9780674730854>
- Rondhi, M., Pratiwi, P.A., Handini, V.T., Sunartomo, A.F., Budiman, S.A., 2018. Agricultural Land Conversion, Land Economic Value, and Sustainable Agriculture: A Case Study in East Java, Indonesia. *Land* 7(4), 148-167. <https://doi.org/10.3390/land7040148>
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk, *Econometrica* 47 (2), 263–292. <https://doi.org/10.2307/1914185>
- Moretti, E., 2010. Local Multipliers. *Amer. Econ. Rev: Papers and Proceedings* 100 (2), 1-7. <https://eml.berkeley.edu/~moretti/multipliers.pdf>

【Appendix 1】

IMPORTANT [Survey for Adults +18 years old]

1. HOUSEHOLD No.: RESIDENT No.:

1.6- HOME ADDRESS (Locate on map)

1.1- NUMBER OF RESIDENTS (per household):

1.7- LIVING PREFERENCES (Please choose one from each column):

1.2- PERSONAL INFORMATION

Housing Neighborhood

Age:

A- Attractiveness A- Crime Rate

Gender (Male, Female):

B- Space Supply B- Green Areas Supply

Marital Status (Single, Married):

C- Building State C- Surroundings

Vehicle Ownership (Type, How many):

D- Buy / Rent a House D- Social neighborhood

Approximate Income per month:

E- Centrality (near Job)

1.3- PRIMARY EMPLOYMENT:

F- Education facility

1.4- SECONDARY EMPLOYMENT (if any):

1.5- TRAVEL SCHEDULE (24-Hour itinerary of a common weekday):

Example				
Number	Time	Activity	Location (Exact Place / Map Location / Address / Coordinates)	Travel Mode (how?)
1	7:30 am	Work	"B" District, in front of market, street "...", no. 25-5	bus
2	12:00 pm	Lunch	"C" District, restaurant, 1 block north of work	walk
3	5:00 pm	Friend's house	"E" District, street "...", no. 10-2	taxi
4	7:00 pm	Home	"A" District, street "...", no. 109-1	bus
Weekday Travel Schedule				
Number	Time	Activity	Location (Exact Place / Map Location / Address / Coordinates)	Travel Mode (how?)

2. LIFESTYLES (ARELLANO online: <https://www.arellano.pe/estilos-de-vida/los-estilos-de-vida-en-latinoamerica/>)

2.1- What is your main occupation now?

- (A) Dependent worker (I work for a company, I like to buy the latest technology products, and I am not the owner)
- (B) Independent worker (I have my own business or company)
- (C) my home (I dedicate myself to my home most of the time)
- (D) Student (I study most of the time)
- (E) Eventful (I work from time to time)
- (F) Pension
- (G) unemployed (more than 6 months)
- (H) live from rents / receive donations

2.2- What is the last degree of study you reached?

- (A) primary school
- (B) incomplete middle school
- (C) complete middle school
- (D) incomplete high school
- (E) complete high school
- (F) short career (less than 2 years)
- (G) higher technical career incomplete
- (H) higher technical career complete (more than 3 years)
- (I) superior university career incomplete
- (J) superior university career complete
- (K) master's degree (2 years)
- (L) doctoral degree (PhD)
- (M) Post bachelor courses (specialized courses)
- (N) none (didn't study)

2.3- Now, how do you perceive?

Your personal laboral situation:

- (A) 0 (very bad)
- (B) 1
- (C) 2
- (D) 3
- (E) 4
- (F) 5
- (G) 6
- (H) 7 (very good)

Personal economic situation:

- (A) 0 (very bad)
- (B) 1
- (C) 2
- (D) 3
- (E) 4
- (F) 5
- (G) 6
- (H) 7 (very good)

2.4- How much do you agree or disagree with the following phrases?

I like to buy the latest technologies (...)

- (A) 3 (strongly disagree)
- (B) 2 (very much disagree)
- (C) 1 (a little disagree)
- (D) 1 (a little agree)
- (E) 2 (very much agree)
- (F) 3 (strongly agree)

I usually buy products at lower prices (...)

- (A) 3 (strongly disagree)
- (B) 2 (very much disagree)
- (C) 1 (a little disagree)
- (D) 1 (a little agree)
- (E) 2 (very much agree)
- (F) 3 (strongly agree)

2.5- According to the next phrases: how often do you?

Keep up with the latest fashion (...)

- (A) Never
- (B) A few times
- (C) Sometimes yes, sometimes no
- (D) Almost always
- (E) Always

Blame your personal situation to the government

- (A) Never
- (B) A few times
- (C) Sometimes yes, sometimes no
- (D) Almost always
- (E) Always

Participate in promotions or in offers

- (A) Never
- (B) A few times
- (C) Sometimes yes, sometimes no
- (D) Almost always
- (E) Alwa

LIFESTYLES (VALS online: <http://www.strategicbusinessinsights.com/vals/presurvey.shtml>)

1. I am often interested in theories.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

2. I often ask people's advice about clothes, vacations and other spending decisions.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

3. A good negotiator doesn't just get food in the bowl, but the bowl itself.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

4. I love to make things I can use every day.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

5. I follow the latest trends and fashions.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

6. Just as the Bible says, the world literally was created in six days.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

7. I like being in charge of a group.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

8. I like to learn about art, culture, and history.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

9. I often crave excitement.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

10. I am really interested in only a few things.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

11. I would rather make something than buy it.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

12. I dress more fashionably than most people.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

13. The federal government should encourage prayers in public schools.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

14. I have more ability than most people.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

15. I consider myself an intellectual.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

16. Only a fool gives more than they get.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

17. I like trying new things.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

18. I am very interested in how mechanical things, such as engines, work.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

19. I like to dress in the latest fashions.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

20. There is too much sex on television today.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

21. I like to lead others.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

22. People who think too much annoy me.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

23. I like a lot of excitement in my life.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

24. I must admit that my interests are somewhat narrow and limited.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

25. I like making things of wood, metal, or other such material.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

26. I want to be considered fashionable.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

27. Religion is the most important way to know what's morally correct.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

28. I like the challenge of doing something I have never done before.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

29. No matter how much evil I see in the world, my faith in God is strong.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

30. I like to make things with my hands.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

31. I am always looking for a thrill.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

32. I like doing things that are new and different.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

33. I like to look through hardware or automotive stores.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

34. I would like to understand more about how the universe works.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

35. I need to get the news everyday.

- Mostly disagree
- Somewhat disagree
- Somewhat agree
- Mostly agree

36. Sex:

- Male
- Female

37. Age:

- 18–24
- 25–29
- 30–34
- 35–44
- 45–54
- 55–64
- 65–74
- 75 or older

38. Marital Situation/History:

- Never married
- Now married
- Engaged
- Separated or divorced
- Widowed

39. What is the highest level of formal education you have completed?

- Grade 8 or less
- Grades 9–11
- High School or equivalent
- 1 to 3 years of college or technical school
- College graduation (4 years)
- Attended or completed graduate school

40. What was your total household income before taxes for the most recent calendar year (January through December)? By *your household*, we mean all persons living in your primary home who share basic finances with you.

(Please include income received by all members of your household and from all sources, including salaries, pensions, interest, dividends, bonuses, capital gains, and profits.)

- less than \$5,000
- \$5,000 – \$9,999
- \$10,000 – \$14,999
- \$15,000 – \$19,999
- \$20,000 – \$24,999
- \$25,000 – \$29,999
- \$30,000 – \$34,999
- \$35,000 – \$39,999
- \$40,000 – \$44,999
- \$45,000 – \$49,999
- \$50,000 – \$59,999
- \$60,000 – \$74,999
- \$75,000 – \$99,999
- \$100,000 – \$149,999
- \$150,000 – \$199,999
- \$200,000 – \$249,999
- \$250,000 or more

41. Are you a student currently enrolled at an institute or university?

- Yes
- No

Appendix 2

User's Guide

Contents

Section I: How to source regional metrics for automated data supervision?.....	171
Section II: How to collect candidate training samples?.....	174
Section III: How to perform sub-models of the OLUTM model?.....	176

Section I

How to source regional metrics for automated data supervision?

Section I provides a detailed sequence of steps to source regionally-consistent metrics for automatic supervision of candidate training samples. The objective is to retain high-quality samples for land cover classification on-demand. Metric sets correspond to pre-defined regions in an ecozone of interest. Section II will provide a detailed set of steps to collect candidate training samples. Existing regional metrics, mapping products, and regional classification models can be accessed through Mendeley Repository v.4 <https://data.mendeley.com/datasets/mzp3k6fmtz>.

1. Extract a spaceborne ICESat-2 ground-track acquisition over a wide and heterogeneous area of interest:
 - a. Extraction: North (latitude_20_m) from upper terrain/land end to East (longitude_20_m) lower terrain/land end. Write id's North to South to avoid confusion.
 - b. Copy/paste columns North, one beneath the other in a single column.
 - c. Copy/paste columns East, one beneath the other in a single column.
 - d. No sorting, no erasing.
 - e. Extract LiDAR Data: Canopy_H_20m from upper North to lower East end ids.
 - f. Extract LiDAR Data: Canopy_Mean from upper to lower end fids.
2. Save LiDAR data to .csv file.
3. Erase LiDAR Data with 3.4E** values for: Canopy_Mean, Canopy_H_20m, Longitude, and Latitude.
4. Upload point data to GIS and upload a pre-existing land cover product.
5. Clean LiDAR point data to avoid autocorrelation problems: Eliminate points landing ≥ 45 m away from inter-class perimeters (perimeters = class borders).
6. Sample the land cover product and retain labels.
7. Upload Google Earth Satellite.cn to manually inspect landings over the land cover product and inspect whether >50% of the pixel corresponds the label. Only consider agreement for labels with >50% of the pixel covered by that class.
8. Upload a forest height product.
9. Sample the forest height product.
10. Then, use the following filters to refine ‘agreed/disagreed’ interpretations allotting a binary value for each filter:
 - a. Class-Height thresholds (see step 11 below) to filter the vertical structure of each class using ICESat-2 canopy_mean values, or h_canopy_20 values where no canopy_mean values exist.
 - b. Use forest height values to validate ICESat-2 canopy heights in range of +/- 6 m.
11. A variation of discrete heights sourced from FAO (https://www.fao.org/3/x0596e/x0596e01f.htm#p665_54535) were used; (1) Dense short vegetation (shrubs, percentage of bare areas were available in the land cover product, but more than 95% of our samples returned a 0% bare cover. Therefore, we did not factor in this measure) ≤ 5 m, (2) Open Tree Cover (coniferous, broadleaf, mixed woody plants, wetlands) > 5 m and < 12 m, (3) Dense Tree Cover (closed dry and closed wetland) > 12 m, (4) built up (this class was vetted on the basis that most structures in rural areas have up to 2 levels and pixels are combined with canopy cover) < 8 m, (5) water (open water bodies) = 0 m, and (6) cropland (cultivated and managed land) < 5 m for agriculture and > 5 meters for managed forests.
12. Create a new column and add the three columns with 0's or 1's (thematic agreement column, class height agreement column, and the lidar vetting column).
13. Extract level 3 samples from the dataset. L3 having three 1's and L0 having zero 1's.
14. Erase level 0 and level 1 to eliminate spurious data units.
15. Retrieve level 3 samples distributed every 500 m to validate mapping outcomes and save for later.
16. Create a new EXCEL file and paste the prevailing level 3 samples.

17. Add level 2 samples with positive thematic agreements and positive class height supervisions to the new EXCEL file. Discard the rest of samples.
18. Name this file ‘metrics dataset.csv’.
19. Upload .csv file to GIS and save as .shp file.
20. Upload the .shp file to Google Earth Engine in assets folder. GEE Access: (https://code.earthengine.google.com/?accept_repo=users/martingarciafry/gf).
21. Sample multitemporal data in May, June, July, and August: Verify target year in a classification region of interest and then run the code (alternatively use your own code) to sample atmospherically-corrected and cloud-masked Sentinel-2 mosaics.
22. Sample raster bands during: (i) ‘Cultivation Season’ (March/April), (ii) ‘growing season_1’ (May/June), (iii) ‘growing season_2’ (July/August), and (iv) ‘harvest season’ (October).
23. Export and then download multitemporal datasets from Google Drive.
24. Stack these datasets using each point’s corresponding id value. Note that repeated Lon/Lat values will be displayed.
25. Upload the updated metric dataset as .csv file to GIS.
26. Append ancillary data with nearest neighbor sampling. Here, I used ‘sample raster values’ to append ancillary data to each sample unit:
 - a. First, sample DEM elevation data (Ground height) for choice of DEM product. Here, we used the MERIT DEM (http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/).
 - b. Second, sample Topographic Wetness Index (TWI). Here we used: Geomorpho90m, a MERIT DEM derived data product to measure topographic-related metrics. (<http://www.spatial-ecology.net/dokuwiki/doku.php?id=topovar90m>).
 - c. Third, append the Forest Height data product: we used the Global Forest Cover Change product with height metrics in Band 1 (<https://glad.umd.edu/dataset/gedi>).
 - d. ‘Raster>Analysis>Aspect’ is estimated using the DEM model in TIFF format (~90m). (Tutorial: <https://www.youtube.com/watch?v=B-5RQ9o9EyU>).
 - e. ‘Raster>Terrain Analysis>Slope’. Or use GDAL>Slope (Tutorial: <https://www.youtube.com/watch?v=7eIFvZ4fU6k>). Slope is also estimated using the DEM model in QGIS with TIFF format (~90m). We used a WGS 8 UTM CRS projected from the default CRS [4326].
 - f. Fourth, sample the topographic solar radiation layer estimated with ‘GRASS>r.sun.insoltime’ for the Day 200 of a target year close to July (due to the apparent lack of clouds while seeking Sentinel-2 images), using the Elevation layer, Aspect layer, and Slope layer in QGIS 3.21. (Tutorial: <https://www.youtube.com/watch?v=0z2trThOYaQ>).
27. Calculate the mean temporal surface reflectance value corresponding to each point in the metrics dataset using ‘QGIS>raster calculator>mean’, or using formulas in EXCEL. The appended data has static temporal values that will be used later.
28. Estimate the following indices for land surface phenology stages (Use B8 with 10 m bands (i.e., B2, B3, and B4) and use the B8A with near- and short-wave infrared bands (i.e., B5, B6, B11, and B12):
 - a. NIR_Green (B8-B3/B8+B3).
 - b. Normalized Difference Vegetation Index (B8-B4/B8+B4).
 - c. SWIR_1_SWIR_2 (B11-B12/B11+B12).
 - d. Tasseled Cap Greenness (-0.3599*B2-0.3555*B3-0.4734*B4+0.6633*B8-0.0087*B11-0.2856*B12).
 - e. Tasseled Cap Wetness (0.2578*B2+0.2305*B3+0.0883*B4+0.1071*B8-0.7611*B11-0.5308*B12).
 - f. Normalized Difference water index_I (B3-B8/B3+B8).
 - g. Normalized Difference water index_II (B8A-B11/B8A+B11).
 - h. BN (B2-B11/B2+B11)
 - i. SAVI (B8-B4/B8+B4+0.5) x1.5
 - j. Finally, using a moving-average time-series of the NDVI column for Cropland samples, estimate the first, third, and last quartile for NDVI, Band 3 (Blue), Band 8 (NIR), and Band 11 (SWIR 1) to classify periods of change.

29. In EXCEL, regressions will reveal which features (static or temporal) are sensitive to the independent variables of choice. We followed literature guidelines to select representative features for treed classes (NDVI), wetlands (SAVI), Built-up areas (BN), water (NDWI_I), and cropland (SAVI). Upper 95th percentile sensitive features should be selected as metric contenders (5-8 metrics per class).
30. After selecting features, proceed to extract data bounds using ‘Data Analysis>Descriptive Statistics’ and obtain the max, min, and standard error for each feature per class.
31. Use the standard error to calculate the min/max range values (min. + st.e., max – st.e.).
32. These become the official MEAN metrics for supervising training data with a filtering tool, available in the Mendeley Repository: <https://doi.org/10.17632/mzp3k6fmtz.4>.
33. We recommend using multitemporal raw metrics (avoiding Step 27) because labeled agreements are well supervised with fused canopy heights.

For reference, the filtering tool uses a simple Python programming language with the imported Numpy library to classify candidate training units. This tool supervises class-segmented datasets and saves a compact classified dataset for user-based discriminations. All output files in .csv format are saved locally in the same location where the tool was saved. Access the Filtering tool here:

https://github.com/martingarciafry/GCC_model

Section II

How to collect candidate training samples?

Section II provides a sequence of detailed steps to collect candidate training samples for land cover classification models. The resulting dataset is used to filter unwanted, unreliable, or spurious data samples. Existing regional metrics, mapping products, and land cover models can be accessed through our Mendeley Repository v.4 <https://data.mendeley.com/datasets/mzp3k6fmtz>.

1. For year ‘x’ and region ‘y’: download ICESat-2 ATLAS08 Spaceborne LiDAR data:
 - a) Extraction: North (latitude_20_m) from upper terrain/land end to lower terrain/land end.
Copy/paste columns one beneath the other in a single column.
 - b) Extraction: East (longitude_20_m) from fid (North) to fid (South).
 - c) No sorting, no erasing.
 - d) Extract LiDAR Data: Canopy_H_20m from upper to lower end ids.
 - e) Extract LiDAR Data: Canopy_Mean from upper to lower end ids.
 - f) (Optional) Extract LiDAR Data: Slope from upper to lower end ids.
2. Save LiDAR data to CSV file.
3. Erase LiDAR Data with 3.4E** values for: Canopy_Mean, Canopy_H_20m, Longitude, and Latitude.
4. Repeat steps 1-3 for each strong beam Ground Track (GT) (upon desired training data size). Here, we used one GT as metric samples and pre-processed eighteen GTs to calibrate training data samples.
5. Merge Data: all training samples in a single .csv file.
6. Upload .csv file to GIS: Save as point .shp file.
7. Upload land cover product to GIS.
8. Clean LiDAR point data to avoid autocorrelation problems; erase points that are ≥ 45 meter away from class perimeter (perimeter= inter-class border).
9. Append and sample the land cover label classifications in band 1.
(<https://glad.umd.edu/dataset/global-land-cover-land-use-v1>)
10. Save as SHP file and upload the file to Google Earth Engine in assets folder. GEE Access:
(https://code.earthengine.google.com/?accept_repo=users/martingarciafry/gf).
11. Run the code for a target year to sample atmospherically-corrected and cloud-masked Sentinel-2 image collections.
12. Sample raster bands during: (i) ‘Cultivation Season’ (March/April), (ii) ‘growing season_1’ (May/June), (iii) ‘growing season_2’ (July/August), and (iv) ‘harvest season’ (October) of a target year.
13. Download multitemporal datasets from Google Drive and then stack them using each point’s corresponding id value. Note that repeated Lon/Lat values will be displayed.
14. Once a compact dataset is compiled, add a ‘class code’ column to identify land cover labels with own values (1 – 7). We used: Dense Short Vegetation (1), Open Tree Cover (2), Dense Tree Cover (3), Wetland (4), Built up (5), Water (6), and Cropland (7) based on a desired study of land surface phenology and in accordance with FAO’s land classifications.
15. Now, upload ancillary data to QGIS: MERIT Digital Elevation Model, Geomorpho90m - Topographic Wetness Index, and the Global Forest Cover Height product.
16. Append ancillary data with nearest neighbor sampling: Here we used ‘sample raster values’ to append data for each sample unit:
 - a) First, sample DEM elevation data (Ground height) for choice of DEM product. Here, we used the MERIT DEM (http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/).

- b) Second, sample Topographic Wetness Index (TWI). Here we used: Geomorpho90m, a MERIT DEM derived data product to measure topographic-related metrics. (<http://www.spatial-ecology.net/dokuwiki/doku.php?id=topovar90m>).
 - c) Third, append the Forest Height data product: we used the Global Forest Cover Change product with height metrics in Band 1 (<https://glad.umd.edu/dataset/gedi/>).
 - d) ‘Raster>Analysis>Aspect’ is estimated using the DEM model in TIFF format (~90m). (Tutorial: <https://www.youtube.com/watch?v=B-5RQ9o9EyU>).
 - e) ‘Raster>Terrain Analysis>Slope’. Or use GDAL>Slope (Tutorial: <https://www.youtube.com/watch?v=7eIFvZ4fU6k>). Slope is also estimated using the DEM model in GIS with TIFF format (~90m). We used a WGS 8 UTM CRS projected from the default CRS [4326].
 - f) Fourth, sample the topographic solar radiation layer estimated with ‘GRASS>r.sun.insoltime’ for the Day 200 of a target year close to July (due to the apparent lack of clouds while seeking Sentinel-2 images), using the Elevation layer, Aspect layer, and Slope layer in QGIS 3.21. (Tutorial: <https://www.youtube.com/watch?v=0z2trThOYaQ>).
17. Save LiDAR point data with appended ancillary data as .csv file with ids.
18. Estimate the following indices for land surface phenology stages (Use B8 with 10 m bands (i.e., B2, B3, and B4) and use the B8A with near- and short-wave infrared bands (i.e., B5, B6, B11, and B12):
- k. NIR_Green (B8-B3/B8+B3).
 - l. Normalized Difference Vegetation Index (B8-B4/B8+B4).
 - m. SWIR_1_SWIR_2 (B11-B12/B11+B12).
 - n. Tasseled Cap Greenness (-0.3599*B2-0.3555*B3-0.4734*B4+0.6633*B8-0.0087*B11-0.2856*B12).
 - o. Tasseled Cap Wetness (0.2578*B2+0.2305*B3+0.0883*B4+0.1071*B8-0.7611*B11-0.5308*B12).
 - p. Normalized Difference water index_I (B3-B8/B3+B8).
 - q. Normalized Difference water index_II (B8A-B11/B8A+B11).
 - r. BN (B2-B11/B2+B11)
 - s. SAVI (B8-B4/B8+B4+0.5) x1.5
 - t. Finally, using a moving-average time-series of the NDVI column, estimate the first, third, and last quartile for NDVI, Band 3 (Blue), Band 8 (NIR), and Band 11 (SWIR 1) to classify cropland periods of change.
19. Three important quality filters were used to classify data (A filtering tool can be found in our Mendeley repository to automate this part:
- a) First, spectral metric upper/lower bounds using band combinations and static ancillary data for manually-agreed labels only.
 - b) Second, class height thresholds are determined to filter each class unit using ICESat-2 values corresponding to canopy_mean, or h_canopy_20 where no canopy_mean values exist.
 - c) Third, Forest height values are used to validate ICESat-2 canopy heights if these are within a range of 6 m.
20. A variation of discrete heights sourced from FAO (https://www.fao.org/3/x0596e/x0596e01f.htm#p665_54535) were used; (1) Dense short vegetation (shrubs, percentage of bare areas were available in the land cover product, but more than 95% of our samples returned a 0% bare cover. Therefore, we did not factor in this measure) ≤ 5 m, (2) Open Tree Cover (coniferous, broadleaf, mixed woody plants, wetlands) > 5 m and < 12 m, (3) Dense Tree Cover (closed dry and closed wetland) > 12 m, (4) built up (this class was vetted on the basis that most structures in rural areas have up to 2 levels and pixels are combined with canopy cover) < 8 m, (5) water (open water bodies) = 0 m, and (6) cropland (cultivated and managed land) < 5 m for agriculture and > 5 meters for managed forests.
21. Once the dataset has been classified, exclude the least reliable labels (level 0 and level 1) to ascertain a minimum degree of quality training labels.
22. Finally, use the provided training data set template.csv to eliminate unnecessary data columns and export the .csv file to a .geojson file for model training.
23. Access the model here: https://github.com/martingarciafy/GCC_model

Section III

How to perform sub-models of the OLUTM model?

Section III provides a detailed series of steps to perform sub-models of the OLUTM model. We encourage users to combine this guide with supplementary information provided in the main text. Future applications of the OLUTM model should consider understanding the basic concepts behind each of these steps. Further, additional information can be found in our shared repository in Mendeley Data: <https://data.mendeley.com/datasets/t9p23k3pyn/3>

1. Household aggregation method

1.1 Classifying Survey Data (Daily Travel Schedules)

1.1.1. Jobs/Expertise Locations: Each person is assigned with a land-use state according to their job/expertise from the survey and can be abbreviated as follows;

CBD – Job in a Central Business District
C – Commercial Job
SI – Job in a Small Industry
LI – Job in a Large Industry
E – Job in Education

1.1.2. The number of “states” is a function of the number of occupants which can be calculated with equation:

$$n_{householdstates} = \binom{n+r-1}{r} = \frac{(n+r-1)!}{r!(n-1)!}$$

Where “r” is the number of occupants and “n” is the number of states which can be related to a single household of singles and couples.

- | | | | |
|-----------|-----------|------------|-------------|
| 1) CBD+C | 2) CBD+SI | 3) CBD+E | 4) LI+C |
| 5) C+SI | 6) C+E | 7) SI+E | 8) SI+LI |
| 9) LI | 10) E | 11) SI | 12) E+LI |
| 13) C | 14) CBD | 15) CBD+LI | 16) C+C |
| 17) LI+LI | 18) SI+SI | 19) E+E | 20) CBD+CBD |

There are twenty (20) states which can be associated with a single individual, or couple.

1.1.3. Households are composed of 1,2,3....n working adults, which affect the number of daily trips.

Each household is assigned with one of the twenty states and the number (1 - n) is added to households with more than 2 adults as identifiers.

Land use compatible families are arranged into groups to organize and determine the proximity of the group to a desired polar growth node as shown in the next figure:

CBD+CBD	CBD+C	CBD+CBD	CBD+C	CBD+CBD
E+E	E+C	CBD+C	C+S	S+S
E+E	E+C	C+C	C+S	S+S
E+E	E+C	C+L	C+S	S+S
LI+LI	C+C	LI+LI	C+S	LI+LI
Commercial Compatible Families				
CBD	SI	E	Compatible Singles	
Li+Li	Si+Li	Li+Li	C+Li	E+Li
Si+Si	Si+E	Li+E	C+E	C+C
Si+Si	Si+E	E+E	C+E	C+C
Si+Si	Si+E	CBD+E	CBD+E	C+C
CBD+SI	CBD+E	CBD+CBD	CBD+C	CBD+CBD
Education Compatible Families				
SI	CBD	C	Compatible Singles	
Si+Si	Si+C	C+C	C+C	Li+Li
Si+Li	Li+C	Li+C	Li+C	Li+CBD
Si+Si	Li+Si	Li+Li	Li+CBD	CBD+CBD
Si+E	Li+Si	Li+E	Li+E	Li+CBD
E+E	Li+E	E+E	Li+Li	E+E
Large Industry Compatible Families				
C	E	SI	Compatible Singles	
C+C	Si+C	Si+Si	Li+Li	Li+Si
CBD+C	CBD+Si	CBD+Si	CBD+Si	CBD+Li
C+C	CBD+C	CBD+CBD	CBD+Li	Li+Li
E+C	CBD+C	CBD+E	CBD+E	CBD+Li
C+C	CBD+E	E+E	E+E	E+Li
Small Industry Compatible Families				
E	C	CBD	Compatible Singles	
C+C	Si+C	Si+Si	Li+Li	Li+Si
CBD+C	CBD+Si	CBD+Si	CBD+Si	CBD+Li
C+C	CBD+C	CBD+CBD	CBD+Li	Li+Li
E+C	CBD+C	CBD+E	CBD+E	CBD+Li
C+C	CBD+E	E+E	E+E	E+Li
Central Business District Compatible Families				
C	E	LI	Compatible Singles	

*Each household group is defined by a square of four alternatives within each colored land use block. These admit compatible singles whose land use match the ones already defined within a group.

1.1.4. Classifying survey data (Lifestyle Segmentation Data)

1.1.4^a. From Lifestyle segmentation

Each person has been assigned with a lifestyle segment according to the initial survey questionnaire using VALS™ or “Lifestyles” by Arellano:

VALS (*Developed Countries*)

1. Innovators (I) 2. Thinkers (T) 3. Believers (B) 4. Achievers (A) 5. Strivers (S)
6. Experiencers (E) 7. Makers (M) 8. Survivors (S2)

Arellano Marketing (*Developing Countries*)

1. Sophisticated (S) 2. Progresist (P) 3. Modern (M) 4. Formalist (F) 5. Conservative (C)
6. Austere (A)

1.1.4^b. The number of “states” is a function of the number of adult occupants per household.

In a community-based neighborhood unit, a cluster of houses that meet each other’s needs and engage in proactive activities consists in no more than 30 houses for a 100 x 100 meter neighborhood area (丁目).

Results from the survey in Appendix 1 can provide household (HH) compositions:
(30 houses x 0.6; 1-adult HH’s) + 2*(30 x 0.3; 2-adult HH’s) + 3*(30 x 0.1; +3-adult HH’s)
=100%, 45 Adults in total
30 HH’s, 10,000 sqm, 180 sqm / HH, 5,400 sqm built area, 4,600 sqm unbuilt area (46%)

VALS:

1 Adult HH: $30 \times 0.6 = 18$ H.H.

“Lifestyles” Arellano:

1 Adult HH: $30 \times 0.3 = 9$ H.H.

2 Adult HH: $30 \times 0.3 = 9$ H.H.
3 Adult HH: $30 \times 0.1 = 3$ H.H.

2 Adult HH: $30 \times 0.6 = 18$ H.H.
3 Adult HH: $30 \times 0.1 = 3$ H.H.

Each state is assigned with a weight: 0.6, 0.3, 0.1 based on a 1 adult, 2 adults, +3 adult household, respectively. Neighborhoods ideally consists of diversified lifestyles with weight ratios for each according to national samples from same or similar surveys.

1.1.5. Make clusters with 18 (1 Adult) HH, 9 (2 Adult) HH and 3 (+3 Adults) HH using the weight ratios assigned to each segmentation in the cluster (a thorough explanation of this process is available in [Garcia and Murao 2020](#)):

To capture the weight of a given segmentation, the total amount of adult segment groups represents a percentage of the total number of adults filling out the survey. Each percentage is further multiplied by the total number of adults in each cluster of 30 HH's to obtain the ratio.

Example, according to US National Data Survey in 2017:

(I): $9\% \times 45 = 4$ Adults	(T): $15\% \times 45 = 7$ Adults	(B): $12\% \times 45 = 5$ Adults
(A): $13\% \times 45 = 6$ Adults	(S): $13\% \times 45 = 6$ Adults	(E): $13\% \times 45 = 6$ Adults
(M): $11\% \times 45 = 5$ Adults	(Su): $14\% \times 45 = 6$ Adults	Total = 45 Adults

For the US, a cluster of 45 Adults equivalent to 1 neighborhood cluster should contain:

18 singles: 18	9 Couples: 18	3 Triplets: 9 = 45 Adults
----------------	---------------	----------------------------------

Segment groups:

(4) Innovators, (7) Thinkers, (5) Believers, (6) Achievers, (6) Strivers, (6) Experiencers, (5) Makers, (6) Survivors = **45 Adults**

Source: <http://www.strategicbusinessinsights.com/vals/free/2018-04valsbystate.shtml>

2. Population growth Forecast

2.1. The annual census population count for the last 10 years in the smallest municipal area is collected using statistical data from the local government.

2.2. Population growth rates are estimated using UN Population growth rate indicator system: “*The rate of population growth, r, between two time points, t1 and t2, is calculated as an exponential rate of growth, conventionally expressed in percentage units per year: $r = 100 \ln(P2/P1)/(t2-t1)$* ”.

Source: https://www.un.org/esa/sustdev/natinfo/indicators/methodology_sheets/demographics/population_growth_rate.pdf

2.3. Once the population growth (%) is available, project the population ten years from your t1 count.

2.4. The density for the interest district (people / km²) is also estimated.

3. Dasymetric Redistribution

3.1. Chapter 5, Supplementary information, provides a detailed sequence of steps necessary to obtain the population density of a target area inside a local municipal district. The total population count (census value) is redistributed across an area (administrative/census unit) asymmetrically using ancillary data, such as land

cover and building footprint area, to allocate unequal weights to each destination pixel at a target spatial resolution. Pixels can be formed aggregating land cover samples from a 30 m spatial resolution to a 1 km² spatial resolution

4. Resettlement Site Selection

This step is the most important step so far and requires a multidisciplinary view on urban and regional planning.

4.1 Since land is required, an environmental and vulnerability assessment is prior to candidate settlement selection.

According to laws abiding for land ownership for such cases, or land already separated in advance for catastrophic events (best case scenario), governments have already assessed a preliminary hazard map involving weather and ground compositional fluctuations using historical data.

4.2 Another important prerequisite for long-term development is to select a site annexed to the city's infrastructure via local and arterial roads. Also and preferably, in between the city (as the main polar node) and a regional pole. This does not mean to locate a site in hazard prone areas, but rather safely in-between two poles of constant stock flow exchange. Such exchange is known as economic corridors. These corridors traditionally nurture new urban development.

4.3. The last requisite for new siting or land tenure in preparation for future resettlement is the necessity for expansion which is required on a long term basis. This is a matter of utmost importance to the benefactors and beneficiaries because urban land use expansion can detriment an area or flourish healthy progress and development. To this end, Chapter 4 estimates land use growth as a product of urban resettlement in rural areas.

5. Surveying the Local Road Network

Now that a candidate resettlement site has been selected with all the appropriate requisites, a road network analysis becomes the first step to estimating the effectiveness of a site. Even semi-urban infrastructure requires this analysis. Also, it is assumed that a hosting city is present at the regional scale and therefore, node analysis is necessary to lay down a data-informed recovery plan.

5.1. Density

A radius from the site's centroid to the farthest destination within the activity-based schedule survey is drafted and named hereinafter the *Area of Interest (AOI)*.

5.1.1. Within the AOI, every economically important node and/or pole is identified.

According to graph theory, nodes are points or intersections in a network where lines or pathways intersect or branch. In transportation research, nodes are places where inflowing resources meet profitably creating building densities, activities, and people.

5.1.2. Nodes and poles within the AOI are catalogued for analysis.

5.1.3. Measure road widths connecting all the selected nodes and assign a type using Table 5.3.

5.1.4. Assign an influence radius to each intersection using Table 5.3.

5.1.5. Calculate the influence area for each of the selected nodes.

5.1.6. The total population of the AOI estimated in sub-section 3.1 is subdivided using each of the node's area of influence as a weight parameter adding up to 100% of the population destined to live in the AOI 10 years in the future. If nodal influence areas coincide, apply a weight definition process by which the sum of their corresponding road widths is compared and a new weight is appointed.

5.1.7. A building enclosure ratio from Table 5.3 is assigned and the maximum building height of each road is estimated. This will provide with the amount of surface area available

for new development during the next ten years (The enclosure ratios used are limited to the surveyed areas). Note: Minimum housing area required is 90m².

5.1.8. If a selected site is located within the influence area of a node, apply ratios to periphery roads.

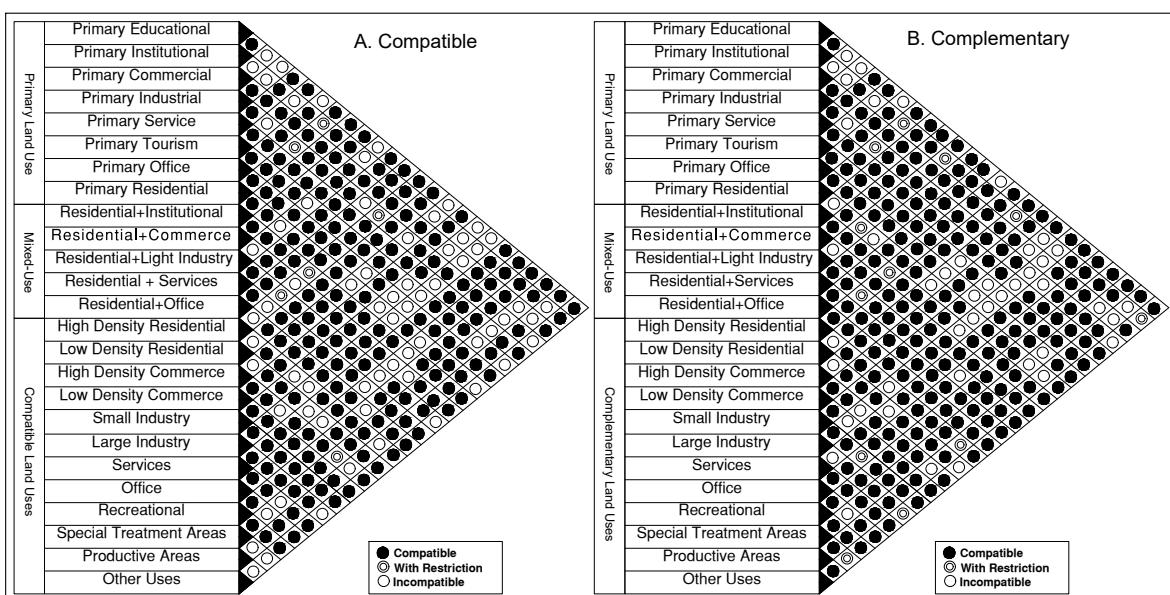
6. Land Use Allocations

6.1. Allocate areas for new households destined to live in the AOI. Each household has 1, 2, or 3 adults per HH. Therefore, assign local standard areas for each type of HH. We used 70m² (1-adult), 100 m² (2-adult), and 120 m² (+3-adult). The remaining area is used for non-residential development.

6.2. Compatible Land uses: Activities that work together or in synergy by need or incentive.

6.3.. Road hierarchy is used to classify land sharing a side with roads. Subsequently, existing structural activities categorize each node of the AOI with land use categories: Housing, Commerce, Small Industry, Large Industry, Office, Service, or multi-purpose. Structural activities are services, firms, or other land use that attracts people from outside the district (inter-district nodes) or outside the prefecture (metropolitan nodes). Using this knowledge, new structural activities are required in nodes where district and inter-district structural activites do not exist.

6.4. Using the remaining area forecasted for development, assign compatible land uses to the structural activities in each nodal influence area. Use matrix 1 below.



Matrix 1. Compatible and complementary land use allocations

6.5. Complementary Land Uses

6.5.1. Complementary Land Uses: Activities that add value to neighboring activities.

6.5.2. Classify road hierarchies and apply enclosure ratios for housing in local roads and collector roads when site is not found within the nodal's area of influence.

6.5.3. Include neighborhood activities from Matrix 2.

Complementary 3.3	Activities	Complementary Land Use										Min. Area	
		Dining Hall	Administration	Meeting Hall	Kinship	Grey Water	Biodegradable	Local Market	Elementary	Bus	Bicycle	Parking Lot	
Recreational / 1 Ha	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	250m ²				
Small Industry / 1 Ha	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1,800m ²
Waste Treatment (288m ³ / day / 10 Ha)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	380m ²				
Commerce / 10 Ha	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1,600m ²						
Education / 10 Ha	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	150m ²							
Public Transportation / 10 Ha	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	-							
Private Transportation / 10 Ha	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	-									

Matrix 2. Community-based land use activities.

7. Accessibility

- #### 7.1. Identify Access routes and periphery roads to the candidate site.

- ## 7.2. Catalogue block sizes along boundary roads Table 5.1.

- ### 7.3. Is the site in a nodal influence area?

If yes, use road hierarchy, building density, and intersections to determine a permeability level using Table 5.1.

If no, use road hierarchy, building density, and intersections to determine a permeability level using Table 5.1.

- 7.4. For privacy and security purposes, connectivity between periphery and interior residential permeability has been considered for accessibility and by-pass transit.

- 7.5. Based on density levels, road connectivity, and accessibility levels choose a street design that will satisfy the aforementioned requirements.

- 7.6. Finally, determine road width based on density classification levels:

If under influence area, apply enclosure ratios.

Otherwise, apply minimum ratios based on high residential density of 60 households per hectare.

**Note: This will maximize land usage and supply demand for economic activity within the area. For this reason, node densification with new households and new jobs is extremely important for this plan.*

8. Household allocations

- 8.1. Households from our synthetic data set were used to fill each lot of 120 m² in the simulated settlement.

- 8.2. HH's are clustered using a clustering method detailed in Garcia and Murao 2020 and applied here too.

- 8.1. These HH clusters represent displaced people clustered using their travel attributes (e.g., a majority of people travel south to a popular destination or travel east to the ocean, etc.). Therefore, each cluster is placed

- 8.3. Placement factors used in this study were: Ages over 50-years-old and a south destination. The majority of people in each cluster with any of these attributes would be placed closer to the settlement's entrance/exit.

- 8.4. Next, synthetic households within each cluster were assigned to particular lots using the household allocation model in Chapter 5. As a result, the simulated settlement was ready for travel simulation.