

**TSUNAMI RISK AND VULNERABILITY:
REMOTE SENSING AND GIS APPROACHES FOR SURFACE
ROUGHNESS DETERMINATION, SETTLEMENT MAPPING
AND POPULATION DISTRIBUTION MODELING**

Dissertation
der Fakultät für Geowissenschaften
der Ludwig-Maximilians-Universität München



eingereicht von

Muhammad Rokhis Khomarudin

im Juli 2010

1. Gutachter: Prof. Dr. Ralf Ludwig
2. Gutachter: Prof. Dr. Günter Strunz

Datum der Disputation: 7. Dezember 2010

ACKNOWLEDGMENTS

First, I would like to sincerely thank my supervisors, Prof. Dr. Ralf Ludwig and Prof. Dr. Günter Strunz, for their support, guidance and care in finalizing this PHD thesis. Special thanks go to Dr. Kai Zoßeder, Dr. Joachim Post, Dr. Torsten Riedlinger, Matthias Mück and Stephanie Wegscheider for three and a half years of intensive discussions — I learned a great deal from all of you. My appreciation also goes to other United Nations University, Institute for Environmental and Human Security (UNU-EHS) PhD students, Sumaryono, Widjo Kongko, and Widodo S. Pranowo for fruitful collaboration in carrying out the research. Thanks also to Evalyne Katabaro from UNU-EHS for the invaluable help on administrative matters.

I am very fortunate to have my wife, Retno Kustiyah, who always gives me support, encouragement and delicious Indonesian food during my stay in Germany. Thanks for your patience and love. My lovely son, Fathan Muhammad Alif, you are my inspiration to keep my spirits up. To my mother, my mother-in-law and all my family in Pekalongan and Batu, Indonesia, thank you for providing me every day 'doa' for our success.

I would also like to thank Prof. Stefan Dech and Dr. Harald Mehl for having facilitated my work at DLR and also to Ibu Ratih Dewanti, Ibu Komala and Pak Agus Hidayat from LAPAN for providing their permission for me to study in Germany. It has been a great opportunity for me to study and learn about the science and culture. Thanks also go to all my colleagues, too many to mention — to whom I would just like to say, *vielen dank*.

This research work has been performed in the framework of the German Indonesian Tsunami Early Warning System (GITEWS project), which is carried out by a large group of scientists and engineers from the German Research Centre for Geosciences (GFZ, consortium leader) and its partners from the Alfred Wegener Institute for Polar and Marine Research (AWI), the German Aerospace Center (DLR), the GKSS Research Centre, the German Marine Research Consortium (KDM), the Leibniz Institute for Marine Sciences (IFM-GEOMAR), the United Nations University (UNU), the Federal Institute for Geosciences and Natural Resources (BGR), the German Agency for Technical Cooperation (GTZ), as well as Indonesian and other international partners. Funding is provided by the German Federal Ministry for Education and Research (BMBF), Grant 03TSU01.

This research has been carried out under a PhD scholarship provided by UNU-EHS, and with support from the German Aerospace Center (DLR) and the National Institute of Aeronautics and Space (LAPAN) of Indonesia.

SUMMARY

The research focuses on providing reliable spatial information in support of tsunami risk and vulnerability assessment within the framework of the German-Indonesian Tsunami Early Warning System (GITEWS) project. It contributes to three major components of the project: (1) the provision of spatial information on surface roughness as an important parameter for tsunami inundation modeling and hazard assessment; (2) the modeling of population distribution, which is an essential factor in tsunami vulnerability assessment and local disaster management activities; and (3) the settlement detection and classification from remote sensing radar imagery to support the population distribution research.

Regarding the surface roughness determination, research analyses on surface roughness classes and their coefficients have been conducted. This included the development of remote sensing classification techniques to derive surface roughness classes, and integration of the thus derived spatial information on surface roughness conditions to tsunami inundation modeling. This research determined 12 classes of surface roughness and their respective coefficients based on analyses of published values.

The developed method for surface roughness classification of remote sensing data considered density and neighborhood conditions, and resulted in more than 90% accuracy. The classification method consists of two steps: main land use classification and density and neighborhood analysis. First, the main land uses were defined and a classification was performed applying decision tree modeling. Texture parameters played an important role in increasing the classification accuracy. The density and neighborhood analysis further substantiated the classification result towards identifying surface roughness classes. Different classes such as residential areas and trees were combined to new surface roughness classes, as "residential areas with trees". The density and neighborhood analysis led to an appropriate representation of real surface roughness conditions. This was used as an important input for tsunami inundation modeling.

By using Tohoku University's Analysis Model for Investigation Near-field Tsunami Number 3 (TUNAMI N3), the spatially distributed surface roughness information was integrated in tsunami inundation modeling and compared to the modeling results applying a uniform surface roughness condition. An uncertainty analysis of tsunami inundation modeling based on the variation of surface roughness coefficients in the Cilacap study area was also undertaken. It

was demonstrated that the inundation modeling results applying uniform and spatially distributed surface roughness resulted in high differences of inundation lengths, especially in areas far from the coastline. This result showed the important role of surface roughness conditions in resisting tsunami flow, which must be considered in tsunami inundation modeling.

With respect to the second research focus, the population distribution, a concept of population distribution modeling was developed. Within the modeling process, weighting factor determination, multi-scale disaggregation and a comparative study to other methods were conducted. The basis of the developed method was a combination of census and land use data, which led to an improved spatial resolution and accuracy of the population distribution. Socio-economic data were used to derive weighting factors to distributing people to land use classes. Moreover, in case of missing input data, an approach was developed that allows for the determination of generalized weighting factors. The approach to use specific weightings, where possible and generalized ones, where necessary, led to a flexible methodology with respect to the achievable accuracy and availability of data. A comparative study was performed by comparing this new model with previously developed population distribution models. The newly developed model showed a higher accuracy.

The detailed population distribution information was a valuable input for the vulnerability assessment being the main data source for human exposure assessment and an important contribution to evacuation time modeling.

In support of the population distribution research, settlement classification using TerraSAR-X imagery was conducted. A current classification method of speckle divergence analysis on SAR imagery was further developed and improved by including the neighborhood concept. The settlement classification provided highly accurate results in dense urban areas, whereas the method needs to be further developed and improved for rural settlement areas.

Finally, it has been shown how the results of this research can be applied. These applications cover the integration of surface roughness conditions into the tsunami inundation modeling and hazard mapping. The contributions to tsunami vulnerability assessment and evacuation planning were shown. Additionally, the results were integrated into the decision support system of the Tsunami Early Warning Center in Jakarta.

CONTENTS

Summary	I
Contents	III
List of Figures	VII
List of Tables	XII

CHAPTER 1: INTRODUCTION

1.1. Motivation	1
1.2. Research Questions	4
1.3. Research Objectives	5
1.4. Research Benefits.....	6
1.5. Thesis Structure.....	6

CHAPTER 2: RESEARCH FRAMEWORK

2.1. GITEWS Framework.....	8
2.2. Risk and Vulnerability Framework	10
2.3. Research Contribution to Tsunami Risk assessment	12
2.4. Study Area	13

CHAPTER 3: STATE OF THE ART OF RESEARCH

3.1. Remote Sensing and Tsunami Risk and Vulnerability	19
3.1.1. Overview	19
3.1.2. Damage assessment	20
3.1.3. Derivation of relevant information for risk and vulnerability ..	20
3.1.4. Monitoring of reconstruction and rebuilding in the post disaster phase	21
3.1.5. Remote sensing classification techniques	21
3.1.6. Summary of remote sensing and tsunami risk and vulnerability	23
3.2. Surface Roughness Determination	23
3.2.1. Overview	23
3.2.2. The role of vegetation on reducing tsunami impact	25
3.2.3. Estimation of the surface roughness coefficient.....	29
3.2.4. Roughness coefficient in tsunami modeling	33
3.2.5. Remote sensing approaches for surface	

roughness classification	35
3.2.6. Summary of surface roughness determination	36
3.3. Population Distribution Modeling	37
3.3.1. Overview	37
3.3.2. Top-down approaches	38
3.3.3. Bottom-up approaches	52
3.3.4. Summary of population distribution modeling	56
3.4. Settlement Classification Using Remote Sensing Data	57
3.4.1. Overview	57
3.4.2. Settlement classification using optical imagery	57
3.4.3. Settlement classification using SAR imagery	59
3.4.4. Using data fusion of optical and SAR Imagery	62
3.4.5. Summary of settlement classification using remote sensing data	63
3.5. Conclusion of State of the Art of Research	63

CHAPTER 4: METHODOLOGY

4.1. Data Requirement	67
4.2. Surface Roughness Determination	68
4.2.1. Surface roughness classes and their coefficient estimation.....	68
4.2.2. Surface roughness classification.....	69
4.2.3. Accuracy Assessment	75
4.2.4. Integration into tsunami inundation modeling	77
4.3. Population Distribution Modeling	82
4.3.1. General concept	82
4.3.2. Weighting factor determination.....	84
4.3.3. Accuracy Assessment	87
4.3.4. Multi-scale disaggregation.....	88
4.3.5. Comparative study.....	90
4.4. Settlement Classification by Using Radar Imagery Based on Speckle Divergence and Neighborhood Analysis.....	91
4.4.1. Classification procedure	91
4.4.2. Accuracy Assessment	94
4.3.4. Transfer of the approach to other areas	94

CHAPTER 5: RESULTS

5.1. Data Collection	95
5.1.1. Satellite imagery and spatial information	95
5.1.2. Demographic and socio-economic statistics	98
5.1.3. Reference data	99
5.2. Surface roughness determination	103

5.2.1. Surface roughness classes and their coefficient estimation	103
5.2.2. Surface roughness classification by remote sensing	104
5.2.3. Integration of surface roughness for tsunami inundation modelling	114
5.3. Population Distribution Modeling	117
5.3.1. Result of the population distribution modeling concept.....	117
5.3.2. Weighting factor determination	119
5.3.3. Result from questionnaire analysis.....	126
5.3.4. Error of population distribution (E_{PD})	127
5.3.5. Comparative study	128
5.3.6. Multi-scale disaggregation for population distribution in Sumatra, Java and Bali.....	129
5.3. Settlement Classification	131
5.4.1. Pre-processing.....	131
5.4.2. Classification result	132

CHAPTER 6: DISCUSSIONS

6.1. Data Collection	136
6.1.1. Satellite imagery and spatial information	136
6.1.2. Demographic and socio-economic statistics.....	137
6.1.3. Reference data	137
6.2. Surface Roughness Determination	138
6.2.1. Surface roughness classes and their coefficient estimation	138
6.2.2. Surface roughness classification by remote sensing	140
6.2.3. Integration of surface roughness for tsunami inundation modeling	143
6.3. Population Distribution Modeling	144
6.3.1. Weighting factor determination	144
6.3.2. Accuracy of population distribution	145
6.3.3. Multi-scale disaggregation of population distribution.....	146
6.3.5. Commuters	146
6.4. Settlement classification	148

CHAPTER 7: APPLICATION OF THE RESEARCH

7.1. Overview	150
7.2. Hazard Assessment	150
7.3. Human Exposure.....	152
7.4. Evacuation Planning.....	154
7.5. Integration into the Decision Support System.....	156
7.6. Conclusions	158

CHAPTER 8: CONCLUSIONS

- 8.1. Surface Roughness Determination 159
- 8.2. Population Distribution Modeling 160
- 8.3. Settlement Classification..... 161
- 8.4. Application of the Research 162

- References** 163
- Curriculum Vitae** 177

List of Figures

1.1	The occurrence of tsunami with their earthquake moment magnitudes and intensities along the Sunda trench (Source: ITDB/WLD, 2007)	1
2.1	GITEWS Early Warning System framework	9
2.2	The GITEWS Work Packages	9
2.3	The risk and vulnerability concept from Bogardi, Birkmann and Cardona (the BBC Framework)	11
2.4	The GITEWS risk and vulnerability framework, the blue boxes are high lightening the research contribution	11
2.5	The elements of research	14
2.6	The research study area, Cilacap as pilot region (below) and the broad scale area (above) for the implementation of population distribution	15
2.7	The bathymetry, the subduction zone, and the historical tsunami and earthquake occurrences in the Indian Ocean around the Cilacap region (Source: ITDB/WLD, 2007)	17
3.1	The disaster management cycle and remote sensing contributions	19
3.2	The potential of vegetation on reducing tsunami impact in Pangandaran (Source: Latief and Hadi, 2007))	26
3.3	Observation pictures from the effect of tsunami 2004 in Nias Island and Simeulue (Source: Kongko, 2005)	27
3.4	Effects of forest width and density to tsunami reduction (source: Harada and Kawata, 2004)	28
3.5	Examples of photograph matching (Source: Brisbane city council, 2003)	30
3.6	The maximum run-up differences (m) by changing roughness coefficient (Source: Myers and Baptista, 2001)	33
3.7	The top-down and bottom-up approaches on population distribution by using remote sensing and GIS	38
3.8	General relationship between (a) zone boundaries and population-weighted centroids, and (b) population-weighted	

centroids and gridded population estimates (Source: Martin, 2006)	40
3.9 The smoothing process of pycnophylactic interpolation until 25 iterations (Source: National Center for Geographic Information and Analysis University of California Santa Barbara based on Tobler, 1979)	42
3.10 Fundamental approaches to population distribution mapping, modified from Sleeter and Gould (2007)	44
3.11 Population distribution maps: the LandScan (left), and CIESIN population grids (right)	50
4.1 The steps of surface roughness classification using optical satellite imagery	70
4.2 Confusion matrix accuracy assessment calculation (modified from Congalton and Green, 1999)	76
4.3 The 16 reliable source models of hypothetical tsunamis. Sc.11 is the worst case source	79
4.4 Locations of point samples to analyze the influence of difference roughness classes on tsunami flow-depth and velocity	80
4.5 Synthetic example of the disaggregation method used to improve the spatial population data	82
4.6 The boundary effect problem in the population distribution	84
4.7 Potential number of people engaged in different land use activities during the day- and night-times in the village	86
4.8 Multi-scale disaggregation concept to provide the population distribution in west coast of Sumatera, south coast of Java and Bali	89
4.9 The steps of settlement classification by using TerraSAR-X	92
4.10 The threshold decision for potential settlement and real settlement	93
4.11 Illustration the neighbourhood analysis for settlement classification (a) real settlement detection (b) potential settlement and (c) the settlement mask by neighbourhoods a kernel size of 3 x 3 pixels	94
5.1 Satellite imagery data collection for this research: (a) SPOT-5 multi-spectral (24 June 2004); (b) SPOT-5 Panchromatic (24	

June 2004); (c) TerraSAR-X for Padang (04 April 2008); (d) TerraSAR-X Cilacap 1 (24 January 2009); (e) TerraSAR-X Cilacap 2 (04 February 2009); and (f) TerraSAR-X Cilacap 3 (15 February 2009)	96
5.2 Land use data collection in difference scales (1: 100,000 scale in the red box, 1: 50,000 scale in the pink box, and 1: 25,000 scale in the yellow box.)	97
5.3 Statistical data on population distribution at the village level for the entire coast of Sumatra, Java and Bali	98
5.4 Example of land cover reference polygons for the accuracy assessment of the surface roughness classification	99
5.5 The reference data for settlement classification, Cilacap (left) and Padang (right)	100
5.6 The surface roughness classes and their relation to the tsunami protection probability (from very low to high)	104
5.7 Result of the spatial enhancement of a SPOT-5 scene (right), based on a multi-spectral (left) and panchromatic imagery (middle)	105
5.8 Results of image variable calculation of a SPOT-5 scene based on NDVI (left), SNDVI (middle) and ENDVI (right)	105
5.9 Result of textural analysis of a SPOT-5 scene: panchromatic imagery (left) and the texture imagery (right)	106
5.10 The decision tree model for main land use classification	109
5.11 The result of decision tree classification of SPOT-5 data for main land use classes	110
5.12 Surface roughness classification in Cilacap District	112
5.13 Difference of tsunami velocity based on different roughness conditions: uniform roughness coefficient (top), spatially distributed roughness values (low)	114
5.14 Absolute difference of maximum velocity and maximum water flow depth (%) between tsunami inundation modeling using uniform roughness coefficient and spatially distributed roughness values	115
5.15 Coefficient of variation (COV) for the maximum velocity (top) and maximum flow depth (low)	116
5.16 Coefficient of variation in selected samples for maximum flow	117

	depth (blue line) and maximum velocity (red line)	
5.17	The result of population distribution modeling for day- and night-time (lower left and right), based on dasymetric mapping census data (top left) and land use as ancillary data (top right) for Cilacap District	118
5.18	Population distribution based on the administrative boundary only (left) and the improved population distribution, with boundary effects (middle) and without boundary effects (final result, right)	119
5.19	Distribution of potential number of people engaged in different land use activities in villages in the west coast of Sumatra, the south coast of Java, and Bali	120
5.20	Average of coefficient of variation (%) of potential number of people engaged in different land use activities by the category of village (above) and those distributions for settlement (bottom)	122
5.21	Generalized weighting factor for day-time population distribution	124
5.22	Generalized weighting factor for night-time population distribution	125
5.23	The percentages of people carrying out activities inside and outside their villages in Cilacap in the day-time; and (b) the percentages of the different commuter categories carrying out activities outside their villages	127
5.24	RMSE error of population distribution in Cilacap for various land use classes, based on an accuracy assessment	128
5.25	Error comparison of three models for population distribution	129
5.26	Result of population distribution in the broad scale area — the west coast of Sumatra, the south coast of Java, and Bali — after applying the multi-scale disaggregation method	130
5.27	TerraSAR-X image intensity (left) and after speckle divergence analysis (right)	131
5.28	Settlement classification results and real settlement by using TerraSAR-X (see chapter 4.4.1)	132
5.29	Result of speckle divergence and neighborhood analysis for settlement detection by using TerraSAR-X on Cilacap District	133
5.30	Result of speckle divergence and neighborhood analysis for	134

settlement detection using TerraSAR-X of the Padang District	
6.1 Image before (a) and after (b) tsunami from LANDSAT TM	140
6.2 Transformation of NDVI by (a) sigmoid and (b) exponential function	141
6.3 The histogram samples of NDVI for three vegetation classes (tree, shrubs, and crop field)	141
6.4 The heterogeneous patterns of land use in the study area: (a) tree and shrub; (b) crop field and shrubs	142
6.5 Potential additional of people in the Cilacap District (red represents the high potential additional of people due to infrastructural facilities)	147
6.6 Example showing the additional number of people in villages based on The sum of the commuters (here, 80 people) within the municipality	148
6.7 Examples of settlements in the rural area of Cilacap District (settlements surrounded by vegetation)	149
7.1 Tsunami inundation maps for Padang District based on (a) uniform roughness and (b) spatially distributed roughness	151
7.2 Tsunami hazard maps for Kuta, Bali	152
7.3 Human exposure maps of Kuta, Bali for day-time population	153
7.4 Number of exposed people in the hazard zone and changes during the day- and night-time	154
7.5 Evacuation time map for major warning tsunami in Bali (Source: Strunz <i>et al.</i> , 2010)	155
7.6 A Planning map recommendations for the need of tsunami evacuation buildings (Source: Strunz <i>et al.</i> , 2010)	156
7.7 Integration of key parameters of risk assessment into the decision support system (Source: Raape <i>et al.</i> , 2008)	157

List of Tables

3.1	The relation between tsunami intensity and tsunami damage	25
3.2	Methods of roughness coefficient estimation (modified from Sellin <i>et al.</i> , 2003)	29
3.3	Roughness coefficients for different land use/cover classes in several research publications	31
3.4	The roughness coefficient estimation by Koshimura <i>et al.</i> (2009)	32
4.1	Data requirements for the research	67
4.2	Analysis of image variables	73
4.3	Steps of the semi-automatic classification	75
4.4	Population error matrix based on information of classified and reference land use (LU) data	75
4.5	The numerical parameters of tsunami model	77
4.6	The reliable source models of hypothetical tsunamis	78
5.1	Compilation of all data used in this research	101
5.2	The roughness coefficients assigned to surface roughness classes based on literature review	103
5.3	The accuracy of SPOT-5 data image variables to classify the main classes of surface roughness by using a decision tree model	107
5.4	The accuracy of decision tree model by using relevant SPOT-5 variable (four spatially enhanced spectral bands of SPOT-5, NDVI, and Texture)	108
5.5	Accuracy assessment of remote sensing classification for main land use roughness (number of pixels)	111
5.6	Accuracy assessment of the remote sensing classification of surface roughness classes	113
5.7	The classification accuracy resulting from a standard maximum likelihood approach	113
5.8	Details of coefficient of variation (%) of potential number of people engaged in different land use activities by the category of village	123

5.9	A comparison of the weighting factors in Cilacap District between the generalized weighting factors from the model and the reference values from the questionnaire results	126
5.10	Accuracy assessment of speckle divergence and neighborhood analysis for Cilacap District	132
5.11	Accuracy assessment of speckle divergence and neighborhood analysis for Padang District	134

CHAPTER 1: INTRODUCTION

1.1.Motivation

The devastating Sumatra earthquake of 26 December 2004 was the second largest ever detected rupture in the Earth's crust. Only a few minutes after its detection, the first tsunami waves hit the coastlines of northern Sumatra and some time later, Malaysia, Thailand, Sri Lanka, India and Somalia were also affected. Due to the location of the earthquake epicenter, Aceh Province, Indonesia, complained of the largest number of casualties and damages.

This tremendous disaster caused about 230,000 deaths and initiated – in addition to major relief efforts – intensive tsunami research and the emerging need for tsunami early warning systems that would mitigate tsunami impact.

Due to the great potential for further tsunami events, Indonesia has taken centre stage in the current tsunami research. Figure 1.1 shows the tsunami occurrences from 1859 to 2007 with their earthquake magnitudes and intensities along the Sunda Trench (ITDB/WLD, 2007).

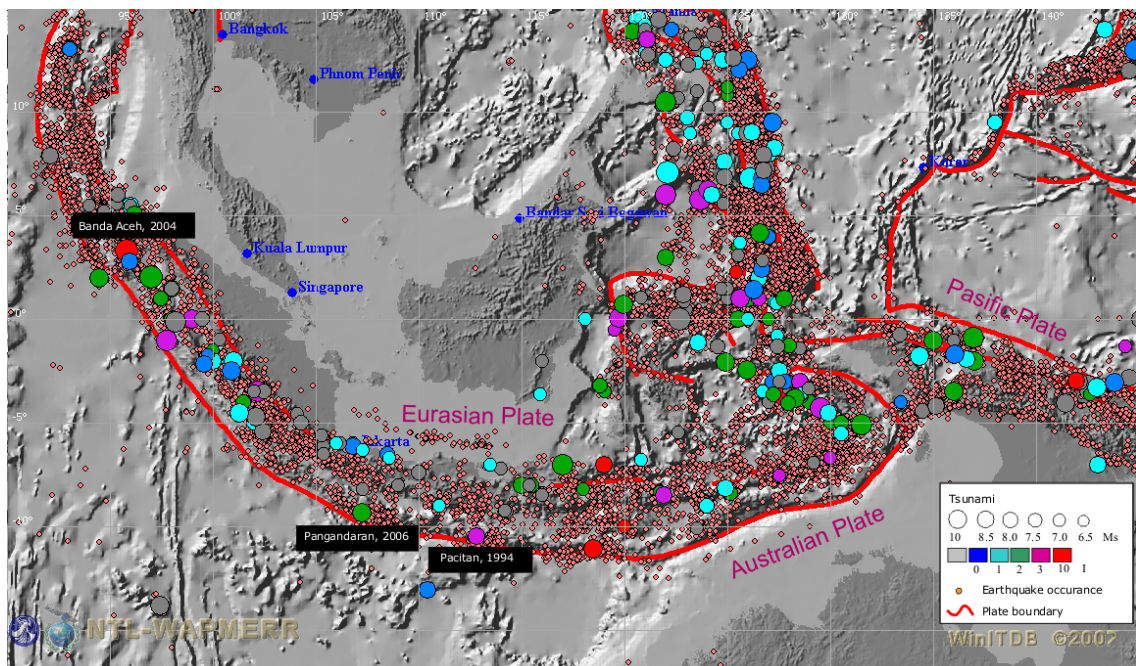


Figure 1.1 The occurrence of tsunami with their earthquake magnitudes and intensities along the Sunda trench (Source: ITDB/WLD, 2007)

Based on these geological conditions and historical data, Indonesia is a highly susceptible region to tsunami hazards. At the same time, its population is increasingly growing in the coastal areas. In particular, the islands of Sumatra and Java have experienced a tremendous population growth since the 1970s. Java is well known for being the most populated areas in the world. Due to these conditions, a strategy on mitigating the consequences of a tsunami impact is strongly required.

Risk assessment plays an important role in designing strategies for disaster risk reduction and disaster response. Knowledge about exposed elements, and their susceptibility, coping and adaptation mechanisms is a precondition for the development of people-centered warning structures (Post *et al.*, 2008).

Overall, risk assessment has two dimensions:

- the hazard assessment;
- the assessment of vulnerability.

Both components have a significant influence on the level of tsunami risk and provide the basis for the tsunami risk assessment.

In terms of the tsunami impact, the hazard assessment is often covered by numerical inundation modeling, which indicates a probability for the spatial distribution of the maximum inundation area depending on the location of the source (Geist and Parson, 2006; Annaka *et al.*, 2007; Burbidge *et al.*, 2008), and Power *et al.*, 2007). It is clear that there is a correlation between the distance to the coast and the degree of impact. In addition, terrain altitude is an important influence factor. Moreover, in tsunami inundation modeling, a number of additional parameters must be taken into account, including surface roughness. Surface roughness of potentially affected areas plays an important role in mitigating or increasing tsunami intensity, by influencing the tsunami flow, tsunami direction, inundation length, and velocity of water. In addition, detailed information on surface roughness and the specific influences for tsunami inundation will provide decisive findings to reduce the long-term tsunami risk to human life and property by efficient urban planning. Therefore, it is necessary to quantify the surface roughness. The availability of up-to-date spatial data is crucial to achieve this goal. While in situ field surveys are time-consuming and expensive, remote sensing provides an efficient possibility to derive the relevant surface parameters. The main challenge in using remote sensing is the development of an appropriate classification methodology that would allow for the automatic or semi-automatic derivation of these parameters from the remote sensing data.

Since the consideration of surface roughness in tsunami research is very limited, the first part of this thesis is devoted to research on how to describe the relevant surface roughness class to tsunami inundation modeling and to elaborate a concept for a roughness classification based on medium- and high-resolution satellite imagery.

Vulnerability assessment as the second dimension of risk assessment plays a crucial role in tsunami disaster reduction strategies. Vulnerability can be defined as *"The conditions determined by physical, social, economic, and environmental factors or processes which increase the susceptibility of a community to the impact of hazards"* (ISDR, 2004). With respect to this definition, vulnerability is trans-disciplinary and multi-dimensional, covering social, economic, physical, political, engineering and ecological aspects and dimensions (Post *et al.*, 2009). Detailed and up-to-date information on population distribution in the potential hazard zone is the most important factor in mitigating the impact of natural disasters, e.g. by evacuation planning.

The available information about population distribution is mostly based on statistical data and is related to administrative boundaries, such as village, municipal, district, province and national borders. In most countries and also in Indonesia, population distribution data are available for villages as the smallest administrative level. However, for an efficient disaster management, it is essential to have a good insight into the spatial distribution of the population at risk. The available census data on the village level are too coarse for a planning process at the local level. The question then is: How can population data in census districts be disaggregated to smaller geographical or mapping sectors that better satisfy the demands for information for emergency actions and disaster mitigation management? (Hofstee and Islam, 2004).

In terms of remote sensing and GIS, spatial improvement on population distribution is not a new task. However, applying current research findings to tsunami early warning and disaster management issues is unsatisfying because of the static information of most approaches. The crucial assessment unit in quantifying human response capability related to tsunamis is time, but current research results considering temporal analyses do not meet the demands for tsunami-related topics. For vulnerability assessment, it is necessary to have spatially explicit information on population distribution as well as on its temporal distribution, i.e. the changes during the day- and night-time, which is crucial information needed for evacuation planning.

Consequently, the second part of this thesis is devoted to improve the spatial and temporal resolution of the population distribution data in coastal areas of Indonesia. In addressing this research problem, it focuses on a sub-national

and local-scale assessment related to different levels of tsunami disaster management. Results and conclusions from this part seek to provide key information for effective warning and evacuation planning, and will be implemented in the tsunami risk and vulnerability assessment framework.

A key parameter in disaggregating population data based on a dasymetric mapping approach (see chapter 3) is the availability of comprehensive, detailed and up-to-date land use data, which is always a problem, especially in large and fast developing countries like Indonesia. Previous research has demonstrated the potential of remote sensing in mapping settlement areas. With the availability of high-resolution synthetic aperture radar (SAR) sensor systems, e.g. of the German TerraSAR-X Earth observation satellite, new opportunities of settlement mapping are available. SAR systems are capable of acquiring data at day and night, independently of weather or environmental conditions. Hence, sophisticated and efficient methodologies for settlement classification using TerraSAR-X are needed for the mapping and updating of land use and settlement areas.

In the third part of the thesis, current research findings of settlement detection using TerraSAR-X will be reviewed. Methods will be further developed or improved and applied to coastal areas in Indonesia. The research findings shall contribute to assessing risk and vulnerability before an expected disastrous tsunami event, in coordinating emergency actions during an event, or in managing disasters after the event.

Risk assessment aims at revealing the two dimensions of hazard and vulnerability assessment, and at deducing options for action related to adaptation and mitigation measures. This research focuses on the development of new methodologies based on remote sensing and GIS techniques, which shall help to provide improved geo-spatial data contributing to reduce tsunami risk at coastal areas of Indonesia and to cope with inevitable tsunami impact.

1.2. Research Questions

The research is devoted to provide decisive parameters for an efficient tsunami risk assessment in order to reduce the impact of tsunami on the coastal areas of Indonesia. For this purpose, different research aspects and approaches are examined. The research questions that will be answered are stated below.

a) Surface roughness determination

1. How can surface roughness be appropriately described and quantified in order to be used in tsunami inundation modeling and hazard assessment?
2. How can surface roughness parameters be derived by remote sensing?
3. How can the derived surface roughness parameters be integrated into tsunami modeling and hazard assessment?

b) Population distribution modeling

1. How can adequate population distribution data be derived or be spatially improved based on available statistical data?
2. Is it possible to derive further information about the spatial and temporal changes, especially the different population distribution during the day- and night-time?

c) Settlement classification using remote sensing

1. How can settlement areas be efficiently mapped using remote sensing technologies?

1.3. Research Objectives

Based on the above research questions, the research objectives are as follows:

a) Surface roughness determination

1. Determine an appropriate surface roughness parameterization for tsunami inundation modeling and hazard assessment.
2. Develop a new remote sensing for surface roughness classification.
3. Integrate the derived roughness coefficients into tsunami modeling and tsunami hazard mapping.

b) Population distribution modeling

1. Develop a new or improved concept on how to derive adequate population data based on the disaggregation of available statistical data.

2. Determine day and night population distribution by modeling and integrating socio-economic factors into the developed approach.
3. Analyze the transferability and accuracy of the developed population distribution model.

c) Settlement classification using remote sensing

1. Further develop and improve methods for the determination of settlement areas with a focus on SAR remote sensing data.

1.4. Research Benefits

This research results shall contribute to tsunami risk mitigation and reduction at coastal areas in Indonesia. By integrating this research into the risk and vulnerability assessment of the German Indonesia Tsunami Early Warning System (GITEWS) project, a concerted focus on decisive research needs in the tsunami context is ensured. Major benefits of this research are:

- the development of new remote sensing methods and GIS models;
- contribution to the risk and vulnerability assessment of the German Indonesia Tsunami Early Warning System (GITEWS) project and the tsunami disaster management in Indonesia;
- remote sensing-based mapping of surface roughness and derivation of settlement areas as a crucial contribution to tsunami modeling and hazard assessment;
- an improved and automated approach for population distribution modeling based on GIS approaches as input for vulnerability analyses;
- contribution to evacuation planning and disaster risk reduction.

1.5. Thesis Structure

This thesis consists of eight chapters.

- Chapter 1 explains the motivation of the research, presents the research questions and objectives, points out the research benefits, and describes the structure of the thesis.

- Chapter 2 provides the research framework. It shows the contribution of the research in the context of the GITEWS project and within the risk and vulnerability assessment part. All research components are clearly represented and described in this chapter in order to point out the coherent structure of the thesis. Finally, all research topics are linked to the respective study area, and relevant spatial conditions are explained in detail.
- Chapter 3 explains the current status on roughness coefficient determination and classification research, the research development on population distribution modeling by using remote sensing and GIS, and the current status of remote sensing settlement classification research. It explains the current gaps in this specific field of research.
- Chapter 4 explains the data required the various steps in the research process, and the approach to answer the research questions. It concretizes the objectives of this research.
- Chapter 5 shows result of data collection, the research findings on the surface roughness and its coefficient estimation, the improvement on surface roughness classification by remote sensing, the improvement on population distribution modeling, and the results of settlement classification.
- Chapter 6 discusses and interprets these research findings with respect to the state of the art of the research. It also discusses the implication of the results in the practical/theoretical aspect of this thesis.
- Chapter 7 explains the application of the research, the integration of its results on the tsunami modeling, and the risk and vulnerability assessment.
- Chapter 8 concludes the main research findings and their potential consequences.

CHAPTER 2: RESEARCH FRAMEWORK

This research is part of the risk and vulnerability assessment within the German Indonesian Tsunami Early Warning System (GITEWS) project. It contributes to three major components of the project: (1) the provision of spatial information on surface roughness as an important parameter for tsunami inundation modeling and hazard assessment; (2) the modeling of population distribution, which is an essential factor in tsunami vulnerability assessment and local disaster management activities; and (3) the settlement detection and classification from remote sensing radar imagery to support the population distribution research.

Detailed explanations of the GITEWS framework, the risk and vulnerability framework, research elements and the study area are stated below.

2.1. GITEWS Framework

The GITEWS project is a research collaboration between the Federal Ministry of Education and Research (BMBF) of Germany and the Indonesian Ministry of Research and Technology (RISTEK), aimed to develop an effective tsunami early warning system for Indonesia. The Joint Declaration of BMBF and RISTEK was signed on 14 March 2005. Its concept integrates terrestrial observation networks of seismology and geodesy with marine measuring processes and satellite observation. Several German and Indonesian institutions are involved and work together to set up end-to-end early warning system, which also contains the capacity-building provision for Indonesian institutions. The system itself was successfully launched on 11 November 2008 by the Indonesian President. The overall concept of the project is shown in Figure 2.1.

GITEWS is designed as a comprehensive project consisting of several work packages and is conducted by a consortium of nine German research institutions in close cooperation with Indonesian partner institutions. The German Aerospace Center (DLR) is responsible for the Early Warning and Mitigation Center (EWMC), in which the risk and vulnerability assessment is integrated as an essential component (see Figure 2.2).

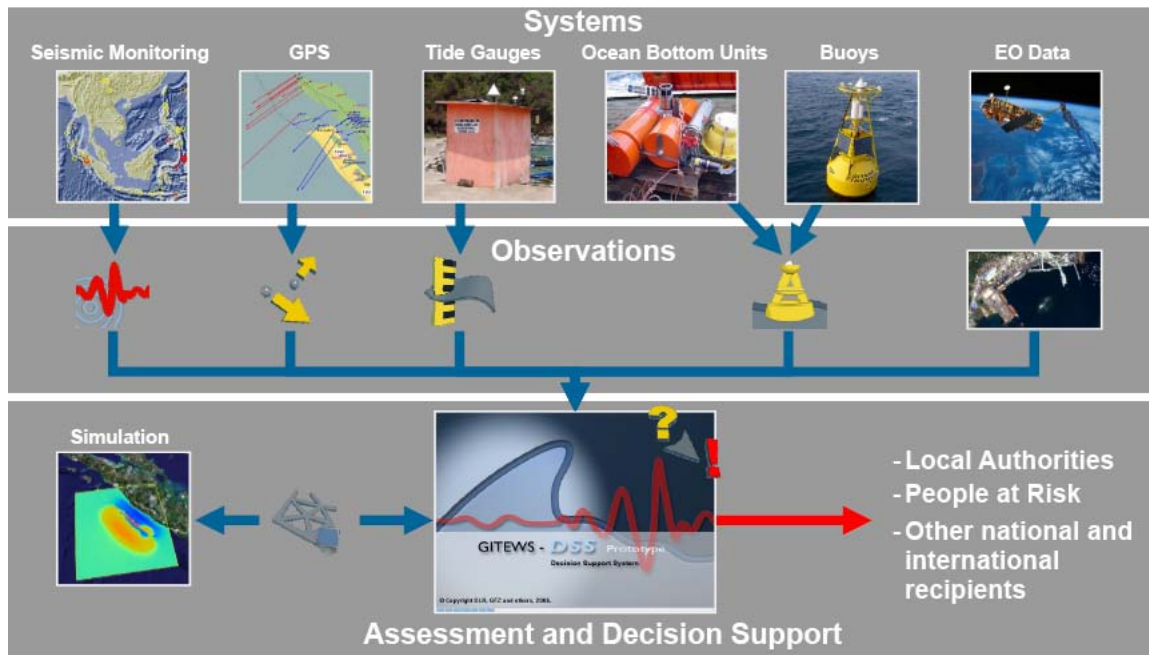


Figure 2.1 GITEWS Early Warning System framework

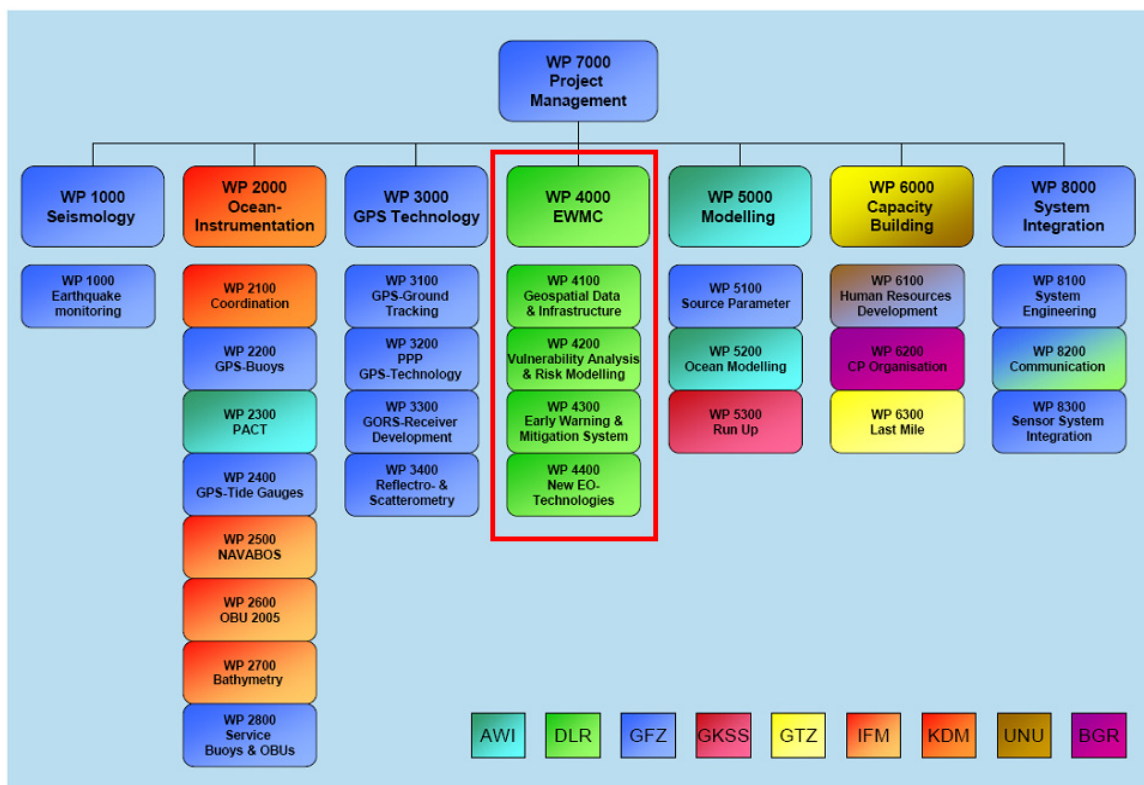


Figure 2.2 The GITEWS Work Packages

2.2. Risk and Vulnerability Framework

Risk management serves as a general guideline for developing disaster preparedness and adaptation strategies based on a continuous strategy of risk and vulnerability assessment (ISDR, 2004). This implies that risk and vulnerability assessment is an integral part in the development of an effective end-to-end early warning system, which significantly contributes to disaster risk reduction to low-frequency but extreme events like tsunamis.

Risk and vulnerability assessment in the GITEWS context contributes to two main tasks:

- enhancing crisis management capacities (e.g. emergency assistance) during an early warning scenario;
- developing disaster risk reduction strategies such as adaptation and mitigation measures (e.g. evacuation and land use planning).

The BBC framework has been adopted in order to implement effective early warning disaster response and recovery (see Figure 2.3). It shows the importance of addressing the potential intervention tools that could help to reduce vulnerability in the social, economic and environmental spheres. The framework integrates social, economic and environmental aspects into the vulnerability assessment, thus reflecting the “three pillars” of sustainable development (Birkmann, 2006).

The risk and vulnerability assessment within GITEWS mainly focuses on vulnerability and risk factors of people exposed to tsunami hazard, in terms of loss of life, injury and loss of livelihood. Accordingly, the GITEWS vulnerability assessment aims at developing socio-economic indicators and assessment tools for the continuous improvement of intervention tools, such as early warning communication, evacuation planning, disaster response and rehabilitation.

Regarding the BBC-framework, the risk and vulnerability assessment in the GITEWS project is translated into a coherent conceptual model, integrating the above-mentioned aspects.

Generally, the concept was to establish monitoring and quantification of the spatial vulnerabilities within the timeline of disasters (see Figure 2.4). The properties or deficiencies related to warning (e.g. people’s ability to understand a warning), to evacuation (ability to respond immediately), as well as to emergency relief and recovery are stated accordingly (Post *et al.*, 2008). The concept is explained more details in Figure 2.4.

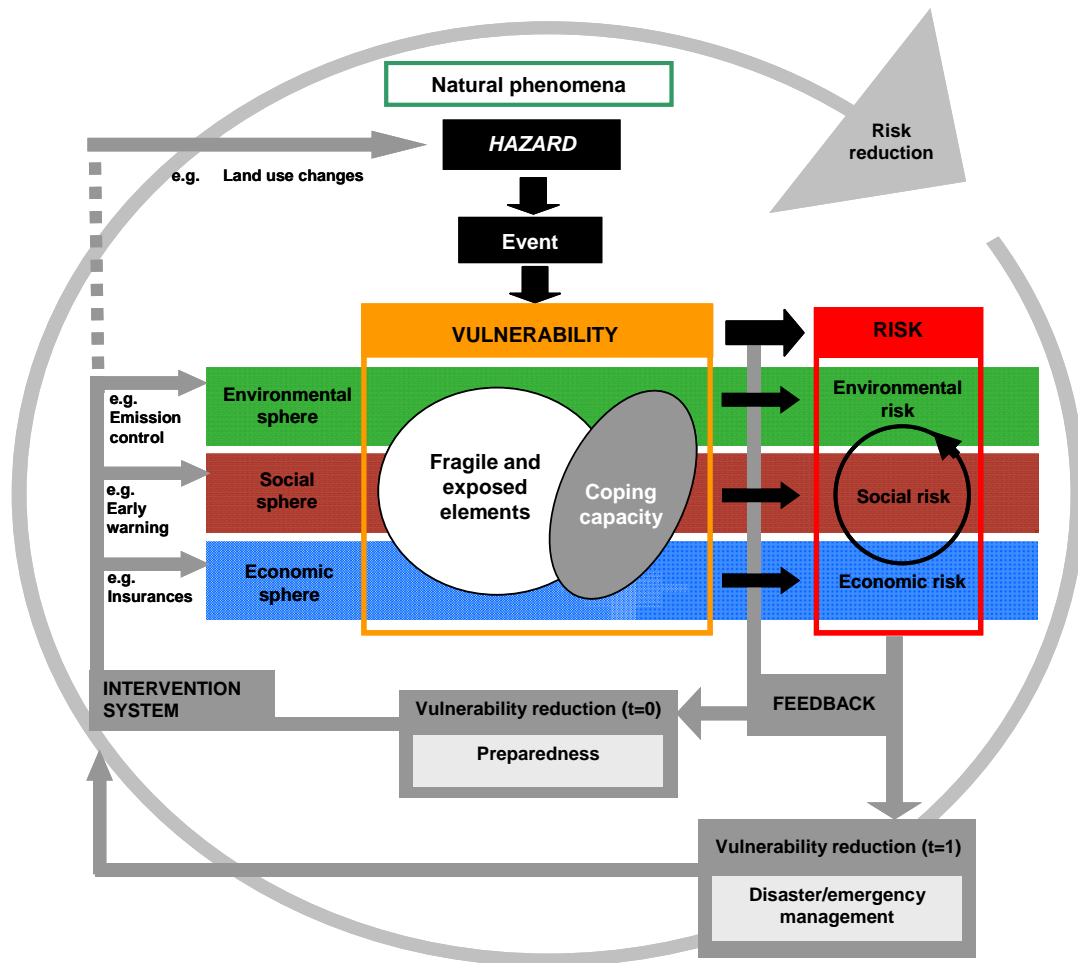


Figure 2.3 The risk and vulnerability concept from Bogardi, Birkmann and Cardona (the BBC Framework)

Early Warning and Response Phases	Preparedness Measures	Topics for Risk Assessment	HAZARD VULNERABILITY
1 Warning Decision	Hazard Assessment	What is the potential hazard impact on land?	
2 Warning Decision	Exposure estimation	Are there exposed people and critical infra.?	
3 Warning Dissemination	Warning chain development	Do people receive and understand a warning?	
4 Anticipated Response	Awareness / sensitization	Do people respond to warnings and evacuate?	
5 Evacuation	Evacuation strategy	Who & how many people are able to reach a safe place?	
6 Emergency Relief	Contingency planning	Who & how many people need relief assistance?	
7 Sustainable Recovery	Recovery planning	Who is able to recover from the impacts?	

Figure 2.4 The GITEWS risk and vulnerability framework, the blue boxes are highlighting the research contribution

The aim of the hazard assessment is to achieve a better understanding of tsunami and the potential impacts on society in order to enable local and national decision-makers and other stakeholders to become better prepared for future tsunami events. In case of a tsunami, the decision whether a region should be warned requires sound information on the hazard zone and expected impact, on the spatial distribution of exposed population and on critical infrastructures. The immediate response component is represented by analyzing people's ability to receive and understand a warning, their decision to evacuate and their capability to do so. Here, mainly socio-economic indicators are relevant that describe the intrinsic social properties of the population exposed to tsunamis.

The component recovery is not directly linked to the early warning chain, but rather to the post-disaster management phase, since it deals with developing vulnerability indicators for disaster relief and rehabilitation aspects in the aftermath of a tsunami disaster. Hence, vulnerability assessment comprises information on the number of people exposed, critical facilities and immediate response capability. The vulnerability information is then combined with the hazard information based on the decision tree logic to calculate the overall tsunami risk (Post *et al.*, 2008).

2.3. Research Contributions to Tsunami Risk Assessment

As highlighted in Figure 2.4, this research contributes to the hazard and exposure component in order to provide decisive input parameters for tsunami risk assessment.

Detailed information about surface roughness is indispensable for tsunami inundation modeling and hazard assessment as a precondition for calculating the exposed population of the coastal areas. The surface roughness parameters derived in this thesis will be implemented accordingly in the risk and vulnerability framework above.

A crucial factor in a people-centered risk assessment is detailed information about population distribution in the potential tsunami hazard zone, since disaster management strategies differ widely between uninhabited and highly populated areas. New and improved methodologies for settlement classification and population disaggregation will strongly enhance the available population data and consequently provide key information for tsunami warning decision support, effective warning and warning chain planning, as well as for evacuation and contingency planning.

Key elements in this research are Remote Sensing (RS) and Geographical Information Systems (GIS) (see Figure 2.5).

Remote sensing data are the precondition for determining surface roughness and settlement detection. In this research, both optical and radar sensors are used. Optical, multi-spectral SPOT 5 data were used for roughness coefficient classification, while TerraSAR-X radar data were used for settlement classification.

GIS modeling was applied for processing and analyzing population data. A dasymetric mapping concept was developed by combining available census data and ancillary data such as land use in order to improve the spatial population distribution. Statistical data, field survey results, reference maps and questionnaires were used to validate the gained modeling results. GIS technology provides efficient tools for the combination and analysis of different input data and is an essential working basis in this research. An accuracy assessment was conducted to check the quality of the GIS modeling and remote sensing classification techniques. As shown in Figure 2.5 and mentioned above, quality controlled results are implemented in the respective parts of the GITEWS risk and vulnerability assessment.

2.4. Study Area

The GITEWS risk and vulnerability assessment is conducted at two scales: at the broad scale and at the local level, each fulfilling a different purpose.

While the broad-scale assessment provides products in the context of early warning for an entire exposed coastal zone, the local assessment aims at contributing to the development of local, specific disaster preparedness, adaptation and mitigation strategies (e.g. urban planning and evacuation planning for priority areas).

In this research, the Cilacap region was selected as the study area at the local level for modeling and analyzing surface roughness, settlement classification and population distribution, while the latter, the population distribution modeling was also implemented for the coast of Sumatra, Java and Bali (broad-scale level, see Figure 2.6).

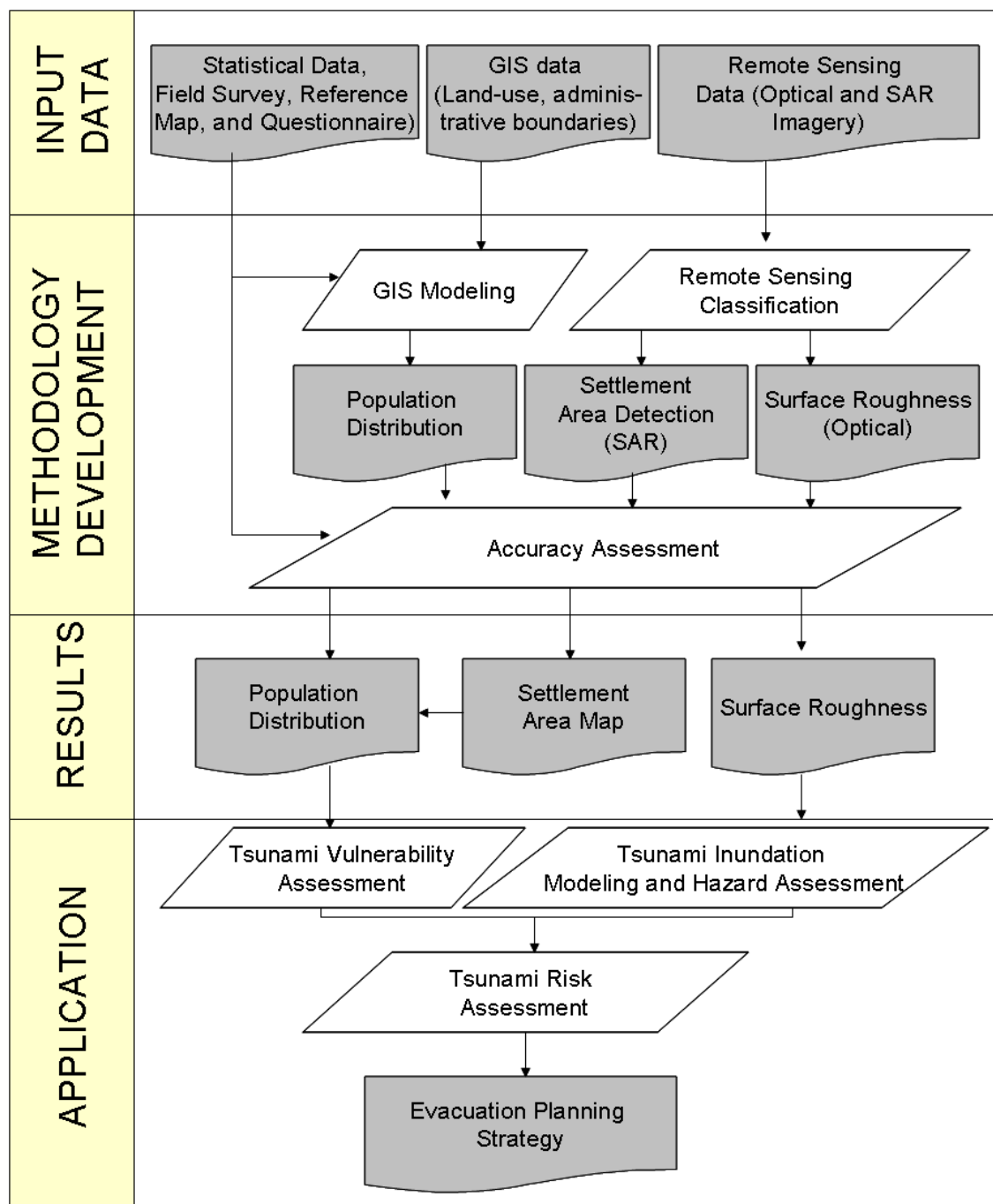


Figure 2.5 The elements of research

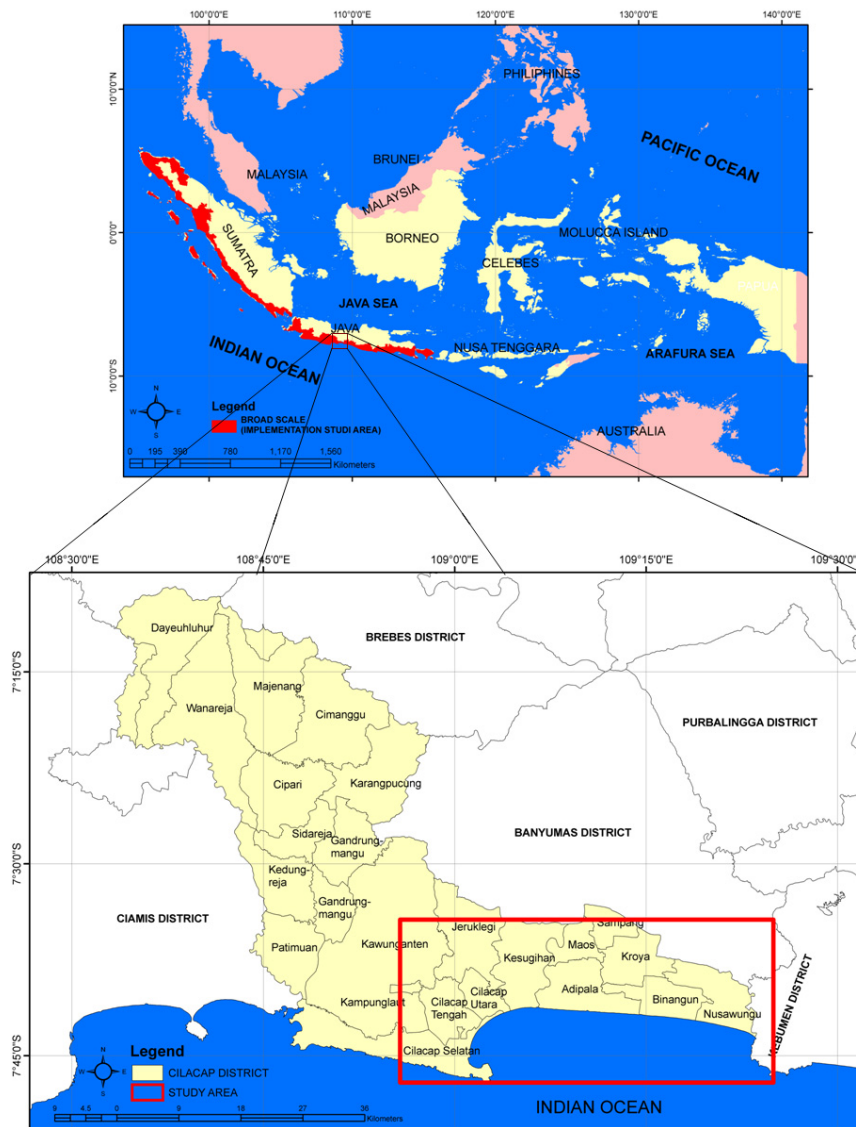


Figure 2.6 The research study area, Cilacap as pilot region (below) and the broad scale area (above) for the implementation of population distribution

Cilacap was selected as a pilot region in this research for the following reasons:

- Cilacap is prone to tsunami disasters. The 2006 Tsunami confirmed this condition, and based on seismic analysis and historical data; there is a high probability for tsunami occurrence in this region in the future.
- Cilacap is the largest city on south coast of Java. Several critical infrastructures, such as industrial areas, oil tanks, power plants, ports and hospitals, are located close to the beach.

- A great number of people are living or working near the beach, mostly in the agriculture, fisheries and tourism sector.
- In addition to Padang and Bali, Cilacap is a GITEWS pilot area, and therefore, the results of this research strongly support project needs.

2.4.1. Geographic and socio-demographic characteristics

Cilacap is the largest district in the Province of Central Java, Indonesia. The area of Cilacap region covers about 225,000 ha, and consists of 24 municipals (BPS, 2005). Cilacap city lies between 108°4'30"-109°30'30" East and 7°30'00"-7°45'20" South; it bordered by Banyumas District in the north, Kebumen District in the east, Ciamis District in the west, and the south part of this region directly faces the Indian Ocean.

Based on BPS data, the number of registered people in 2005 is approximately 1.7 million, with an average growth rate of 0.52% per year. The average population density in Cilacap is about 803 people per km².

Most employees in Cilacap District work in the agriculture sector followed by the services, trade, industry and transportation sectors (BPS, 2005). The main source of income in Cilacap is from the agriculture sector.

2.4.2. Land use structure

The variation of land use conditions in Cilacap District is high. In the west part of this district, there is Nusakambangan Island, with a hilly surface and very dense vegetation. The center part of Cilacap District is characterized by a varied settlement structure with high, medium and low density, surrounded by rural landscapes such as cropland and plantation. The flat area is crossed by various rivers and interrupted by small hilly zones.

2.4.3. Tsunami exposure

The coastal area of Cilacap is directly exposed to the Indian Ocean. Due to tectonic and bathymetric conditions, elevation on land and the recorded historical earthquake and tsunami events, the region of Cilacap is prone to tsunami hazard (see Figure 2.7).

The region of the plate boundary between the Australian plate and the Sunda plate has a high seismic activity. Based on the integrated tsunami database

(ITDB), between 1859 and 2007, four tsunamis and 289 earthquakes with moment magnitudes of more than 5 were detected around the Cilacap region. Based on Indonesian National Coordinating Agency for Disaster Management (BAKORNAS PBP) 2006, the 2006 Tsunami with a moment magnitude (M_w) of 7.7 killed around 650 people, injured around 520 people, and around 30 people are missing in this area and caused a lot of damage and economic loss. The event affected Ciamis, West Java (Pangandaran), Cilacap and Kebumen Districts, up to the famous beach in Yogyakarta, Parangtritis. In the district of Cilacap alone, 157 people were killed, 10 were missing, and 10 were injured in Adipala, Binangun, and Nusawungu (BAKORNAS PBP, 2006). The strongest recorded earthquake was in 1943, with a moment magnitude (M_w) of 8.1; however, no tsunami was reported.

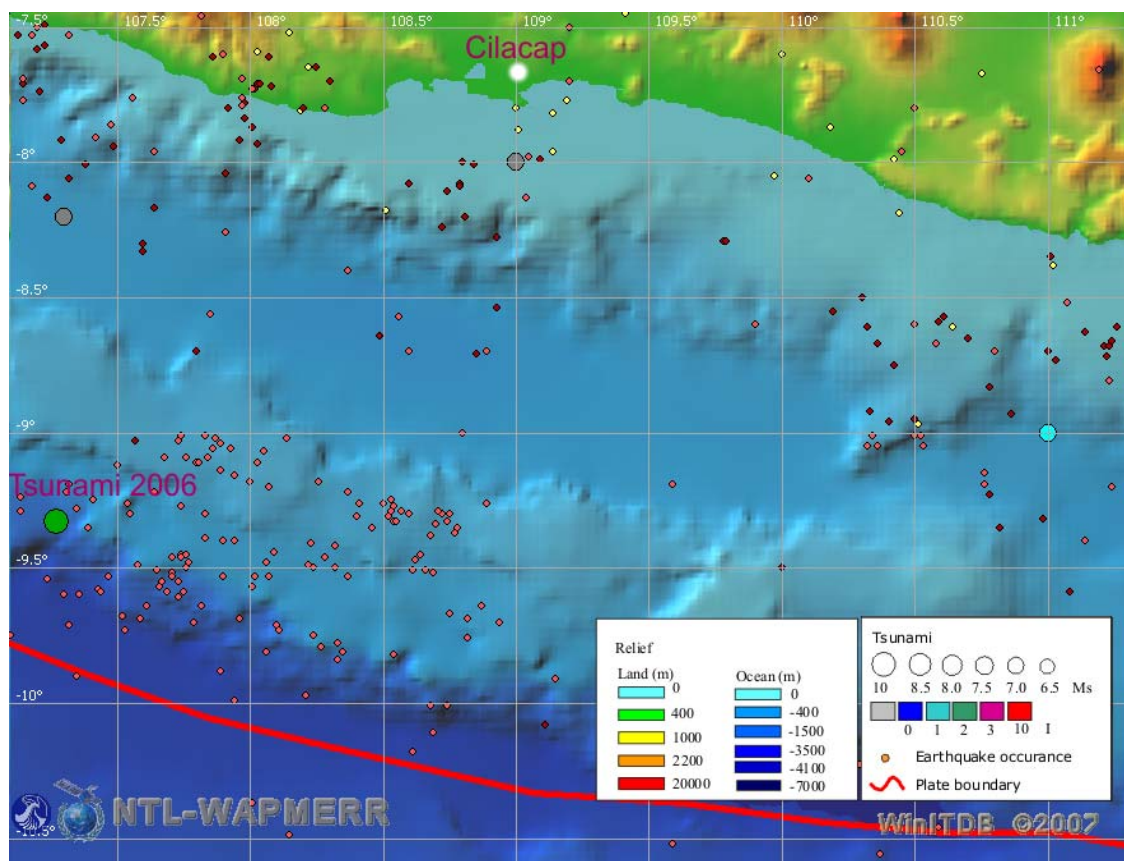


Figure 2.7 The bathymetry, the subduction zone, and the historical tsunami and earthquake occurrences in the Indian Ocean around the Cilacap region (ITDB/WLD, 2007)

CHAPTER 3: STATE OF THE ART OF RESEARCH

This chapter is designed to analyze the current status on research and development on surface roughness determination, population distribution mapping, and settlement classification. The current status is critically described and needs for further research works are identified.

In the first sub-chapter, a review of the current status of research on the application of remote sensing and GIS in the field of vulnerability and risk assessment as well as remote sensing classification techniques is given.

In the second sub-chapter, the current status of research on surface roughness in relation to tsunami flow resistance is reviewed. It provides an overview of the factors affecting flow resistance and illustrates how vegetation is one important factor. The role of vegetation and its characteristics to reduce tsunami impact, the analysis of surface roughness, including the use of remote sensing techniques, and the estimation of surface roughness coefficients are explained.

The third sub-chapter summarizes the current status of research on the application of remote sensing and geographical information systems (GIS) in the field of population distribution modeling. Top-down and bottom-up approaches are explained.

The fourth sub-chapter describes the current status of research on settlement classification using remote sensing data. The chapter includes optical as well as SAR imagery applications, and the fusion of both kinds of data.

The final sub-chapter concludes the state of the art of research and explains the needs for further research, to which this thesis will contribute.

3.1. Remote Sensing and Tsunami Risk and Vulnerability

3.1.1. Overview

For almost 40 years, remote sensing has been used for monitoring the Earth's surface, its conditions and changes. In the field of disaster management, in particular, remote sensing can help to analyze areas that are prone to natural and man-made hazards and potential damages. Risk and vulnerability assessments are an important part of disaster management and can be supported by remote sensing for pre-disaster analyses.

Regarding tsunami risk and vulnerability assessment and modeling, remote sensing techniques have been used for the following applications throughout the disaster cycle (see Figure 3.1):

- Damage assessment and rapid mapping to support the emergency response phase immediately after a disaster has occurred
- Contributions to vulnerability and risk assessment in the pre-disaster phase by deriving relevant information such as land use, settlement areas and buildings, elevation, etc.
- Monitoring of reconstruction and rebuilding in the post-disaster phase.

Specific remote sensing classification techniques are required to fulfill the demands of tsunami risk vulnerability analysis. Hence, this sub-chapter reviews the current state of the art of such techniques.



Figure 3.1 The disaster management cycle and remote sensing contributions

3.1.2. Damage assessment

Shortly after the Indian Ocean Tsunami 2004, high-resolution images acquired before and after the tsunami and damage assessments based on such images had been released to the relief organizations and centers. This was very helpful to describe the impact and extent of tsunami occurrence, for example, in Banda Aceh. It showed that remote sensing data is useful for disaster analysis; its application is not limited to tsunami disasters, but also for other kinds of natural and man-made disasters such as floods, landslides, forest fires and oil spills.

In general, there are two main goals using remote sensing data for analyzing damage: rapid mapping assessment (e.g. Belward *et al.* 2007) and mapping the affected hazard impact zone (e.g. McAdoo *et al.* 2007; Iverson and Prasad, 2007; Chatenoux and Peduzzi, 2007). The latter is very important in risk and vulnerability assessment in order to know the important parameters for tsunami hazard mapping. For example, the tsunami run-up is influenced by the topography of a region (McAdoo *et al.* (2007), the geomorphologic conditions, mangroves, and coral reefs influenced the characteristics of the tsunami inundation in Aceh (Chatenoux and Peduzzi, 2007), and the landscape analysis is used to model the tsunami damage in Aceh Province (Iverson and Prasad, 2007).

The damage assessment research is vital for improving the tsunami hazard mapping, and for analyzing what kinds of parameters influence the tsunami run-up and inundation. The role of vegetation on tsunami impact was examined through the analysis of the damaged areas by remote sensing and GIS, (e.g. Danielsen *et al.* (2005), Parish and Yiew (2005), Chang *et al.* (2006) and Olwig *et al.*, 2007).

3.1.3. Derivation of relevant information for risk and vulnerability

Several components needed for tsunami risk and vulnerability assessment, such as elevation, land use, vegetation condition and population distribution can be derived from remote sensing data. Willige (2008) assessed tsunami hazard in the northern Aegean Sea by using the Digital Elevation Model (DEM) of the Shuttle Radar Topographic Mission (SRTM) and land use from LANDSAT TM. She analyzed the potential flooding area by tsunami using GIS modeling. Another application has been shown by Post *et al.* (2009). They used land cover, slope and population data to analyze evacuation planning strategies to minimize the impact of tsunami hazards using GIS. Both studies have shown the important role of remote sensing in tsunami risk and vulnerability modeling.

Moreover, the development of remote sensing technology has the potential to identify and assess objects, even on a detailed scale, e.g. the condition of buildings. The direct assessment of the vulnerability of objects such as buildings is not possible (e.g. Mueller *et al.*, 2006). However, relevant parameters can be derived from remote sensing data, which can be used as input for vulnerability assessment as well as to identify buildings for potential vertical evacuation as a basis for developing an appropriate evacuation planning strategy.

Another important component that can be derived from remote sensing data, especially for tsunami and flood hazard mapping, is surface roughness condition. Several studies have been undertaken; this issue is addressed in detail in chapter 3.2.

3.1.4. Monitoring of reconstruction and rebuilding in the post-disaster phase

Remote sensing data is also used to monitor the progress of post-disaster environmental and infrastructure rebuilding activities in a cost-effective manner (Friesecke, 2006). An example of this application was shown by Shofiyati *et al.* (2005), who assessed the land use damaged by the Indian Ocean Tsunami 2004 for Aceh, Indonesia, especially in agriculture. The information was used for soil reclamation and spatial planning in the coastal area. Another application was shown by Adams *et al.* (2004), who developed a remote sensing technique to monitor the reconstruction of buildings collapsed during earthquakes in Boumerdes, Algeria, and Bam, Iran.

Both studies demonstrated the effectiveness of using remote sensing data for the monitoring of reconstruction and rebuilding activities in a disaster-damaged area. Remote sensing can thus significantly support disaster management activities in the recovery phase of the disaster cycle (Figure 3.1).

3.1.5. Remote sensing classification techniques

The quality of tsunami risk and vulnerability analysis by using remote sensing depends on the accuracy of the classification technique. Until now, the accuracy of such techniques is already well developed, and research is still ongoing to achieve further improvements. Current developments on remote sensing classification techniques are reviewed below.

Remote sensing classification has made great progress over the past decades mainly in the following three areas (Lu and Weng (2007):

- development of advanced remote sensing classification techniques;
- integration of multiple remote-sensing features in the classification process;
- integration of ancillary data and knowledge into the classification procedures.

Conventional methods such as the maximum likelihood classification to the more advanced object-based classification have been used and developed with great success. This advancement is in line with the progress of remote sensing sensor technology that led to high spatial resolution to below 1 m. Object-based image classification, for example, was developed for high resolution image classification (e.g. Blaschke, 2009). From this comprehensive literature review it can be concluded that this methodology represents a significant trend in remote sensing and GIS science. In contrast, in low- and medium- resolution remote sensing data such as NOAA AVHRR and MODIS, one pixel usually contains several objects of the earth surface and heterogeneous information (mixed pixels). Spectral mixture analysis (SMA) was recognized as an effective method for dealing with the mixed pixel problem (Lu and Weng, 2007).

The integration of multiple remote sensing features including spectral, spatial, multi-temporal, and multi-sensorial information are important for improving the classification accuracy. One example of this application is data fusion of multi-spectral and panchromatic bands ("pan-sharpening"). Using multi-temporal imagery is another useful method to classify objects such as water, bare land or paddy field. As well known, paddy fields have three important phases: the water phase, the vegetative phase and the bare phase. By using a single image, a paddy field in the water phase will be classified as water, or in the vegetative phase, as vegetation. The use of multi-temporal imagery can improve the classification results of such features (e.g. Xiao *et al.*, 2006). Hyper-spectral data can help to classify objects in more detail due to narrow spectral bands. As a result of specific spectral properties, it is possible to distinguish between different types of vegetation, phytoplankton in the ocean, and soil (e.g. Belluco *et al.*, 2006).

The integration of ancillary data such as digital elevation models, population density, road network, soil type, temperature, or precipitation, can also improve the accuracy of the classification results. For example, Stathakis and Kanellopoulus (2008) incorporate global elevation ancillary data for land use classification using granular neural networks. Liu *et al.* (2008) integrated GIS

data to monitor mangrove forest changes with decision tree learning. They demonstrated that the use of the decision tree method on a combined dataset of multi-temporal Landsat TM Images and GIS data can be effective in delineating spatial and temporal distribution of mangrove forest. Moreover, classification accuracy can also be increased by including human knowledge, or “knowledge-based classification” (e.g. Hung and Rid, 2002; Judex *et al.*, 2006).

3.1.6. Summary of remote sensing and tsunami risk and vulnerability

Remote sensing methodologies are useful during all phases of the disaster cycle (Figure 3.1), especially for damage assessment, mapping of risk-relevant information and monitoring of post-disaster reconstruction and rebuilding. For these applications, the choice of an appropriate remote sensing classification technique is important to achieve adequate accuracy of the results.

There are still open issues and challenges in improving these techniques for several purposes, namely, for determining the surface roughness condition, the population distribution, and the settlement classification. Hence, sections 3.2 to 3.4 will review the state of the art of these components in order to assess in detail the pending issues of current methods and necessary improvements.

3.2. Surface Roughness Determination

3.2.1. Overview

Surface roughness can be defined as the smoothness and coarseness of the earth’s surface related to water flow resistance. Smooth earth surfaces, such as open land and water, have low resistance, and coarse earth surfaces, such as dense vegetation and dense residential area, have high resistance. In tsunami inundation and flood modeling, water resistance is described by a roughness coefficient; Manning’s roughness coefficient is the most widely used.

Manning’s roughness formula is stated as follows:

$$V = \frac{kR^{2/3}S^{1/2}}{n} \quad (3.1)$$

Where:

V is the cross-sectional average velocity (m s^{-1})

- k is a factor to keep the equation dimensionally correct ($m^{1/3} s^{-1}$)
- n is the roughness coefficient (independent of units)
- R is the hydraulic radius (m)
- S is the slope of the water surface or the linear hydraulic head loss (dimensionless).

In applying Manning's roughness formula, the greatest difficulty lies in determining coefficient n , since there is no exact method of selecting the n value (Chow, 1959). The estimation of the surface roughness coefficient will be further discussed in sub-chapters 3.2.3. Before that, the factors affecting the roughness coefficient shall be first described.

Cowan (1956) formulated as follows:

$$n = (n_0 + n_1 + n_2 + n_3 + n_4)m \quad (3.2)$$

Where:

- n is the roughness coefficient
- n_0 is a basic n value for straight, uniform, smooth channel in the natural material involved
- n_1 is a value added to n_0 to rectify the surface irregularity
- n_2 is a value for variation in shape and size of the channel cross section
- n_3 is a value for obstruction
- n_4 is a value for vegetation condition
- m is a correction factor for meandering channel.

This formula was developed for flood modeling of channel and river flow. For tsunami events that struck lowland coastal areas, it could be modified as suggested by Morin *et al.* (2000):

$$n^2 = n_0^2 + n_4^2 \quad (3.3)$$

This formula shows that the vegetation condition is the most important for determining a roughness coefficient. Recent developments also include building density for calculating the roughness coefficient (Koshimura, 2009). The roughness coefficient formula based on building condition is as follows (Dutta *et al.*, 2007 and Koshimura *et al.*, 2009):

$$n = \sqrt{n_0^2 + \frac{C_d}{2gd} * \frac{\theta}{100 - \theta} * D^{4/3}} \quad (3.4)$$

Where:

- n is the roughness coefficient
- n_0 is the roughness value due to land use except building
- θ is the building/house occupancy ratio in the computational grid (23 m)
- C_d is the drag coefficient
- d is the horizontal scale of houses measured by using GIS
- D is the modeled flow depth
- g is the gravitational acceleration constant.

Based on these formulas, the two main components affecting the roughness coefficient are vegetation and building conditions. Regarding building condition, it is obvious that building density is the most important characteristic. For vegetation, some further analysis is needed, and thus, the following section is dedicated to exploring the role of vegetation and its characteristics in reducing tsunami impact.

3.2.2. The role of vegetation on reducing tsunami impact

Several studies showed that vegetation could contribute to protecting coastal areas against tsunami waves (Kathiresan and Rajendran, 2005; Vermaat *et al.*, 2006; Danielsen *et al.*, 2005; Iverson and Prasad, 2007). However, the role of vegetation for tsunami protection very much depends on, among other factors, the tsunami wave height. Empirical studies show that the protective role of vegetation is mostly effective if the tsunami wave height is below approximately 7 m. If the tsunami wave height is higher, the protective vegetation will be destroyed in most cases and will no longer have wave reduction effects (Harada and Kawata, 2004) (Table 3.1). In the Indian Ocean Tsunami 2004, for example, run-up heights were in some areas peaked at more than 20 m and the coastal vegetation, e.g. near Banda Aceh, was completely destroyed (Hiraishi, 2005).

Table 3.1 The relation between tsunami intensity and tsunami damage

Tsunami intensity	0	1	2	3	4	5
Tsunami wave height (m)	1	2-3	4-7	8	16	32
Coastal control forest	Mitigates damage, stops drafts, and mitigates tsunami		There is only partial damage and stop drafts	There is complete damage, no reduction effect		

Source: Harada and Kawata (2004)

In the case of the 2006 Tsunami in west Java, which had lower wave heights (4-10 m), the vegetation proved its ability to protect buildings near the coast (see Figure 3.1). High-resolution satellite imagery acquired before and after the tsunami in west Java shows that houses with no vegetation protection were completely destroyed, while houses with vegetation protection were only slightly damaged. Figure 3.2 shows that the vegetation was able to resist the tsunami flow and decrease the tsunami energy, and hence the building behind the vegetation has been only slightly damaged.

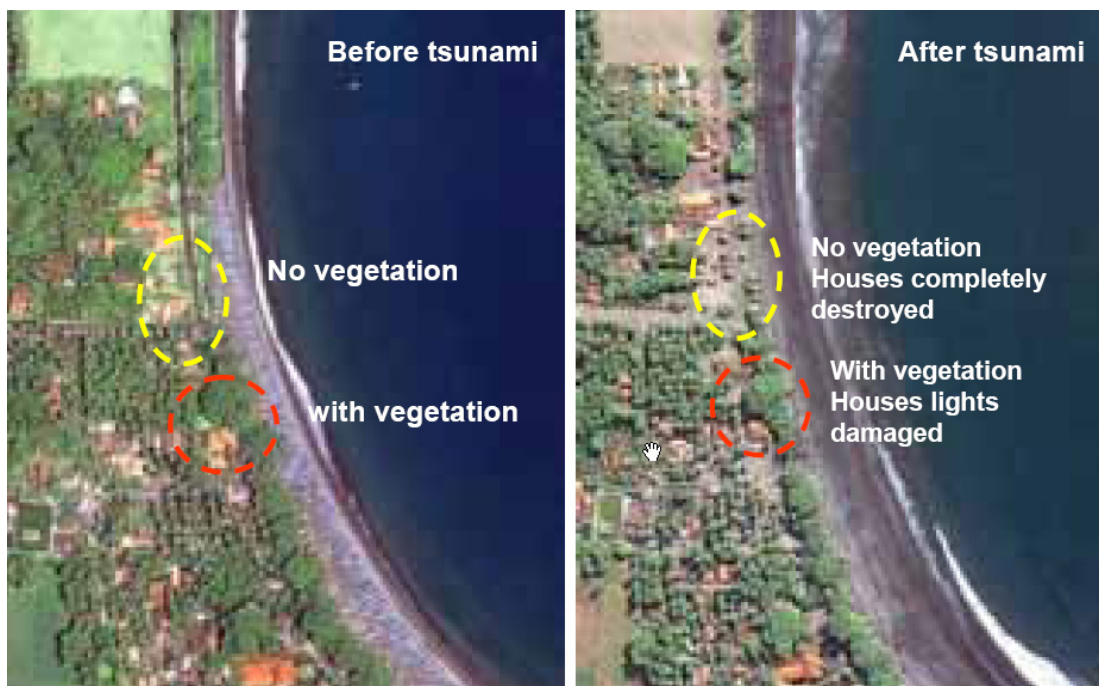
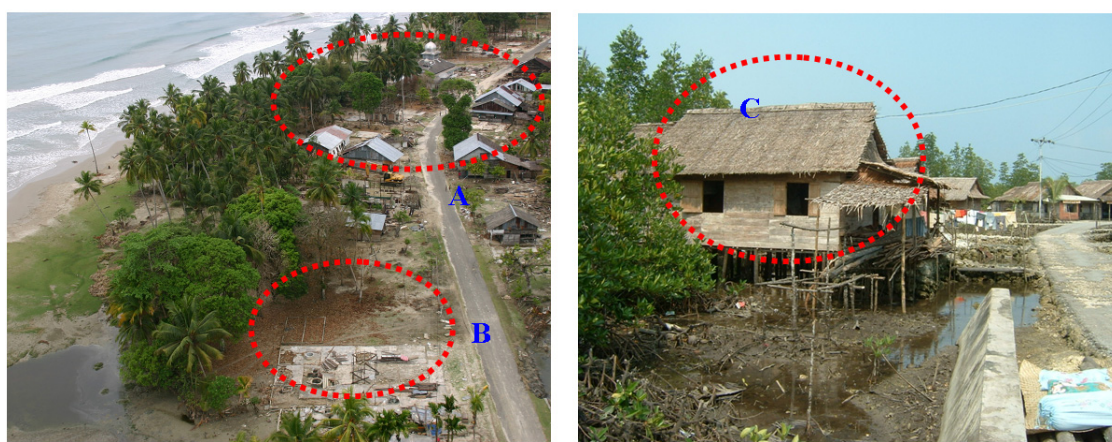


Figure 3.2 The potential of vegetation on reducing tsunami impact in Pangandaran (Source: Latief and Hadi, 2007))

The photographs in Figure 3.3 show additional examples of the protective role of vegetation in reducing the hazard impacts of the Indian Ocean Tsunami 2004 (Kongko, 2005). The left photo shows an example of the west coast of Simeuleu Island, where some houses built of wooden material (houses A) were saved from the tsunami's destructive force due to protection by a dense coconut tree zone of approximately 10 m width. There used to be a similarly built house not far from this location (B), which had no protective vegetation and was completely destroyed. The right photo shows houses built of bamboo material in Muawe-Lahewa (northern Nias Island), which was protected from the tsunami impact by dense mangrove forest.



West Coast of Simeuleu Island

Muawe-Lahewa North Nias Island

Figure 3.3 Observation pictures from the effect of tsunami 2004 in Nias Island and Simeulue (Source: Kongko, 2005)

These examples clearly show the potential of vegetation to reduce tsunami impact, but there is a need to investigate the specific characteristics of potentially protective vegetation. The characteristics of coastal vegetation that play a protective role include width, density and structure of vegetation.

Several studies deal with the correlation of the reduction rate of tsunami impact and the width of a forest zone (Harada and Kawata, 2004; Kongko, 2005), and the forest density, respectively (Harada and Kawata, 2004). The reduction rate was examined in terms of run-up, inundation depth, current and hydraulic force. The forest density proved to have a moderate influence on the reduction of inundation depth and current (Harada and Kawata, 2004), whereas the forest width has a significant influence on the reduction of run-up, current and hydraulic force (Harada and Kawata, 2004; Kongko, 2005). Figure 3.4 shows correlations between forest width and density and reduction rates.

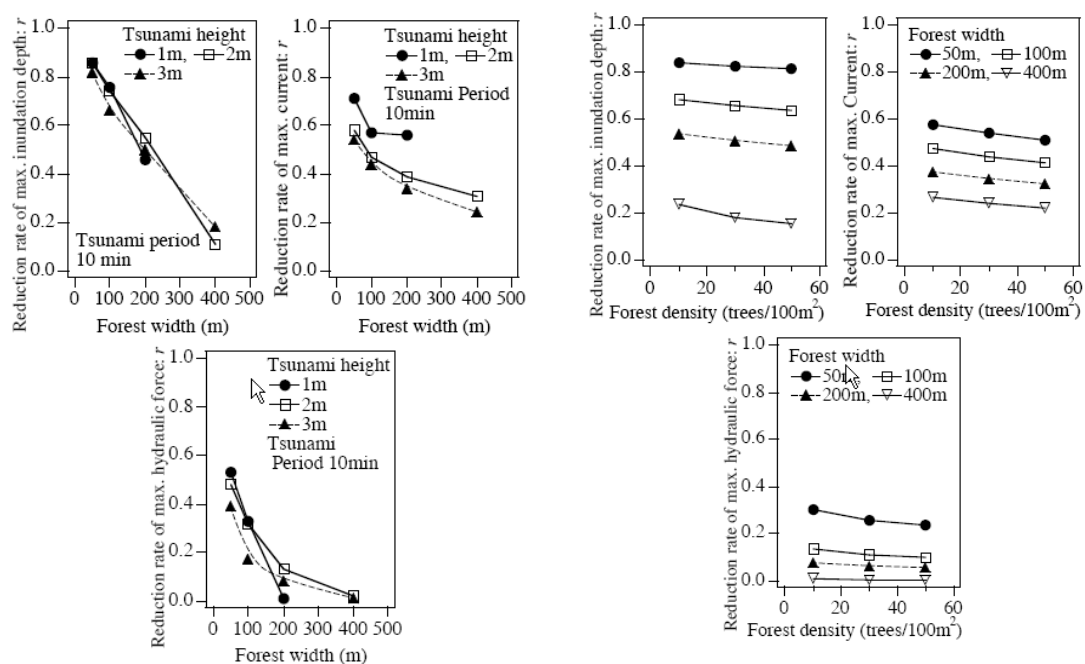


Figure 3.4 Effects of forest width and density to tsunami reduction (Source: Harada and Kawata, 2004)

Moreover, Tanaka *et al.* (2007) analyzed coastal vegetation structures and their functions in tsunami protection in Sri Lanka and the Andaman coast of Thailand. By field measurement, they analyzed the effectiveness of vegetation structure in providing protection from tsunami damage. This research showed that the neighborhood of vegetation types is an important factor regarding tsunami impact reduction. For example, both *Casuarina equisetifolia* and *Pandanus odoratissimus* have only little effect on reducing tsunami impact, but in a combined association of both plant species, the effect is strong.

It may be concluded that width, density and structure of vegetation play important roles in tsunami reduction. This can be considered when recommending the development of natural structures to defend against tsunami waves. Surface roughness is an important input parameter for tsunami modeling, which needs to be quantified properly and integrated in mathematic tsunami modeling. As mentioned, the roughness coefficient is the quantification of the surface roughness condition, and it is most difficult to determine. The following section gives more detailed information about the current state of research to determine the surface roughness coefficient.

3.2.3. Estimation of the surface roughness coefficient

Several studies on roughness coefficient estimation have been undertaken (e.g. Chow, 1959; Arcement and Schneider, 1989; Jarvela, 2005). The applied methodologies can be grouped into four categories:

- Laboratory experiments
- Site visit and photo matching
- Use of roughness coefficients from published research
- A combination of these methods.

Table 3.2 gives an overview of these methodologies and summarizes the data requirements as well as the advantages and disadvantages of each method.

Table 3.2 Methods of roughness coefficient estimation (modified from Sellin *et al.*, 2003)

Method	Data required	Advantages	Disadvantages
Laboratory experiments	The drag coefficient, sufficient water velocity scenario, and appropriate surface roughness miniature.	The most accurate method for flexible surface roughness.	Only applicable to static surface roughness conditions. Each variation of the conditions requires costly recalibration.
Site visit and photo matching	Site visit and photographs.	Quick and widely used, gives an approximate figure for any condition.	Requires experience, no variation of water level.
Use of roughness coefficients in published research	A roughness classes table with roughness coefficient values.	Systematic, familiar.	Apparent accuracy is deceptive.
A combination of these methods	A roughness classes table with roughness coefficient values and formulas from laboratory experiments.	Systematic, familiar.	No deviation of roughness coefficients.

Laboratory experiments

In flood engineering, scientists assess the roughness coefficient usually by experimental research in laboratory and field survey studies. For example, Righetti and Armanini (2002) analyzed the flow resistance on sparsely distributed bushes; Jarvela (2005) investigated the flow resistance of natural vegetation; Wilson (2007) analyzed the flow resistance models for flexible submerged vegetation; and Wilson and Horritt (2002) analyzed the flow resistance models for flexible submerged grass.

This method is the most accurate approach in determining the roughness coefficient and also the conditions of vegetation or buildings can be modeled close to reality. Scenarios of water flow and velocity can be also undertaken to determine the appropriate roughness coefficient.

Site visit and photo matching

A very quick determination of the roughness coefficient is possible by site visit and photo matching. Arcement and Schneider (1989) outline an approach to match photographs and real conditions of a study area: several photographs are provided to guide people in estimating the roughness coefficient. The researcher's experience is required in using this method. Brisbane City Council (2003) also provided several photos to estimate the roughness coefficient. Figure 3.5 shows examples of photographs and the corresponding roughness coefficient values.

The weakness of this methodology is its subjectivity: different people will categorize the roughness coefficient differently according to their experience and knowledge in interpreting the photo. Therefore, this is a fast method to determine the roughness coefficient, but with a relatively high degree of error.



Figure 3.5 Examples of photograph matching (Source: Brisbane city council, 2003)

Use of roughness coefficients from published research

Research based on laboratory experiments usually only measures a limited number of classes of surface roughness and in a costly manner. Sellin *et al.* (2003) therefore suggest using published roughness coefficients such as by Chow (1959) and Arcement and Schneider (1989). However, the problem of this approach is the limited transferability and the inconsistency of published roughness coefficients. For example, Hill and Mader (1987) determined the roughness coefficient for lake, water, river and sea at 0.007 and for cropland or grazing land at 0.015, while Murashima *et al.* (2008) determined 0.025 for lake, water, river and sea, and 0.020 for cropland or grazing land. Table 3.3 shows examples of roughness coefficients from research publications.

Combination methods

A combined approach of surface roughness coefficient determination was developed by Koshimura *et al.* (2009). The combination consists of roughness coefficients from published research and roughness coefficients calculated with a formula developed in laboratory experiments. They modified the surface roughness coefficient from published research by adding the density of a populated area as class of surface roughness. Table 3.4 shows the roughness coefficient estimation by the combination method.

Table 3.3 Roughness coefficients for different land use/cover classes in several research publications

Surface Roughness	Roughness Coefficient	References
Water		
- River, lake	0.007	Hills and Mader (1987); Berrymann (2007)
	0.025	Murashima <i>et al.</i> (2008)
- Shallow water area	0.025	Koshimura <i>et al.</i> (2009)
- Sea/pond	0.025	Latief <i>et al.</i> (2007)
Land		
- Open field without crop	0.015	Hills and Mader (1987); Berrymann (2007)
- Rice field, cropland, arable land	0.020	Murashima <i>et al.</i> (2008); Koshimura <i>et al.</i> (2009); Imamura (2009)
	0.025	Latief <i>et al.</i> (2007)
Vegetation		
- Rare vegetation	0.030	Murashima <i>et al.</i> (2008)
- Relative dense vegetation	0.050	Latief and Hadi (2007)
- Dense vegetation	0.070	Latief and Hadi (2007)
- Dense brush/tree	0.070	Hills and Mader (1987); Berrymann (2007)
- Mangrove	0.060	Latief <i>et al.</i> (2007)
- Forest	0.150	Latief <i>et al.</i> (2007)
Developed Area		
- Road	0.016	Latief <i>et al.</i> (2007)
- Developed area	0.030	Hills and Mader (1987)
- Built-up areas	0.035	Berrymann (2007)
- City center	0.100	Hills and Mader (1987); Berrymann (2007)
	0.120	Sugimoto <i>et al.</i> (2003)
- Building	0.150	Latief <i>et al.</i> (2007)
- Rare density residential area (1-20%)	0.040	Murashima <i>et al.</i> (2008); Imamura (2009)
- Medium density residential area (20-50%)	0.060	Murashima <i>et al.</i> (2008); Imamura (2009)
- High density residential area (>50%)	0.080	Murashima <i>et al.</i> (2008); Imamura (2009)

Table 3.4 The roughness coefficient estimation by Koshimura *et al.* (2009)

Surface Roughness Class	Roughness coefficient	Methods
Smooth ground	0.020	Published research
Shallow water area or natural beach	0.025	Published research
Vegetated area	0.030	Published research
Populated area	0.045	Published research
Dense populated area	Equation 3.4.	Equation developed in laboratory experiments

3.2.4. Roughness coefficient in tsunami modeling

Researchers have studied the implementation of the roughness coefficients in the tsunami modeling, e.g. Dao and Tkalich (2007), who examined the sensitivity of tsunami propagation for the 2004 Indian Ocean Tsunami. Their results indicated that in shallow water, the roughness coefficient is an important parameter in the tsunami modeling. Lower roughness coefficients will significantly increase the tsunami wave.

In a previous study, Myers and Baptista (2001) examined the factors influencing the numerical simulations of tsunami, and their implication for hazard mitigation. One of the important factors in their research is roughness. They analyzed the differences in run-up modeling results when changing the roughness coefficient. The original roughness coefficient was 0.0275, and the modifications were 0.015 and 0.035. This research demonstrated the important influence of the roughness coefficient on tsunami modeling results: in the three segments of the research along the Okushiri coastline, the tsunami run-up differs from – 6 to +6 m by changing the roughness coefficient (see Figure 3.6).

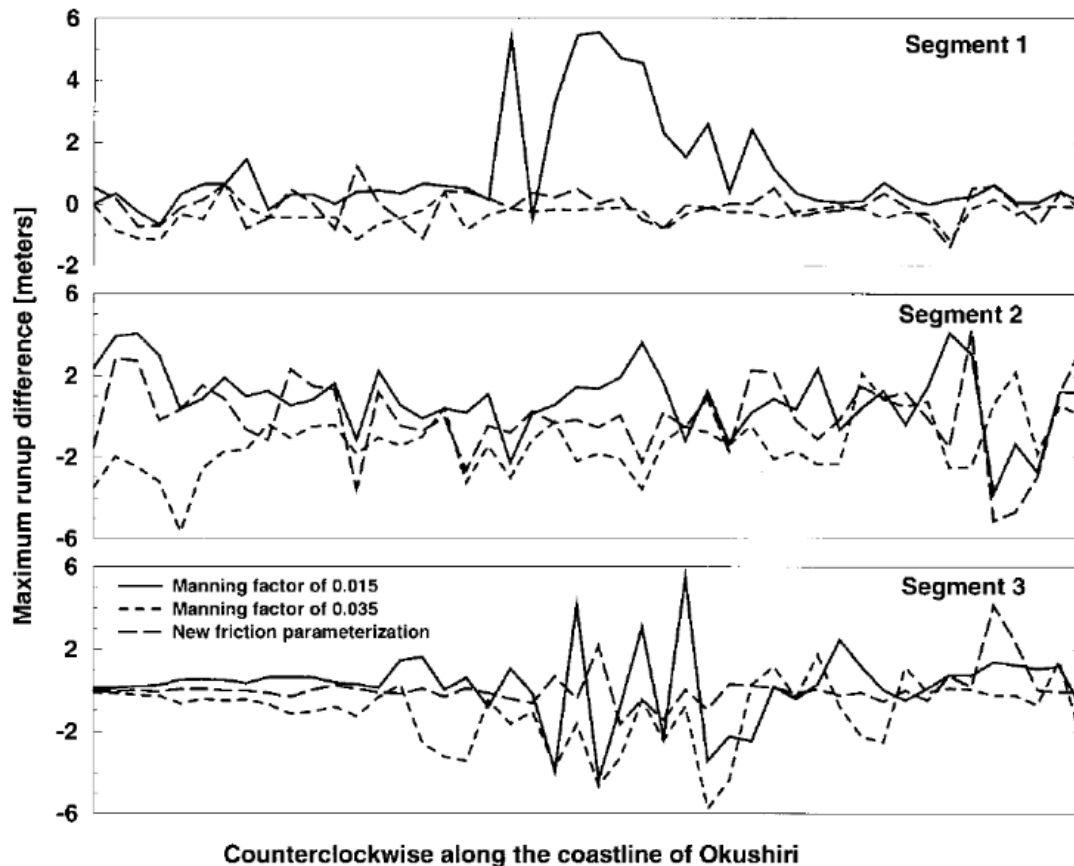


Figure 3.6 The maximum run-up differences (m) by changing roughness coefficient (Source: Myers and Baptista, 2001)

Another study related to the effect of roughness coefficients on the tsunami modeling was published by Gayer *et al.* (2008), who analyzed the importance of the roughness coefficients in tsunami modeling. Gayer *et al.* (2008) revealed the importance of using spatially differentiated roughness maps. Similar to findings of Myers and Baptista (2001), the roughness coefficient was varied. A higher coefficient resulted in considerably reduced inundation area and water velocity.

Based on the results of the mentioned research, it is clearly shown that the roughness coefficient is a significant factor in tsunami modeling. Modifying the roughness coefficient in this modeling will lead to significantly different results of tsunami inundation. However, most of the studies use a uniform roughness value on the tsunami inundation and the modeling scenario merely adjust it with another value. In fact, surface roughness conditions in an inundation area differ spatially. Instead of applying uniform roughness in the modeling, it is therefore necessary to implement spatially differentiated surface roughness

conditions. Remote sensing data can provide information about the spatial distribution of surface roughness conditions and are thus a tool to fill this gap.

3.2.5. Remote sensing approaches for surface roughness classification

For remote sensing applications on surface roughness classification, there are two main activities:

- Mapping of the tsunami inundation area to empirically analyze the effects of surface roughness in order to reduce tsunami impact;
- Classification of land use to derive surface roughness conditions related to tsunami or flood modeling.

For the first application, Chang *et al.* (2006) and Olwig *et al.* (2007) showed the potential of dense vegetation in Thailand and Indonesia to reduce the tsunami flow. Comparisons of pre- and post-tsunami satellite images provided information on inundated areas and the vegetation conditions at particular regions. The lesson learned from the remote sensing data show the reduction of tsunami flow by wood vegetation (Olwig *et al.*, 2007), and mangroves (Chang *et al.*, 2006). Both studies showed the effect of surface roughness qualitatively on the tsunami reduction without a quantitative estimation of the roughness coefficient.

For the second application of surface roughness classification, a variety of studies have been undertaken using remote sensing data, such as ASTER (Chiang *et al.*, 2004), IKONOS-2 (Van der Sande *et al.*, 2003; Koshimura *et al.*, 2009; SAR (Mason *et al.*, 2009), Schumann *et al.*, 2007), Landsat TM (Melese, 2003), airborne laser scanning (Straatsma and Baptist, 2008) and LIDAR (Murashima *et al.*, 2008). Almost all of these studies classify the land use class by using conventional remote sensing classification techniques. The estimation of the roughness coefficients is based on published research results.

The surface roughness classes contain specific characteristics, i.e. density and neighborhood, which means that the classification methodology should be different to the usual technique. The parameters describing the vegetation density and the neighborhood characteristics should be taken into account in the remote sensing classification methodology. Therefore, there is a need for developing more advanced classification methods that take into account these additional parameters.

A new application of surface roughness classification by using remote sensing is introduced by Koshimura *et al.* (2009), which considers a density analysis. Their study indicates the importance of building density on roughness coefficient estimation, but it still does not consider neighborhood classes, which are a desirable improvement of the remote sensing classification technique for surface roughness classification. The study by Koshimura *et al.* (2009) is further limited by the methodology of classification used. The classification by visual interpretation and manual digitization is time-consuming and ineffective. It needs to be improved by automatic or semi-automatic classification.

3.2.6. Summary of surface roughness determination

Research on the role of vegetation in protecting against or mitigating tsunami impact proved that coastal vegetation — depending on its condition (width, density, and structure) — may be able to reduce the impact of a tsunami. The flow resistance of the surface can be described by a roughness coefficient. To quantify the influence of surface roughness conditions on tsunami inundation modeling, and thus to perform sensitivity analyses, it is necessary to include surface roughness coefficients. There are several approaches to estimate the roughness coefficient; one uses published values of roughness coefficients. The accuracy of the thus estimated coefficient is often deceptive; this shortcoming shall be solved in this research. Previous research on remote sensing techniques to classify surface roughness conditions needs further improvement, since in general; only land use conditions are classified. Analyses of vegetation characteristics showed that the components vegetation density and neighborhood should also be taken into consideration. Hence, a new classification methodology is required since an integration of surface roughness coefficients in tsunami inundation modeling is necessary to improve the modeling results.

3.3. Population Distribution Modeling

3.3.1. Overview

Research on population distribution with respect to remote sensing and GIS can be categorized into two categories: top-down and bottom-up approaches. Top-down approaches usually disaggregate global information to a detailed scale, whereas bottom-up approaches generate global information out of detailed scale information.

Top-down approaches use statistical data and disaggregate it spatially to derive the population's distribution. To increase the spatial resolution, ancillary data such as land use are used. There are two main methods — spatial interpolation and dasymetric mapping.

- Spatial interpolation disaggregates tabular data or statistic data to geospatially distributed information through various interpolation approaches.
- Dasymetric mapping is used to redistribute the number of people in an administrative unit to smaller units by using ancillary data such as land use.

Bottom-up approaches analyze the conditions of texture, impervious surface and other structures that can be derived from remote sensing data to estimate and extrapolate population distribution. Statistical data as well as sampling data from surveys are used for regression and correlation analyses to estimate the number of people in an administrative unit. To date, estimations of the number of people in an administrative unit on small scales such as country or province yields good results with high accuracy; for large scales, however, this approach still needs some methodological improvements. Figure 3.7 summarizes the state of the art on population distribution modeling by using remote sensing and GIS.

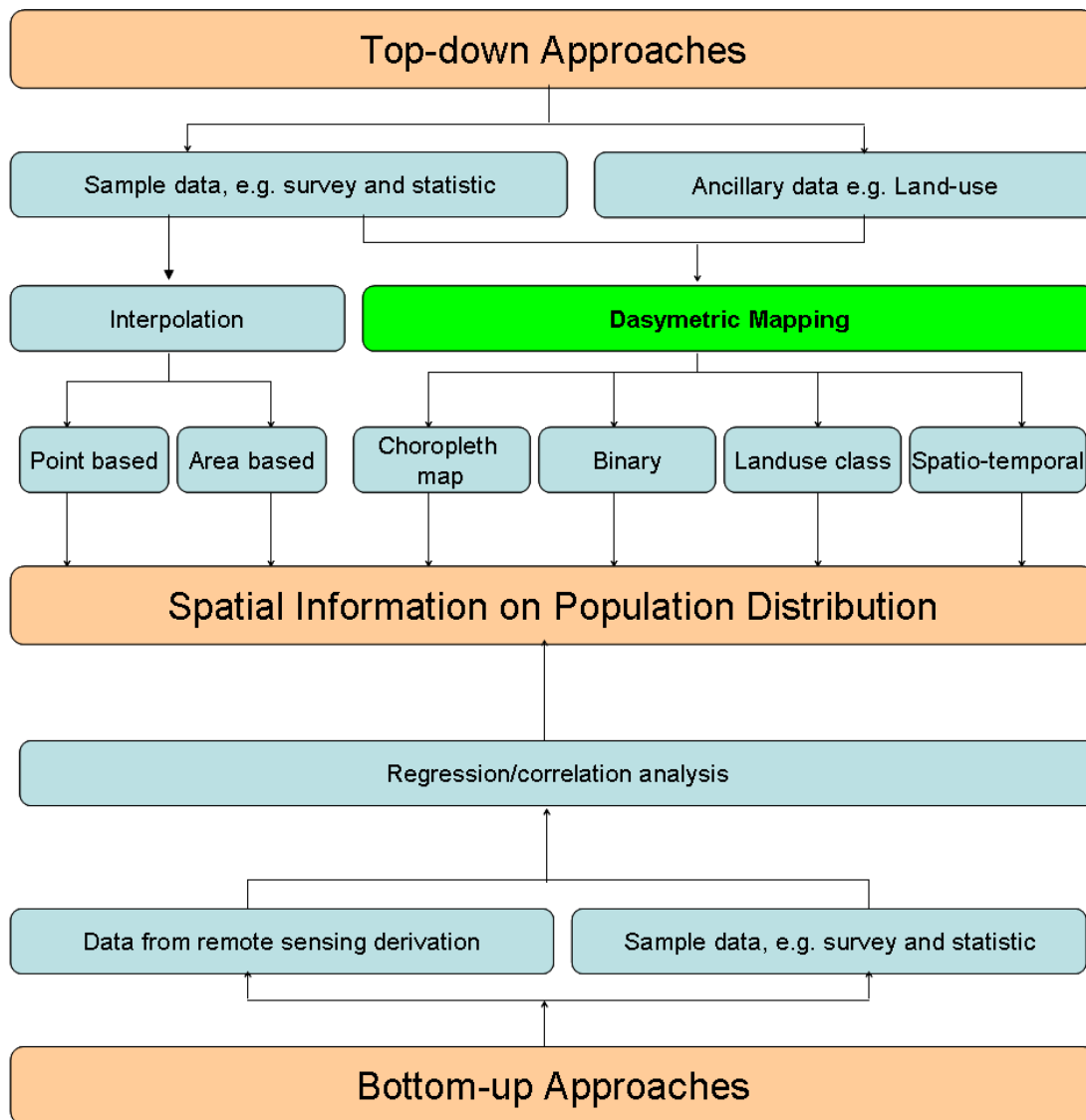


Figure 3.7 The top-down and bottom-up approaches on population distribution by using remote sensing and GIS

3.3.2. Top-down approach

Spatial interpolation

The main issue of spatial interpolation is the procedure of estimating the value of properties at unsampled sites within the area covered by existing observation. Spatial interpolation methods to be used for population distribution can be categorized in two groups:

- Point-based interpolation

- Area-based interpolation.

Traditionally, the point-based method is applied to isarithm, i.e. contour line mapping, and the area-based method, to isopleths mapping (Lam, 1983). The point-based interpolation method is derived from available data such as number of people from census data to spatial information of population distribution. Several point-based interpolation methods have been developed, for example, kriging, inverse distance weighting (IDW), natural neighbor and minimum curvature. Numerous algorithms for point interpolation have been developed in the past, but none of them is superior to any others for all applications (Lam, 1983).

A widely quoted point-based interpolation method for population distribution mapping is centroid zone (Martin and Bracken, 1991). This study describes a simple algorithm that uses census data centroid with a spreading function to allocate people to neighboring grid squares (Langford *et al.*, 1991). This method allows congregating the number of people to the centroid, and when the location of cell far from the centroid has fewer numbers of people. Normally, the centroid data that can be used for this method are geo-referenced postcode data. The estimation of the number of population is considered from the centroid zone of the available data; the formula is stated as follows:

$$\hat{P}_i = \sum_{j=1}^c P_j W_{ij} \quad (3.5)$$

Where:

\hat{P}_i is the estimated population of cell i

P_j is the population at centroid j

W_{ij} is the unique weighting of cell i with respect to centroid j

c is the total number of data centroids.

The unique weighting of cell i with respect to centroid j is calculated by the distance between the center of cell i and centroid j. This value disaggregates the total number of population recorded in the centroid to the grid cell. The formula of this weighting is stated as follows:

$$W_{ij} = \left(\frac{k^2 - d_{ij}^2}{k^2 + d_{ij}^2} \right)^\alpha \quad (3.6)$$

Where:

k is the initial kernel width; this determines the maximum width of the kernel, and a value should be chosen that is representative of the width of a typical reporting area, for available centroid location

d_{ij} is the distance between the center of cell i and centroid j

α offers control over the shape of the distance decay function within the extent of the spatial kernel, although alternative distance decay functions could be readily specified, the simple $\alpha = 1$.

In order to preserve the total of population, then

$$\sum_{i=1}^{k_j} W_{ij} = 1 \quad (3.7)$$

Where:

k_j is the number of cells within the window.

Figure 3.8 illustrates the point-based interpolation by Martin (2006).

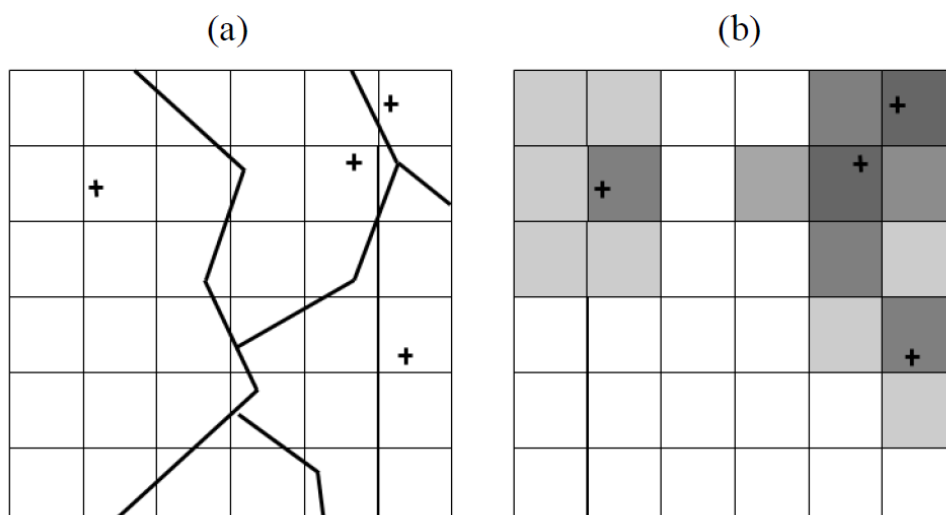


Figure 3.8 General relationship between (a) zone boundaries and population-weighted centroids, and (b) population-weighted centroids and gridded population estimates (Source: Martin, 2006)

Although this method is well developed, Wu *et al.* (2005) revealed some problems for the population distribution analysis:

- The centroid determination is unclear. This is the main source of errors on the population distribution model by this method. The centroid in the Martin and Bracken (1991) method usually has the highest number of people, but this is not true in reality. Using the geo-referenced postcode data does not ensure an aggregated location of people.
- The arbitrary assumption on the interpolation methods is inadequate to explain the population distribution in the real world.
- Most of those methods do not preserve the total value of population distribution in the source of data.

The second interpolation method for population distribution is the area-based method. Tobler's (1979) pycnophylactic interpolation is probably the most widely quoted area-based interpolation method (Wu *et al.*, 2005). This procedure generates a smooth surface from polygon-based data, which preserves their mass or volume property (Comber *et al.*, 2008). This procedure is based on the iterative weighting average of nearest neighbors. It is a mathematical smoothing process that considers the conditions of neighboring polygons. Figure 3.9 illustrates the steps of pycnophylactic interpolation with 25 iterations.

Although the results of population distribution mapping by both spatial interpolation techniques are smooth and nice-looking, the accuracy of the estimation is low and the population distribution not close to reality. Population distribution depends on human activity and has dynamic characteristics; hence, its methodology should consider these. For tsunami early warning, temporally differentiated information on population distribution (day- and night-time) is necessary. Therefore, dasymetric mapping is selected in this research as the most appropriate method.

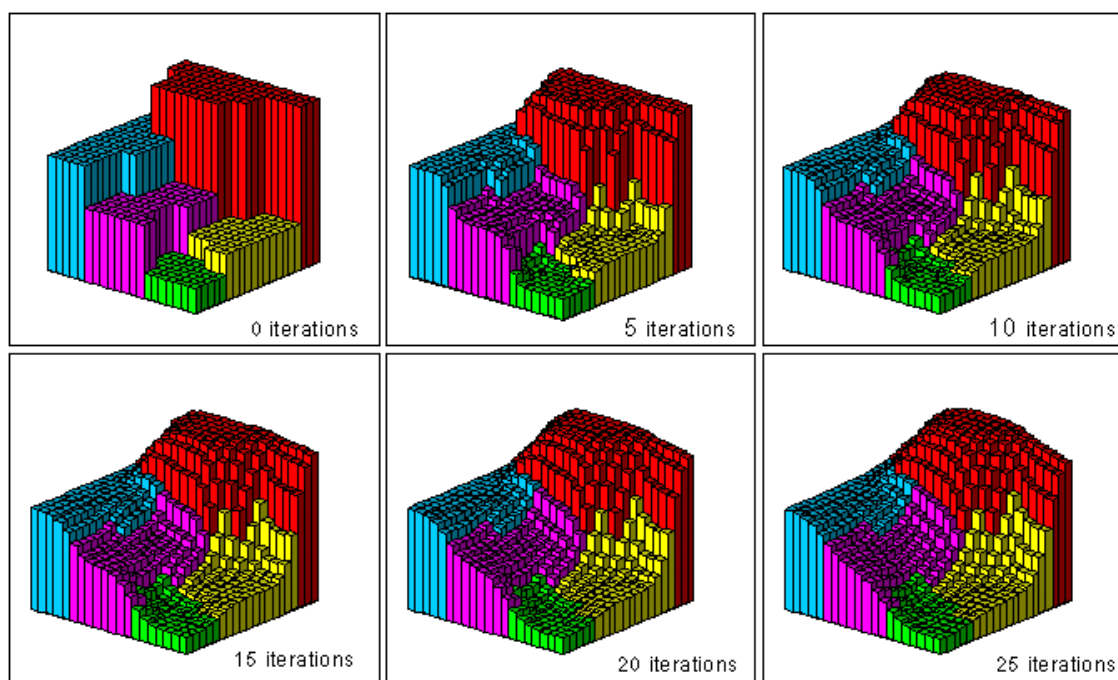


Figure 3.9 The smoothing process of pycnophylactic interpolation until 25 iterations (Source: National Center for Geographic Information and Analysis University of California Santa Barbara based on Tobler, 1979)

Dasymetric mapping

Dasymetric mapping is an area-based thematic mapping method that generates more detailed spatial information by combining global area information such as administrative units with ancillary data. This makes this method different to areal interpolation. Spatial improvement of population distribution data is one example of this technique. It is a fundamental technique to indicate where people are located by considering where they are engaged in activity. It disaggregates population data (e.g. census) to geographical features such as settlement and land use to have fine spatial information. Dasymetric mapping usually employs land use/land cover data extracted from remote sensing data (Liu, 2004). The method is well developed and used by many applications in the research of population distribution mapping.

The method was first developed and named by the Russian Cartographer Pyotr Semenov, who in the 1920s, developed and published a multi-sheet population density map of European Russia at a scale of 1:420 000 (Bielecka, 2005). The dasymetric method was created and used due to the need for accurate visualization methods of population data. Today, dasymetric mapping is widely used for producing finer resolution population information with automatic tools such as GIS and remote sensing.

There are mainly four approaches used to map the population distribution by dasymetric mapping:

- The Choropleth map. One very simple and pragmatic technique is to divide the number of people from census data by the area of administrative units.
- Binary method. A more advanced approach is to distribute the census population data to inhabited settlement areas only (Wright, 1936; Langford and Unwin, 1994; Fisher and Langford, 1996; Holt *et al.*, 2004; Wu *et al.*, 2005; Mennis, 2003; Reibel and Agrawal, 2007; Tatem *et al.*, 2007).
- Land use class method. This method has been improved by the techniques to disaggregate census population data by using land use classes (Holloway *et al.*, 1997; Sleeter and Gould, 2007; Bielecka, 2005; Gallego and Peedell, 2001). Also road networks or terrain characteristics are included in the analysis to make sure to exclude areas unlikely used for activity like steep slopes (Dobson *et al.*, 2000; Balk *et al.*, 2005, and Schneiderbauer and Ehrlich, 2007).
- Spatio-temporal method. The most complex procedure on dasymetric mapping is spatio-temporal population distribution (Sleeter and Wood, 2006; Ahola *et al.*, 2008). It shows the differences of population distribution by time, e.g. day and night.

The illustration of the four methods on dasymetric mapping is shown in Figure 3.10. Figure A represents the total number of people, aggregated by census delineated unit; Figure B represents settlement (red) versus uninhabited (green) area with population evenly distributed within the settlement area; and Figure B (below) represents an urban 3-class method where population is distributed in high (red), medium (pink), and low (yellow) with low, medium, and high density of settlement as weighting. Figure C represents the population distribution based on land use classes and areal weighting, which consider, e.g. people activities on agricultural areas (light green). The last approach considers the population distribution by time, such as day and night (Figure D). It uses land use classes and models, e.g. that people during the nighttime stay mainly in the settlement area, and during the daytime, are mostly in the agriculture area.

In general, the main challenge in dasymetric mapping is determining the appropriate weighting factors for the land use and land cover (LULC) classes. The weighting factor is difficult to assess and is determined in many different ways. For example, the weighting factors that are applied by Holloway *et al.*

(1997) estimate that urban areas naturally have the main share of people's habitations, while other land use classes like agriculture, open land or forest areas only account for a minor part. Dobson *et al.* (2000) and Schneiderbauer (2007) use similar approaches, also taking into account topographic constraints (e.g. terrain slopes).

The above-mentioned studies provide weighting factors with respect to logical considerations and to the proportion of the land use classes, but they do not consider people's behavior based on their activity. In the focus of natural hazard risk assessment, an improvement of these methodologies is needed to provide an appropriate assessment of the weighting factors to derive the information about the people at risk as precisely as possible. For this task, statistical data about human activities in certain administrative units (villages) can be consulted to derive appropriate and reliable weighting factors.

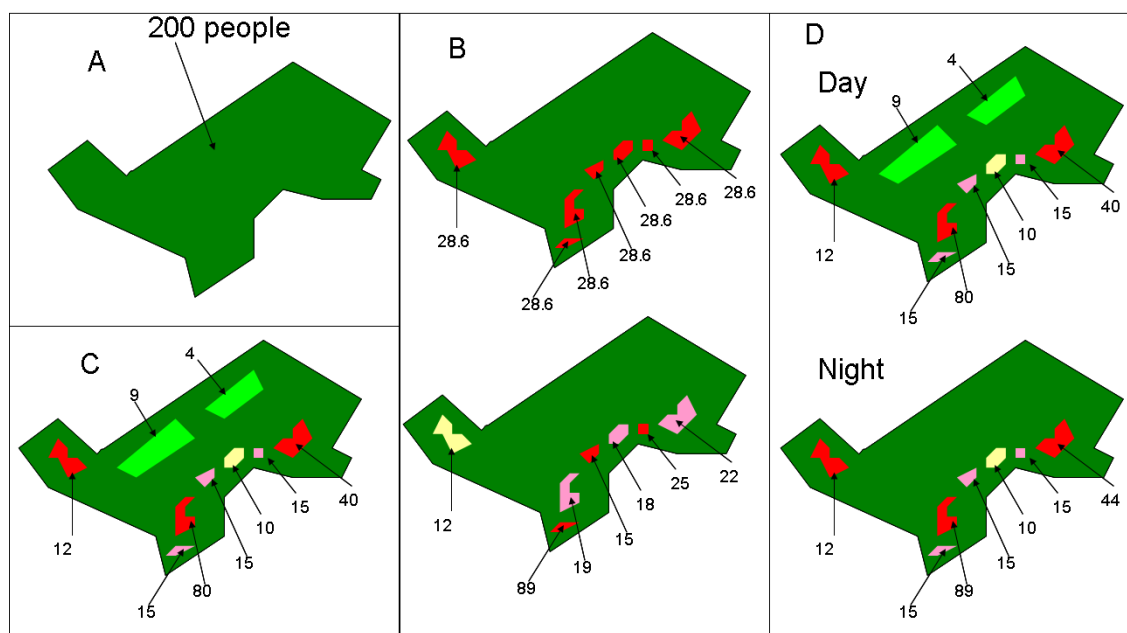


Figure 3.10 Fundamental approaches to population distribution mapping, modified from Sleeter and Gould (2007)

To better understand dasymetric mapping and outline its problems, several methods used for population distribution mapping are reviewed:

- the Mennis (2003) approach, which represents the binary method;
- the Gallego and Peedell (2001) approach, which represents the land use data method, and the Schneiderbauer and Ehrlich (2007) and

Dobson *et al.* (2000) approaches, which represent the land use data method with additional ancillary data;

- the Sleeter and Wood (2006) and Ahola *et al.* (2008) approaches, which represent the spatio-temporal method.

Mennis (2003) proposes a comprehensive method to disaggregate numbers of people from statistical data, taking into account the use of inhabited areas. He determines the proportion of human activities in several land use classes as low and high by sampling the number of people who are carrying out their activities in inhabited classes based on the "block group". For Mennis (2003), the hierarchy unit from large to small is the county, the block group, and the grid cell. The grid cell has three classes — non-settlement, low density settlement and high density settlement. The population in a grid cell of an inhabited area is calculated as follows:

$$POP_{ubc} = (f_{ubc} \times POP_b) / n_{ub} \quad (3.8)$$

Where:

POP_{ubc} is the population assigned to one grid cell of settlement class in block group b and in county c

POP_b is the population of block group b

f_{ubc} is total fraction of each settlement class in block group b and in county c

n_{ub} is the number of grid cells of settlement class in block group b.

By this formula, if the number of people in the block group b in a county c is 100 people, the fraction for each settlement class (f_{ubc}) — non-settlement, low, and high — is 0.022, 0.195, and 0.783, and the number of grid cells of that settlement class is 1, 3, 6, respectively. Then, the population density is 2.2, 6.5, and 13.5 people per grid cell, respectively.

The crucial point of population distribution mapping in this method is the fraction for settlement class (f_{ubc}). This fraction was developed from the analysis number of samples in a county. The step-by-step formula is stated as follows:

1. Calculate the population density fraction:

$$d_{uc} = p_{uc} / (p_{hc} + p_{lc} + p_{nc}) \quad (3.9)$$

Where:

- d_{uc} is the population density fraction of settlement class in county c
- p_{uc} is the population density (persons/10,000 m²) of settlement class in county c
- p_{hc} is the population density (persons/10,000 m²) of settlement class h (high) in county c
- p_{lc} is the population density (persons/10,000 m²) of settlement class l (low) in county c
- p_{nc} is the population density (persons/10,000 m²) of settlement class n (non-settlement) in county c.

Hypothetically, if the sample of population density of non-settlement, low- and high-density settlement were 1, 3, and 6 persons/10,000 m², respectively, then the population density fraction of settlement class is 0.1, 0.3, and 0.6, respectively.

2. Calculate the area ratio of settlement:

$$a_{ub} = (n_{ub} / n_b) / 0.33 \quad (3.10)$$

Where:

- a_{ub} is the area ratio of settlement class in block group b
- n_{ub} is the number of grid cells of settlement class in block group b
- n_b is the number of grid cells in block group b.

Hypothetically, if the number of grid cells of non-settlement, low and high density settlement in a block group b was 1, 3, and 6, respectively, then the area ratios of these classes are 0.3, 0.9, and 1.8, respectively. The total fraction for each settlement class can then be calculated.

3. Calculate the total fraction for each settlement class:

$$f_{ubc} = (d_{uc} x a_{ub}) / ((d_{hc} x a_{hb}) + (d_{lc} x a_{lb}) + (d_{nc} x a_{nb})) \quad (3.11)$$

where:

d_{hc} is the population density fraction of settlement class h (high) in county c

a_{lc} is the population density fraction of settlement class l (low) in county c

d_{nc} is the population density fraction of settlement class n (non-urban) in county c,

a_{hb} is the area ratio of settlement class h (high) in block group b

a_{lb} is the area ratio of settlement class l (low) in block group b

a_{nb} is the area ratio of settlement class n (non-urban) in block group b.

The Mennis (2003) approach is a robust method; it can be expanded to further land use classes and even implemented to day- and night-time population distribution analysis if information is available on the temporal variation of population density.

Nevertheless, there are some limitations:

- The weighting factor (total fraction) is derived from the observation analysis in a county (administrative unit above block group). Consequently, this results in an error because there are different characteristics of people's activities in every unit level.
- In fact, the area ratio of settlement of one third each is rarely found in the reality (formula 3.10). The formulation of the weighting factor therefore causes an error because of this assumption.
- A boundary effect in the population distribution occurs based on the administrative unit disaggregation. Two adjacent pixels divided by a boundary may have different population density values. This may not represent reality properly because both pixels might actually form of a single land use class.

Other binary methods of population distribution (e.g. Wright, 1936; Langford and Unwin, 1994; Fisher and Langford, 1996; Holt *et al.*, 2004; Wu *et al.*, 2005; Reibel and Agrawal, 2007; and Tatem *et al.*, 2007)) that disaggregate population data from an administrative unit to uninhabited (non-settlement) and inhabited (settlement) areas have similar limitations because they ignore other land use classes as potential location of people's activities

The named limitations of the binary methods need to be de-restricted, which will be carried out by using land use/cover data, deriving reliable weighting factors and mitigating boundary effects.

Gallego and Peedell (2001) proposed a land use data method to map the people density in Europe using CORINE land cover data. They revealed that the population in a commune is an accumulating resulting from multiplying the population density in one land cover type with its area. The formula to calculate the population is stated as follows:

$$Y_m = U_c W_m \quad (3.12)$$

$$X_m = \sum_c S_{cm} Y_m \quad (3.13)$$

$$X_m = \sum_c S_{cm} U_c W_m \rightarrow W_m = \frac{X_m}{\sum_c S_{cm} U_c} \quad (3.14)$$

$$Y_m = U_c \frac{X_m}{\sum_c S_{cm} U_c} \quad (3.15)$$

Where:

- X_m is number of people in an administrative unit (e.g. village)
- Y_m is density of population for land cover type c in village m
- S_{cm} is area of land cover c in village m
- U_c is the weighting factor for each land cover type
- W_m is an adjustment factor to ensure that the total of population in all villages matches the value of the administrative unit.

To derive the weighting factors, initial values of proportional coefficients for each land cover type are used to calculate the population distribution disaggregation from a higher administrative level. By this disaggregation process, the accumulation of the people number in communes is calculated and then compared to the statistical population number in each commune. The iterative process is performed and until the errors become stable. The values of

the proportional coefficients for each land cover type (U_c) when the errors became stable, are then used for combining population figures from census data and land cover type.

As well as that of Mennis (2003), this is a robust method to distribute the population distribution number from the administrative unit to the land use classes. Also, it could be developed towards a spatio-temporal analysis by adjusting the weighting factor. Nevertheless, there several difficulties in this approach emerge:

- The iterative process is time-consuming and determining the initial value as weighting is not closer to the reality.
- The derived weighting factors are transferred from the higher to the lower administrative level. The proportional coefficient for each land cover type (U_c) is identical for all communes. But an error arises here due to different commune characteristics, which increase the more the characteristics of people activities differ.
- The explanation of population disaggregation by the formula is not transparent. In particular, the adjustment factor to ensure that the total of population in each village matches the administrative unit (Formula 3.12) does not indicate the real condition of population distribution.
- This method does also not eliminate the boundary effect.

To achieve a more realistic population distribution map, it is necessary to find a solution for the problems mentioned above. Schneiderbauer and Ehrlich (2007) improved the method by including further land use/cover categories. They analyzed the improvement of population distribution in Zimbabwe. The research collected available ancillary data such as land use, urban area, settlement, road network and elevation. A comprehensive weighting factor determination was developed using expert knowledge to point out convergent locations of activity of people. But due to the lack of calibration data, the result of population distribution analysis is difficult to evaluate.

The similar method was applied by Dobson *et al.* (2000). They also disaggregated the number of people on the district level to ancillary data such as road network, slope, land cover and nighttime lights. Similar to Schneiderbauer and Ehrlich (2007), the derivation of the weighting factor is based on expert knowledge without considering people's day- and night time activities. By using the nighttime light data, there is also the problem of

underestimating population density in the urban area and overestimating in the rural area (Sutton et al. (1997)). Dobson *et al.* (2000) produced LandScan population grid data, which is available globally. Another similar product was developed by Center for International Earth Science Information Network, Columbia University (CIESIN), USA, using the method of Balk *et al.* (2005). This method underlined that the ancillary data such as nighttime light and road network has some limitations and uncertainties resulting in a potentially high error rates. As with the LandScan product, this set of data is available globally. The example of LandScan and CIESIN products is shown in Figure 3.11.

In developing countries with limited high-resolution population data, LandScan and CIESIN products are useful to describe the situation where people are located. However, for tsunami evacuation planning, for example, these data sets are not sufficient; more detailed population numbers based on people's activity in a certain area are needed to minimize errors in evacuation capacity, and contingency or shelter planning due to inaccurate population data.

Research on population distribution by binary and land use/cover methods has not considered temporal distribution variations. As mentioned, an early warning system needs a population distribution by time period, since the taken measures may vary depending on the time of the day, which would make the system more effective. Therefore, a method considering time is needed, i.e. spatio-temporal population distribution.

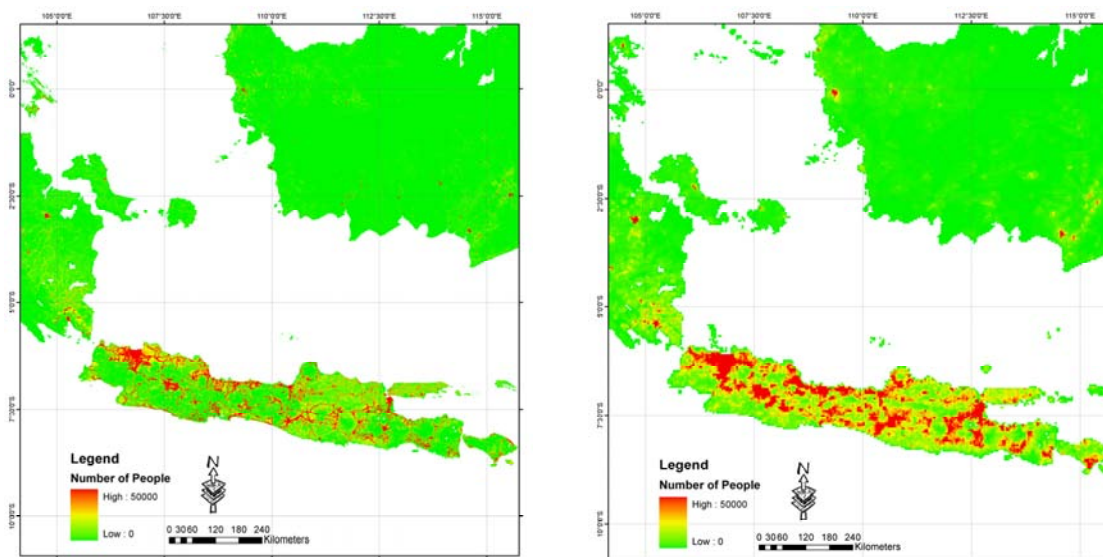


Figure 3.11 Population distribution maps: the LandScan (left), and CIESIN population grids (right)

Sleeter and Wood (2006) and Ahola *et al.* (2008) focused on this kind of research. Their studies represent a new approach of population distribution by dasymetric mapping, which takes into consideration both spatial and temporal distribution. Sleeter and Wood (2006) adopted the Mennis (2003) method and included commuters in the block census data in a coastal community in Oregon, USA. For the night-time population, they used the total population in the block of a census and then implemented the Mennis (2003) method to distribute this number to the urbanization classes.

For the day-time population, this method modifies the total number of block census population by taking into account the number of persons who stay in the block during the day and the number of employees coming to that block (commuters), which are recorded in the InfoUSA business database. This research does not consider outbound commuters due to a lack of data. By using the InfoUSA business database, the number of people in each block during the day-time can be calculated. In a developing country like Indonesia, since there is no data of this kind available, it is difficult to adopt these methods. The development of improvements of this methodology needs to consider the availability of data in the study area.

Ahola *et al.* (2008) have an available database in their study area that enables them to estimate the population during different times of the day. The methodology of this research uses the kernel density estimation similar to the point-based interpolation method developed by Martin and Bracken (1991). By using expert knowledge weighting factor adjustment, this research has a problem explaining the real situation of population distribution. Like the method of Martin and Bracken (1991), this method has some limitations. Hence, an improvement is necessary to reduce the uncertainties and converge the population distribution modeling closer to reality.

Based on the review of the top-down approaches on population distribution mapping, dasymetric mapping is more reliable than spatial interpolation since it calculates the population distribution closer to real situation. Although the approach is robust, enhancement of the methodology is still needed in some parts.

First, the weighting factor to disaggregate the number of people from census to ancillary data must take into account the spatial and temporal variability. The characteristics of human activities in different land use classes and different geographical regions differ. For example, in agricultural areas in Europe, there are less people working in this land use class due to a high degree of mechanization. During the winter time, there are no activities in agricultural area. In a tropical country like Indonesia, the activities of people in agriculture

areas will be higher than in Europe. The growing season is almost done throughout the whole year, and due to the limited degree of mechanization and cultivation of labor-intensive crops like rice, more people engage in agricultural activities. The weighting factors of agriculture in Europe must therefore be different from those in Indonesia.

In examining available formulas in current research on population distribution mapping using a dasymetric method, there is a lack of knowledge on the distribution of population data from the census. Hence, a modification of the distribution formula is necessary to consider and diminish the described limitations. The elimination of the boundary effect should also be taken into consideration when modifying the formula.

Another issue is related to the needs for risk and vulnerability assessment. For risk and vulnerability assessment, the mapping of population distribution by time (e.g. day- and night-time, workdays and holidays) is important. Sleeter and Wood (2006) and Ahola *et al.* (2008) revealed that a main problem of population distribution mapping by time is caused by commuters, i.e. the number of people who leave one administrative unit for work and return there at night. Additionally, in some regions, the temporal variations due to tourists should be taken into account.

Based on the explanation above, there are three topics that should be improved on the population distribution modeling by remote sensing and GIS:

- development of appropriate weighting factors for disaggregating population data from census data to land use classes;
- modification of the formula in order to comprehensively explain the steps of population disaggregation, considering the reality of where people are located and including the elimination of the boundary effect;
- and including commuters and tourists in the population distribution modeling during day- and night-time.

3.3.3. Bottom-up approach

In contrast to the previously described approaches, it is also possible to use bottom-up approaches as a method to estimate the population density distribution from the variables derived from remote sensing data. This method comprises collection of census data, image classification and analysis, statistic analysis including linear regression and multivariate regression, and finally, the population distribution mapping itself. It was inspired by Tobler (1969), who

demonstrated that population can be estimated with a high degree of accuracy by measuring the size of human settlements observed from satellite photography generated by the Gemini manned space light program (Elvidge *et al.*, 1996). Various studies use different kinds of remote sensing derived variables, which can be categorized as follows:

- using nighttime lights (Elvidge *et al.*, 1997; Sutton *et al.*, 2001; Lo, 2001; Briggs *et al.*, 2007; Zhuo *et al.*, 2009);
- using impervious surfaces such as settlement or residential area, and urban morphology (Mesev, 1998; Chen, 2002; Lu *et al.*, 2006; Mubareka *et al.*, 2008);
- using multiple image variables such as image texture, NDVI and other indices (Li and Weng, 2005; Liu *et al.*, 2006);
- using land use/cover (Yuan *et al.* 1997; Pozzi and Small, 2001; Min *et al.*, 2002; and Tian *et al.*, 2005).

Nighttime lights

Elvidge *et al.* (1997) used data from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS), which is able to detect cities and towns during nighttime in order to estimate population density. They estimated the population in 21 countries, and received good correlation between the number of people and the "lit" area of the image. The results indicate that DMSP-OLS data can successfully be used to define and update the spatial distribution of human population on a global basis. However, it was noted that this product is not able to identify human settlements that lack electric power, as is often the case in developing countries.

Based on the Elvidge *et al.* (1997) research, Sutton *et al.* (2001) estimated the population number of more than a thousand cities from the lit area of the images. They analyzed six population aggregates: national, all low-income countries, all middle-income countries, all high-income countries, population counts of the world, and selected cities. With the exception of selected cities, the population estimation error in each aggregate was sufficiently low. This means that this research allows an estimation of urban population for every country in the world, but that significant error must be taken into consideration for city-specific population estimation.

Another study that used nighttime light was carried out by Lo (2001). He developed a model to estimate population density in China using area, volume

and percentage of light, and pixel mean as independent variables. Like Sutton *et al.* (2001), this study was divided into three spatial levels — province, county and city level. The estimation of population was satisfying for the province level, but this research concluded, as did Sutton *et al.* (2001), that the error was generally too high for counties and cities level.

The latest research on estimating population density using nighttime light data was published by Zhuo *et al.* (2009). Similar to Lo's results, this research achieved a good correlation between population and digital number of nighttime image, but the results were not sufficient for the urban population.

In summary, of all studies using nighttime lights for population estimation, it can be concluded that sufficiently accurate results can be produced for the national and province level, but high errors result for population estimations of the county or city level and for developing countries.

Impervious surfaces

Other studies derive impervious surfaces from remote sensing data to estimate population distribution. One example is Mesev (1998), who correlates low, medium, and high densities of residential areas and tower blocks with population distribution. He found that such impervious surfaces have good correlation with the population density.

The same is concluded by Chen (2002), who found that areal census dwelling data had higher correlations with areas of different residential densities (derived from Landsat TM) than with the aggregated whole residential area at an individual census zone level. Lu *et al.* (2006) also used Landsat TM-derived residential impervious surfaces for population estimation and a good correlation, with population data could be achieved. Mubareka *et al.* (2008) detected settlement locations from Landsat TM and correlated them to population data. The result is a reasonable estimate of settlement population numbers for rural settlements (<1000 inhabitants) and a slightly weaker correlation in urban areas.

In addition to the population estimation, the latter three publications also focused on the classification technique to identify impervious surfaces or settlements from remote sensing data. The decision to apply appropriate classification technique and thus the quality of the classification considerably affects the population estimation.

Multiple image variables

Li and Weng (2005) estimated the population in the city of Indianapolis, Indiana, USA, by using multiple image variables derived from Landsat ETM, such as NDVI, SAVI and band ratio. This research explored the correlation between the image variables and the population density. By using multiple regressions, this research shows that the integration of textures, temperatures and spectral response substantially improves the accuracy of population density estimation. Li and Weng (2005) also revealed that a reasonably accurate population density estimation can be provided by combining multiple variables from remote sensing data derivation.

A similar study regarding the use of multiple image variables was undertaken by Liu *et al.* (2006). By using image texture from high-resolution images and multiple regressions, they found that the coefficient determination of the regression model is below 60%. This means that estimating the amount of population using the texture of high-resolution images is insufficient due to mixed pixels.

The application of multiple image variables as in Li and Weng (2005) suggested improved accuracy of the population density result. Although the application of multiple variables already resulted in good accuracies, Li and Weng indicate that using the remote sensing technique is still challenging due to the complexity of urban landscape and population distribution. In particular, in detailed scale, the technique of remote sensing classification can still be improved for better results on population distribution.

Land use/cover

The last category of remote sensing derived variables to estimate population density is land use/land cover. Encouraged by the result of dasymetric mapping, where population distribution is linked with population activity, Yuan *et al.* (1997) applied the multivariable regression to examine the correlation between land cover types and population counts by an enumeration district. The resulting correlation is high, and could be implemented to reliable estimates of the population at the district level. The same result was produced by Min *et al.* (2000), who achieved good correlation between habitation type (land use) and population density in several regions in China.

Another example of using land use/cover to estimate the population density was done by Pozzi and Small (2001). Although they were not clear on how they did the actual estimation of the population density from land use data, they

showed that there is a consistency between land use/cover and population density in suburban areas: consistent spectral characteristics of suburbs were related to the degree of vegetation cover. A good result on estimation of population density by using land use/cover was also presented by Tian *et al.* (2005). The result indicated that land use data is sufficient to estimate the population density in China and proved that population density is correlated to where people carry out their activity.

It can be concluded that at the current state of research, good correlations between population density and the remote sensing derived data can be produced. The estimation of population from remote sensing data can be done at a global scale and at the national, province and district levels. For estimations at the local or city level, the research of population estimation still must be improved especially concerning the remote sensing techniques to classify land use or impervious surfaces.

3.3.4. Summary of population distribution modeling

Research on population distribution modeling by using remote sensing data has made great progress in the past years, especially dasymetric mapping. This methodology is allegedly the approach with the highest accuracy. Although this method is not new, there are still some open issues and potential improvements, especially regarding the weighting factor determination. The weighting factor determination for different times of the day as well as for different land use classes and the commuters are recognized as still open problems. This research is designed to provide a solution. The modification of the population distribution formula, socio-economic data analysis, and multi-scale disaggregation are required actions towards this solution.

The bottom-up approach of using data derived from remote sensing to estimate population distribution is a promising approach. To date, insufficient accuracy at detailed scales still poses a challenge for this research. Due to conditions and characteristics of settlements, several researches indicate the need and challenge to identify these precisely. Hence, the settlement classification using remote sensing data is a necessary task.

3.4. Settlement Classification Using Remote Sensing Data

3.4.1. Overview

Information about settlements is a very important input parameter for population distribution mapping. Research on settlement classification is necessary since settlements have specific characteristics. The combination of buildings, roads, vegetation and open land, which all have different spectral characteristics, pose specific challenges on remote sensing data. Various classification approaches have been developed applying either optical or synthetic aperture radar (SAR) or a fusion of both data. The applications vary from pure settlement classification (e.g. Zha *et al.*, 2003) to climate change research (e.g. Weng, 2001).

To analyze the state of the art and the gaps in the field of settlement classification methodologies, this sub-chapter reviews research in the following categories:

- optical imagery
- SAR imagery
- data fusion of optical and SAR imagery.

3.4.2. Settlement classification using optical imagery

The research on settlement classification using optical data can be categorized into three types of classification approaches:

- common techniques such as maximum likelihood classification (supervised and unsupervised);
- indices and threshold classification;
- knowledge-based classification with use of ancillary data.

Common techniques

Common remote sensing classification techniques such as supervised and unsupervised maximum likelihood classification are widely used for urban growth and change detection applications (e.g. Li and Zhao, 2003). Several studies (e.g. Weng, 2001) show a good accuracy at settlement and land use classification using Landsat TM imagery. Supervised and unsupervised

maximum likelihood mostly is not used to classify purely settlement, but for overall land use/cover classifications. For example, Weng (2001) classifies settlement and also land cover to evaluate urban expansion in Zhujian Delta, China.

It is widely used since it is usually implemented in image processing software available on the market. But this method occasionally cannot differentiate objects that have similar spectral reflectance, such as urban area and bare land. In addition, especially in mountainous areas, settlement classification occasionally makes problems. Hence, other methods are needed, such as index development and knowledge-based classification.

Indices

Several classification approaches have been developed, which are based on indices. Zha *et al.* (2003), for example, use the Normalized Difference Built-up Index (NDBI) on Landsat TM imagery for automatic classification of urban areas.

A new method for extraction of built-up areas using Landsat TM was developed by Xu (2007). He created the Index-based Built-up Index (IBI), which is based on three thematic indices: NDBI, the Soil Adjusted Vegetation Index (SAVI) and the Modified Normalized Difference Water Index (MNDWI). With IBI, built-up areas receive higher values than other land uses. The formulas of those indices are stated as follows:

$$NDBI = \frac{(MIR - NIR)}{(MIR + NIR)} \quad (3.16)$$

$$SAVI = \frac{(NIR - Red)(1 + l)}{(NIR + Red + l)} \quad (3.17)$$

$$MNDWI = \frac{(Green - MIR)}{(Green + MIR)} \quad (3.18)$$

$$IBI = \frac{\left[NDBI - \frac{(SAVI + MNDWI)}{2} \right]}{\left[NDBI + \frac{(SAVI + MNDWI)}{2} \right]} \quad (3.19)$$

Where:

MIR is the mid-infrared band of Landsat ETM (canal 5)

NIR is the near-infrared band of Landsat ETM (canal 4)

Red is the red band of Landsat ETM (canal 3)

l varies between 0 and 1 and its weighting is dependent on vegetation cover or soil moisture conditions (Huete, 1988). The default *l* factor of 0.5 was used for all images.

Green is green band of Landsat ETM (canal 2).

Xu *et al.* (2007) obtain good accuracy of the settlement classification result, but a preliminary investigation of the research at hand showed that IBI index fails to discriminate settlement areas and open or bare land. Improvement of this methodology is needed; texture analysis of the panchromatic band and knowledge-based classification may enhance the accuracy.

Knowledge-based classification

By using knowledge-based classification and ancillary data, the accuracy of settlement classification can be significantly increased. Mubareka *et al.* (2008) detected settlement areas using ancillary data: they used a digital elevation model to support the detection of settlements in mountainous areas. Also, other ancillary data such as road density and road coverage can be applied to increase the accuracy of settlement classification (e.g. Zhang *et al.* (2002) and Epstein *et al.* (2002)).

3.4.3. Settlement classification using SAR imagery

Radar imagery has the advantage of being almost independent of day- and night-time and weather conditions. Therefore, SAR imagery is the most appropriate data for tropical countries like Indonesia where there is a high degree of cloud coverage throughout the year. In particular, for settlement classification, the methodologies are still under development. The most common techniques are:

- Textural analysis (e.g. Dekker, 2003; Acqua and Gamba, 2003)
- Multi-temporal analysis (Acqua *et al.*, 2003; Pellizzeri *et al.*, 2003)
- Statistical model (e.g. Tison *et al.*, 2004)

- Spatial index analysis (e.g. Stasolla and Gamba, 2008)
- Object-oriented classification (e.g. Thiel *et al.*, 2008; Esch *et al.*, 2010).

Textural analysis

The most common radar imagery classification for earth surface detection is by using textural analysis. Dekker (2003) is one example, using ERS SAR imagery. Modern texture measures such as histogram measures, wavelet energy, fractal dimension, lacunarity and semivariogram were analyzed. The result improved the accuracy of the classification compared to widely used texture measures such as mean intensity, variance or entropy. Despite this improvement, the classification was still not satisfying, because of the low separability of some classes, especially of urban area and forest, and urban area and industry/greenhouse.

Acqua and Gamba (2003) propose to use the Histogram Density Index (HDI) to identify urban characteristics of high-resolution imagery. However, like Dekker (2003), the approach failed to classify main parts of urban areas like city centers, residential and suburban areas.

The current methods of texture analysis therefore seem to be insufficiently developed for the detection of urban areas. The low separability of some features is the main barrier of using textural analysis for settlement classification.

Multi-temporal analysis

Being dissatisfied with the previous result, Acqua *et al.* (2003) modify the methodology with multi-temporal analysis to improve the urban characterization. The main problems of urban area classification by radar imagery are multiple bouncing, layovers and shadowing. Multi-temporal imagery produces a multi-angle look on urban areas and thus allows identifying different details of the same object. Hence, the discrimination of different classes in urban areas is facilitated. Acqua *et al.* (2003) again apply the HDI and using multi-temporal imagery, it provides better classification results than in the previous study. Additionally, the extraction of streets is also improved by this approach.

Another comprehensive research on settlement classification by using multi-temporal analysis was done by Pellizeri *et al.* (2003). They compared statistical analysis and the neural kernel-based approach to classify urban area. The

result from the analysis shows that the neural kernel-based approach slightly outperforms the statistical approach when classifying mono-temporal imagery. By extending the data to multi-temporal and multi-frequency, both approaches achieve similar accuracy of classification. This shows that multi-temporal data can slightly improve the accuracy of single data classification method.

In summary, using multi-temporal SAR imagery leads to a better classification result than mono-temporal imagery because it minimizes some of the problems of SAR imagery over urban areas.

Statistical model

Tison *et al.* (2004) revealed that classes can be discriminated based on their statistical properties, which require an accurate statistical model. Hence, they investigated various statistic distribution models such as Weibull, Log-Normal, Nagatani-Rice and Fisher, and analyzed how these distributions fit the urban area characteristics in SAR imagery. The Fisher distribution proved the most appropriate distribution for urban characteristics in their study area Toulouse, France.

This is an important result related to settlement classification using SAR imagery. It might be used as a reference for future applications of this kind. But one has to be aware that the best fitting statistical distribution for one region is not necessarily transferable to any other.

Spatial index analysis

Stasolla and Gamba (2008) propose spatial indices to extract settlements from high-resolution SAR imagery. After reviewing several indices, they defined three spatial indices for their purposes, Moran's Index, Geary's Index, and the Getis-ord Index. The basic principal of these indices is a local indicator of spatial association (neighborhood).

Settlements being inhomogeneous objects were the reason to apply these indices. Hence pixel-by-pixel classification is inappropriate for classification. This method resulted in good overall accuracy, but omission and commission errors were still high. Considerable improvement of the methodology is especially needed on indexing thresholding, which indicates a problem for transferability to other area.

Object-based classification

The latest research on settlement classification opts for an object-based approach. Thiel *et al.* (2008) and Esch (2010) describe the new methodology to detect settlement area using TerraSAR-X. They propose a pre-processing method called "speckle divergence analysis", with a comparably straightforward approach of analyzing local statistics. They achieve more appropriate noise suppression compared to establish filtering routines, which preserve true structures with texture and contour information (Thiel *et al.*, 2008). This filter is based on a sigma probability of the Gaussian distribution of speckle noise. It removes speckle and suppresses noise, but preserves the texture, structure and contour information. This is indispensable since many regions are characterized by small-scale structures. The speckle divergence image is used for object-based settlement classification. The result of this classification is satisfying and has good accuracy. It is promising to implement the methodology for analyzing the study area of this research, with an added improvement.

3.4.4. Using data fusion of optical and SAR imagery

Fusions of different types of satellite imagery are useful sometimes for increasing the accuracy of the classification result. The potential of one satellite image could fill the deficiencies of another image. One example is the fusion of radar and optical imagery where the optical imagery provides information about the spectral properties of objects, and radar imagery provides high resolution texture information. Additionally, the problem on cloud cover in optical imagery is minimized. Research on settlement classification using fusion imagery was undertaken by Gamba and Acqua (2007) and Corbane *et al.* (2008).

Gamba and Acqua (2007) classified settlement area by fusion of SAR and optical imagery in two study areas, Al-Fashir, Sudan, and Pavia, Italy. They used the method of multi-scale texture analysis to examine the different textures of both satellite imageries. The rule-based fusion algorithms, spatial analysis with multi-width texture, and Markov Random Field (MRF) classifier have been applied for the classification process. The authors used two different data in their research: first, a combination of ENVISAT ASAR data and SPOT 5 for Al-Fashir, Sudan, and second, a combination of ERS 1/2 and ENVISAT ASAR and Landsat TM for Pavia, Italy. In the first application, although the misclassifications of rock/bare soil is still remains, the joint image with rule-based on multi-texture reduced the problem significantly. Future research must be undertaken to eliminate the misclassification. The second application also shows improved accuracy by using data fusion.

Corbane *et al.* (2008) analyzed the potential combination of SAR and optical imagery for operational rapid urban mapping. They use Radarsat-1/ENVISAT and SPOT 4/5 imagery. There are two study areas, one in Bucharest, Romania, the other one in Cayenne, French Guiana. Unsupervised K-means and rule-based fusion are applied for the settlement classification. The method has successfully increased the accuracy of the results compared to the single use of SAR or optical imagery.

3.4.5. Summary of settlement classification using remote sensing data

Remote sensing settlement classification has made great progress, both for optical and radar data and their fusion. Settlement classification with radar imagery is the most challenging. Several studies on application of radar imagery have been conducted using techniques such as textural analysis, multi-temporal analysis, statistical model, spatial indexes and object-based classification.

Improvements of these approaches are still needed since the characteristics of settlement areas are different throughout the world. The latter application provides comprehensive results and is promising for application in Indonesia. But the threshold determination for settlement classification must be improved and the conversion of the object-based procedure to a pixel-based one is a potential challenge in this methodology.

3.5. Conclusion of the State of the Art of Research

This chapter summarized the state of the art of research of three components that are important for tsunami vulnerability and risk assessment: surface roughness determination, population distribution mapping, and settlement classification. Remote sensing and GIS) play an important role in these research fields.

Several studies proved that vegetation plays an important role in mitigating tsunami impact along the coast. The degree of mitigation depends mainly on parameters such as width, density and structure of the coastal vegetation. The inclusion of such surface roughness conditions into tsunami models significantly influences the inundation modeling results and makes the modeled inundation area more reliable for planning. In order to be able to integrate the surface roughness conditions into a tsunami model, it is necessary to quantify these conditions, which is done by determining surface roughness coefficients.

There are diverse approaches to determine these coefficients, starting from complex and costly but accurate laboratory experiments through to quick but subjective site visits with photo matching, and the use of coefficients published by other researchers, which is a widely used method albeit with limited transferability and inconsistencies. Each of these approaches has its advantages and disadvantages. It is necessary to choose an appropriate approach and implement required adjustments to improve it.

In addition, the different surface roughness conditions along a coast should be reflected in spatially distributed roughness coefficients. Information about the present surface roughness conditions without extensive field surveys is possible by classification from remote sensing data. But previous studies dealing with this task were not able to fully capture the different surface roughness classes. The analysis of the role of vegetation on tsunami mitigation indicates that density and neighborhood classes are important parameters that should be taken into account when deriving surface roughness conditions from remote sensing data. The development of an approach to accurately identify the different surface roughness classes is a rewarding challenge to bring tsunami hazard modeling forward.

For tsunami risk and vulnerability assessment, it is essential to determine an accurate population distribution, at best differentiated for different times of day. The review of the research status showed that dasymetric mapping outperforms mere spatial interpolation for population distribution modeling. Since currently developed formulas to perform distribution modeling are unsatisfying, it is necessary to modify and improve them. The weighting factors are crucial for the distribution modeling. In order to derive reliable weighting factors, it is desirable to integrate statistical socio-economic data. Thus, it would be possible to obtain a more reliable and realistic pattern of people's spatial and temporal distribution as achieved by current distribution models.

Moreover, knowledge of the precise location and extent of settlement areas is an important input for the dasymetric population mapping. In developing countries, in particular, detailed, consistent, complete and up-to-date land use/land cover data are often not available; Here again, remote sensing data are a good tool to derive such data. The review of the methodologies to classify settlements from various types of remote sensing data highlighted their advantages and disadvantages. Radar imagery is more appropriate for tropical regions than optical data due to frequent cloud coverage. The review also revealed that the methodologies for classification of radar data in particular still have deficiencies and need to be further developed and improved.

The gaps in research in these fields were outlined in this chapter — surface roughness determination, population distribution mapping and settlement classification. By contributing to fill these gaps, this research will effectively support the improvement of tsunami hazard assessment as well as of risk and vulnerability analyses. Consequently, it will better serve aims of disaster management.

CHAPTER 4: METHODOLOGY

This chapter describes the developed procedures and methodologies to answer the research questions formulated in the introduction chapter with regard to the research objectives in surface roughness determination, population distribution modeling and settlement classification.

In the first sub-chapter, the basic data requirements for this research are mentioned. It is a brief explanation of the type, the potential source, and the usages of the data in this research.

In the second sub-chapter, the developed concept of surface roughness determination taking into account object density and neighborhood is presented. The improvements of the method to estimate of surface roughness classes and its coefficients, the new remote sensing methodology for surface roughness classification, and integration of the spatially distributed surface roughness information into tsunami modeling and hazard mapping are presented. The latter procedure examined the importance of spatially distributed surface roughness information in tsunami modeling/hazard mapping.

In the third sub-chapter, for population distribution modeling, the new improvement concept on dasymetric mapping with regard to spatio-temporal differentiation by combining socio-economic data is presented. The procedure for analyzing socio-economic data to determine the proportion of people's location with respect to different land use is also explained. The transferability of weighting factors for other Indonesia regions is analyzed to prove the applicability of the model.

In the last sub-chapter, the further developed methodologies on settlement classification using radar imagery and the procedure of implementation in the study area are presented. The statistical analysis for threshold determination is also explained as well as the training sample collection process.

In addition, in order to demonstrate the quality of the methodologies, every research topic is combined with an accuracy assessment and a comparative study of available methodologies.

4.1. Data Requirements

Several datasets are required for analyzing the surface roughness determination, population distribution modeling, and the settlement classification. Table 4.1 is a list of data requirements, including the type of data, potential sources and use of these data. Satellite imagery – either optical or radar - is used as primary data of the research. In addition, secondary data such as socio-economic data is used to support the model development.

Table 4.1 Data requirements for the research

No.	Type of data	Potential Sources	Usage
1.	Optical Imagery (e.g. SPOT 5)	Satellite data providers	Surface roughness determination
2.	Radar Imagery (e.g. TerraSAR-X)	Satellite data providers	Settlement classification
3.	Digital Surface/ Terrain Model	National mapping agencies or commercial data providers —e.g. Intermap, United States Geological Survey (USGS)	Tsunami modeling implementation input
4.	Bathymetry	Oceanographic mapping agencies or open sources —e.g. GEBCO, National Coordinating Agency for Survey and Mapping (Bakosurtanal), Agency for the Assessment and Application of Technology (BPPT)	Tsunami modeling implementation input
5.	Roughness coefficients	Publications in scientific journals, guidance documents	Surface roughness class and its coefficient determination
6.	Population census data and growth rate	National statistic agencies — e.g. Indonesian Statistical Bureau (BPS)	Population distribution modeling
7.	Administrative boundaries	National statistic or mapping agencies — e.g. Indonesian Statistical Bureau (BPS), Bakosurtanal	Population distribution modeling
8.	Land use data	National mapping agencies or land use data derived from remote sensing — e.g. Bakosurtanal, National Institute of Aeronautics and Space (LAPAN)	Population distribution modeling
9.	Socio-economic data (potential of village, occupation)	National statistic agencies — (e.g. Indonesian Statistical Bureau (BPS)	Population distribution modeling
10.	Reference maps	National mapping agencies, e.g. Bakosurtanal, or field survey	All research fields

4.2. Surface Roughness Determination

There are some limitations of remote sensing classification in previous studies for determining surface roughness, especially regarding the consideration of density and neighborhood, which were outlined in chapter 3. Density and neighborhood are important characteristics that contribute to surface roughness with respect to their function to resist tsunami flow. Hence, this research proposes a concept on density and neighborhoods classification for surface roughness determination. The steps of this concept are stated as follows:

- **Surface roughness classes and their coefficient determination.** Defining the surface classes and their associated roughness coefficients is the first step in the overall approach. The determined surface roughness classes are used as basis for this process.
- **Surface roughness classification.** This step includes the pre-processing of the satellite imagery, analyses of relevant image variables and the classification of the land use classes, including density and neighborhood classification.
- **Accuracy assessment.** This process evaluates the quality of the classification result.
- **Integration into Tsunami Modeling/Hazard Mapping.** This is an important step to demonstrate the usefulness of the results for tsunami modeling/hazard mapping.

4.2.1. Surface roughness classes and their coefficient estimation

The estimation of the surface roughness classes is a very important and decisive step prior the remote sensing data classification. The chosen approach is based on the analysis of the research publications and the assigned roughness coefficients. Instead of relying on one publication only, a synopsis of the most relevant investigations has been made. Moreover, this approach also allows for the estimation of the variance and uncertainty of the used coefficients.

The following steps were performed to determine the surface roughness classes and their coefficients:

1. Review of publications that contain surface roughness classes and their coefficients.

2. Analysis of these publications. The analysis focuses on literature published in peer-reviewed journals or cited by several scientists.
3. Analysis of variability and transferability of a surface roughness class considering density and neighborhood classes. For a better understanding of the variability of surface roughness in the study area, field surveys were performed in 2007 and 2008.
4. Determination of the surface roughness class based on publications, results of field survey analyses, and the information derived from remote sensing data.
5. Analysis of the minimum, maximum, average, and variance of the coefficient for each surface roughness class from the selected literature.

4.2.2. Surface roughness classification

The capability of remote sensing data to provide spatial information on surface roughness classes makes it an important tool for this research. The chosen classification technique and its accuracy are the core of the methodology. This sub-chapter explains step by step the surface roughness classification using remote sensing data, from pre-processing to classification of surface roughness and determination of coefficients. Additionally, the accuracy assessment to check the quality of the result is explained.

SPOT 5 multi-spectral and panchromatic 2.5 m imagery were used in this research to investigate the surface roughness conditions in the study area. The steps to perform the roughness coefficient classification are as follows:

- Pre-processing of satellite images including spatial enhancement, image variable calculation and textural analysis.
- Analysis of relevance of variables to identify effective parameters to be used in the remote sensing classification.
- Classification of main land use types as basis of surface roughness classification.
- Analysis of density and neighborhood; these classes are part of the surface roughness classes.

The summary of the classification steps to derive the spatial distribution of surface roughness conditions is shown in Figure 4.1. This figure also shows the

single steps of pre-processing, analyzing the relevance of variables, main land use classification, and density and neighborhood classification.

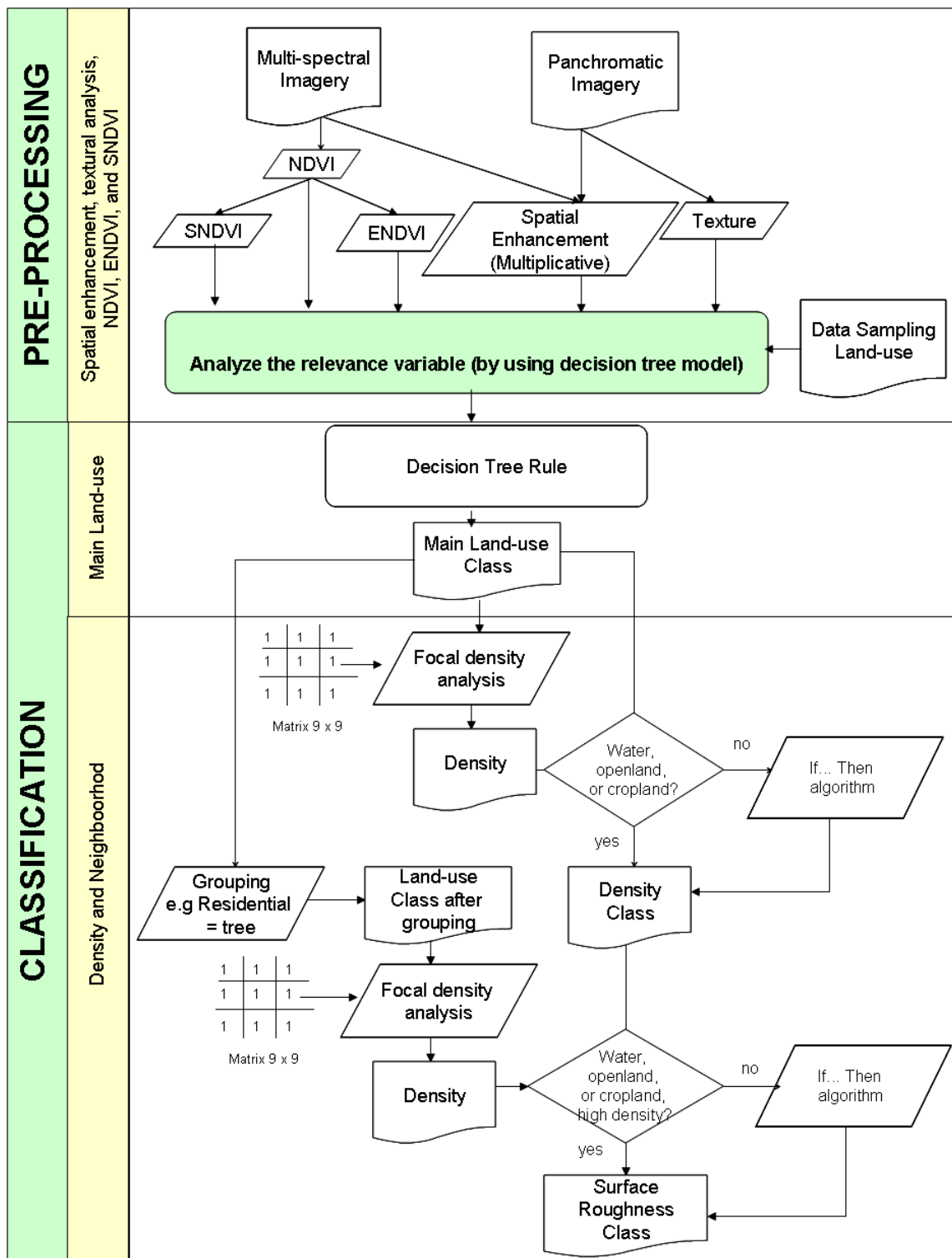


Figure 4.1 The steps of surface roughness classification using optical satellite imagery

Pre-processing

First, the satellite imagery, in this case SPOT 5 imagery (multi-spectral and panchromatic), must be pre-processed. The pre-processing consists of:

- spatial enhancement;
- image variables calculation;
- and textural analysis.

The spatial enhancement combines the high resolution panchromatic image with the lower resolution multi-spectral image to produce high spatial resolution multi-spectral imagery (pan-sharpening). This method applies a multiplicative algorithm that integrates two raster images, e.g. SPOT 5 multi-spectral image (10 m) and SPOT 5 panchromatic (2.5 m) images, to increase the intensity of the image and the spatial resolution of the multi-spectral data.

Additionally, several indices such as the Normalized Differenced Vegetation Index (NDVI), the sigmoid NDVI (SNDVI) and exponential NDVI (ENDVI) are calculated from the multi-spectral image. The following are the formulas of NDVI, SNDVI and ENDVI:

$$\text{NDVI} = (\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED}) \quad (4.1)$$

$$\text{SNDVI} = 1 / (1 + e^{-\text{NDVI}}) \quad (4.2)$$

$$\text{ENDVI} = \text{EXP}(\text{NDVI}) \quad (4.3)$$

Where:

NDVI is the Normalized Differenced Vegetation Index

SNDVI is the Sigmoid Normalized Differenced Vegetation Index

ENDVI is the Exponential Normalized Differenced Vegetation Index

RED is the sensor's red band (e.g. SPOT 5 band 2 (610-680 nm))

NIR is the sensor's near-infrared band (e.g. SPOT 5 band 3 (780-890 nm)).

The next step is a textural analysis of the panchromatic image. It maximizes the function of the panchromatic image as a high spatial resolution image. The average, standard deviation, and coefficient of variation of kernel size of 9 x 9 pixels are calculated using the following formulas:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n x_i \quad (4.4)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (4.5)$$

$$c_v = \frac{\sigma}{\bar{x}} \quad (4.6)$$

Where:

- s is the standard deviation of the digital numbers in a kernel
- N is the total of pixels in a kernel
- x_i is the digital number of each pixel in a kernel
- \bar{x} is the average digital number in a kernel
- c_v is the variation coefficient of digital numbers in a kernel.

These results of the pre-processing (spatial enhancement, image variable calculation, and the texture) are then stacked to a single file, which is used as input for analyzing relevant variables for the land use classification.

Analysis of image variables

After the pre-processing, the relevance of variables for the classification process is analyzed to find the most effective combination of parameters. The decision tree method is used to decide which image variables are the most appropriate for the classification. Therefore, training samples with land use known from field surveys and visual interpretation of satellite data are analyzed. For every sample of land use class, the layer stack created during the pre-processing is used, which contains the four spatially enhanced spectral bands of SPOT 5, NDVI, SNDVI, ENDVI, and the texture of the panchromatic image. Diverse parameter combinations are then examined and tested. The image variable combination with the highest classification accuracy of the training areas is chosen for the classification of the complete remote sensing imagery. Table 4.2 shows the analyzed parameter combinations.

Table 4.2 Analysis of image variables

No.	Image variable combinations
1.	Only NDVI
2.	Only ENDVI
3.	Only SNDVI
4.	Only Texture
5.	NDVI and Texture
6.	Four spatially enhanced spectral bands of SPOT 5
7.	Four spatially enhanced spectral bands of SPOT 5 and NDVI
8.	Four spatially enhanced spectral bands of SPOT 5, NDVI and SNDVI
9.	Four spatially enhanced spectral bands of SPOT 5, ENDVI and SNDVI
10.	Four spatially enhanced spectral bands of SPOT 5, NDVI and texture
11.	Four spatially enhanced spectral bands of SPOT 5, NDVI, SNDVI, ENDVI and texture.

Main land use classification

Having identified the most effective image variable combination for a remote sensing data set, the surface roughness classification is performed in two steps: main land use classification and density and neighborhood classification.

For the classification of main land use classes, the non-parametric decision tree method is implemented here. The thresholding method of the decision tree analysis facilitates the classification process and is well integrated in modeler software to create an automatic classification procedure.

Decision tree classifiers belong to the group of supervised and hierarchical classifiers where subsets of classes are processed at multiple stages. The internal node includes more than one class, while the leaf node includes only one class. The decision tree has a number of advantages over the maximum likelihood (ML) and artificial neural network (ANN) algorithms, being computationally fast and easy to apply (Chen and Wang, 2007).

Similar to supervised classifiers, training samples of land use are required to then develop the decision tree rules. The number of main land use classes influences the result of the determination of surface roughness coefficients. After setting up the decision tree rules for every main land use class, the main land use classification can be performed, which is then followed by density and neighborhood analyses.

The main land uses for surface roughness classification based on remote sensing data are "water", "open field", "crop field", "shrubs", "tree" and "residential area".

Density and neighborhood analyses

The density of land use features was calculated by using a kernel size of 9 x 9 pixels (~23 m spatial resolution) based on Koshimura (2009). It resulted in three categories of densities — rare, medium and high density. These categories were classified as follows (Murashima *et al.*, 2008):

- Rare density : 0–20 %
- Medium density : 20–50 %
- High density : more than 50%.

The result is used for the neighborhood classification. A class of rarely dense residential area may contain a combination of trees and residential area, or vice versa. Hence, the classification of neighborhood is performed as follows particularly as an example for residential area with trees:

- Selected main land use classes such as the class of residential and tree are grouped into one class (grouping).
- The focal density analysis is performed and is categorized into two classes, high (> 80%) and low (< 80%); this is called the “density class” (after grouping).
- The result of the density class (before grouping) is then overlaid with the new density class (after grouping).
- The class of residential with trees can be identified where low or medium density residential/tree class (before grouping) meet the high density class (after grouping).
- Where low or medium density residential/tree class (before grouping) meets the low-density class (after grouping), the original class is not changed.

The same process is also performed for classifying the neighborhood of trees and shrubs, and other neighborhood classes.

Neighborhood analysis

The processes of the main land use classification by decision tree and the density and neighborhood analyses are packed into a GIS procedure. After manually selecting training samples, this package can be applied and the

classification process will be performed automatically. The processing steps of the semi-automatic model are shown in Table 4.3.

Table 4.3 Steps of the semi-automatic classification

No.	The process	Type of process
1.	Training sample selection	Manual
2.	Decision tree modeling	Automatic
3.	Transfer of decision tree model to image processing software	Manual
4.	Process of main land use classification	Automatic
5.	Process of density and neighborhood classification	Automatic

4.2.3. Accuracy assessment

An accuracy assessment is performed to check the quality of the classification. To conduct a proper accuracy assessment, references from the field are needed. Two field surveys were conducted in the study area in 2007 and 2008; 87 reference locations — reliable, equally distributed throughout the study area and of adequate size — were selected and captured with GPS. The confusion matrix accuracy assessment by Story and Congalton (1986) is used in this research. This methodology allows identifying the omission and commission error of every surface roughness class. The concept of calculating the overall accuracy, omission and commission errors is illustrated in Table 4.4 (Story and Congalton, 1986; Naesset, 1996; Liu *et al*, 2007).

Table 4.4 Population error matrix based on information of classified and reference land use (LU) data

Classified	Reference				
	LU 1	LU 2	...	LU k	Total
LU 1	n_{11}	n_{12}	...	n_{1k}	n_{1+}
LU 2	n_{21}	n_{22}	...	n_{2k}	n_{2+}
.
.
.
LU k	n_{k1}	n_{k2}	...	n_{kk}	n_{k+}
Total	n_{+1}	n_{+2}	...	n_{+k}	N

The overall accuracy, commission error and omission error are then calculated as follows:

$$\text{Overall Accuracy, } OA = \frac{\sum_{i=j}^k n_{ij}}{N} \times 100\% \quad (4.7)$$

$$\text{Omission error, } OE = \frac{n_{+j} - n_{jj}}{n_{+j}} \quad (4.8)$$

$$\text{Commission error, } CE = \frac{n_{i+} - n_{ii}}{n_{i+}} \quad (4.9)$$

Where:

- i is classified land use data
- j is reference land use data
- LU is land use class
- N is total of pixel reference and classified class
- n_{ii}, n_{jj} is number of pixels in a certain land use, which is classified as the same land use of reference
- n_{+j} is total number of pixel a certain reference land use
- n_{i+} is total number of pixel a certain classified land use.

Figure 4.2 provides a numerical example of the confusion matrix accuracy assessment calculation.

		Reference Data				
		F	U	W	Row Total	
Classified Data	F	40	9	8	57	Land Cover Categories F = Forest U = Urban W =Water
	U	1	15	5	21	
	W	1	1	20	22	
	Column Total	42	25	33	100	

Omission Error	Commission Error	Overall Accuracy
F = (42-40)/42 = 5 %	F = (57-40)/57 = 30%	= (40 + 15 + 20)/100
U = (25-15)/25 = 40%	U = (21-15)/21 = 29%	= 75/100 = 75%
W = (33-20)/33= 39%	W = (22-20)/22= 9%	

Figure 4.2 Confusion matrix accuracy assessment calculation (modified from Congalton and Green, 1999)

4.2.4. Integration into tsunami inundation modeling

The implementation of the surface roughness classification into tsunami modeling is important to show the role of surface roughness for tsunami inundation. The settings of the tsunami model have focused on the scenario of the differences between spatially distributed and the uniform surface roughness condition. The following are the explanation of tsunami modeling setup and the procedure of analyzing the tsunami inundation result. In the latter procedure, the difference and variance analysis are presented.

Tsunami model setup

The Tohoku University's Numerical Analysis Model for Investigation Near-field Tsunami No. 3 (TUNAMI N3) with 2D Nonlinear Shallow Water Equation (NLSWE) is used as the tsunami model in this research (Imamura *et al.*, 2006). Further details of the tsunami modeling setup parameters are shown in Table 4.5.

Table 4.5 The numerical parameters of tsunami model

No.	Parameter	Parameter conditions
1.	Simulation time	90 minutes (5400 seconds)
2.	Time steps	Varying from 0.2–1 second corresponding to the domain
3.	Number of domain	Six domains of nested grids
4.	Grid spacing	Domain 1 to 6, ~1850 m, ~616 m, ~205 m, ~ 68 m, ~ 23 m, and ~7.6 m
5.	Manning roughness	Four scenarios (uniform roughness, 0.025 and spatially distributed roughness (minimum, average, and maximum))
6.	Source model of hypothetical tsunami	Epicenter of sub fault: 108.68 E, 9.21 S
7.	Boundary condition	Domain 1 – open boundary. Domain 2-6 – incoming flux and water level boundaries
8.	Result recording	Flow depth and current velocity in some points as stations as well as the spatial distribution

Kongko *et al.* (2009) defined 16 reliable source models of hypothetical tsunamis in the Cilacap region; Table 4.6 provides a list of them. For this research, the source model number 11 (sc.11, depth=25 km), which is indicated as a worst case scenario of all hypothetical tsunami sources, was used.

The most credible tsunami scenarios are located close to the coastal area of Cilacap. Figure 4.3 shows the distribution of these locations in the maps. The largest dot is the source of hypothetical tsunami in this research.

Table 4.6 The reliable source models of hypothetical tsunamis

No	Epicenter of subfault			Angle Parameter			Area (km ²)	Slip (m)	Mw
	Lon (deg)	Lat (deg)	Depth (km)	Strike	Dip	Rake			
1	108.008	-9.777	5	289°	10°	95°	25,000	5.98	8
2	108.404	-9.908	5	289°	10°	95°	25,000	5.98	8
3	108.799	-10.038	5	289°	10°	95°	25,000	5.98	8
4	108.876	-9.820	10	289°	10°	95°	25,000	5.98	8
5	108.484	-9.681	10	289°	10°	95°	25,000	5.98	8
6	108.092	-9.549	10	289°	10°	95°	25,000	5.98	8
7	108.169	-9.342	15	289°	12°	95°	25,000	5.98	8
8	108.558	-9.472	15	289°	12°	95°	25,000	5.98	8
9	108.955	-9.606	15	289°	12°	95°	25,000	5.98	8
10	107.320	-9.222	10	289°	10°	95°	25,000	5.98	8
11	108.680	-9.210	25	289°	15°	95°	25,000	5.98	8
12	110.003	-10.020	10	280°	10°	95°	25,000	5.98	8
13	111.147	-10.219	10	280°	10°	95°	25,000	5.98	8
14	112.835	-10.477	10	280°	10°	95°	25,000	5.98	8
15	109.047	-9.335	25	289°	15°	95°	25,000	5.98	8
16	112.835	-10.477	10	289°	15°	95°	25,000	5.98	8

Note: The yellow row represents this research scenario.

Source: Kongko *et al.* (2009)

The crucial step in the tsunami modeling implementation is the use of different surface roughness scenarios. The results of the surface roughness classification were used as input. Four different scenarios of surface roughness coefficients were used, which are minimum, average and maximum coefficients corresponding the result from the determined surface roughness classes plus a uniform coefficient of 0.025 (Kongko *et al.*, 2009). By retaining the same hypothetical tsunami source, bathymetry and elevation data, differences in the tsunami modeling results can be assigned to the effects of different surface roughness conditions. The results of these roughness scenarios can be analyzed in terms of the need for using spatially resolved data of surface roughness conditions.

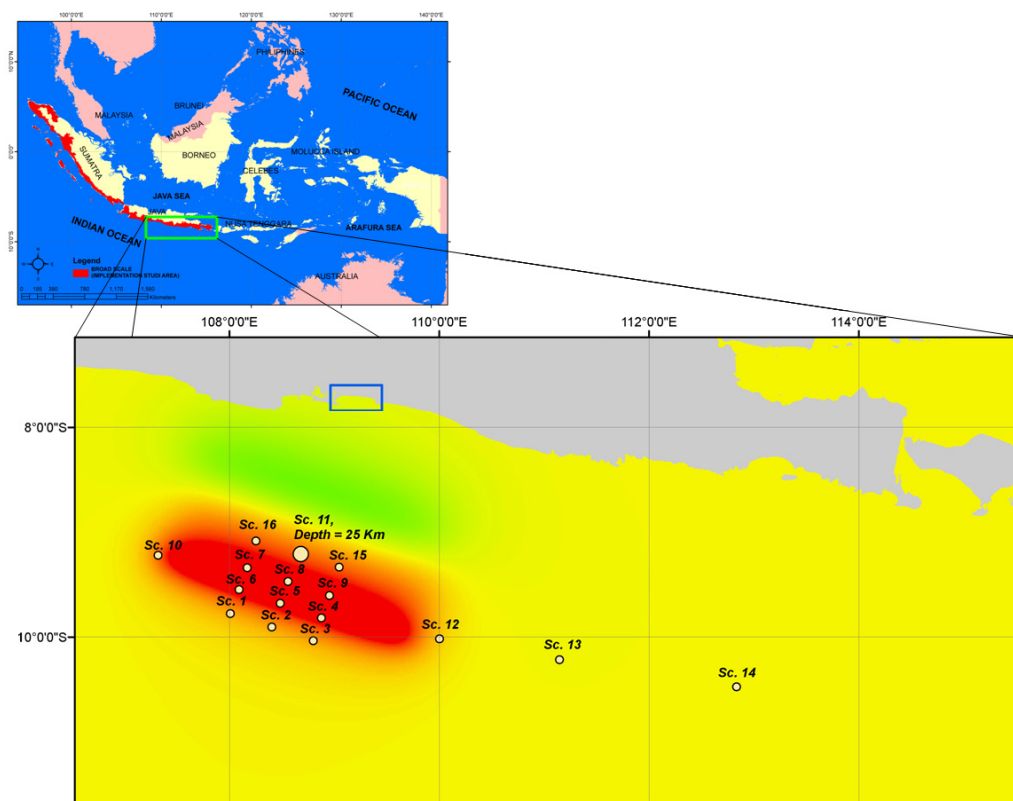


Figure 4.3 The 16 reliable source models of hypothetical tsunamis, Sc.11 is the worst case source

Analysis of Tsunami Inundation Modeling

There are two important output parameters of the tsunami modeling used for analyzing the influence different surface roughness classes, i.e. the tsunami flow depth and tsunami velocity. The analysis is conducted both point- and spatial-based.

The point-based analysis reveals differences of tsunami flow-depth and velocity of the tsunami modeling at 35 point samples (see Figure 4.4). These point samples are located in the modeled inundation area and vary in their distance to the coast (near, middle and far). The influence of difference roughness classes on tsunami flow depth and velocity is shown for each sample point. Additionally, this analysis proves the considerable influence of surface roughness condition on the modeled inundation area.



Figure 4.4 Locations of point samples to analyze the influence of difference roughness classes on tsunami flow-depth and velocity

For the area-based analysis, the absolute differences of flow depth and velocity from the uniform and spatially distributed roughness scenario are calculated. The formula is stated as follows:

$$\Delta h (\%) = \left| \frac{h_{n=U} - h_{n=S}}{(h_{n=U} + h_{n=S}) / 2} \right| \times 100\% \quad (4.10)$$

$$\Delta u (\%) = \left| \frac{u_{n=U} - u_{n=S}}{(u_{n=U} + u_{n=S}) / 2} \right| \times 100\% \quad (4.11)$$

Where:

$\Delta h (\%)$ is the absolute difference of flow depth modeling for different surface roughness condition scenarios

$\Delta u (\%)$ is the absolute difference of velocity modeling for different surface roughness condition scenarios

$h_{n=U}$ is the flow depth modeling for uniform surface roughness condition scenarios

$h_{n=S}$ is the flow depth modeling for spatially distributed surface roughness condition scenarios

$u_{n=U}$ is the velocity modeling for uniform surface roughness condition scenarios

$u_{n=S}$ is the velocity modeling for spatially distributed surface roughness condition scenario.

The area-based analysis shows the spatial variety of tsunami flow-depth, velocity and inundation extent. The spatial information coefficient of variation can be shown spatially based on this analysis. By using this analysis, the variance of tsunami inundation by either flow depth or velocity is shown. The following are the formulas to calculate the coefficient of variation of flow depth and velocity.

$$COVh (\%) = \frac{s_h}{\bar{h}} \times 100\% \quad (4.12)$$

$$COVu (\%) = \frac{s_u}{\bar{u}} \times 100\% \quad (4.13)$$

Where:

$COVh$ is the coefficient of variation of flow depth modeling for different scenarios of surface roughness condition

$COVu$ is the coefficient of variation of velocity modeling for different scenario of surface roughness condition

s_h is the standard deviation of flow depth modeling for different scenario of surface roughness condition

s_u is the standard deviation of velocity modeling for different scenario of surface roughness condition

\bar{h} is the average of flow depth modeling for different scenario of surface roughness condition

\bar{u} is the average of velocity modeling for different scenario of surface roughness condition.

4.3. Population Distribution Modeling

4.3.1. General concept

The basic concept of population distribution in this research is shown in Figure 4.5. The disaggregation of the population from CENSUS data to land use classes is the core of the modeling approach. Statistical analysis of the people's activities is used to allocate disaggregation weights.

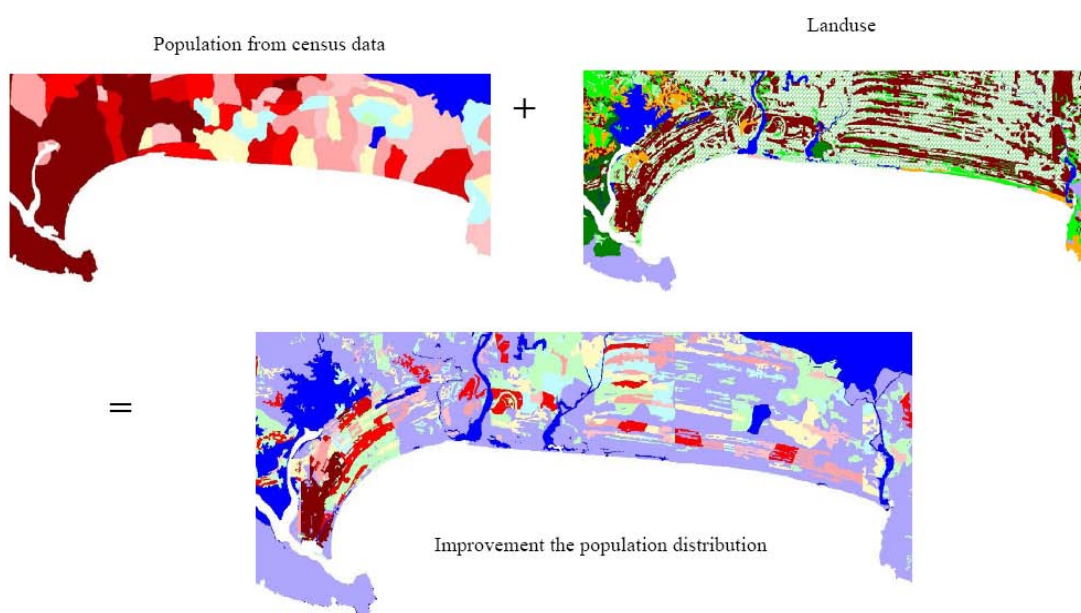


Figure 4.5 Synthetic example of the disaggregation method used to improve the spatial population data

The formulas for population disaggregation are given below:

$$X_d = \sum_{i=1}^n P_i \quad (4.14)$$

$$P_i = \sum_{j=1}^n P_{ij} \quad (4.15)$$

$$P_{ij} = \frac{S_{ij}}{\sum_{i,j=1}^n S_{ij}} \cdot W_i \cdot X_d \quad (4.16)$$

Where:

X_d is the number of people in administrative unit

P_i is the number of people in land use i

P_{ij} is the number of people in polygon j in land use i

S_{ij} is the size of polygon j in land use i in administrative unit

W_i is the weighting of land use i ; it is different during the day and night

$$\text{time} \sum_{i=1}^n W_i = 100\% \cdot$$

This set of formulas describes the distribution of the population of one administrative unit with several types of land use (Equation 4.14). The population of one land use class is the accumulation from several polygons within one administrative unit that have the same land use (Equation 4.15). For example, one land use could have more than one polygon of different sizes inside one administrative unit. Equation 4.16 takes this into account by using the size of the polygons for weighting. In addition to the areal weighting, each polygon is weighted depending on people's activities during the day- and night-time, which characterize a land use class (W_i). The model assumes that the number of outbound commuters is equal to inbound commuters. The determination of the true numbers of in- and outbound commuters is still a problem because it is rarely recorded in the statistical data. To qualify the problem, this research estimates the caused error through an evaluation of a comprehensive questionnaire.

Another problem that must be solved is the effect of dividing administrative boundaries. This problem frequently occurs in dasymetric mapping as this method distributes the number of people from administrative units to land use classes as ancillary data. A spatially connected unit of polygons of one land use class may result in different quantities of people in each single polygon if split due to an administrative boundary. As these polygons actually form a unit, the single polygons should be merged despite their belonging to different administrative units. For example, a hectare of paddy field may be located in two villages, and people carry out their activity in this area. Their activity is not confined to the administrative layer, but solely depends on the land use class. Hence, the population distribution within that paddy field should be equal and independent from the administrative boundary. This research provides a solution that keeps the overall number of people unchanged. The solution is illustrated in Figure 4.6 and equation 4.17. This finding eliminates the boundary effect and smoothes the population distribution map.

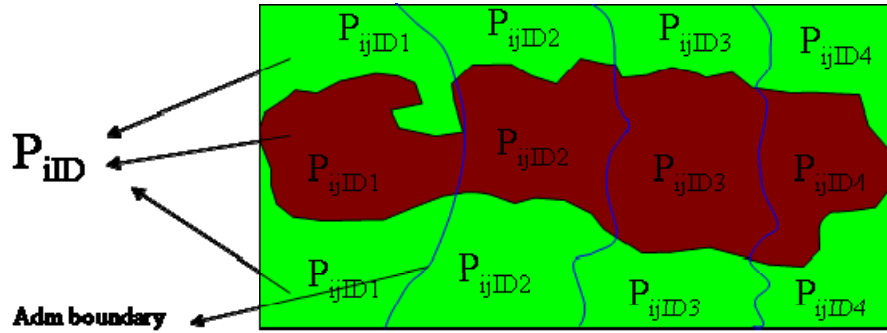


Figure 4.6 The boundary effect problem in the population distribution

$$P_{iID} = P_{ijID1} + P_{ijID2} + \dots + P_{ijIDi} \quad (4.17)$$

Where:

P_{iID} is the number of people in land use with specific ID

$P_{ijID1,2,\dots,i}$ is the number of people in polygon j in land use with specific ID.

Figure 4.6 shows as an example of a unit of settlement polygons (brown color), and two units of agriculture polygons (green color), which are each divided into four polygons due to administrative boundaries (blue lines). The total amount of people in the settlement polygons is the accumulation of the numbers of people in settlement polygons 1, 2, 3, and 4. The same applies for the two units of agriculture polygons. This means that after disaggregating the population within each administrative unit, adjacent polygons of the same land use class are merged and handled as one unit.

The overall procedure of people density distribution is packed in an automatic process model. This model calculates the population distribution based on census data and land use information automatically calculates.

4.3.2. Weighting factor determination

The crucial part of the dasymetric model is the determination of the weighting factors that distribute the population to land use classes during the day- and night-time. This research analyzes the characteristics of socio-economic data of the study area in order to determine people's activities in the land use class and

to derive the weighting factors thereof. There are two steps of weighting factor determination:

- Derivation of the potential number of people engaged in different land use activities by occupation data. It is a knowledge-based initial percentage value for describing where people are located. Furthermore, this value can be used for calculating the percentage of people engaged in land use activity for every village.
- The generalization of the weighting factor. This analysis ensures that weighting factors are specific to every village wherever necessary, or generalized by the village categories such as urban, rural, coastal and non-coastal.

To this end, socio-economic data (potential of village and occupation) are required.

Potential number of people engaged in different land use activities

Available methods developed to provide weighting factors to land use classes do not consider the behavior of people based on their activities. An improvement is needed here to provide an appropriate and objective assessment of the weighting factors to derive information as precisely as possible about people at risk. Therefore, statistical data of people's activities in certain administrative units (villages) are analysed to estimate the weighting factors.

There are two sources of statistical data in Indonesia that provide information about people's activities at the village level and that can be used for the derivation of weighting factors: potential of village data (PODES) and the census data. The PODES data set contains information on the main income sources of the population in a community and the number of workers and non-workers. Additionally, the census data provide information on the percentage of employment in different sectors in each community. These parameters provide an indication of the type, volume and locality of human activities, and can be used to calculate the potential number of people engaged in different land use activities at various times of the day. Figure 4.7 shows the potential number of people engaged in different land use by occupation data. The share of people's activities during the day- and night-time is also shown in this figure. During the night-time, almost all people stay in the settlement area, and during the day-time, they are spread in otherwise inhabitant land use area.

Based on this assumption of potential number of people engaged in different land use activities by occupation data, the calculation of where are people located in the different land use can be done in every village. The general weighting factor is then generated, as described below.

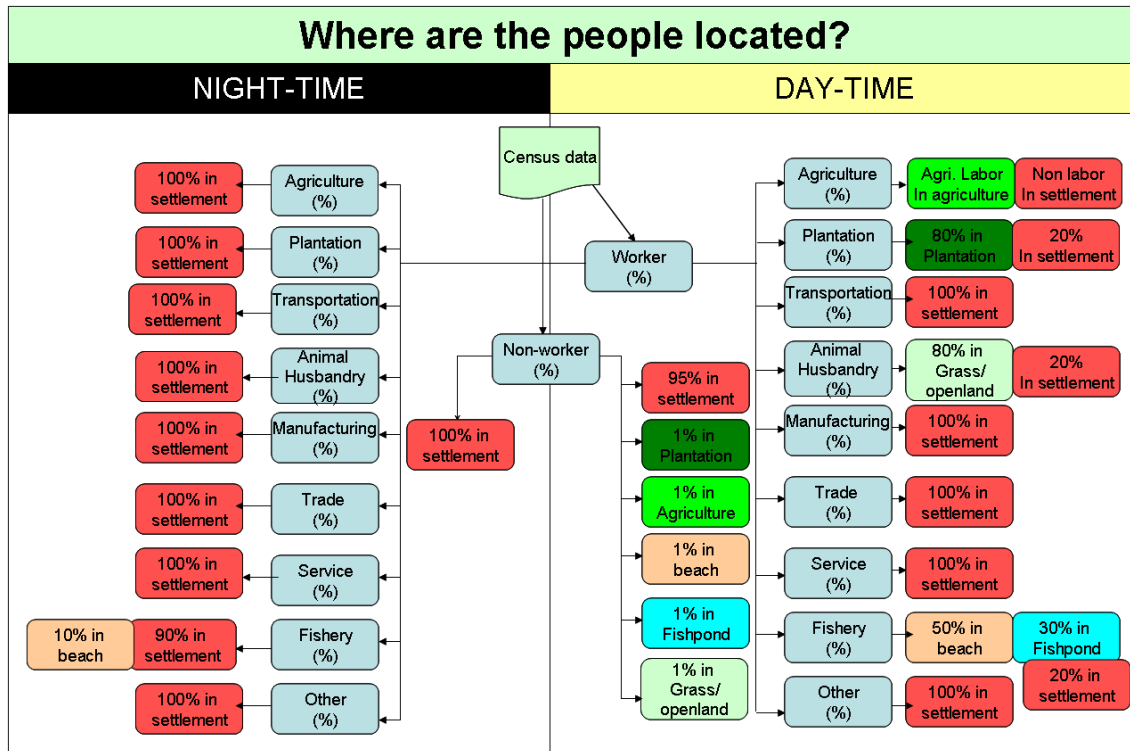


Figure 4.7 Potential number of people engaged in different land use activities during the day- and night-times in the village

Generalization of weighting factor

The first step of weighting factor generalization should answer the question “Could the weighting factors be generalized for the whole coastal area of Indonesia, or are individual weighting factors needed due to different village characteristics?” An analysis of the occupation data answers this question. There are at least 9100 villages with different characteristics in coastal Sumatra, Java and Bali.

There are four possibilities to differentiate and regionally villages category according to their characteristics, as follows:

- urban coastal, urban non-coastal, rural coastal and rural non-coastal areas;

- island (Sumatera, Java, and Bali) plus the first mentioned characteristics;
- potential economics of village;
- municipalities.

The four groups of differentiation are statistically analyzed and the decision on the weighting generalization can be made accordingly. The statistical analyses of the differentiation groups are based on standard deviation and coefficient of variation.

This analysis clarifies if the general weighting factors can be transferred to other areas or not. If it is possible, the uncertainty analysis with regard to the standard deviation of that general weighting factor can be applied.

4.3.3. Accuracy assessment

The accuracy assessment needs reference data of the true population distribution. To this end, a questionnaire was developed and distributed in the study area of Cilacap District. The main questions were:

- What is the type of working sector?
- During which time of the day do the inhabitants usually work?
- In what kinds of land use, do they usually work or stay during their activities?
- Where are these typical sites located: outside or inside their village, how far away?

Through this questionnaire, crucial information about the amount of people in an occupation sector who are working or carrying out their activities in a certain land use during day- or night-time was assessed. Additionally, information on commuters was captured. Based on the questionnaire, the true condition of population distribution in the different land use classes can be formulated as follows:

$$P_{ij}RQ = \frac{S_{ij}}{\sum_{i,j=1}^n S_{ij}} \cdot W_{RQ} \cdot (X_d + (P_I - P_O)) \quad (4.18)$$

Where:

$P_{ij}RQ$ is the number of people in polygon j in land use i in real condition

P_I is the number of inbound commuters

P_O is number of outbound commuters

W_{RQ} is the weighting factor derived from questionnaire.

The error of population distribution in the model can be calculated as follows:

$$RMSE = \sqrt{\frac{\sum (P_{ij} - P_{ijRQ})^2}{n}} \quad (4.19)$$

$$COV = \frac{RMSE}{\bar{P}} \quad (4.20)$$

$$E_{PD}(\%) = \frac{|P_{ij} - P_{ijRQ}|}{P_{ijRQ}} \times 100\% \quad (4.21)$$

Where:

$RMSE$ is the root mean square error

COV is the coefficient of variation

\bar{P} is the average population distribution

E_{PD} is the error of population distribution

P_{ij} is the number of people in polygon land use (model)

P_{ijRQ} is the number of people in polygon land use (questionnaire).

4.3.4. Multi-scale disaggregation

The availability of detailed population information or land use is a general problem in developing countries. There is frequently missing data. As a result, multi-scale disaggregation is developed to integrate the described population distribution model into the entire area in the southwestern coast of Sumatra, the south coast of Java, and Bali, even where data are not available, not completely available, or not sufficiently available. The concept of multi-scale disaggregation is shown in Figure 4.8.

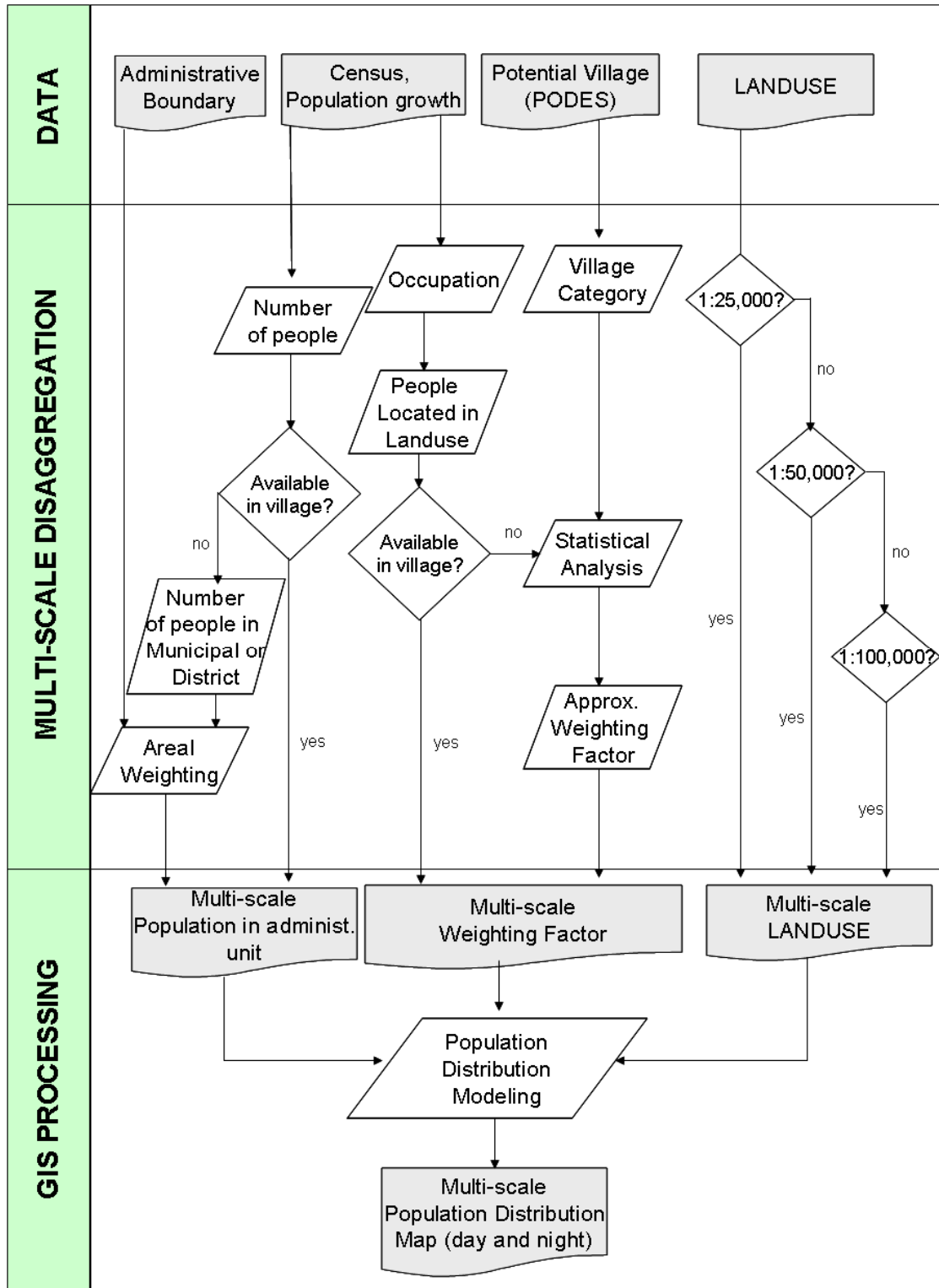


Figure 4.8 Multi-scale disaggregation concept to provide the population distribution in west coast of Sumatera, south coast of Java and Bali

To obtain a good population distribution map with low uncertainty, detailed population data, detailed information statistical data, and detailed land use data should ideally be available. In Indonesia as well as in other developing countries, there is frequently a lack of available data. This is the reason for which multi-scale disaggregation is proposed.

The first needed data set is population data (at the village level). This data is provided by the Indonesian Statistical Bureau (BPS), but the data are sometimes incomplete. To fill the gaps, this research uses the data of the next higher administrative unit, such as the sub-district and the district. By areal weighting, this data is interpolated to the village level.

The second needed data set is statistical data. As described in section 4.2.3, the weighting factors can be derived from this data. This data set is also incomplete in some parts of the study area. This gap is filled by a general weighting factor that is derived by the statistical analysis from the section 4.2.3. The result of general weighting factor can be used for the process of multi-scale disaggregation.

The final, necessary data set is land use data. These data are provided by the Indonesian institutes responsible for mapping, Bakosurtanal and LAPAN. Land use data can be available at scales of 1:25,000, 1:50,000 and 1:100,000. For each area, the largest available scale is used and thus, a best land use data is created.

After preparation of all these data sets, the population distribution model can be applied to provide the population distribution map.

4.3.5. Comparative study

In addition to the accuracy assessment step for assessing the quality of the developed model, the model is also compared to the results of the models from Gallego and Pedell (2001) and Mennis (2003), which were implemented in the same study area of Cilacap. Root mean square error (RMSE), coefficient of variation (COV) and error of population distribution (E_{PD}) are calculated to compare those models.

4.4. Settlement Classification by Using Radar Imagery Based on Speckle Divergence and Neighborhood Analysis

4.4.1. Classification procedure

The settlement classification method is an improvement over the approach of Esch *et al.* (2010). This research proposes the neighborhood analysis from the "real settlement" image using the speckle divergence and intensity image. The process of settlement classification is performed in two steps: pre-processing and classification (neighborhood). To ensure efficiency and quality of the model, an accuracy assessment and transferability analysis are carried out. The steps of the settlement classification are shown in Figure 4.9.

In compact steps, the classification procedure is stated as follow:

- pre-processing (speckle divergence analysis):
- threshold analysis (threshold determination for real settlement and potential settlement):
- neighborhood analysis (classification of settlement areas).

Pre-processing

The pre-processing of the data includes speckle suppression to enhance images degraded by speckle noise. This speckle phenomenon in TerraSAR-X data significantly hampers the analysis of radar data. The approach of speckle suppression consists of determining the difference between the local image heterogeneity and the theoretical, scene-specific heterogeneity of developed speckle.

The calculation of speckle divergence is based on kernel size of 9 x 9 pixels, and the formulas are stated as follows (Esch *et al.*, 2010):

$$D_{x,y} = C_{x,y} - C \text{ with } C_{x,y} = \frac{\sigma_{x,y}}{\mu_{x,y}} \quad (4.22)$$

Where:

- $D_{x,y}$ is the speckle divergence
- C is the theoretical heterogeneity due to developed speckle; it is calculated from the inverse of equivalent number of look (ENL),

- $ENL = L_a + L_r = 1/C$, L_a and L_r defining the effective number of looks in the azimuth and range (stated in TerraSAR X metadata)
- $C_{x,y}$ is the local coefficient of variation defined by the local, it is calculated in kernel size of 9 x 9 pixels
- $\sigma_{x,y}$ is the local standard deviation, it is calculated in kernel size of 9 x 9 pixels
- $\mu_{x,y}$ is the local mean, it is calculated in kernel size of 9 x 9 pixels.

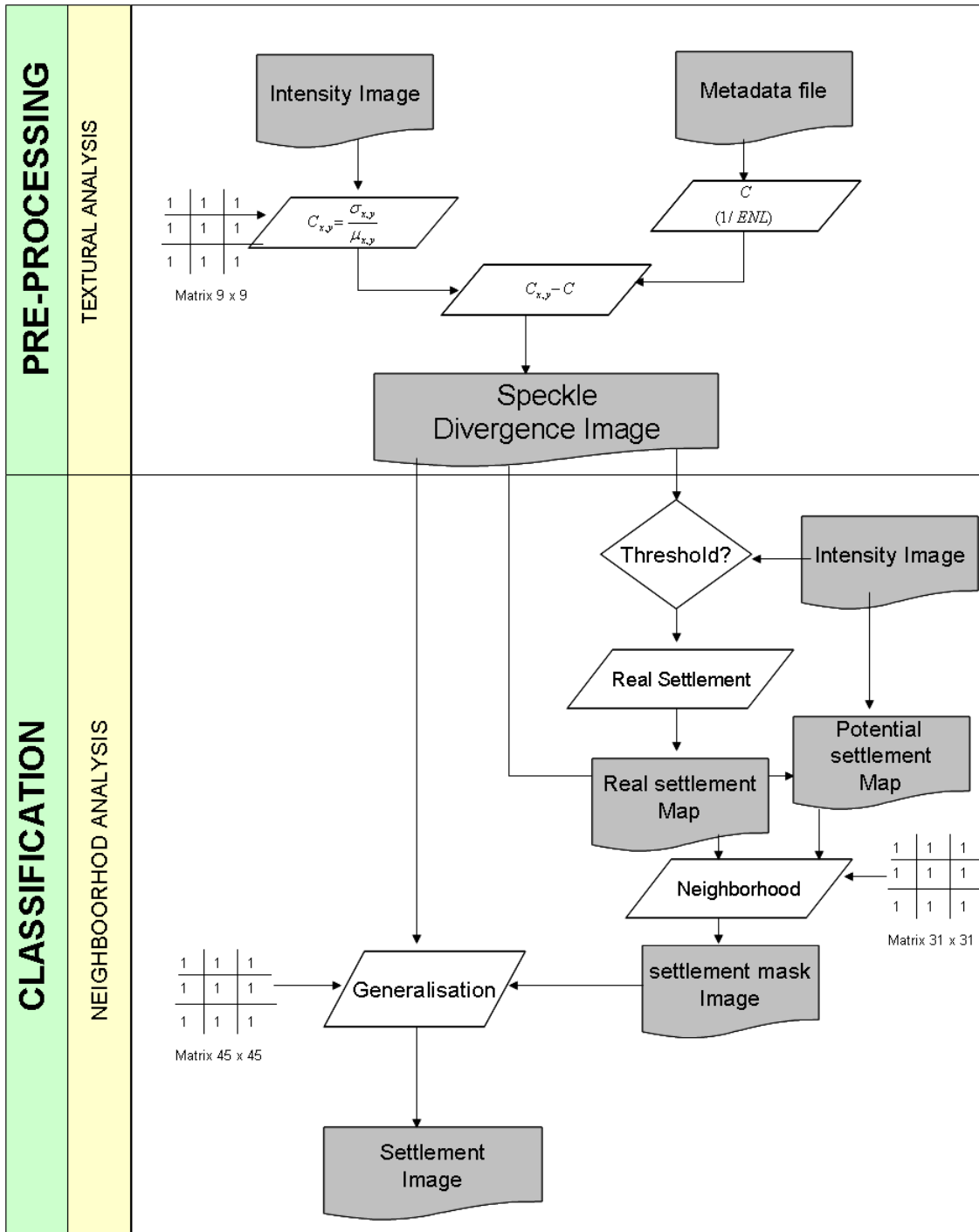


Figure 4.9 The steps of settlement classification by using TerraSAR-X

Threshold analysis

The threshold of real and potential settlement from the intensity and speckle divergence image is defined by analyzing samples of pixels in settlement and non-settlement areas. The threshold can be decided by analyzing the histogram of samples, as show in Figure 4.10.

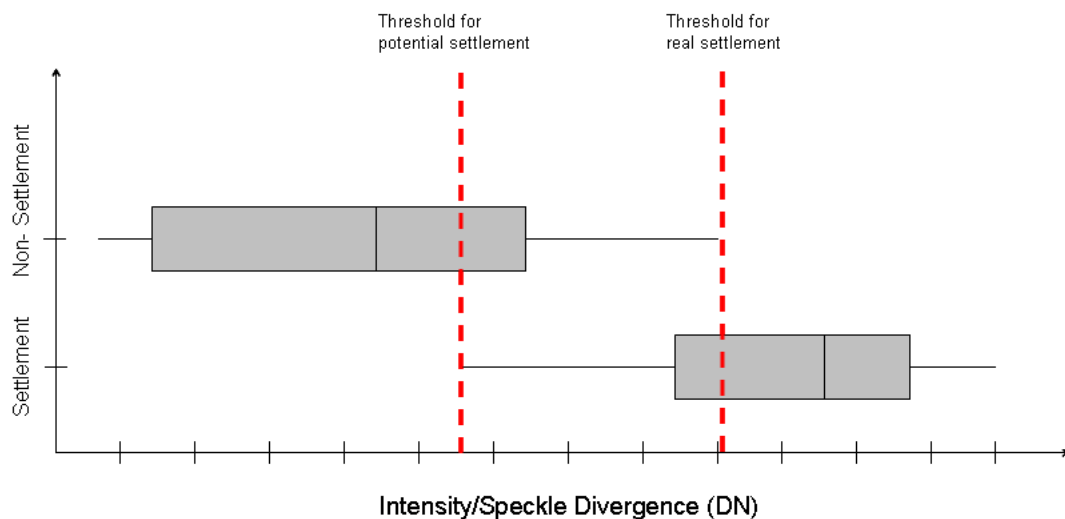


Figure 4.10 The threshold decision for potential settlement and real settlement

Neighborhood analysis

The settlement classification can be performed by generating a simple threshold of the speckle divergence, but this method cannot eliminate false settlement detections in hilly areas that have similar speckle divergence values. To avoid this error, it is proposed to use neighborhood analysis for the detection of settlement areas from TerraSAR-X imagery. The steps are as follows:

- identification of "real settlement" pixels from speckle divergence and intensity images by using the threshold;
- identification of "potential settlement" pixels from speckle divergence and intensity images by using the threshold;
- a combination of "real settlement" and "potential settlement" pixels to identify settlement areas by conditional parameters (settlement mask).

The neighborhood analysis for settlement classification is shown in Figure 4.11:

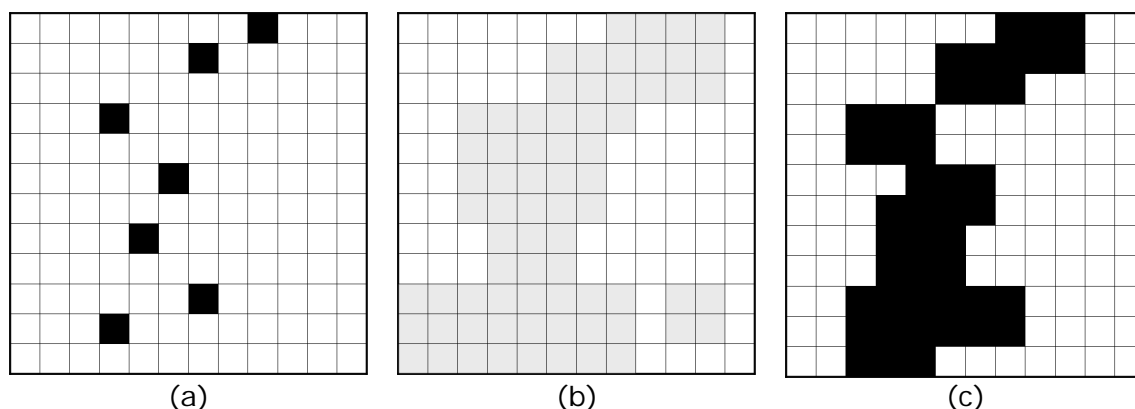


Figure 4.11 Illustration the neighbourhood analysis for settlement classification (a) real settlement detection (b) potential settlement and (c) the settlement mask by neighbourhoods a kernel size of 3 x 3 pixels

The resulting settlement mask is used to mask the intensity image to develop an “urban scatter” image. Then, the density of this image is analyzed by calculating the focal mean of a kernel size of 25 x 25 pixels and accordingly the “urban scatter” by a kernel size of 45 x 45 pixels. The threshold of the density image can be generated by analyzing several pixels to detect the settlement. Finally, the settlement classification is generalized by combining the settlement mask and the speckle divergence image, and filtering by a 45x45 matrix.

4.4.2. Accuracy assessment

Accuracy assessment is carried out by using a higher resolution reference map (e.g. from national mapping agencies) or by in situ mapping in a field survey. By using the same accuracy assessment method as for the surface roughness classification (see section 4.2.3), the quality of the settlement classification methodology is assessed. Furthermore, the omission and commission errors are calculated.

4.4.3. Transfer of the approach to other areas

The same classification methodology is applied to the Padang area. The accuracy assessment is performed accordingly to check the quality and transferability of the methodology, and an error map is created.

CHAPTER 5: RESULTS

This chapter describes the main results of data collection, surface roughness classification and its implementation to tsunami modeling, population distribution modeling and settlement classification using TerraSAR-X data. The results of accuracy assessments and comparative studies are also presented.

5.1. Data Collection

Data collected and processed can be divided into three categories: satellite imagery and spatial information, demographic and socio-economic statistics, and reference data.

5.1.1. Satellite imagery and spatial information

Satellite imagery has been acquired from optical or radar sensors, providing important data for this study (see Figure 5.1). They are preliminarily used for surface roughness determination and settlement classification. For the surface roughness determination, SPOT-5 multi-spectral 10 m and panchromatic 2.5 m spatial resolution have been collected. This imagery has less than 10% cloud coverage. It covers the entire coastal area of the research site. TerraSAR-X imagery was used (see Figure 5.1) for settlement classification based on speckle divergence and neighborhood analysis. Three scenes of TerraSAR-X imagery for the study area in Cilacap and one scene for the implementation area in Padang were used. All TerraSAR-X Imageries were acquired in the strip map mode, which have 2.7 m spatial resolution and were provided by the German Aerospace Center (DLR).

By using both types of data, a comprehensive study on the application of optical and radar imagery is presented to support the assessment of risk and vulnerability. The advantages and disadvantages of these data can be directly proven.

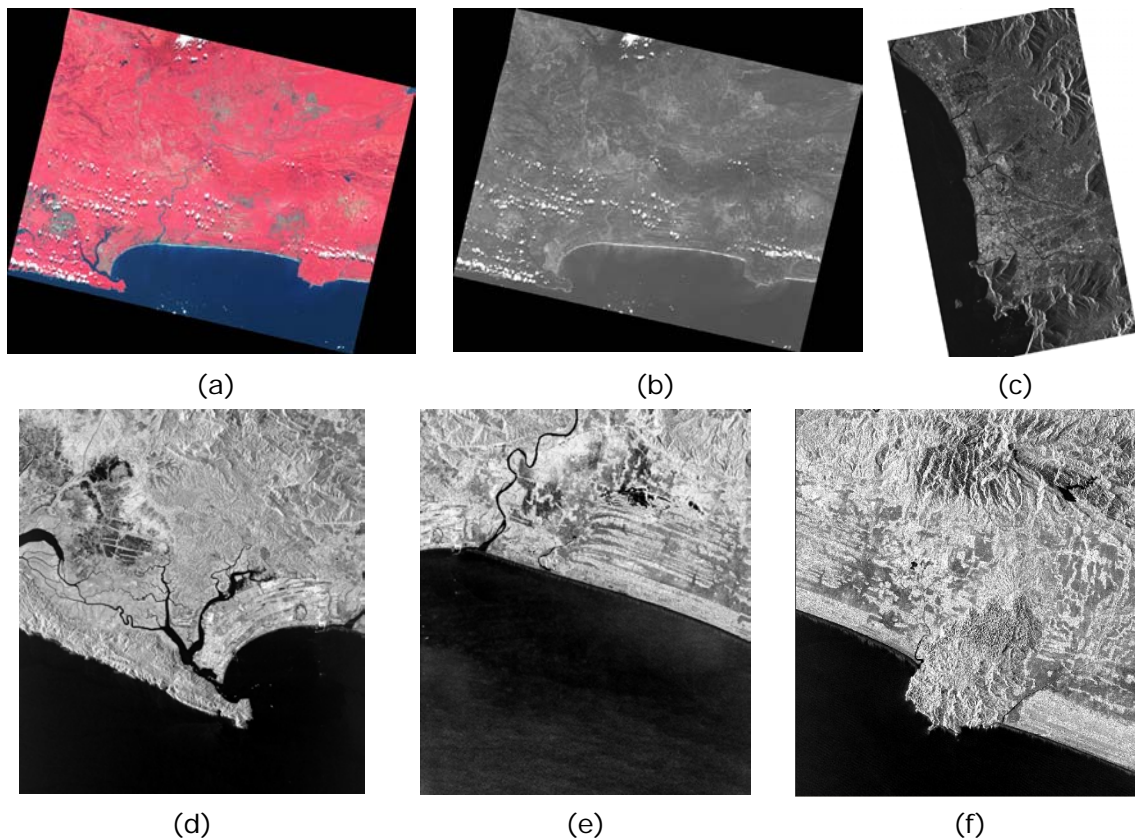


Figure 5.1 Satellite imagery data collection for this research: (a) SPOT-5 multi-spectral (24 June 2004); (b) SPOT-5 Panchromatic (24 June 2004); (c) TerraSAR-X for Padang (04 April 2008); (d) TerraSAR-X Cilacap 1 (24 January 2009); (e) TerraSAR-X Cilacap 2 (04 February 2009); and (f) TerraSAR-X Cilacap 3 (15 February 2009)

For population distribution modeling research, the spatial information of land use has also been collected. These data were provided by the National Coordinating Agency for Surveys and Mapping (Bakosurtanal) and the National Institute of Aeronautics and Space (LAPAN). Several scales of land use maps have been collected for the whole west coast of Sumatra, Java and Bali. For Java and Bali, land use maps were derived from the topographic map at a 1: 25,000 scale. For Sumatra, maps were available at a 1: 100,000 scale, but maps of some of the areas, such as Padang and Aceh, are at a scale of 1: 50,000. Figure 5.2 shows the availability of land use data in the study area of Cilacap and in the implementation area of the entire west coast of Sumatra, Java and Bali.

In this research, the best available data were compiled from the responsible mapping institutions in Indonesia. As a developing country, the availability of detailed maps in Indonesia is a general and frequent problem.

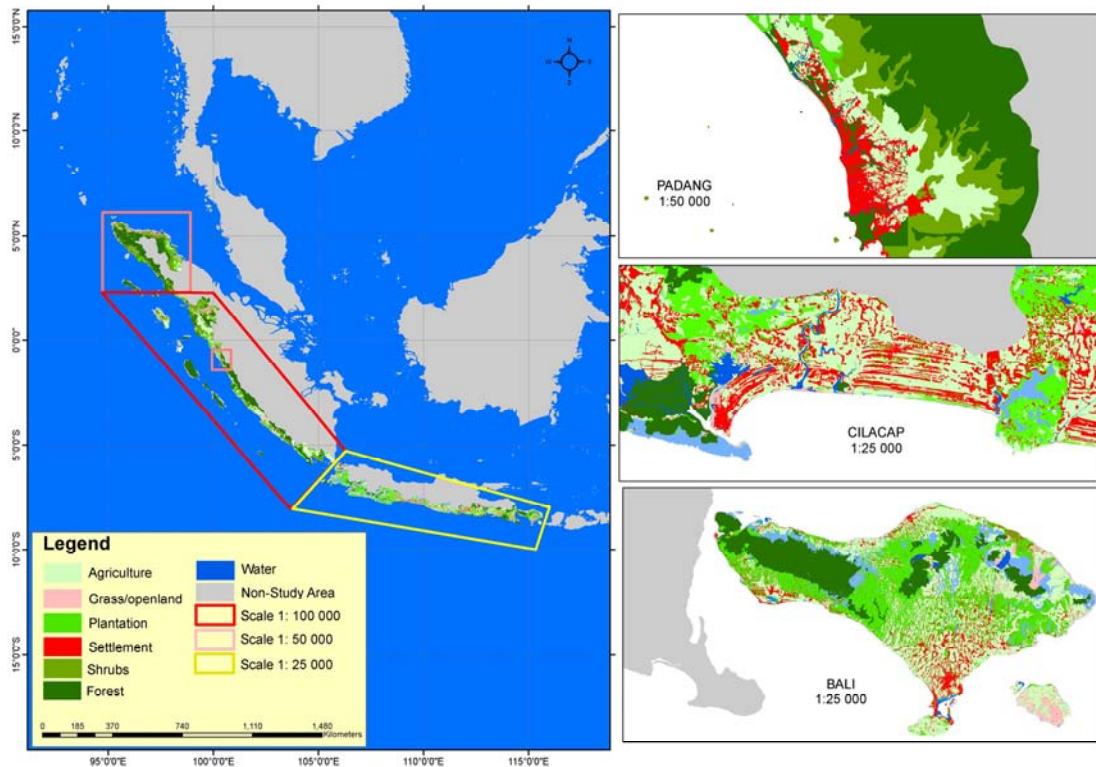


Figure 5.2 Land use data collection in difference scales (1:100,000 scale in the red box, 1: 50,000 scale in the pink box, and 1: 25,000 scale in the yellow box.)

Further important spatial data are bathymetry and terrain/elevation data. The General Bathymetric Chart of the Oceans (GEBCO) bathymetry data with 30 arc-seconds spatial resolution was used. Where available, the more detailed bathymetry data from the National Coordinating Agency for Surveys and Mapping (Bakosurtanal) and observations from the Agency for the Assessment and Application of Technology (BPPT) (single multi-beam sounding) were used. Hence, a best-of bathymetry data set was then used for the tsunami modeling. As elevation data served, the digital terrain model (DTM) with 5 m horizontal spatial resolution, which was derived from airborne SAR interferometry by Intermap was also collected. These data were used as input for tsunami inundation modeling in the study area.

5.1.2. Demographic and socio-economic statistics

Statistical data are important for population distribution modeling and show the number of people in an administrative unit. The main source of socio-economic and demographic statistical data is the Indonesian Statistical Bureau (BPS). It provides census data for 2000 and the population growth rate from 2000 to 2007. These data were collected in order to obtain the number of people at the village administrative level. The population growth rate is used for the future projection of the number of people per village. Figure 5.3 shows the number of people per village based on Census data (BPS, 2000).

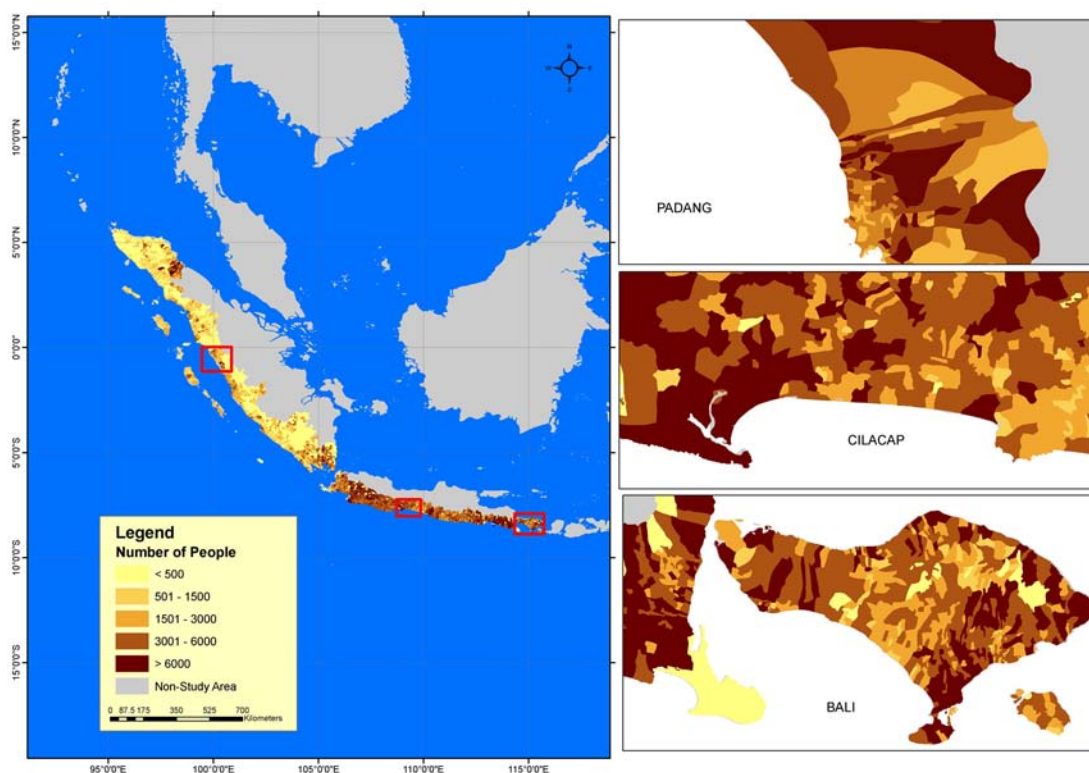


Figure 5.3 Statistical data on population distribution at the village level for the entire coast of Sumatra, Java and Bali

Another source of statistical data is the BPS potential of village data (PODES, 2005) which are updated every five years. It provides socio-economic data explaining the economic characteristics of each village in Indonesia: the main source of income, farmer population, the main commodity, number of infrastructures in the village, type of village (urban, rural, coastal, and non-coastal), total number of people and other information. PODES data are an important source for population distribution modeling.

Another important statistical data is the occupation data for every village, documented in the BPS Census data. The raw census data have been collected and the percentages of people in their respective occupations were calculated. There are ten types of occupations in the census data: food crops, fishery, other agriculture, trade, transportation, plantation, animal husbandry, manufacturing, services and other occupations.

5.1.3. Reference data

Reference data are needed to prove the validity of this research. For the surface roughness determination, the reference for the land use and surface roughness classes was collected by field survey. Figure 5.4 shows the location of field survey and land use delineation used as reference data.

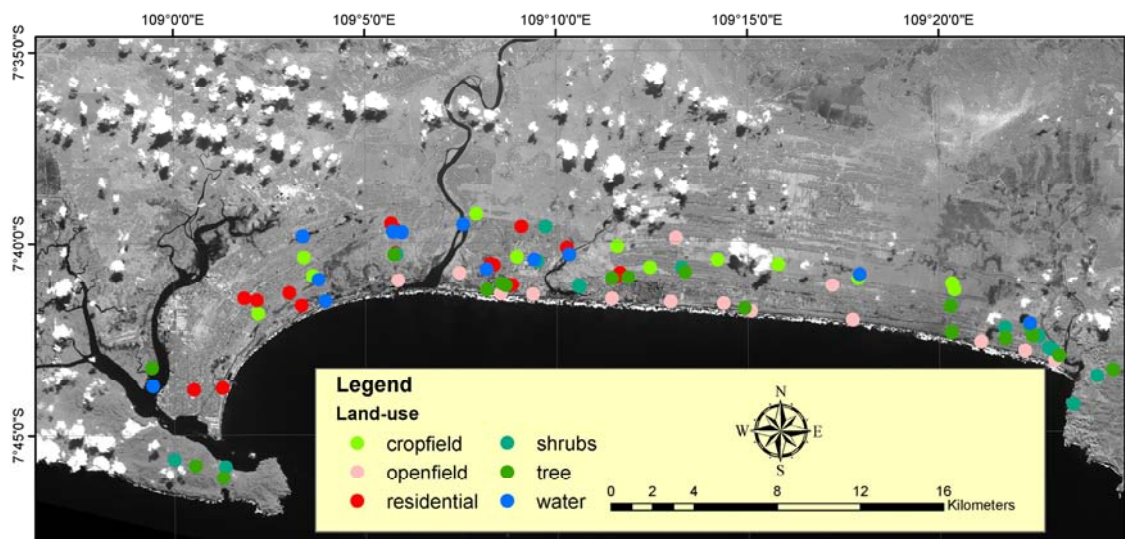


Figure 5.4 Example of land cover reference polygons for the accuracy assessment of the surface roughness classification

For population distribution modeling, the accuracy assessment needs reference data on the true population distribution. To this end, a questionnaire was developed and distributed in the study area of Cilacap District. More than 300 households with a total of 1,119 family members, 207 individuals, and a total 1,326 individuals in 36 villages in Cilacap District responded to this questionnaire. The result of this questionnaire data collection is shown in sub-chapter 5.3.3.

Topographic maps and very high-resolution optical satellite imagery (SPOT-5 and Quickbird) data are used to generate the reference settlement maps for TerraSAR-X classification in Cilacap and Padang; Figure 5.5 shows these maps.

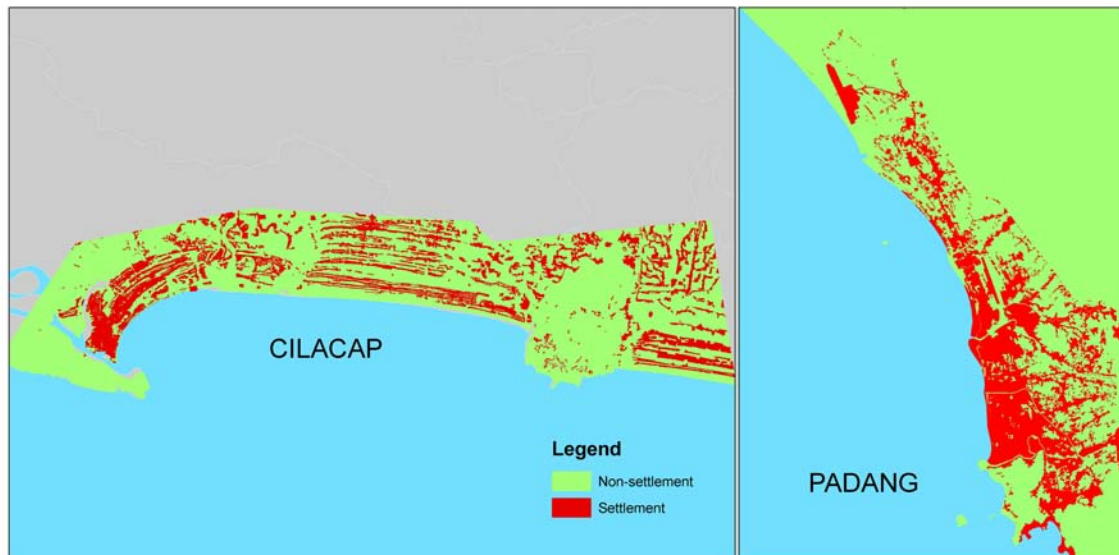


Figure 5.5 The reference data for settlement classification, Cilacap (left) and Padang (right)

The other reference data that have been collected are the tabular value of roughness coefficient in published research (Chow, 1959; Hill and Mader, 1987; and Murashima *et al.*, 2008). This published roughness coefficient values are used further on the determination of surface roughness class and its coefficient estimation (see chapter 5.2.1).

A compilation of all data sets collected and used in this research both for model development and implementation is shown in Table 5.1. It explains type of data, characteristics, resolution/scale, source and usage.

Table 5.1 Compilation of all data used in this research

Type of data	Data Characteristics	Resolution/scale	Source	Usage
Satellite data and other spatial information	SPOT-5 Multi-spectral Cilacap: Acquisition : 24 June 2004	10 m	Data based on SPOTIMAGE2004	Surface roughness determination
	SPOT-5 Panchromatic Cilacap: Acquisition : 24 June 2004	2.5 m	Data based on SPOTIMAGE 2004	Surface roughness determination
	TerraSAR-X Stripmap Mode Cilacap district (3 scenes) Acquisition dates: 1. 24 January 2009 2. 04 February 2009 3. 15 February 2009 Padang district (1 scene) Acquisition date: 1. 03 April 2008	2.7 m	Data based on DLR 2008/09	Settlement Classification
	Cilacap DTM	5 m	Data based on Intermap 2008	Surface roughness determination (tsunami inundation modeling)
	Bathymetry data	±925 m	Data based on GEBCO 2003	Surface roughness determination (tsunami inundation modeling)
	Bathymetry data	250 – 500 m	Data based on Bakosurtanal 2007	Surface roughness determination (tsunami inundation modeling)
	Bathymetry from single multi-beam sounding	0.2 m	Data based on BPPT 2007–2009	Surface roughness determination (tsunami inundation modeling)
	Topographic maps: south coast of Java and Bali)	1 : 25,000	Data based on Bakosurtanal 1999	Population distribution modeling (population distribution disaggregation)
	Land use data: Padang	1 : 50,000	Data based on LAPAN 2005	Population distribution modeling (population distribution disaggregation)
	Land use data (City of	1 : 50,000	Data based on	Population distribution modeling

	Banda Aceh)		LAPAN 2005	(population distribution disaggregation)
	Land use data (west coast of Sumatera, south coast of Java and Bali)	1:100,000	Data based on LAPAN 2005	Population distribution modeling (population distribution disaggregation)
Demographic and Socio-economic statistics	Census 2000	Village level,	Data based on BPS 2000	Population distribution modeling (weighting factor determination)
	Potential of village data (Podes) 2005	Village level	Data based on BPS 2005	Population distribution modeling (weighting factor determination)
	Population growth rate 2000–2007	District level	Data based on BPS 2007	Population distribution modeling (population distribution disaggregation)
	Administrative boundary	Village level	Data based on BPS 2004	Population distribution modeling (population distribution disaggregation)
Reference data	Land use reference polygons	-	Data based on Field survey 2007-2009	Surface roughness determination (accuracy assessment)
	Questionnaire	-	Data based on Field survey 2009	Population distribution modeling (accuracy assessment)
	Published roughness coefficients	-	Chow (1959); Hill and Mader (1987); and Murashima <i>et al.</i> (2008)	Surface roughness determination (estimation of roughness coefficient)
	Settlement reference polygons	-	Data based on SPOTIMAGE 2004 and Topographic map Bakosurtanal, 1999, Digitalglobe 2006	Settlement classification (accuracy assessment)

5.2. Surface Roughness Determination

5.2.1. Surface roughness classes and their coefficient estimation

This study determines 12 classes of surface roughness by using the roughness coefficients in published research and field survey results (see Table 5.1). These classes were derived based on: (1) the land use; (2) the density; and (3) neighborhood patterns. The land use classes are: "water", "field", "shrubs", "trees" and "residential area". The density classes are "low", "medium" and "high" density of object (e.g. vegetation) and the neighborhood pattern is a class which has a combination of two land-use classes such as "residential area with trees" (see section 4.2.2 in *Methodology*). The 12 classes of surface roughness, their coefficient and the deviation are shown in Table 5.2. The roughness classes are grouped into four qualitative classes — "very low", "low", "medium" and "high". This determined surface roughness is used as a basis for developing the new remote sensing classification methodology.

Table 5.2 The roughness coefficients assigned to surface roughness classes based on literature review

No.	Land use	Roughness coefficient	Class of roughness
A.	Water		
1.	River, lake and ocean	0.016±0.009	Very low
B	Field		
2.	Open field	0.030±0.010	Low
3.	Crop field	0.035±0.010	Low
C.	Shrubs		
4.	Low-medium density of shrubs	0.050±0.020	Medium
5.	Shrubs and trees	0.060±0.020	Medium
6.	Dense shrubs	0.100±0.010	Medium
D.	Trees		
7.	Low density of trees	0.060±0.010	Medium
8.	Medium density of trees	0.100±0.020	Medium
9.	Dense tree	0.150±0.050	High
E	Residential		
10.	Low-medium density of residential area	0.050±0.010	Medium
11.	Dense residential area	0.115±0.020	High
12.	Residential area with trees	0.150±0.050	High

Note: The definition of the surface roughness classes is based on the combination of land use, density and neighbourhood patterns.

The surface roughness is an indicator of the protection probability of land use against tsunami threats. Low roughness values thus lead to a low protection probability. The surface roughnesses, which indicate high probabilities of protecting against tsunamis, are "dense tree", "dense residential" area and "residential area with trees" (see Figure 5.6). "River", "lake", "ocean", "open field" and "crop field" have low protection probability (low roughness values, see Figure 5.6).

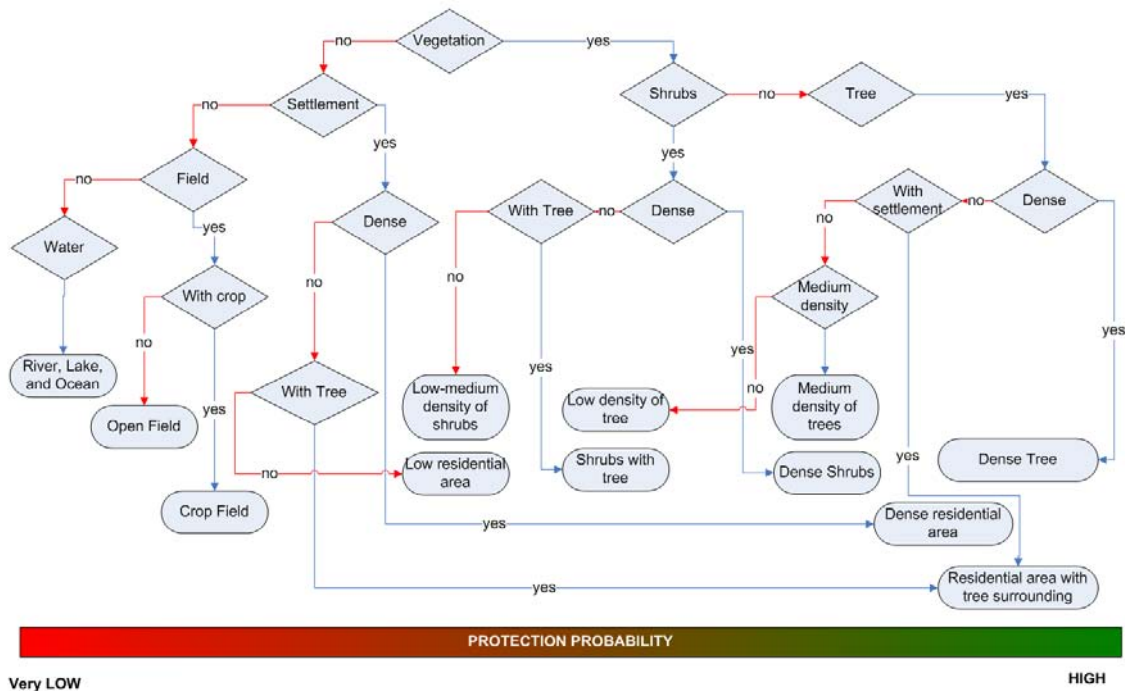


Figure 5.6 The surface roughness classes and their relation to the tsunami protection probability (from very low to high)

5.2.2. Surface roughness classification

Pre-processing result

In order to enhance the spatial resolution, SPOT-5 multi-spectral data were pan-sharpened by using the panchromatic channel and multiplicative algorithm (see chapter 4.2.2.1). The different steps, before and after spatial enhancement, are shown in Figure 5.7. After this pre-processing, a set of SPOT 5 multispectral imagery with 2.5 m spatial resolution was obtained.

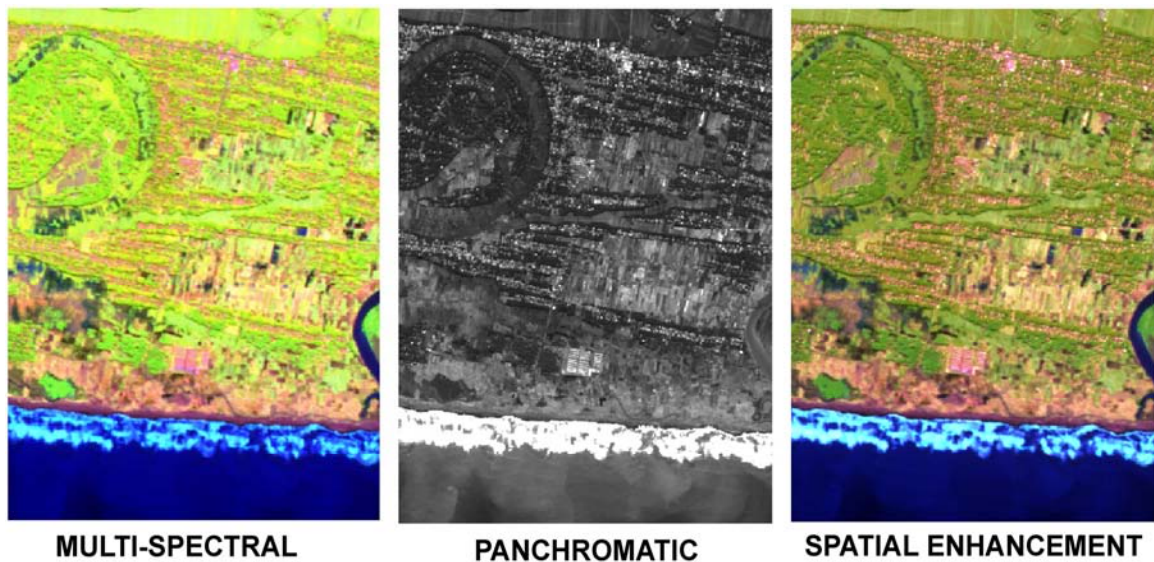


Figure 5.7 Result of the spatial enhancement of a SPOT-5 scene (right), based on a multi-spectral (left) and panchromatic imagery (middle)

The result of image variable calculation is shown in the Figure 5.8. The NDVI, SNDVI and ENDVI have been calculated to differentiate the vegetation cover.

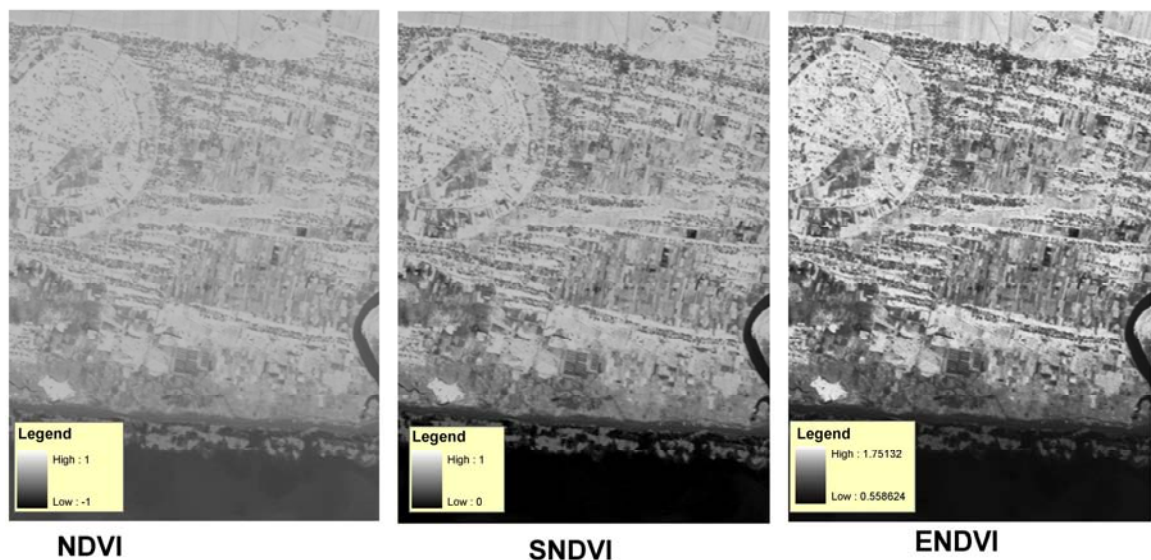


Figure 5.8 Results of image variable calculation of a SPOT-5 scene based on NDVI (left), SNDVI (middle) and ENDVI (right)

The next pre-processing step is the texture analysis (see Figure 5.9), which is important mainly for settlement extraction. Figure 5.9 shows that the settlement areas in the texture imagery become more visible than in the

panchromatic imagery itself. The differentiation of residential areas and open/bare land has significantly improved (see next chapter).

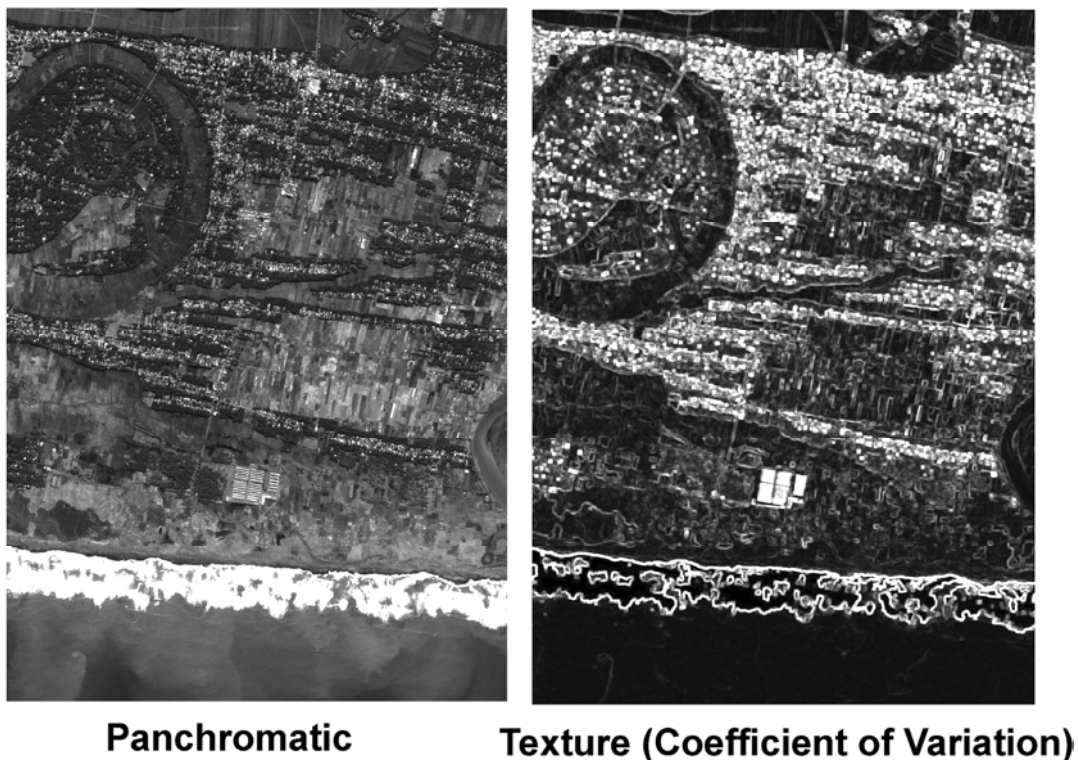


Figure 5.9 Result of textural analysis of a SPOT-5 scene: panchromatic imagery (left) and the texture imagery (right)

Relevant remote sensing variables

Performing a multispectral analysis of SPOT-5 imagery, the following ratios and variables were calculated: NDVI, sigmoid NDVI, exponential NDVI and texture variable. These indices were then used to classify the main land use classes and to derive surface roughness. The result of this analysis together with their internal accuracies is shown in Table 5.3. Based on Table 5.3, the lowest accuracy is found in the combination that uses merely one variable of SPOT-5 imagery (see No. 1-4 in Table 5.3). By using the four channels of SPOT-5 only, the accuracy of classification is no more than 70% (see No. 5 in Table 5.3). The accuracy is higher (over 80%) when using additional variables such as NDVI. Adding texture information, the accuracy is over 90% (Nos. 10 and 11 in Table 5.3). The highest accuracy is obtained when using the combination of four channels of SPOT5, NDVI and texture for land use classification (No. 10. in Table 5.3). Texture significantly increases the accuracy of the classification (from around 80% up to over 90%). The expansion value of NDVI (ENDVI and

SNDVI) did not significantly increase the result of accuracy (see Table 5.3). This result was used in the following step in order to design a decision tree model.

Table 5.3 The accuracy of SPOT-5 data image variables to classify the main classes of surface roughness by using a decision tree model

No.	Remote sensing variable combination	Accuracy of the decision tree model (%)
1.	Only ENDVI	48.14
2.	Only NDVI	48.00
3.	Only SNDVI	48.29
4.	Only Texture	46.00
5.	Four spatially enhanced spectral bands of SPOT-5	68.09
6.	NDVI and Texture	74.00
7.	Four spatially enhanced spectral bands of SPOT-5 and NDVI	83.29
8.	Four spatially enhanced spectral bands of SPOT-5, NDVI and SNDVI	82.86
9.	Four spatially enhanced spectral bands of SPOT-5, ENDVI and SNDVI	81.57
10.	Four spatially enhanced spectral bands of SPOT-5, NDVI and Texture	94.14
11.	Four spatially enhanced spectral bands of SPOT-5, NDVI, SNDVI, ENDVI and Texture	94.43

Table 5.4 shows an accuracy matrix of the decision tree model by using selected relevant variables. The overall accuracy of the model is 94.14%, indicating that the application of this model to the classification procedure can be reached with a high accuracy. The difficulties of this model are the differentiation of "tree", "shrubs" and "crop field" that have a high omission and commission error (> 10%). Mis- and false classification among those three classes occurred accordingly.

Table 5.4 The accuracy of decision tree model by using relevant SPOT-5 variable (four spatially enhanced spectral bands of SPOT-5, NDVI, and Texture)

Classification Reference	Open field	Residential	Sea wave	Shrubs	Tree	Water	Crop field	Total	Omission error (%)
Open field	95	1	0	0	0	3	1	100	5
Residential	2	97	0	0	0	1	0	100	3
Sea wave	0	0	98	0	0	2	0	100	2
Shrubs	0	0	0	88	3	0	9	100	12
Tree	1	0	0	2	97	0	0	100	3
Water	2	0	1	0	0	97	0	100	3
Crop field	2	0	0	10	1	0	87	100	13
Total	102	98	99	100	101	103	97	700	
Commission error (%)	6.9	1	1	12	4	5.8	10.3	Overall	94.14

Land use classification using decision tree model

A decision tree model has been developed using the best combination of image variables (No. 10 in Table 5.3) for land use classification. Figure 5.10 shows this decision tree model, developed by including the automatically derived thresholds to classify the land use classes. This is the basis for further surface roughness classification.

Figure 5.10 shows the classification of the main land use resulting from the decision tree model. In this figure, NDVI is the main variable for differentiating classes: "water", "open fields" and "residential areas" with trees classes such as "crop fields", "shrubs" and "trees". This figure also shows the relevant variables for the differentiation of the main classes. For example, to differentiate residential areas and open field, NDVI channel 1, channel 4 and texture are the most relevant parameters. For residential areas and water, texture information is needed in addition.

Texture clearly plays an important role for the differentiation of residential areas and open fields.

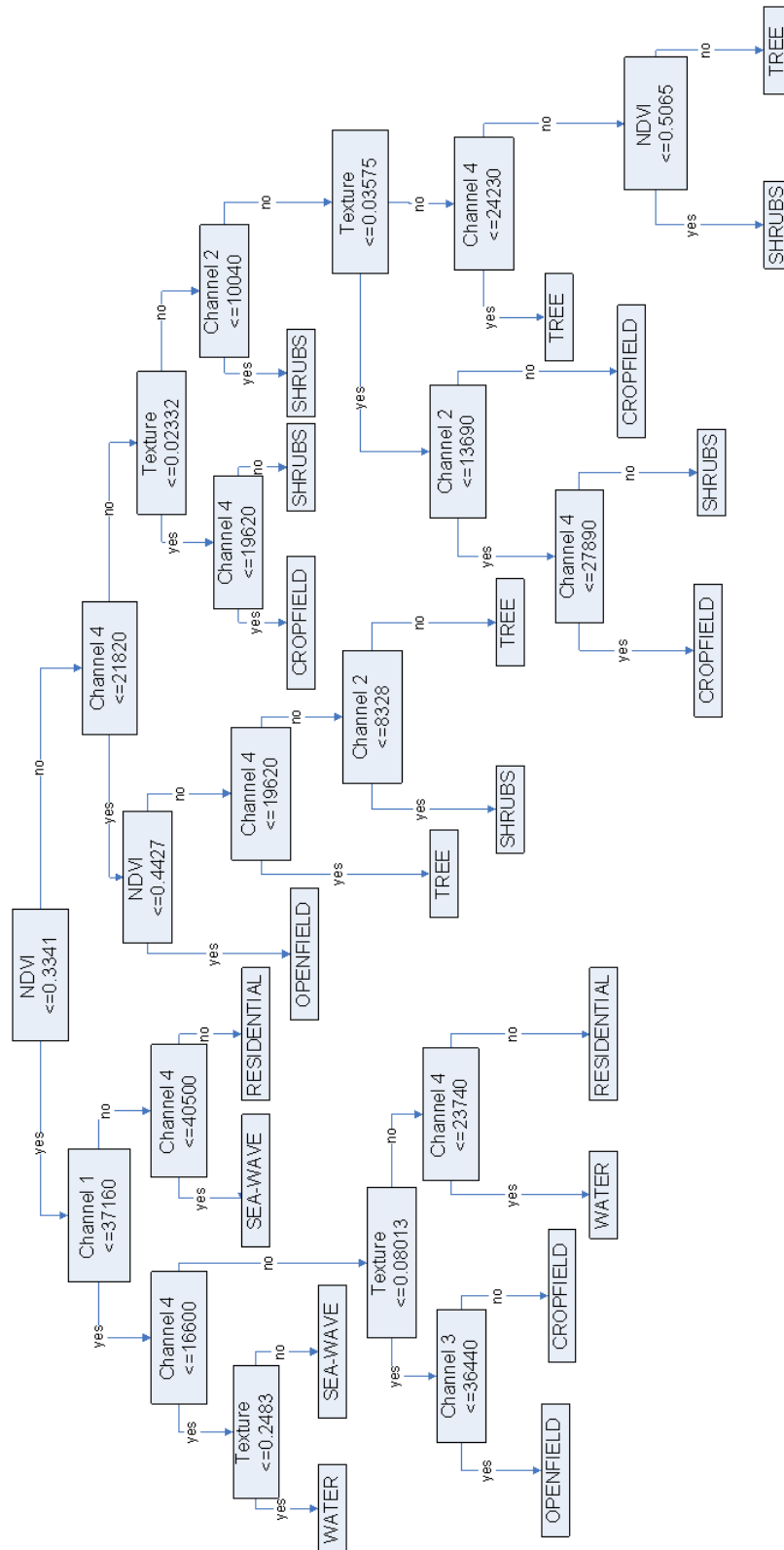


Figure 5.10 The decision tree model for main land use classification.

The result of main land use classification and its accuracy assessment is shown in Figure 5.11 and Table 5.4, respectively. The overall accuracy of main land use classification is 91.6% with the highest omission (25.6%) and commission error (18.8%) for the shrubs. There are still some problems, therefore, with respect to the separability between "shrubs" and "crop fields" or "trees". The confusion matrix in Table 5.4 shows the percentage of shrubs classified as "crop field" or "tree". The highest accuracy is for water classification, with only 0.5% commission error and 0.2% omission error. This method has also a good accuracy on the residential classification, with only 0.7% as commission error and 3.1% as omission error. Open field was classified with only a 1% omission error and a 5% commission error.

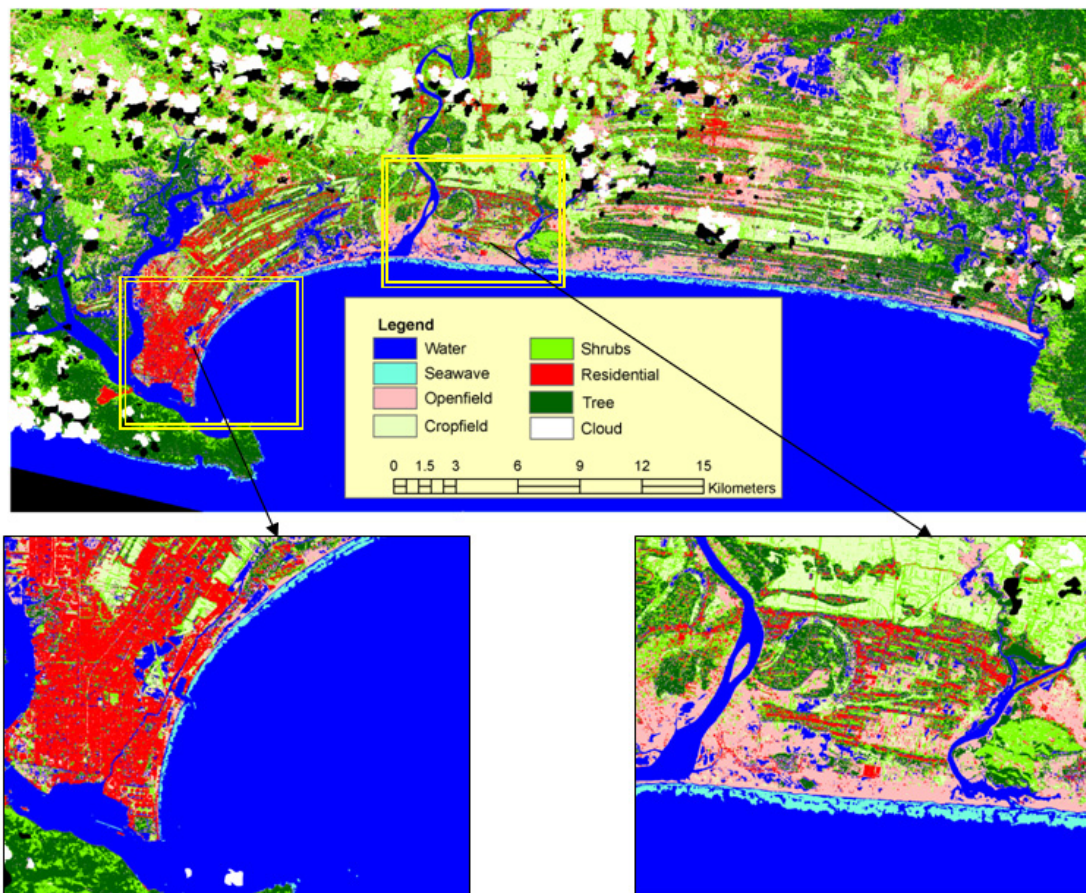


Figure 5.11 The result of decision tree classification of SPOT-5 data for main land use classes

The above-mentioned omission and commission errors are shown more clearly in Table 5.5. The result of classification can be miss (omission error) or false miss (commission error). For example, the misclassification of "shrubs", classified as

“open field”, is around 0.025% (99 pixels); “crop field”, around 5% (20,635 pixels); “tree” around 20% (80,272 pixels); and “residential”, around 0.03% (132 pixels) out of a total of 395,577 pixels-based reference. On the other hand, false classification of “shrubs” also occurred. In “open field”, around 0.003% (12 pixels), “crop field” around 17% (61,789 pixels), “tree” around 0.7% (2,515 pixels), and “residential areas” around 1 % (3,869 pixels) from around 362,625 total pixels of the “shrubs” classification. This example can be used as a guide to interpret the mis- and false classification of the main land use roughness, which is shown in Table 5.5.

Table 5.5 Accuracy assessment of remote sensing classification for main land use roughness (number of pixels)

Classification Reference	Water	Open field	Crop field	Shrubs	Tree	Residential	Total	Omission error (%)
Water	315,757	590	0	0	1	0	316,348	0.2
Open field	72	309,305	877	12	0	2,244	312,511	1.0
Crop field	55	380	357,076	61,789	1,605	16	420,923	15.2
Shrubs	0	99	20,635	294,438	80,272	132	395,577	25.6
Tree	459	8,423	25	2,515	446,654	98	458,176	2.5
Residential	1,040	6,752	87	3,869	135	376,859	388,745	3.1
Total	317,385	325,551	378,701	362,625	528,668	379,351	2,292,282	
Commission error (%)	0.5	5.0	5.7	18.8	15.5	0.7		91.6

Results of surface roughness assessment

The results of roughness classification and its accuracy are shown in Figure 5.12 and Table 5.5 respectively. The coastal area of Cilacap is dominated by very low to low surface roughness values. The zoomed subsets display the urban characteristics (A) and the rural situation (B)

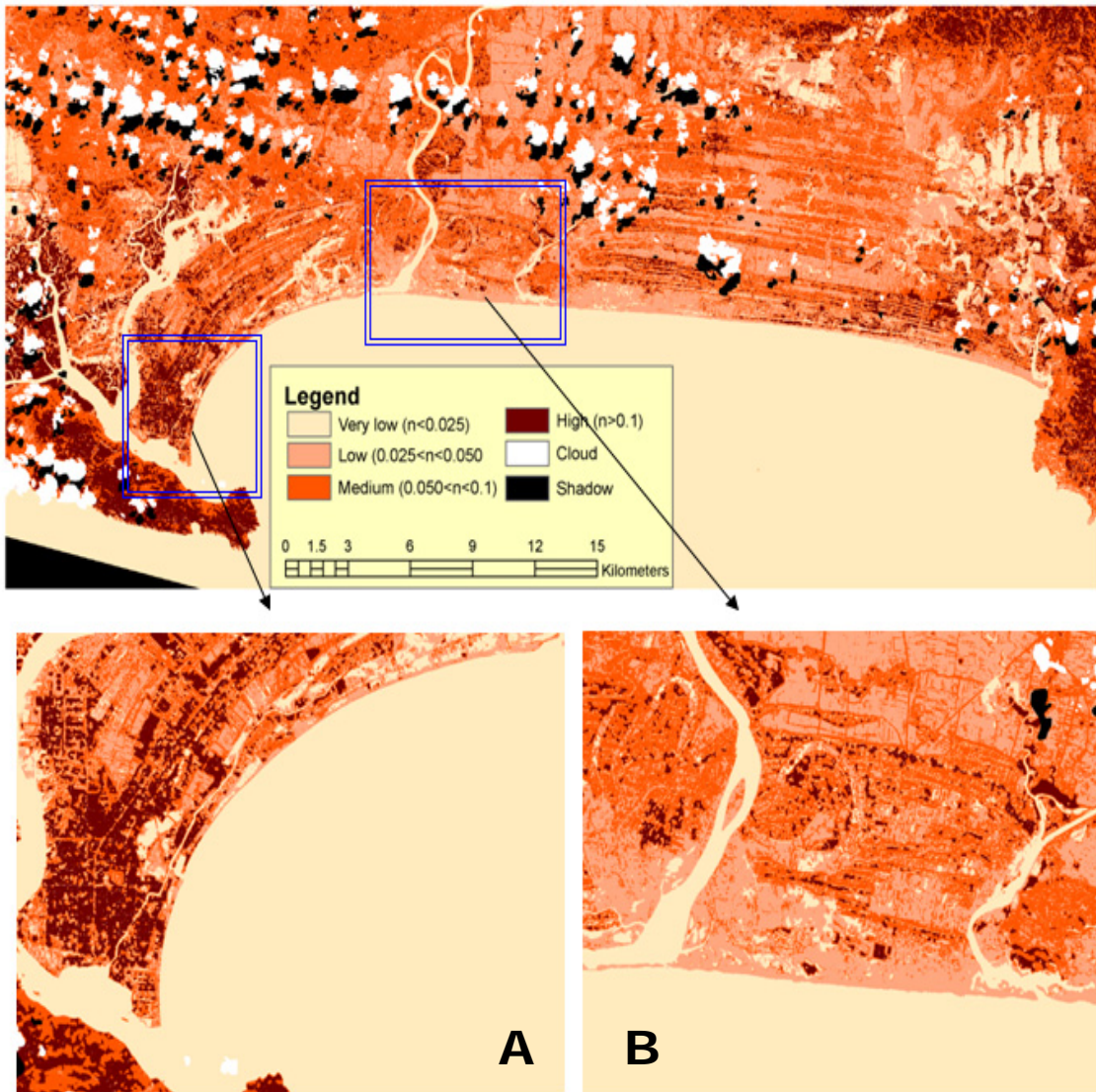


Figure 5.12 Surface roughness classification in Cilacap District

Classification of land use classes and assignment of respective roughness values were achieved with 91.8% overall accuracy. The highest omission and commission values occur in the medium surface roughness class, at 14.4% and 16.4% respectively. Other omission and commission errors in classifying surface roughness are below 10% (see Table 5.5). The overall result shows high accuracy for the surface roughness classification.

Table 5.6 Accuracy assessment of the remote sensing classification of surface roughness classes

Classification \ Reference	Very low	Low	Medium	High	Total	Omission error (%)
Very low	808,319	525	0.0	0.0	808,844	0.1
Low	108	731,061	67,103	0.0	798,272	8.4
Medium	18,468	48,912	682,204	47,656	797,241	14.4
High	1,520	14,886	66,537	757,331	840,275	9.9
Total	828,415	795,384	815,844	804,987	3,244,632	
Commission error (%)	2.4	8.1	16.4	5.9		91.8

The high accuracy of surface roughness classification is mainly due to using the approach that includes the analysis of density and neighborhood characteristics. This accuracy cannot be achieved by using standard classification approaches such as a maximum likelihood classification. Low accuracy values occur especially for classes with a combination of shrubs and open-land (low-medium density of tree) or residential and trees (residential with trees). Table 5.7 shows the resulting accuracy using a standard maximum likelihood approach for analyzing the samples, which contain density and neighborhood characteristics.

Table 5.7 The classification accuracy resulting from a standard maximum likelihood approach

Density and neighborhood characteristics	Accuracy of maximum likelihood model
Combination of shrubs and open land ("low-medium density of shrubs")	60.72
Combination of tree and open land ("low-medium density of tree")	38.49
Combination of residential and open land ("low-medium density of residential")	63.56
High density of shrubs ("Dense shrubs")	94.06
High density of tree ("Dense tree")	95.33
High density of residential ("Dense residential")	90.53
Combination of residential and tree ("Residential with trees")	41.21
Combination of tree and shrubs ("Shrubs and trees")	65.10

5.2.3. Integration of surface roughness for tsunami inundation modeling

The use of the gained spatially distributed surface roughness values is important for tsunami inundation modelling. Surface roughness conditions influence the results of tsunami inundation, especially the velocity of the tsunami wave. Figure 5.13 shows the spatial difference of tsunami velocity using the uniformed roughness coefficient (upper map) and spatially distributed roughness values (lower map). The maximum tsunami velocity mostly decreases when using the spatial distributed roughness values.

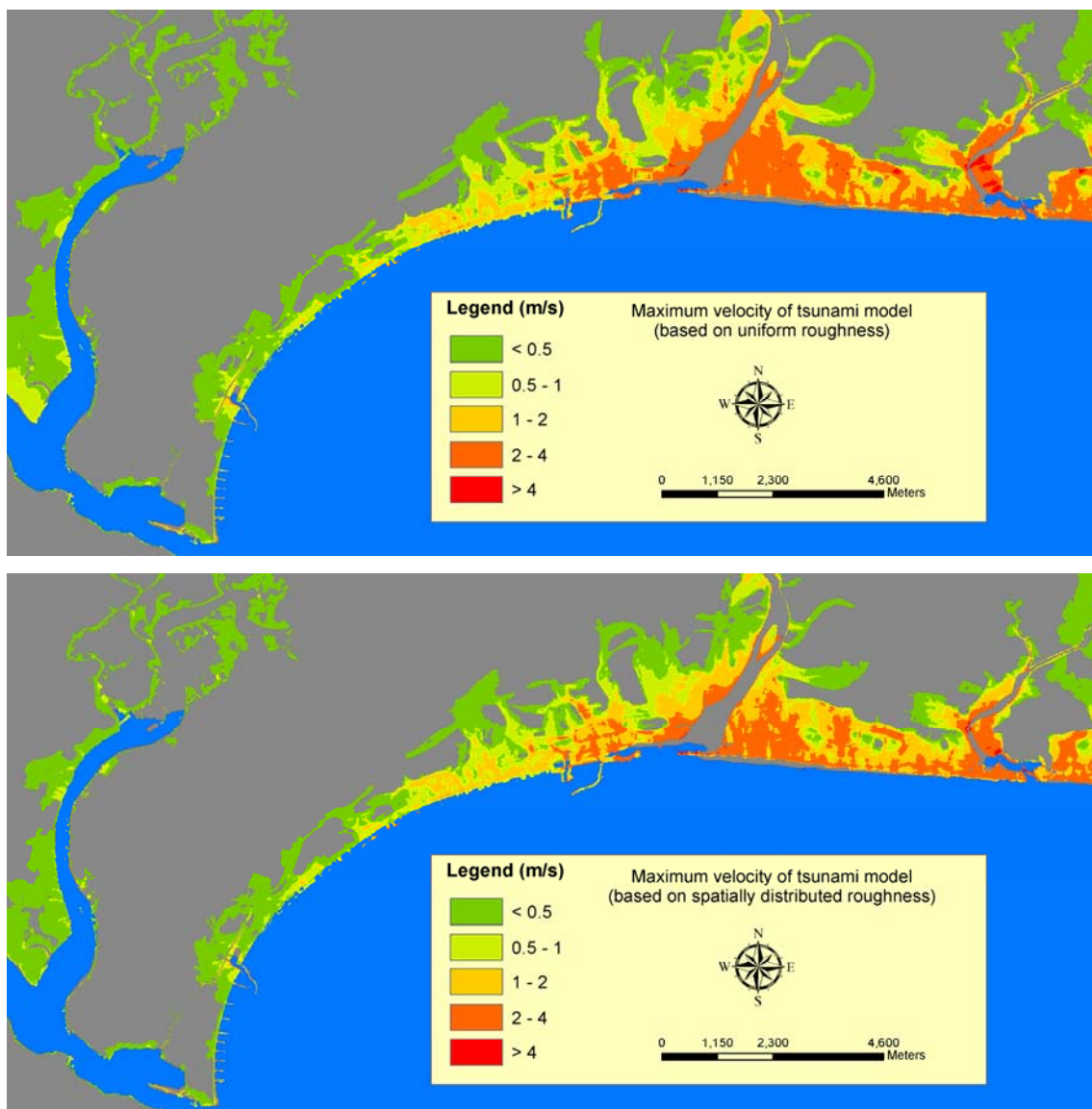


Figure 5.13 Difference of tsunami velocity based on different roughness conditions: uniform roughness coefficient (top), spatially distributed roughness values (low)

Figure 5.14 shows the absolute difference for maximum velocity and for maximum water flow depth (%) between tsunami inundation modeling using uniform roughness coefficient and spatially distributed roughness values. The different representations of surface roughness show greater influence on water velocity than on the water flow depth. Areas far off the coastline show an increasing absolute difference.

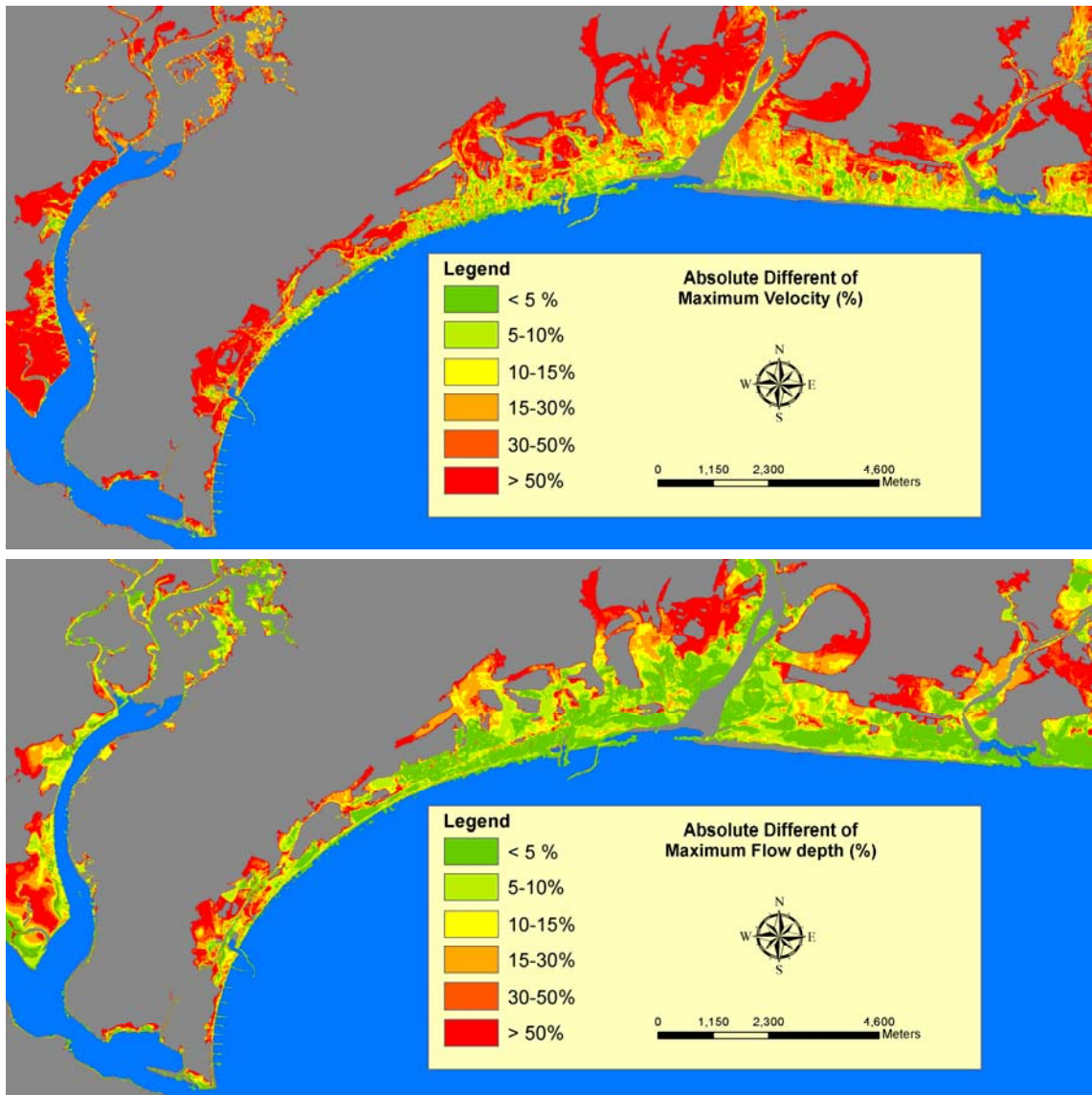


Figure 5.14 Absolute difference of maximum velocity and maximum water flow depth (%) between tsunami inundation modeling using uniform roughness coefficient and spatially distributed roughness values

Further explanation of the importance of spatially distributed roughness to tsunami modeling is shown by the result of the coefficient of variation analysis for the maximum velocity and the maximum flow depth. The coefficient of variation shows the sensitivity of the tsunami inundation modeling result, especially for areas which are far off the coastline (see Figure 5.15).

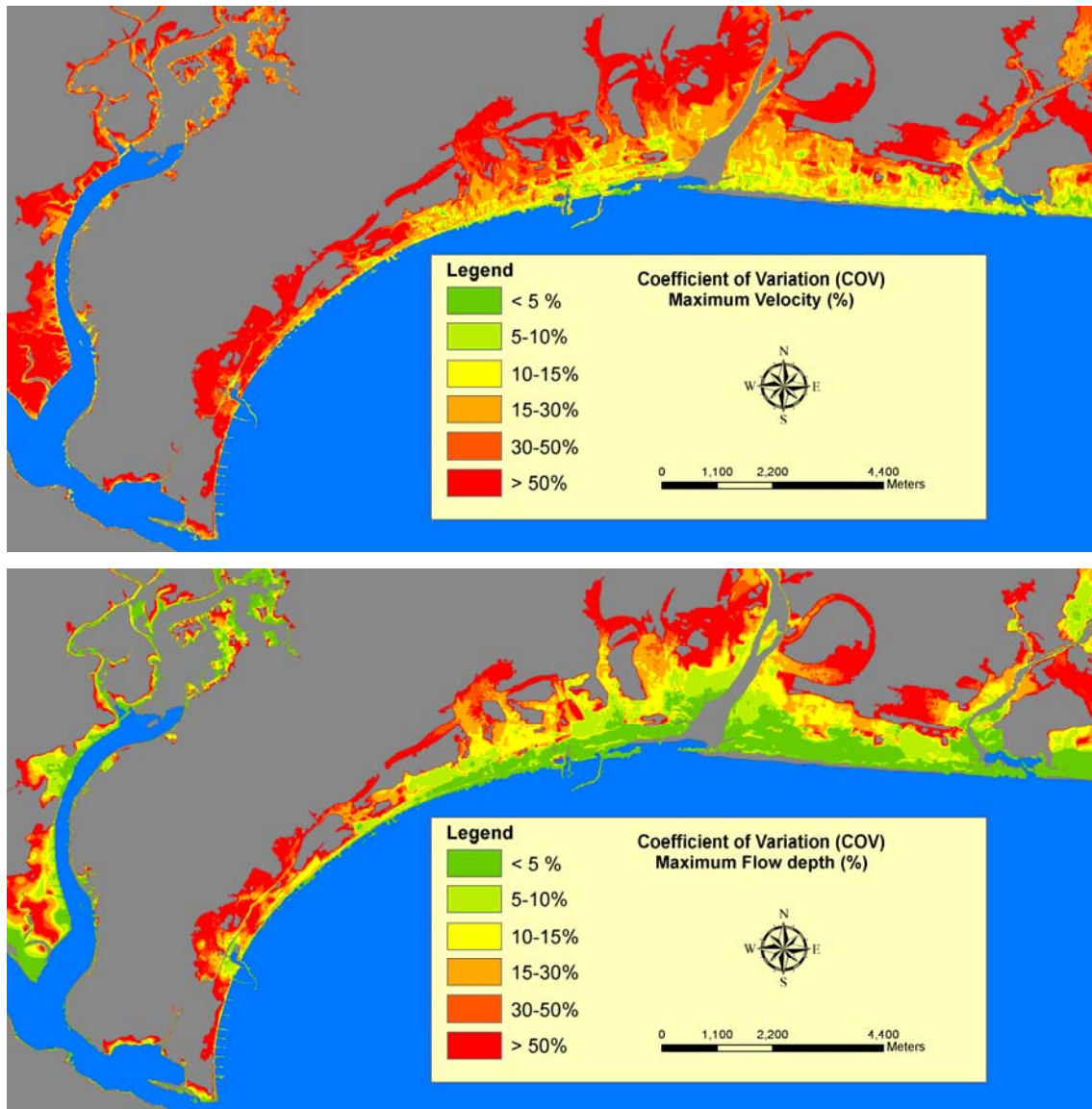


Figure 5.15 Coefficient of variation (COV) for the maximum velocity (top) and maximum flow depth (low)

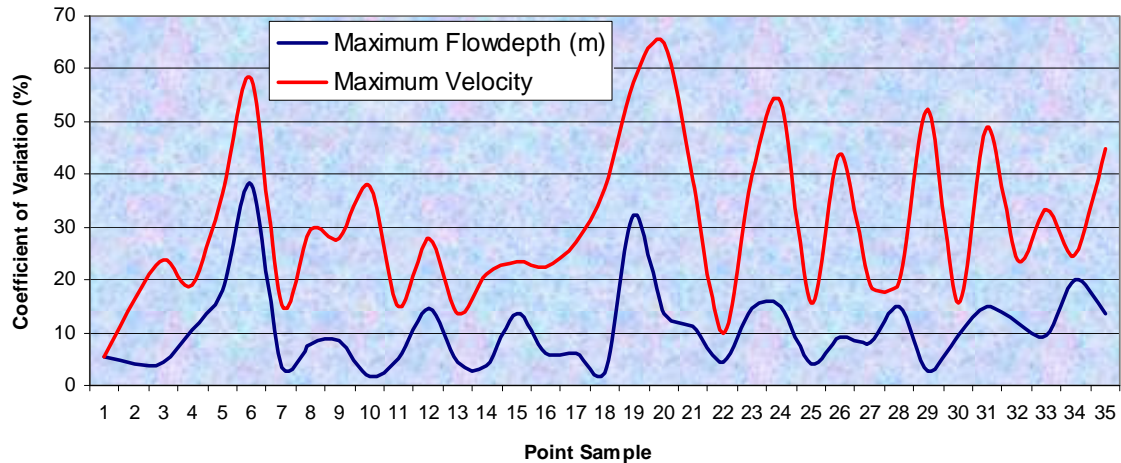


Figure 5.16 Coefficient of variation in selected samples for maximum flow depth (blue line) and maximum velocity (red line)

For 35 point samples in the study area, the coefficient of variation was determined for the maximum velocity and the maximum flow depth (Figure 5.16). This figure shows that the sensitivity of surface roughness is generally higher for the maximum velocity than for the maximum flow depth. The variation of the values depends on the location of the sampling points (see chapter 4.2.5.2 and Figure 4.4) and the surface roughness conditions.

5.3. Population Distribution Modeling

5.3.1. Result of the population distribution modeling concept

The result of different steps in population distribution modeling is shown in Figure 5.17 (see chapter 4.3.1), after combining census with land use data and the differences between population distributions during the day- and night-time. During the day-time, people work in different land use areas, such as settlement and agriculture; during the night-time, most of the people return to their homes, hence the population distribution is concentrated mainly in settlement areas.

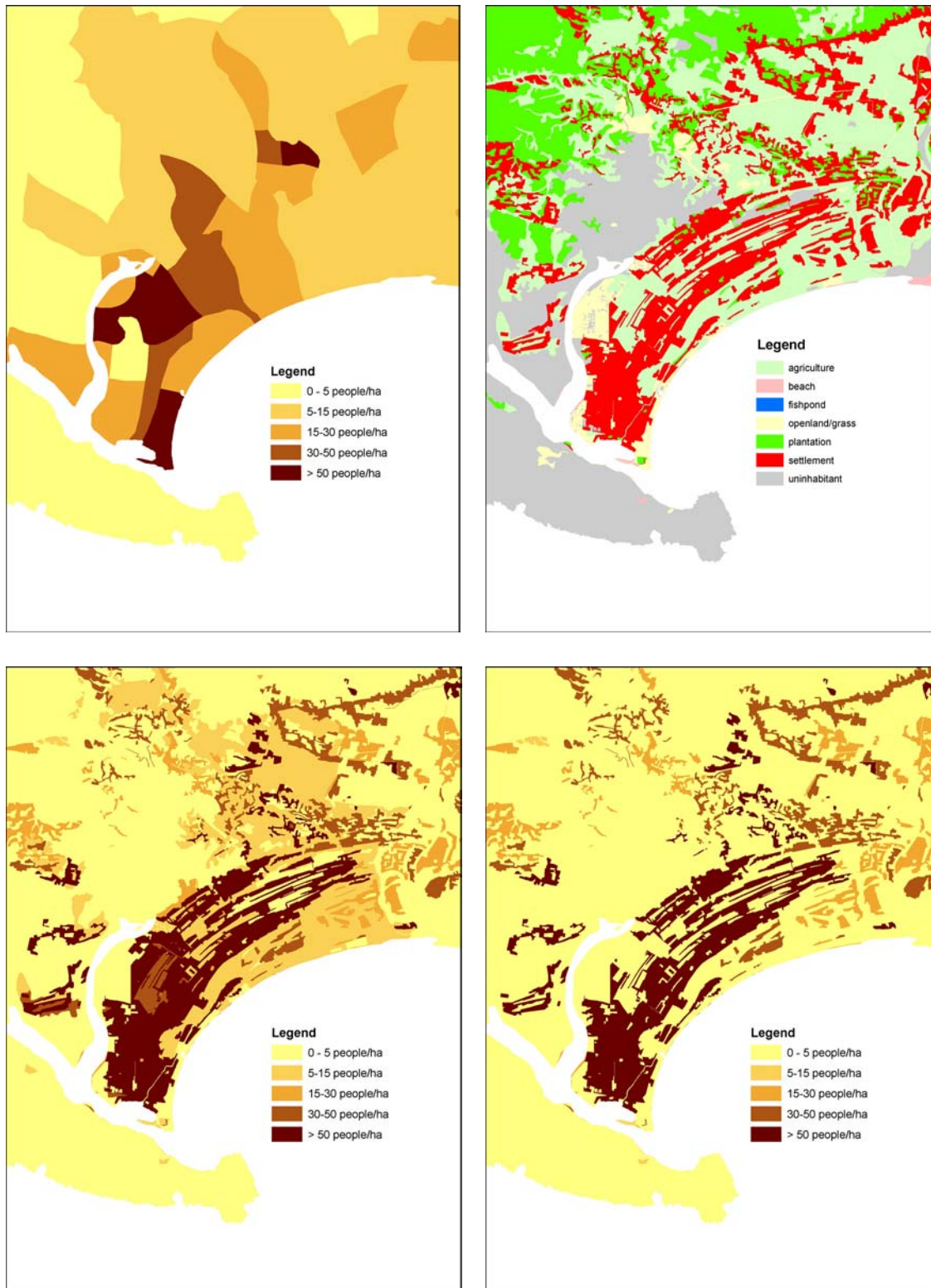


Figure 5.17 The result of population distribution modeling for day- and night-time (lower left and right), based on dasymetric mapping census data (top left) and land use as ancillary data (top right) for Cilacap District

Since intersections occur when combining census and land use data, boundary effects must be taken into account. Therefore, a boundary effect elimination method (see chapter 4.3.1 and figure 4.6) was developed to improve the population distribution result. This result is shown in Figure 5.18 and used for further analysis.

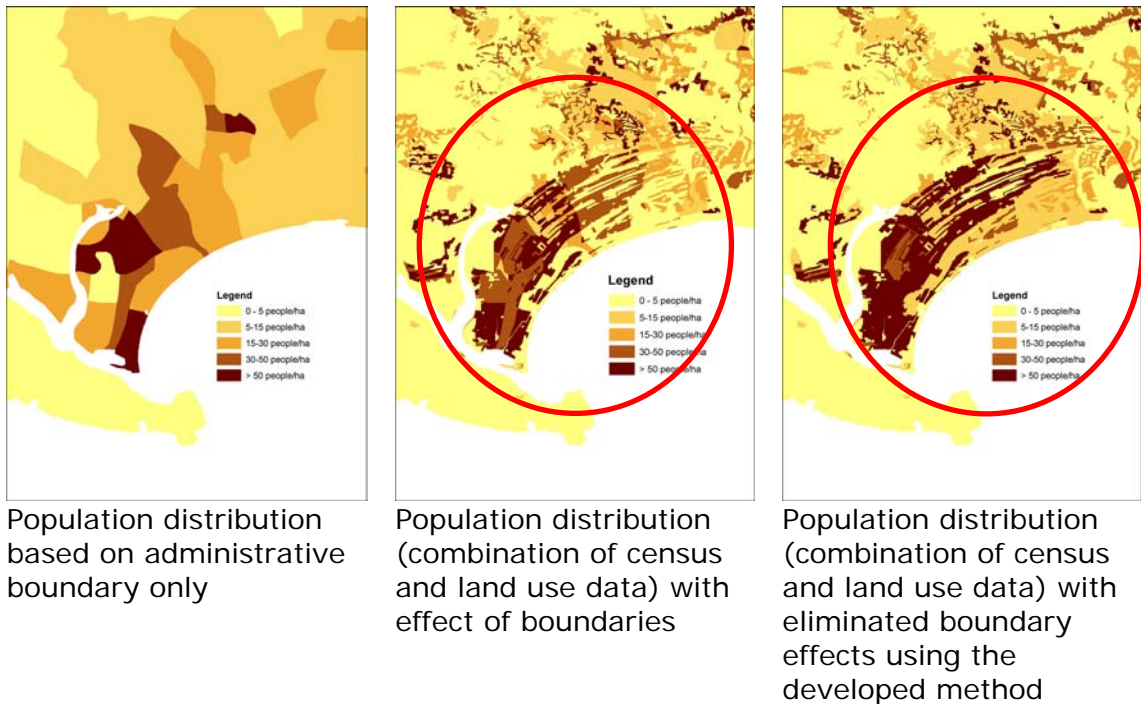


Figure 5.18 Population distribution based on the administrative boundary only (left) and the improved population distribution, with boundary effects (middle) and without boundary effects (final result, right)

5.3.2. Weighting factors determination

For an adequate population distribution modeling, using the given concept shown above, one of the main driving components is the weighting factor. This factor is used to disaggregate the population census values (single number in each administrative unit) using the land use classes as ancillary data (multiple number of people in an administrative unit) for different time periods, day and night. Therefore, it is necessary to: describe the potential number of people engaged in different land use activities, based on a knowledge-based approach; and generalize the weighting factors. The results for both analyses are shown in the following sub-chapters.

Potential number of people engaged in different land use activities

The result of the knowledge-based approach for calculating the potential number of people engaged in different land use activities, using the occupation data, is shown by the box plot distribution in Figure 5.19. This figure shows that “settlement” has the highest weighting factor ($77\% \pm 10.6$) of land use classes. It is followed by “agricultural” ($17\% \pm 10.3$) and “plantation” ($3.2\% \pm 7.2$). Potential number of people engaged in “plantation” activities shows a large range from 0.15 to 61.9 (see Figure 5.19).

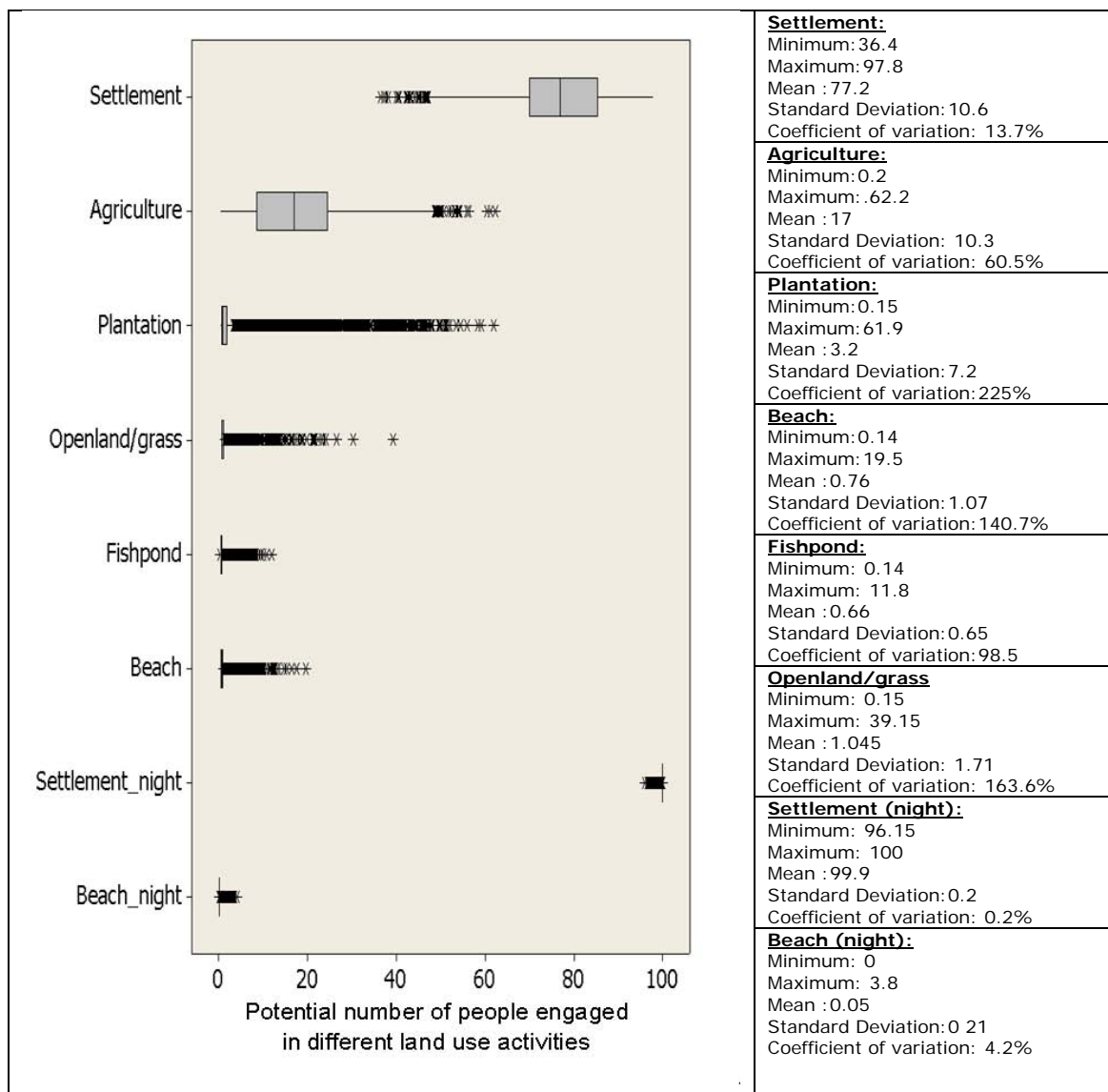


Figure 5.19 Distribution of potential number of people engaged in different land use activities in villages in the west coast of Sumatra, the south coast of Java, and Bali

The coefficient of variation is very high for several land use classes. The selection of appropriate values for the weighting factors leads to high errors in cases with high coefficient of variation. Hence, a generalization of the weighting factor must follow.

Generalization of weighting factor

Due to the variation of potential number of people engaged in different land use activities (see Figure 5.19), the weighting factor cannot be generalized by determining the mean or median value only; otherwise, the generalization would cause high errors on the population distribution mapping. The appropriate approach is to determine a specific weighting factor for every village for the entire west coast of Sumatra, the south coast of Java and Bali. This means that for every village, a separate weighting factor is created for disaggregating the census population data to land use class. For villages where data on people's activity are missing, the approach fails. To overcome this, the analysis of generalization of the weighting factor has been performed by grouping the villages based on their specific characteristics. There are four ways of differentiating and the category of villages according to their characteristics (see Chapter 4):

- urban coastal, urban non-coastal, rural coastal, and rural non-coastal areas (code 1);
- differentiated by island (Sumatra, Java and Bali), and combining the previous mentioned categories (code 2);
- potential economics of villages (code 3);
- municipalities (code 4).

The result of this analysis is shown in Figure 5.20 and the details are provided in Table 5.8. The coefficient of variation for the potential number of people engaged in different land use activities is relatively high (> 50%), except for settlement areas, where the coefficient of variation is below 15%.

In a first step, the generalization is based on the lowest value of the coefficient of variation in the settlements. This is because most of people's activities are concentrated in settlement areas. In this case, the lowest coefficient of variation in the settlement is represented by code 4 (8.56%), followed by code 2 (10.37%), code 1 (10.42%) and code 3 (11.45%).

In a second step, the generalization using code 4 (every municipality has its own weighting factor) revealed the same problem of missing data on people's activity. A possible solution may be to use the generalization based on code 2, but the coefficient of variation is not much different than that for code 1. Therefore, it is assumed that code 1 best fits the generalization. This simple categorization is available for all characteristics of villages in the implemented area.

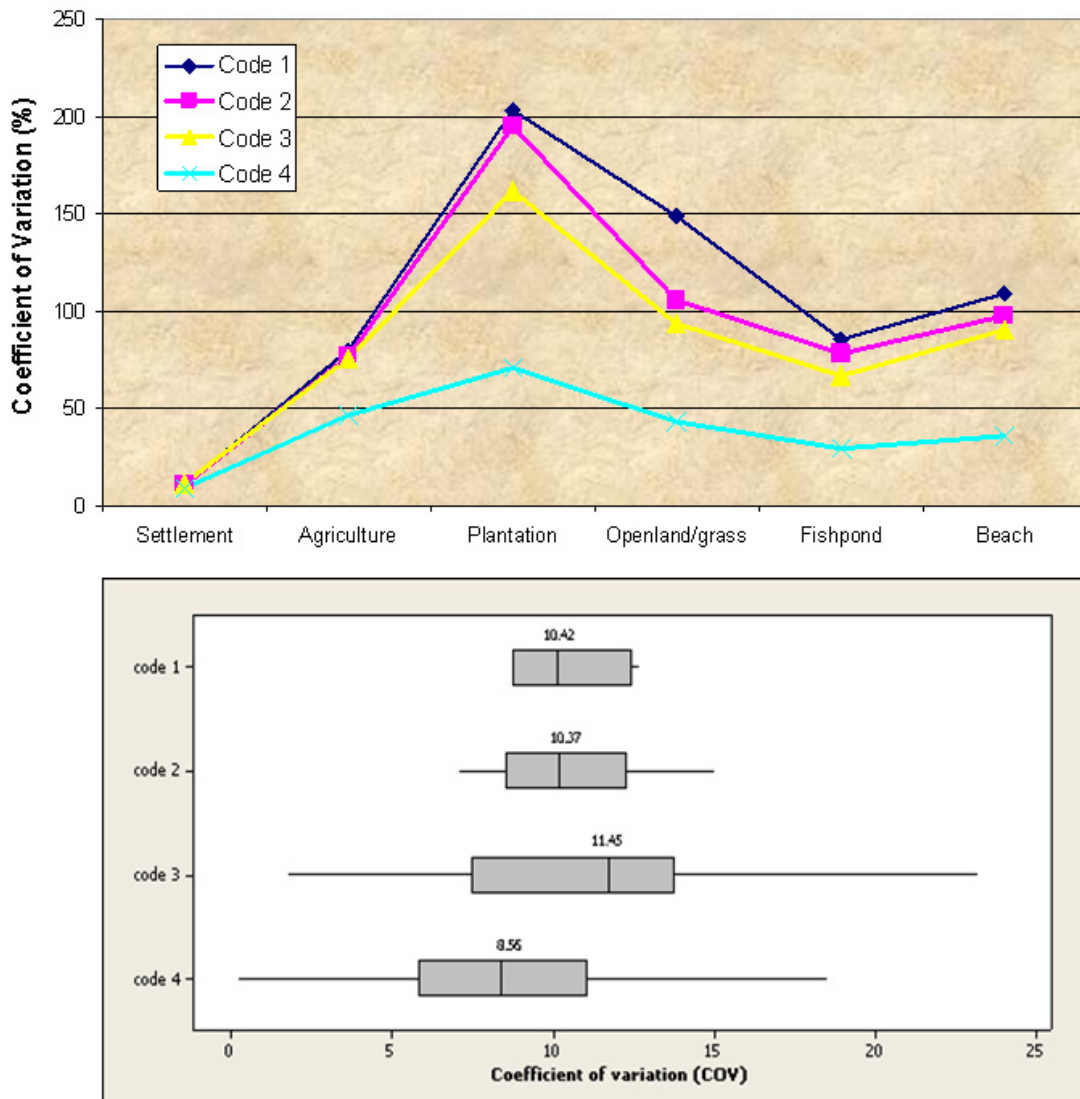


Figure 5.20 Average of coefficient of variation (%) of potential number of people engaged in different land use activities by the category of village (above) and those distributions for settlement (bottom)

Figures 5.21 (day-time) and 5.22 (night-time) show the decision tree for the weighting factor generalization by the category of village based on urban coastal, urban non-coastal, rural coastal and rural non-coastal areas.

Table 5.8 Details of coefficient of variation (%) of potential number of people engaged in different land use activities by the category of village

No.	Categorization	Coefficient of variation of people's activity in land use					
		Settlement	Agriculture	Plantation	Openland /grass	Fish pond	Beach
Urban coastal, urban non-coastal, rural coastal and rural non-coastal areas (Code 1)							
1	Urban coastal	8.7	116.8	168.1	134.1	101.2	116.8
2	Urban non-coastal	8.7	93.6	238.9	96.4	56.5	83.3
3	Rural coastal	11.5	56.8	192.3	196.1	114.5	133.0
4	Rural non-coastal	12.69	50.19	213.28	167.7	69.8	101.4
Differentiated by Island (Sumatra, Java, and Bali) plus the abovementioned categories (Code 2)							
1	Sumatra urban coastal	7.0	121.1	155.1	51.3	93.8	107.8
2	Sumatra urban non-coastal	8.5	121.5	208.5	66.2	65.9	94.4
3	Sumatra rural coastal	11.1	64.7	157.2	107.1	98.6	112.7
4	Sumatra urban non-coastal	12.3	53.1	160.5	127.5	84.3	119.5
5	Java urban coastal	8.5	71.1	182.7	24.5	84.4	94.8
6	Java urban non-coastal	8.8	82.2	215.9	82.8	45.9	68.8
7	Java rural coastal	10.8	36.8	216.8	114.6	89.8	114.1
8	Java urban non-coastal	12.8	46.5	225.2	151.5	36.1	49.3
9	Bali urban coastal	9.5	105.8	181.3	137.5	108.9	135.3
10	Bali urban non-coastal	7.9	101.7	168.0	135.3	41.5	59.9
11	Bali rural coastal	12.3	53.6	273.9	130.7	143.5	159.0
12	Java urban non-coastal	15.0	54.3	192.2	140.4	40.5	56.31
Potential economic of village							
1	Agriculture	13.0	54.6	218.0	167.5	102.9	144.9
2	Mining	13.5	54.1	202.8	180.4	56.8	80.1
3	Industries	13.2	74.1	249.2	119.2	57.0	83.7
4	Electricity	10.2	73.3	31.3	83.7	118.4	149.6
5	Construction	23.2	7	162.0	89.3	10.2	9.1
6	Trade	7.5	122.4	198.9	108.2	68.9	96.2
7	Transportation	10.4	102.8	20.9	32.6	123.2	153.4
8	Finance	1.7	25.3	47.1	0	0	8.3
9	Services	7.2	131.9	255.6	92.7	51.5	73.7
10	Other	14.4	106.2	229.0	63.7	74.7	102.3

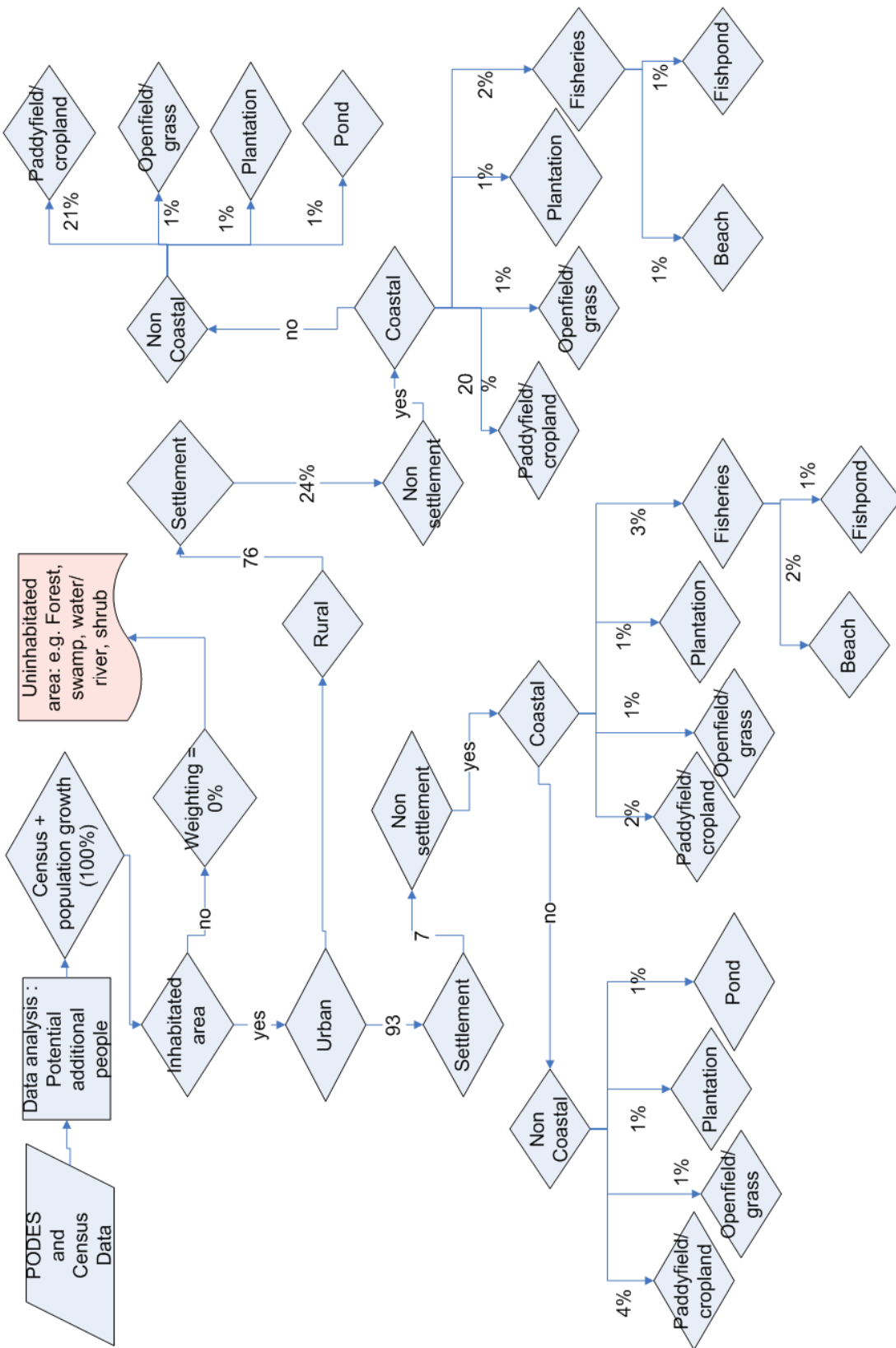


Figure 5.21 Generalized weighting factor for day-time population distribution

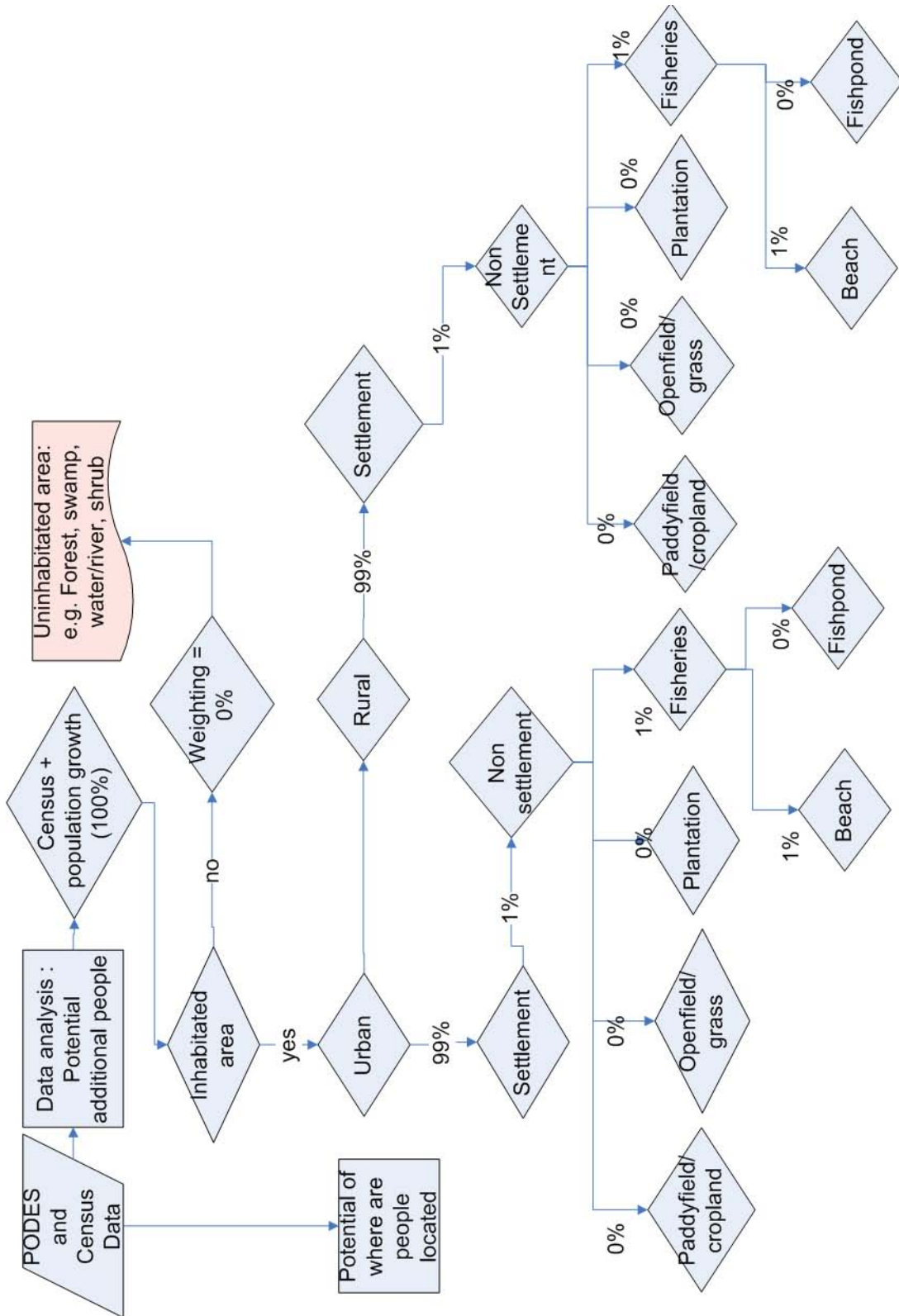


Figure 5.22 Generalized weighting factor for night-time population distribution

5.3.3. Results from questionnaire analysis

The potential number of people engaged in different land use activities during the day- and night-time can be validated based on the questionnaire results for the study area of Cilacap (see Table 5.9).

The results of the questionnaires are comparable to those of the weighting factor generalization. The general weighting factor for settlement in urban-coastal areas is 93.8%, which is very similar to 94.8% based on the questionnaire results. For agricultural land use (in urban coastal areas), the result of the questionnaire shows that people's activity in agriculture is 2% as compared to the modeled weighting factor of 1.6%. The complete comparison is shown in table 5.9.

Table 5.9 A comparison of the weighting factors in Cilacap District between the generalized weighting factors from the model and the reference values from the questionnaire results

Landuse	Urban				Rural			
	Model		Questionnaire		Model		Questionnaire	
	Day	Night	Day	Night	Day	Night	Day	Night
Settlement	93.8	99.9	94.8	99.1	75.9	99.9	80.1	99.3
Agriculture	1.6	0	2	0	20.3	0	13.7	0
Open land/grass	0.7	0	0	0	0.6	0	0.3	0
Plantation	0.7	0	0.7	0	0.7	0	2	0
Pond	0.9	0	0	0	0.5	0	0.1	0
Beach	1.1	0.1	2.4	0.9	0.7	0.1	2.4	0.7
River	0	0	0	0	0	0	1.3	0

The commuters and the actual weighting factor in Cilacap District have been analyzed. Based on the questionnaire, there are five categories of commuters:

- 1) people who carry out their activities outside the village but still in municipal area ("within the municipality")
- 2) people who carry out their activities outside the municipality, but still in the district area ("within the district")
- 3) people who carry out their activities outside the district, but still in the province ("within the province")
- 4) people who carry out their activities outside the province ("outside the province")
- 5) people who carry out their activities by moving in the district, such as pitchmen, bus drivers, city transportation employee, etc. ("mobile").

Based on this categorization, Figure 5.23 shows the percentage of people's activity inside and outside their villages and the percentage of the five commuter categories in Cilacap. During the day-time, most people conduct activities inside the villages (73%), while the remainder (27%) work outside. Of those people carrying out activities outside the villages, 29% stay "within the municipality", 36% "within the district", 2% "within the province", 8% are "mobile" and 25% "outside the province" during the day-time.

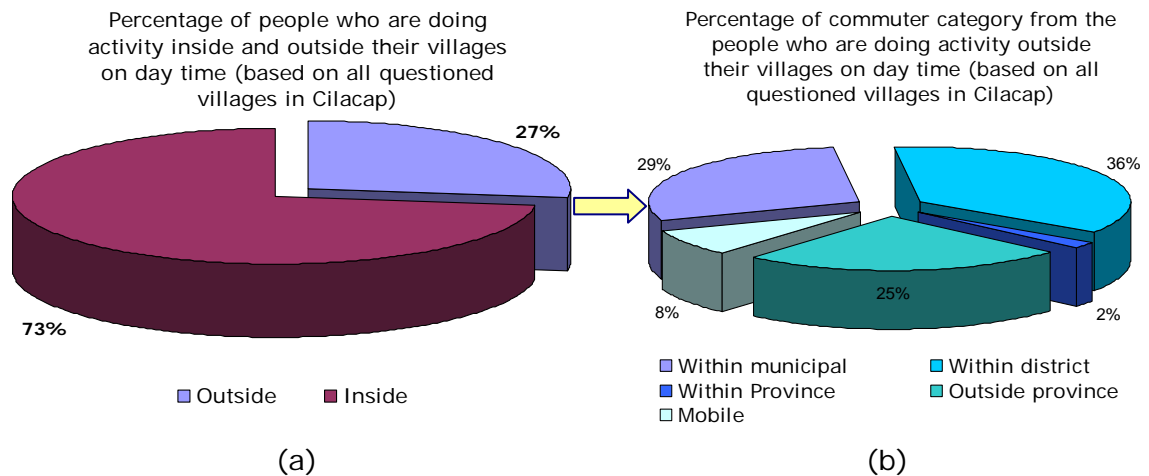


Figure 5.23 (a) The percentages of people carrying out activities inside and outside their villages in Cilacap in the day-time; and (b) the percentages of the different commuter categories carrying out activities outside their villages

The final reference of population distribution can be achieved by combined, the commuter categories and the potential people activity (by using the weighting factor resulting from the questionnaires), this was then used for calculating the error of population distribution (E_{PD}).

5.3.4. Error of population distribution (E_{PD})

E_{PD} has been derived by comparing the population distribution based on the questionnaire results as reference with proposed population distribution model. From this accuracy analysis, the E_{PD} in the Cilacap District for each polygon of land use has been calculated. The median of error is very small, at only 6.3%, with the third quartile is 16.4%. Specifically, the root mean square error (RMSE) of population distribution for settlement is 169 people; 24 people for "agriculture";, two people for "plantations"; two people for "open land/grass";

five people for "fish ponds area"; 16 people for " beach"; and 37 people for "river" (see Figure 5.24).

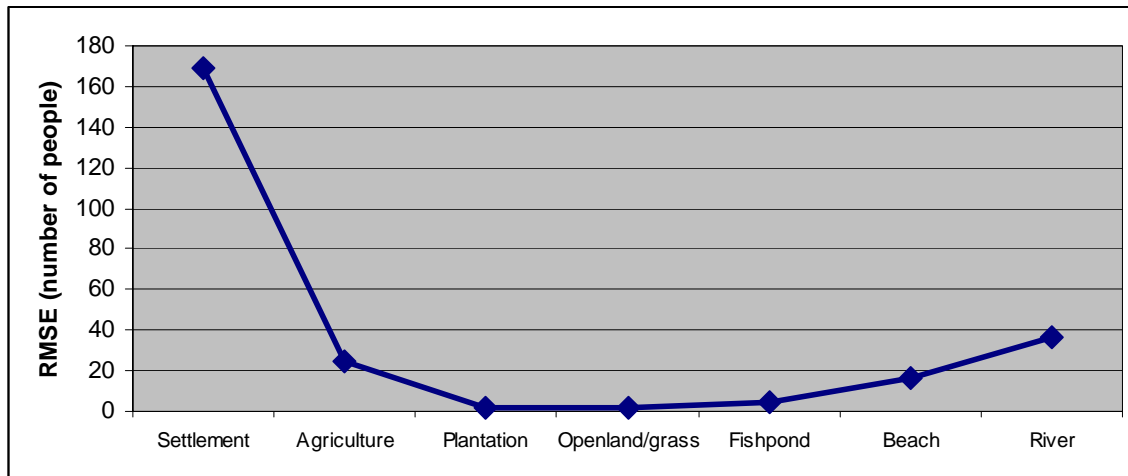


Figure 5.24 RMSE error of population distribution in Cilacap for various land use classes, based on an accuracy assessment

5.3.5. Comparative study

The high quality of the methodology of population distribution (see previous chapter 5.3.4) is also shown by comparing the proposed model with other population distribution models; Gallego and Pedell (2001) and Mennis (2003). The results show that with the developed population distribution model, a higher accuracy is achieved than with other models (see Figure 5.25). The median of the E_{PD} is 6.3% and the third quartile is 16.4%, which is lower than available approaches from Gallego and Pedell (2001) (E_{PD} 11.38%, third quartile 48.4%) and Mennis (2003) (E_{PD} 21.3%, third quartile 108.4%). The higher accuracy is also shown by the RMSE and the coefficient of variation (COV) of the model. The RMSE is 32 and the COV is 0.146. This is better than the results from the Gallego model (RMSE 40, COV 0.185) and the Mennis model (RMSE 53, COV 0.241).

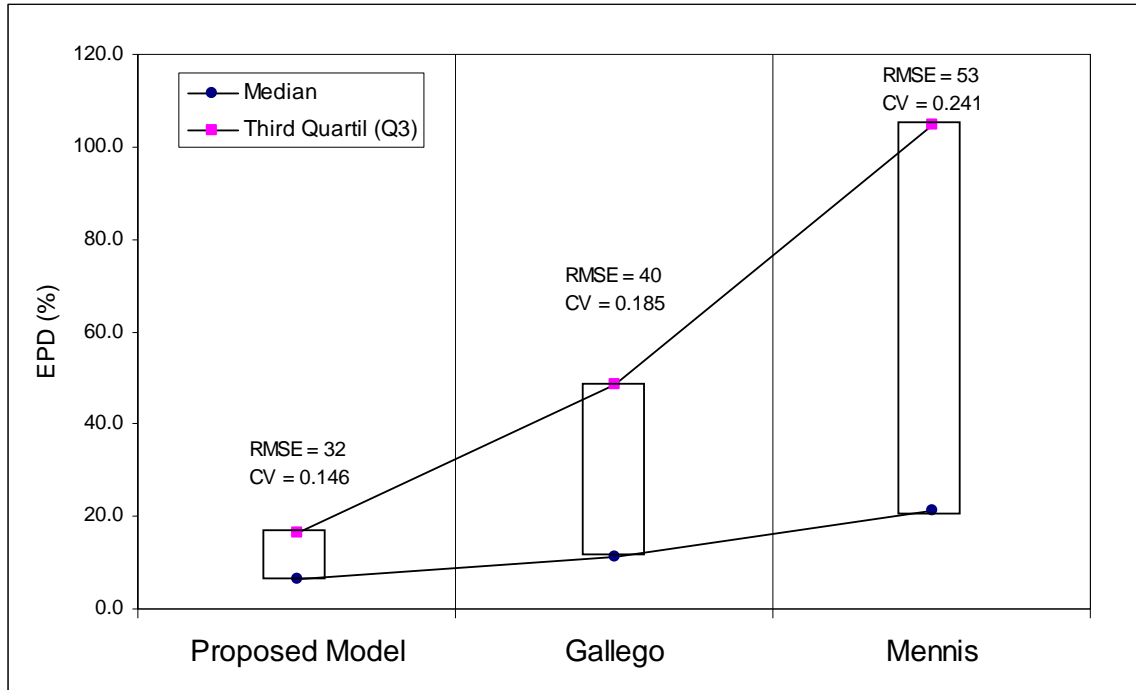


Figure 5.25 Error comparison of three models for population distribution

5.3.6. Multi-scale disaggregation for population distribution in Sumatra, Java and Bali

The population distribution modeling was applied according to the different steps and methods described in the previous sub-chapters. The resulting population distribution is named as “multi-scale disaggregation population distribution” for the west coast of Sumatra, the south coast of Java and Bali; because it integrates the best available data at various scales (Figure 5.26).

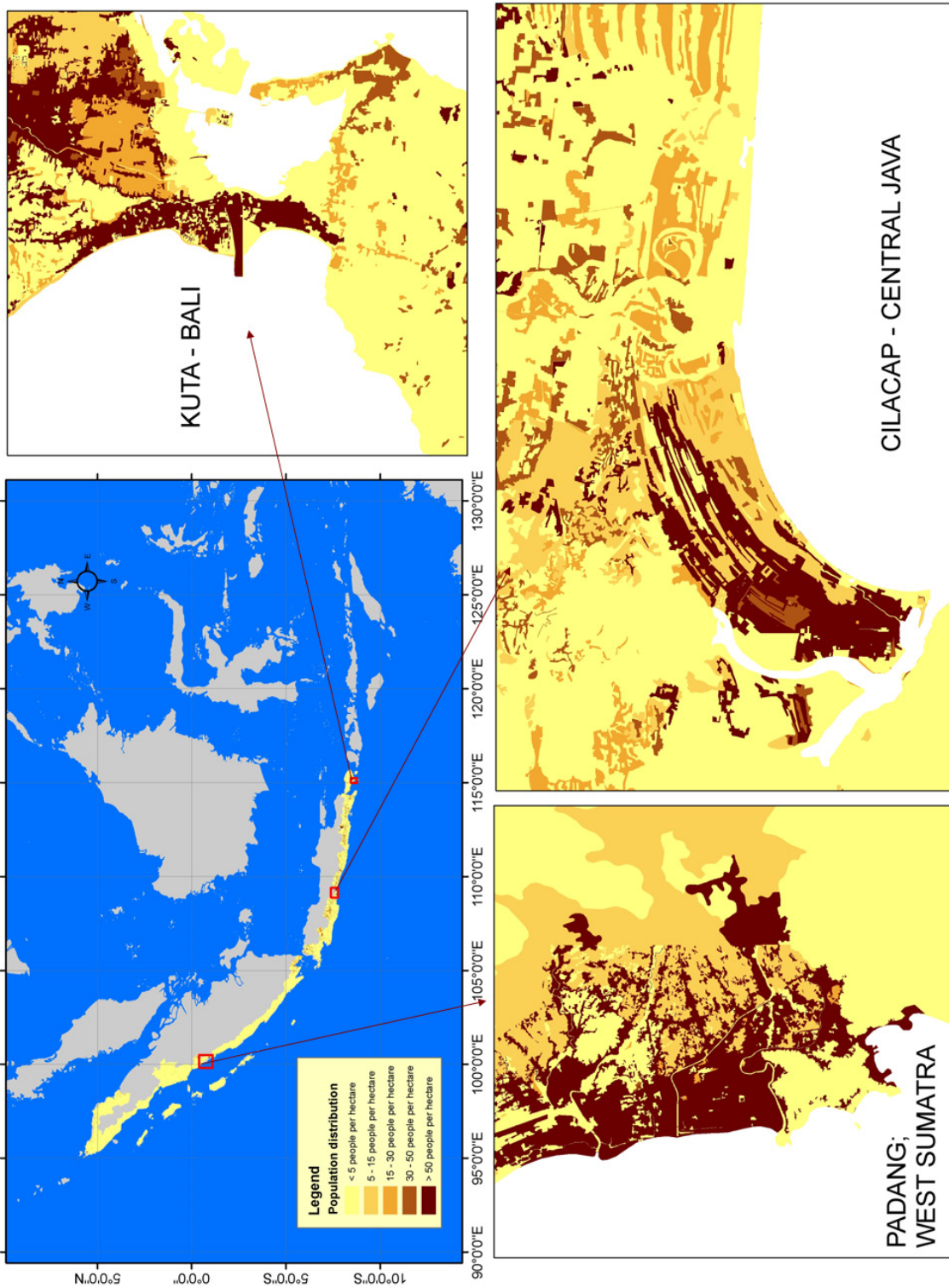


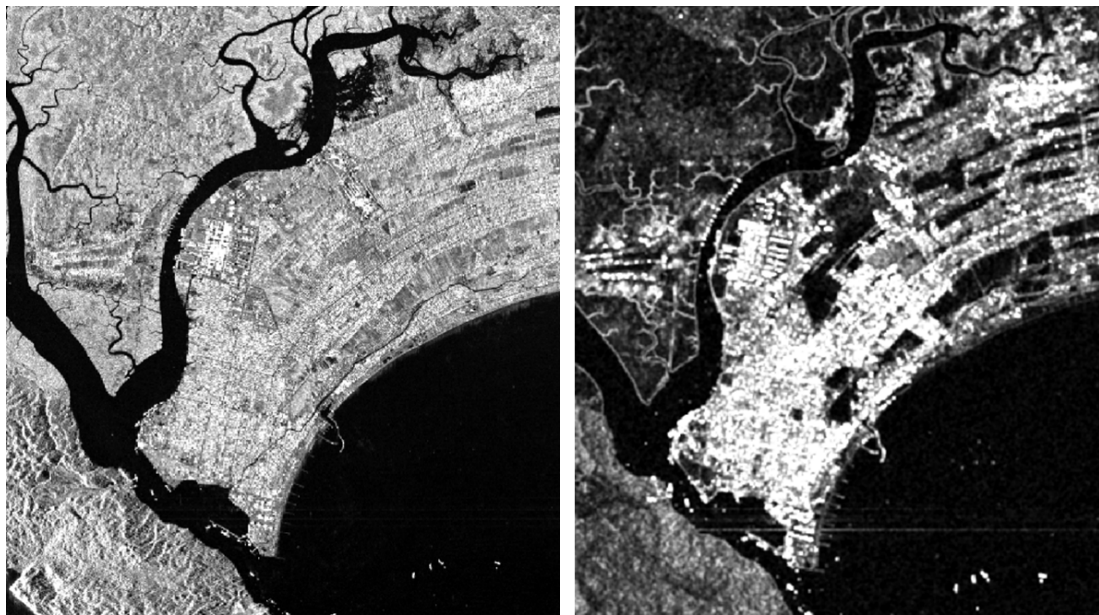
Figure 5.26 Result of population distribution in the broad scale area — the west coast of Sumatra, the south coast of Java, and Bali — after applying the multi-scale disaggregation method

5.4. Settlement Classification

5.4.1. Pre-processing

Speckle divergence

The result of speckle divergence analysis is shown in Figure 5.27. Using this analysis, settlement areas and their characteristics can be distinguished better than using the intensity image. Therefore the result is used further for settlement classification with TerraSAR-X imagery.



TerraSAR-X Intensity

TerraSAR-X After Speckle
Divergence Analysis

Figure 5.27 TerraSAR-X image intensity (left) and after speckle divergence analysis (right)

Neighbourhood analysis

The result of the neighborhood classification to detect the settlement area is shown in Figure 5.28. The classification of settlements is performed automatically and the result is visualized. After this process, the accuracy assessment is carried out to check the quality of the remote sensing classification compared to the settlement areas (see chapter 5.1.3, Figure 5.5).

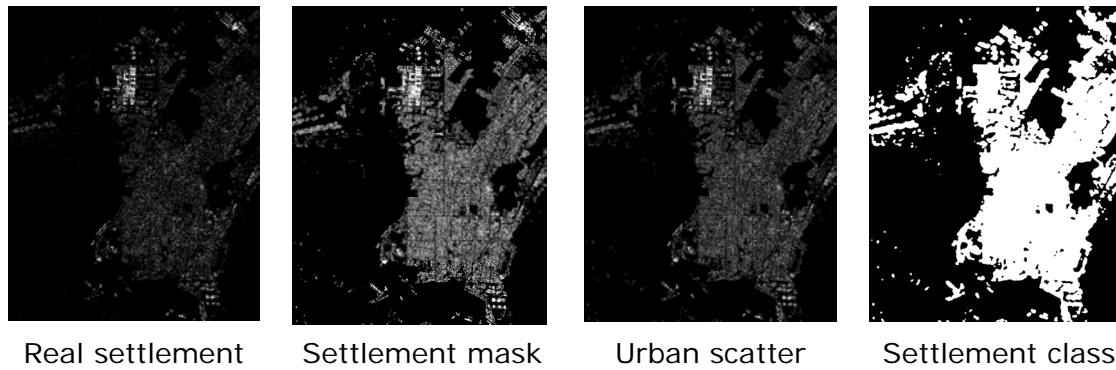


Figure 5.28 Settlement classification results and real settlement by using TerraSAR-X (see chapter 4.4.1)

5.4.2. Classification results

Settlement areas within Cilacap District and Padang District have been classified by using the speckle divergence and neighborhood analysis from TerraSAR-X data. The overall accuracy is 85.5% for Cilacap and 78.14% for Padang. The detailed results of settlement classification in these areas by using TerraSAR-X are described in the following sub-chapters.

Cilacap District

The accuracy assessment of the speckle divergence and neighborhood analysis for Cilacap is shown as confusion matrix in Table 5.10. The commission error of built-up detection in Cilacap area is 27.5% and the omission error is 39.9%. The overall accuracy of settlement classification by using TerraSAR-X in Cilacap area is 85.5%.

Table 5.10 Accuracy assessment of speckle divergence and neighborhood analysis for Cilacap District

Reference \ Classification	Settlement (km ²)	Non-settlement (km ²)	Total (km ²)	Commission error (%)
Settlement (km ²)	101.5	38.4	139.9	27.5
Non-settlement (km ²)	67.5	521.8	589.3	11.5
Total (km ²)	168.9	560.2	729.2	
Omission Error (%)	39.9	6.9	Overall accuracy (%)	85.5

The result of speckle divergence and neighborhood analysis for settlement detection in Cilacap District is visualized in Figure 5.29.

Settlements in rural areas of Cilacap were difficult to detect. The misdetection (blue color) appears more often than the good detection (red color). The occurrence of this error is due to the effect that the settlements in the rural areas are often mixed with vegetation. The settlement with surrounding vegetation is common and typical for almost all rural areas in Indonesia. The vegetation sometimes, covers the roof of a house and is therefore classified as vegetation, not as settlement. In urban areas, the classification of settlements delivers a good detection (see red colors in Figure 5.29).

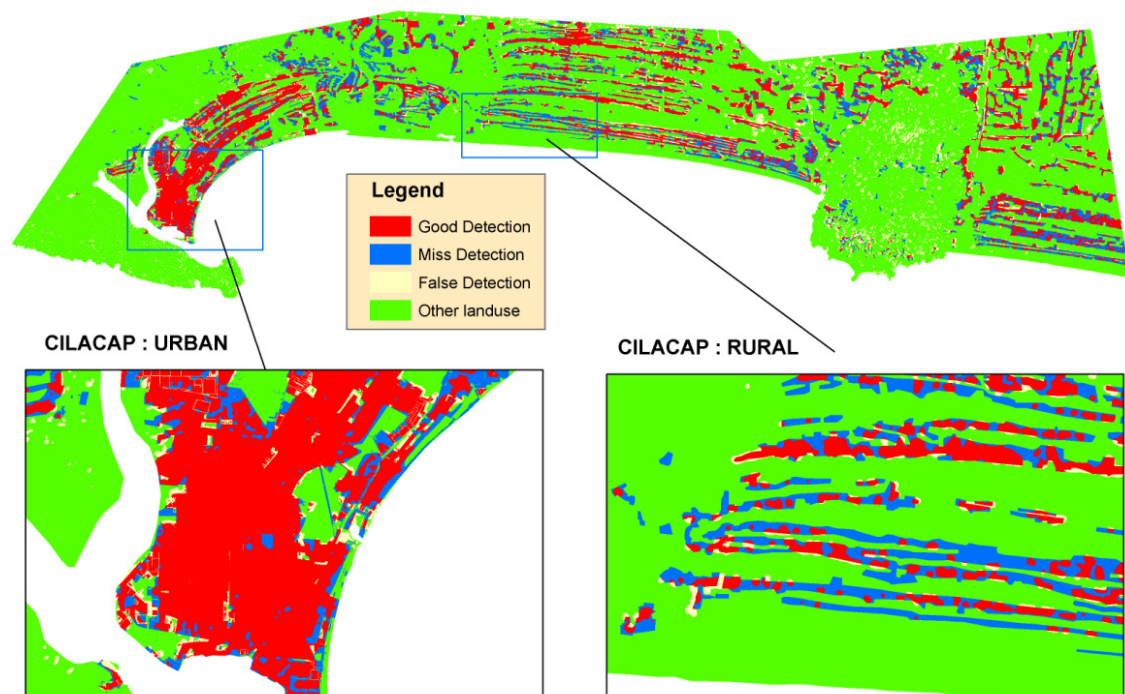


Figure 5.29 Result of speckle divergence and neighborhood analysis for settlement detection by using TerraSAR-X on Cilacap District

Padang District

The same model as that used for Cilacap District to classify settlements using Terra-SAR X data was also implemented to Padang area. The results of classification and accuracy assessment are shown in Figure 5.30.

The results of the accuracy assessment of the neighborhood analysis provides by a confusion matrix shown in Table 5.11. The commission error of built-up detection in Padang area is 21.1% and the omission error is 32.9%. The overall

accuracy of settlement classification using TerraSAR-X imagery is 78.1% for the Padang area.

Table 5.11 Accuracy assessment of speckle divergence and neighborhood analysis for Padang District

Reference classification \	Settlement (km ²)	Non-settlement (km ²)	Total (km ²)	Commission error (%)
Settlement (km ²)	36.8	9.9	46.7	21.1
Non-settlement (km ²)	18.0	62.9	81.0	22.3
Total (km ²)	54.8	72.8	127.7	
Omission error (%)	32.9	13.5	Overall accuracy (%)	78.1

The result of the speckle divergence and neighborhood analysis for settlement detection in the Padang District is shown in Figure 5.30.

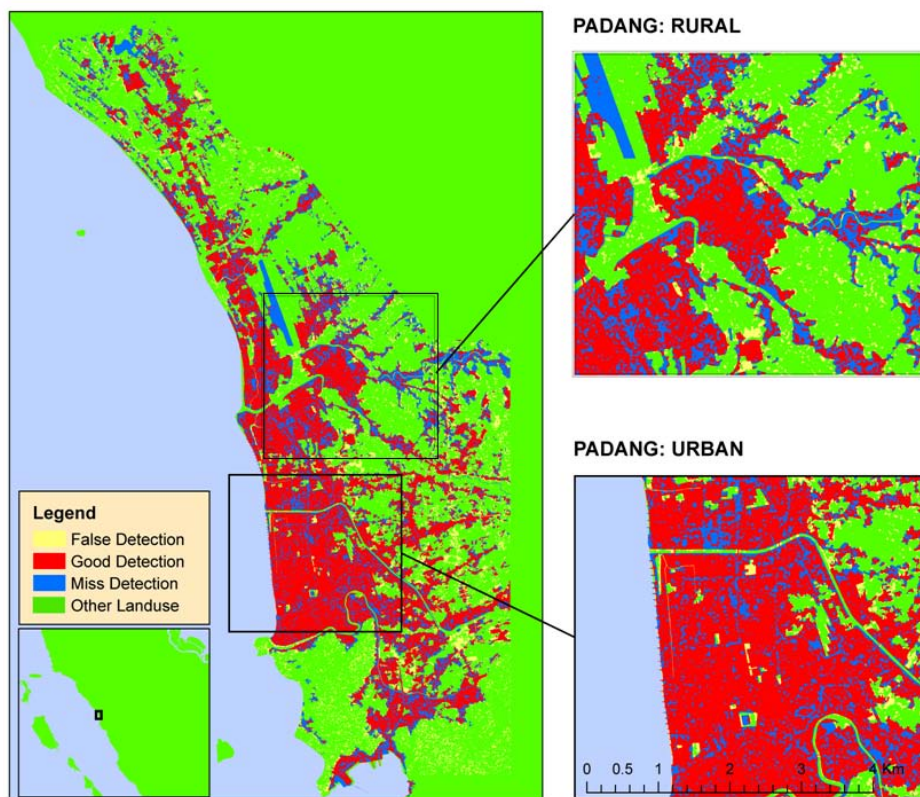


Figure 5.30 Result of speckle divergence and neighborhood analysis for settlement detection using TerraSAR-X of the Padang District

Figure 5.30 shows the spatial distribution of false built-up detection (yellow), settlement misdetection (blue), good built-up detection (red), and other land use classes (green). Similar to the Cilacap region, mis- and false detection in the rural areas of Padang is higher than in the urban areas.

CHAPTER 6: DISCUSSION

This chapter aims at discussing the results of the research carried out to improve tsunami risk and vulnerability assessment according to the result (chapter 5) of the research questions stated in chapter 1. In the framework of an improved tsunami risk assessment, spatially distributed surface roughness determination, population distribution modeling and settlement classification play an important role. Furthermore, recommendations, suggestions and limitations for future research are presented. The structure of this chapter is the same as the previous chapter describing the results.

6.1 Data Collection

6.1.1 Satellite imagery and spatial information

Today, earth observation is a valuable source of timely and spatially distributed information. Remote sensing data, both optical and SAR imagery were crucial for this research. The development of new methods on surface roughness determination and settlement classification required up-to-date satellite imagery. TerraSAR-X data was therefore acquired in the end of 2009 and in the beginning of 2010 to identify the current condition of the study area in Cilacap and Padang District. SPOT-5 imagery was available for 2004 only. In using this data, land use patterns might have been changed in the study area to date. This is insignificant for the image classification development, but must be taken into consideration, because the time shift limits the accuracy assessment.

The multi-spectral bands of SPOT-5 and the additional panchromatic channel were used to derive new variables and improved texture information for classification method development. SPOT-5 was chosen because of the appropriate spatial and spectral resolution for this research, the consistency and availability of data sets in time, and the potential transferability of the methodology. Other optical satellite imagery, e.g. LANDSAT-ETM, ASTER, ALOS and DMC, would have had certain limitations with respect to spatial or spectral resolution, availability or consistency.

TerraSAR-X is a new generation of high spatial resolution SAR imagery, which has several advantages, e.g. extracting urban patterns more clearly. This was

also shown by Esch *et al.* (2010), who stated the importance of settlement classification development using TerraSAR-X imagery.

Land use data was chosen as ancillary data on population distribution modeling, because land use data can help describe where people carry out their activity during the day-time. The data set used was sufficient to perform the population distribution modeling in the pilot areas. For the broad scale approach (the entire west coast of Sumatra, the south coast of Java, and Bali), it was necessary to develop a multi-scale disaggregating method for providing an adequate population distribution data (see chapter 5.3.6).

Bathymetry and topography data are also needed. Dao and Tkalich 2007 stated that these are the most sensitive data for tsunami inundation modeling. Bathymetry and topography data used for the broad scale assessment have sufficient spatial resolution for the tsunami inundation modeling to perform surface roughness sensitivity analysis.

6.1.2 Demographic and socio-economic statistics

The demographic and socio-economic data were used for the population distribution modeling, especially on the weighting factor determination. The complete information on the characteristics of village, number of infrastructures and occupation data is useful for analyzing the potential number of people engaged in different land use activities. For example, farmers perform their activity in agricultural areas, whereas industrial workers or service providers remain in settlements during the day-time. The situation in the night-time is changing according to the structure of the residential areas. It was assumed that most of the people stay at home during night-time. One limitation was missing data in several villages in the implementation area to provide the population distribution in the west coast of Sumatra, the south coast of Java, and Bali. To fill these gaps, the multi-scale disaggregation method (see chapter 5.3.6) was performed.

6.1.3 Reference data

Reference data are crucial for calculating the accuracy assessment of the developed models used. The comprehensive field survey, which was conducted in 2008 and 2009 in Cilacap, is one example of estimating and validating the real conditions of surface roughness, population distribution, and settlement area location. These different reference data types can be described as follows:

1. Several areas in Cilacap have been selected randomly to map the surface roughness conditions in the field in four classes, including recent land use classes. The results were then used in the following accuracy assessment, combining field surveys and remote sensing classification results; it would have been difficult to validate the results without ground truth data.
2. The reference of population distribution was created by disseminating questionnaires in order to derive information on the potential number of people engaged in different land use activities in the coastal area. More than a thousand people responded to the questionnaires, which provide a good basis to explain the true conditions where people are located during the day- and night-time. Since people are dynamic and always moving, the daily conditions vary for the different land use classes.
3. The reference data for settlement detection were derived from the topographic map and further improved manually by using optical satellite imagery (Quickbird and SPOT-5). The resulting polygon reference of settlement areas has been compared with the satellite image classification result for the entire area of Cilacap. By performing this analysis, a differentiation between good, false and miss detection could be shown (see chapter 5.4.2 and Figures 5.29 and 5.30).

6.2 Surface Roughness Determination

The result of surface roughness classes and their coefficient estimation, the remote sensing classification (land use and roughness) and the integration to the tsunami inundation modeling show an improvement to current knowledge. Earth observation analysis can help provide the spatially distributed roughness coefficient. The result of this surface roughness determination is a worthy contribution to the tsunami inundation modeling and hazard mapping.

6.2.1 Surface roughness classes and their coefficient estimation

In a first step, surface roughness coefficients based on a literature review were assigned to land use classes. Roughness coefficients (mean values) published in previous research differ significantly (Chow, 1959; Gayer *et al.*, 2008; Acrement & Schneider, 1989; Hills & Mader, 1987). This leads to substantial variations, which influence the modeling results. To improve this inconsistency, variation coefficients were used (see Table 5.1) and applied to the tsunami inundation modeling (see chapter 5.2.3 and Figures 5.13, 5.14, 5.15 and 5.16).

The published roughness coefficients are usually applied for flood modeling (e.g. Chow, 1959). In tsunami inundation modeling, e.g. the TUNAMI model (Imamura *et al.*, 2006) uses Chow (1959) as a reference for determining the roughness coefficient. Hill and Mader (1987) derived roughness coefficients by empirical models based on the tsunami occurrences in Hawaii and Japan.

In a second step, density and neighborhood relations have been included in the newly developed approach. The density of objects and vegetation and their respective neighborhood relation represents the true condition of surface roughness related to the protection probability. The combination of residential and tree, for example, are better protected against tsunamis, compared to those without vegetation surroundings.

Due to the variation of roughness coefficient values, it is recommended for future research to develop a method using remote sensing and observation data (tsunami flow depth measurement in tsunami inundation area, tsunami velocity and inundation length, flow direction, etc.) from the occurrence of tsunamis. This has not yet been done for lack of sufficient observation data combined with timely satellite imagery. By using earth observation data taken before and after a tsunami event, the simulation of tsunami flow, the corresponding velocity and resistance (roughness parameters) can be analyzed and assessed. Figure 6.1 shows the differences of tsunami flow and inundation in the three marked areas (red cycles). In region 2, which contained dense vegetation before the tsunami struck, the tsunami flow stopped not far from the coastline (approximately 1.3 km). In contrast, in the regions 1 and 3 the tsunami flow continued further inland, because of missing or less dense vegetation cover (approximately 2.7 km from the coastline). Possible variations due to slope can be ignored because the entire area (regions 1-3) has a continuous slope less than 3 degrees.

For transferability purposes, the analysis of several tsunami events is needed for developing and validating the method. In Indonesia, data on the Indian Ocean Tsunami 2004 and the 2006 Tsunami could be a good basis to perform this analysis.

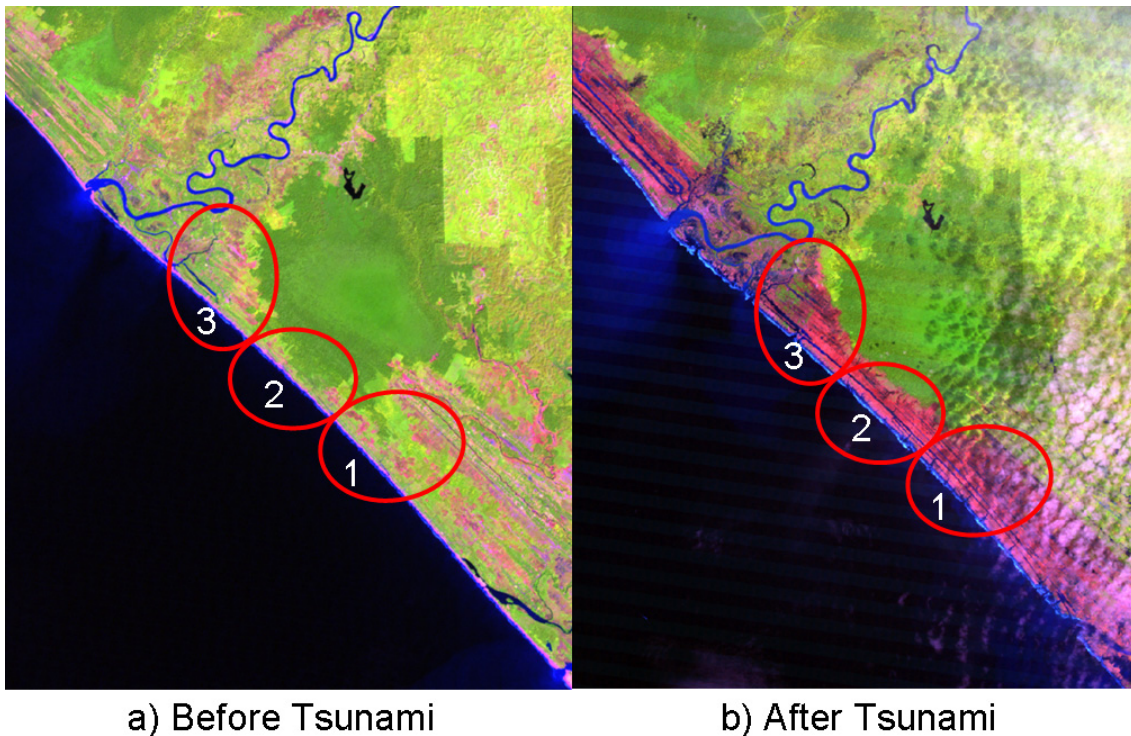


Figure 6.1 Image before (a) and after (b) tsunami from LANDSAT TM

6.2.2 Surface roughness classification by remote sensing

Relevant remote sensing variables

The analysis and combination of remote sensing variables is a key on remote sensing classification for deriving surface roughness classes. As shown in chapter 5.2.2, the appropriate variables for the land use classification in the study area were opposed and combined. The resulting decision tree model accuracy was shown for each variable and also for various combinations. It could be shown that the role of texture as an additional variable for land use classification is particularly important to increase the model accuracy, whereas the exponential NDVI and sigmoid NDVI have only minor influence.

The reason of using NDVI modification was to separate shrubs, crop field and tree, because they have a low separability on spectral reflectance. Based on the theoretical background (Richard and Jia, 2006; Hasan and Akamatsu, 2004), it is assumed that the transformation of NDVI by sigmoid and exponential function should improve the separability and therefore could serve as a relevant variable. Figure 6.2 shows the transformation of NDVI by sigmoid and exponential function.

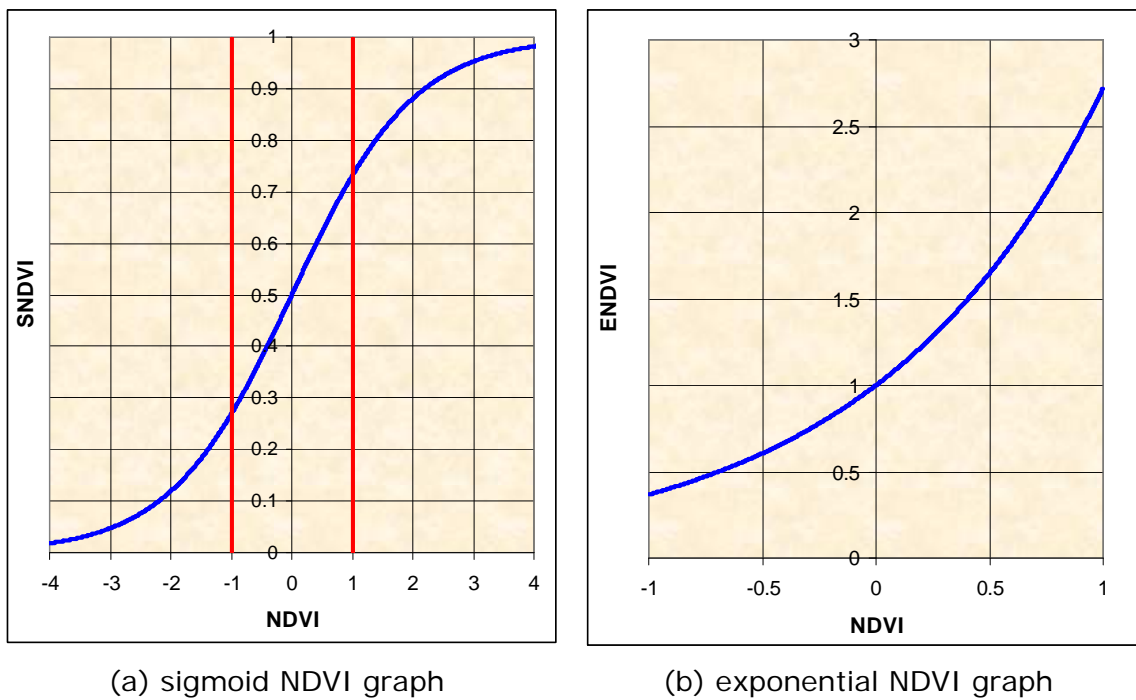


Figure 6.2 Transformation of NDVI by (a) sigmoid and (b) exponential function

Possible reasons that do not show up in the results for the study area could be the limited number of ground truth samples and the low separability of classes through NDVI. Figure 6.3 shows the relevant NDVI characteristics of these classes.

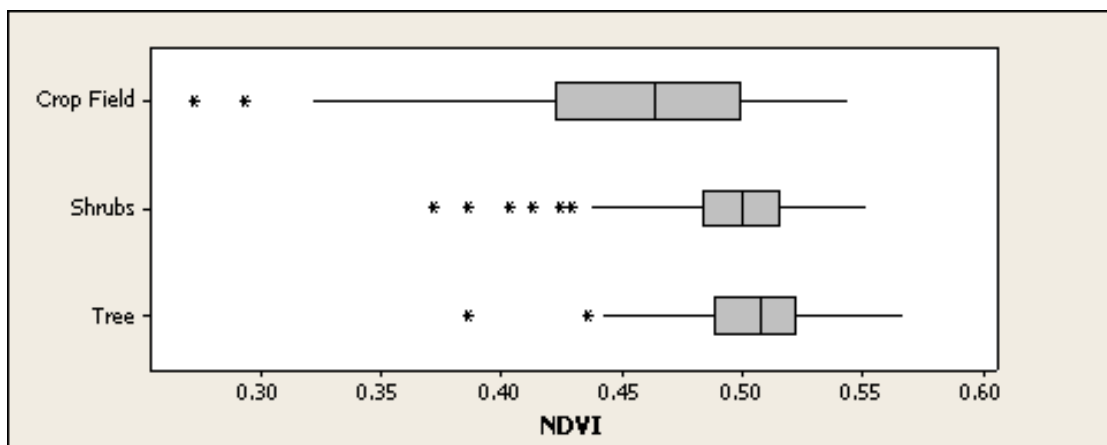


Figure 6.3. The histogram samples of NDVI for three vegetation classes (tree, shrubs, and crop field)

Therefore, it is recommended to use texture for separating "residential area" and "open land" as an addition variable, but not to use sigmoid and exponential NDVI to differentiate "shrubs", "crop land" and "tree" for the Cilacap study area.

Accuracy of surface roughness classification

The overall accuracy of the classification, either on main land use or surface roughness classes, is over 90%. This shows that this approach provides realistic results that can be derived from remote sensing data, using a decision tree model for the land use classification, which then serves as input for the surface roughness classification in the study area.

A limitation of the approach was the weak differentiation of "tree", "shrubs" and "crop field", which showed errors of more than 15%. Possible explanations for this weak point in this Indonesian study area could be the heterogeneous patterns of "tree" and "shrubs" in general and the interrelation between "crop field" and "shrubs" locally (see Figure 6.4).

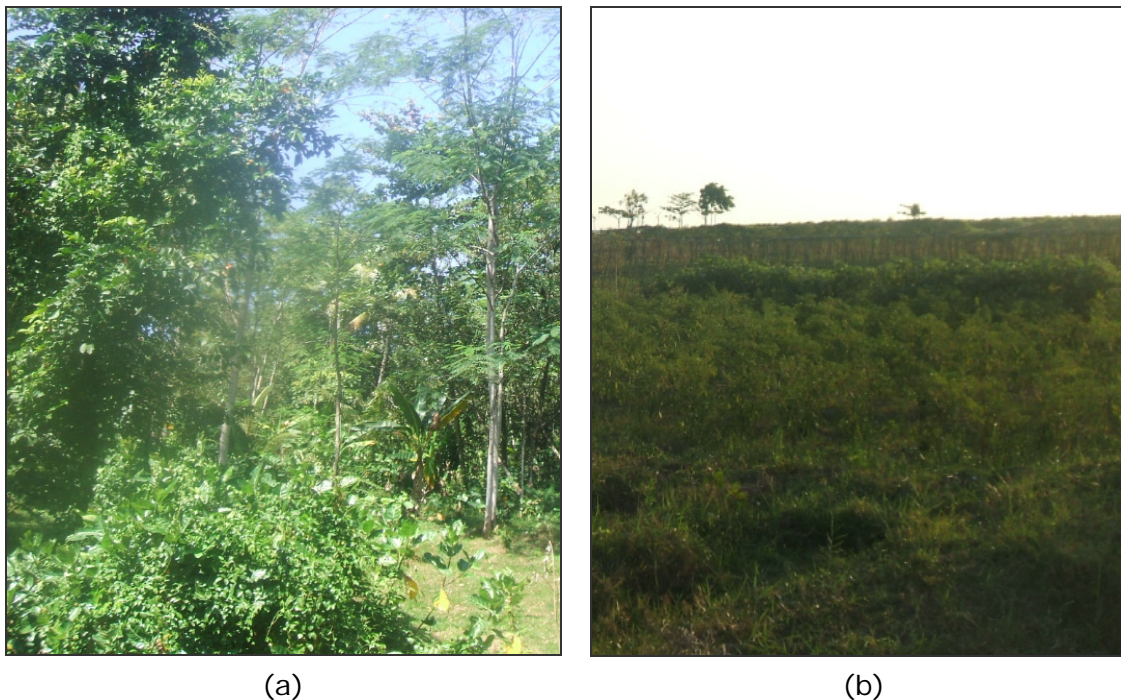


Figure 6.4. The heterogeneous patterns of land use in the study area: (a) tree and shrub; (b) crop field and shrubs

Due to the heterogeneous characteristics of the land use classes in the study area, neighborhood interrelations between the classes and density properties have to be taken into account. As described in chapter 4.2.2, densities of land use features were calculated, using a moving window based on a kernel size of 9 x 9 pixels (~23 m spatial resolution) proposed by Koshimura (2009). The resulting categories of densities are rare, medium and high according to Murashima *et al.* (2008).

The applied method allows combinations of classes, such as “residential area with trees” or “shrubs and tree”. For tsunami inundation modeling in particular, this approach shows clear advantages over the conventional methods, because the roughness coefficients of those classes vary significantly.

6.2.3 Integration of surface roughness for tsunami inundation modeling

The result of the spatially distributed roughness coefficient from remote sensing classification has been implemented to the tsunami inundation modeling for the study area of Cilacap (see chapters 4.2.4 and 5.2.3). It shows that this has a important influence on the inundation results, especially on the sensitivity of tsunami velocity and tsunami flow depth. The results show the high absolute difference and variance of tsunami velocity and flow depth for areas, which are far from the coastline. It proved that there is a significant influence of surface roughness on the resistance of tsunami flow.

The results on the implementation of surface roughness in the tsunami inundation modeling are comparable with Leschka *et al.* (2009) and Gayer *et al.* (2008). They also revealed that the tsunami velocity is more sensitive than tsunami flow depth by changing the spatially distributed roughness coefficient. Generally, high roughness coefficients can decrease the tsunami velocity considerably. The velocity and flow depth are important for analyzing building damages and human stability in flowing water (RESCDAM, 2000). The inaccuracies and errors on the modeling of tsunami velocity and flow depth can influence the analysis of building damage and human stability. Human loss of life caused by tsunami is directly related to the ability to resist the tsunami velocity. Therefore, it is important to calculate not only the inundation area, and the flow depth, but also the tsunami velocity.

Based on this sensitivity analysis, spatially distributed roughness coefficients should be used for tsunami inundation modeling, rather than a uniform roughness parameter. By using the spatially distributed roughness, the result of the tsunami inundation model is closer to the true situation.

6.3 Population Distribution Modeling

A major challenge, not only for tsunami inundation modeling, is the lack of low resolution of population distribution data. In order to improve this situation, it is important to develop a methodology to model the population distribution for regional and local applications.

The proposed model has been described in steps to combine census and land use data for creating an adequate population distribution data set. The process steps of population distribution modeling comprise the development of an improved formula, the weighting factor determination, and the multi-scale disaggregation for population distribution for the coastal areas of Sumatra, Java and Bali.

6.3.1 Weighting factor determination

The result of population distribution shows how to determine weighting factors, which are based on the analysis of socio-economic data. The basic analysis of where people are located by knowing the potential number of people engaged in different land use activities based on occupation data is new in terms of population distribution mapping. It is possible to show that the potential number of people engaged in different land use activities varies between the day- and night-time for every village. This is an important analysis, because the weighting factor plays a role describing the true situation of population distribution. The results are better than in approaches that use an iterative regression process or sampling analysis for determining the weighting factors such as Gallego and Pedell (2001) and Mennis (2003).

By using socio-economic data, the characteristics for each village can be identified and categorized in "fisheries", "agricultural", "plantation" or "business" village. The part of the area, classified as land use, in which people gather during the day- and night-time can be determined based on the above-mentioned categorization. For example, most of the people within a plantation village carry out their daily activities on plantation area and most people within a "business village" carry out their work within the settlement during the day. The data on occupation as part of the socio-economic data plays a role in determining where people are located. During the day-time, a worker in the certain sector (e.g. agriculture) will work in the land use related to that sector of work (e.g. agricultural area). The main activity of the services sector, for example, concentrates on settlement areas.

Based on the results of the analysis, the coefficient of variation for the potential number of people engaged in different land use activities (for the west coast of Sumatra, the south coast of Java, and Bali) is relatively high. This means that every village in this area has its own characteristics of potential number of people engaged in different land use activities. The general weighting factor has been used to solve the problem of missing data, because for some villages, no occupation and socio-economic data are available. The weighting factor had to be categorized in urban and rural areas that either belongs to a coastal or non-coastal class. The main reason for this was to describe the empiric approach, that urban and rural regions as well as coastal and non-coastal villages have different socio-economic characteristics. For example, in agricultural areas, people's activity varies significantly for urban and rural regions (see Figure 5.21).

The weighting factor determination by socio-economic data provides a step forward in population distribution modeling. This makes it possible to model the differences of the distribution of people during the day- and night-time. During the day-time, people are engaged in certain land uses related to their occupation, and during night-time, they return back home to their village. In Cilacap, some people, such as local fishermen, work at night on the beach. It is assumed that commuters should be considered, but due to data limitations, this has not been yet accommodated in the weighting factor determination; it will, however, be discussed in the next section.

Compared to the questionnaire results, the generalized weighting factor shows a good correlation. It can be stated, therefore, that the differentiation of urban and rural for the general weighting factor and the potential number of people engaged in different land use activities in Cilacap is appropriate, although there are minor discrepancies for people's activity in 'rivers' in rural areas in the day-time.

6.3.2 Accuracy of population distribution

The accuracy of population distribution from the proposed model has been examined to show the quality of the model. Similarly, the comparative study of the accuracy has been carried out to show improvement of the model used over the available approaches, such as those of Gallego and Pedell (2001) and Mennis (2003). Based on the accuracy assessment and the comparative study of the population distribution modeling, the proposed model has the lowest error of population distribution (E_{PD}) and a low root mean square error (RMSE). This means that the proposed model is better than the previous model and

closer to the reality of population distribution. The reasons for these limitations were described in chapter 3.3.2, particularly, on the weighting factor determination. As mentioned above, the weighting factor is the important component to drive the disaggregation of census data using land use data. The quality of the weighting factor plays a role in the accuracy of population distribution.

For further analysis, such as the calculation of the number of people in the hazard zone and evacuation planning, the accuracy of population distribution data is important in order to ensure an appropriate capacity of evacuation shelters.

6.3.3 Multi-scale disaggregation of population distribution

The multi-scale disaggregation has been undertaken to solve the problem on the missing of population and occupation data, and the different scale of land use data. This method provides the best available population distribution for the west coast of Sumatra, south coast of Java and Bali. This result is important for the broad scale analysis of tsunami risk and vulnerability. For areas where data are missing or low-scale resolution for land use data are available only, errors appear. The overall discussion on the source of errors and their effects for the population distribution has been presented in Khomarudin *et al.* (2010). The sources of errors in the multi-scale disaggregation of population distribution are from:

- statistical data (number of people, administrative boundaries);
- land use data (thematic, geometric error);
- weighting factor (allocation of population to land use classes).

6.3.4 Commuters

Based on the questionnaire, it was recorded that there is a high number of commuters in the study area (see chapter 5.3.3). This means that the population distribution modeling must consider the commuters as an important component, especially with respect to day- and night-time population distribution. For the study area, it is difficult to use this parameter due to the lack of available data. In the statistical data, the commuters are not recorded for day-time activity, such as in the InfoUSA data, which was used by Sleeter and Wood (2006) for their people distribution mapping.

A method should be developed to accommodate the commuters in the population distribution modeling. The commuter analysis should use the PODES data, which show the number of infrastructures for every village, such as hospitals, schools, government buildings, hotels, supermarkets. Figure 6.5 shows the potential additional of people in Cilacap District based on the infrastructures properties.

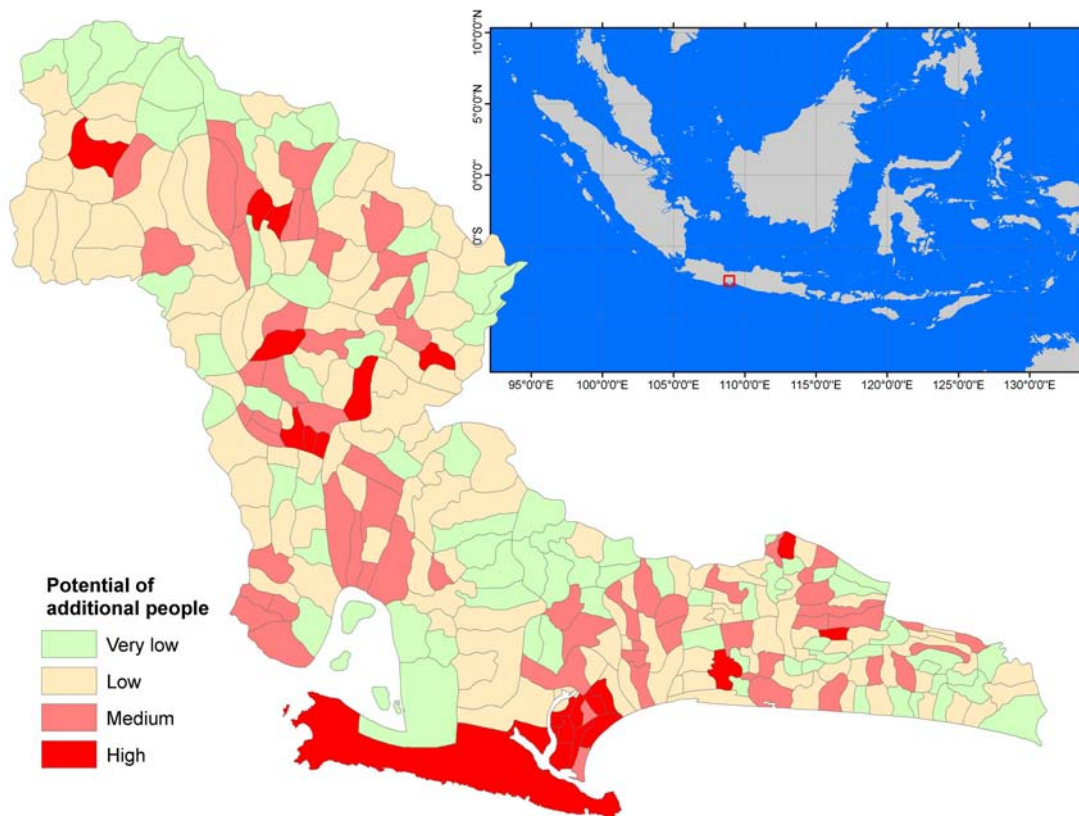


Figure 6.5 Potential additional of people in the Cilacap District (red represents the high potential additional of people due to infrastructural facilities)

The difficulties of this analysis are how to derive the actual number of additional people from the potential additional of people in a village. The first solution is by using the result of the questionnaire, calculating how many commuters are within a municipality and within a district by accumulating the total number of commuters from the villages. Within the municipality, the sums of these numbers are then distributed to every village in the municipality again, based on the number of infrastructure in the respective villages. The village with the highest proportion of infrastructure will have the highest additional number of

people as well. Figure 6.6 shows a calculation of additional people for the villages based on the infrastructure proportion.

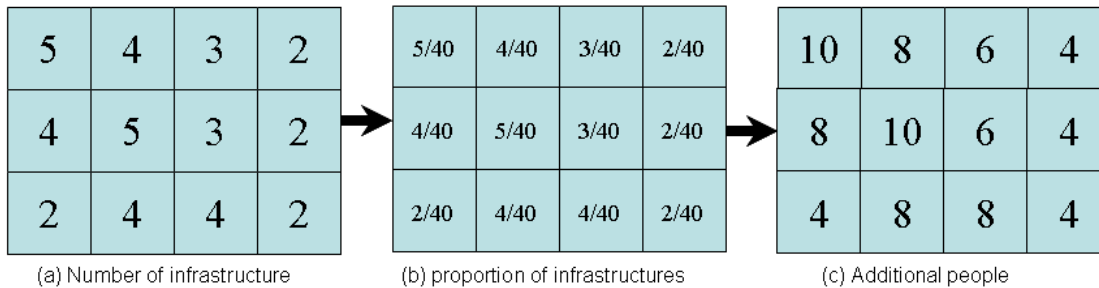


Figure 6.6 Example showing the additional number of people in villages based on The sum of the commuters (here, 80 people) within the municipality

This population distribution result by accommodating the commuters as an important component in the population distribution has been used to calculate the reference of population distribution modeling.

6.4 Settlement Classification

The methodology of settlement classification of TerraSAR-X imagery by using a combination of speckle divergence and the neighborhood analysis has been applied in order to improve the settlement detection, especially for hilly areas. The result of the accuracy assessment for the settlement classification in the study area is over 85%. The high omission and commission errors appear in the result of settlement classification, especially in the rural areas. It is assumed that this is because of the settlement conditions in the rural area, which are usually surrounded by vegetation. The texture of this area in the TerraSAR-X imagery is more likely vegetation than settlement area. Figure 6.7 shows pictures of settlement areas with vegetation surrounding, taken in the rural area of Cilacap.

The transferability of this method has been also successfully investigated in the Padang area, which resulted in an accuracy of over 75%. The omission and commission errors for the rural area are also high, such as for the Cilacap District. The characteristics of settlement areas in the rural parts of Padang are mostly similar to the conditions of the rural areas in Cilacap District.



Figure 6.7 Examples of settlements in the rural area of Cilacap District (settlements surrounded by vegetation)

CHAPTER 7: APPLICATIONS OF THE RESEARCH

7.1. Overview

This chapter shows the application of surface roughness determination and population distribution modeling in the context of tsunami risk and vulnerability assessment. It explains the relevance of using information on surface roughness in the hazard assessment and of enhanced population distribution data for the human exposure assessment, the application of the results in evacuation planning, and the contributions to the decision support system for tsunami early warning.

7.2. Hazard Assessment

The research on determining land use specific roughness coefficients is an important input for tsunami hazard mapping, either through empirical or numerical methods. As mentioned in chapter 5, the different roughness coefficients used as input to tsunami inundation models lead to a better representation of tsunami flow-depth, inundation and velocity.

In addition to the detailed assessment for Cilacap District shown in chapter 5.2.3, the integration of spatially distributed surface roughness to tsunami inundation modeling was also applied to Padang City. Modeled inundation areas obtained when considering uniform roughness values yield a 30% different in inundation area compared to the spatially distributed roughness results. The hazard zone obtained based on uniform roughness coefficient as input for the tsunami inundation modeling is overestimated. Figure 7.1 shows the different inundation results by using uniform and spatially distributed roughness.

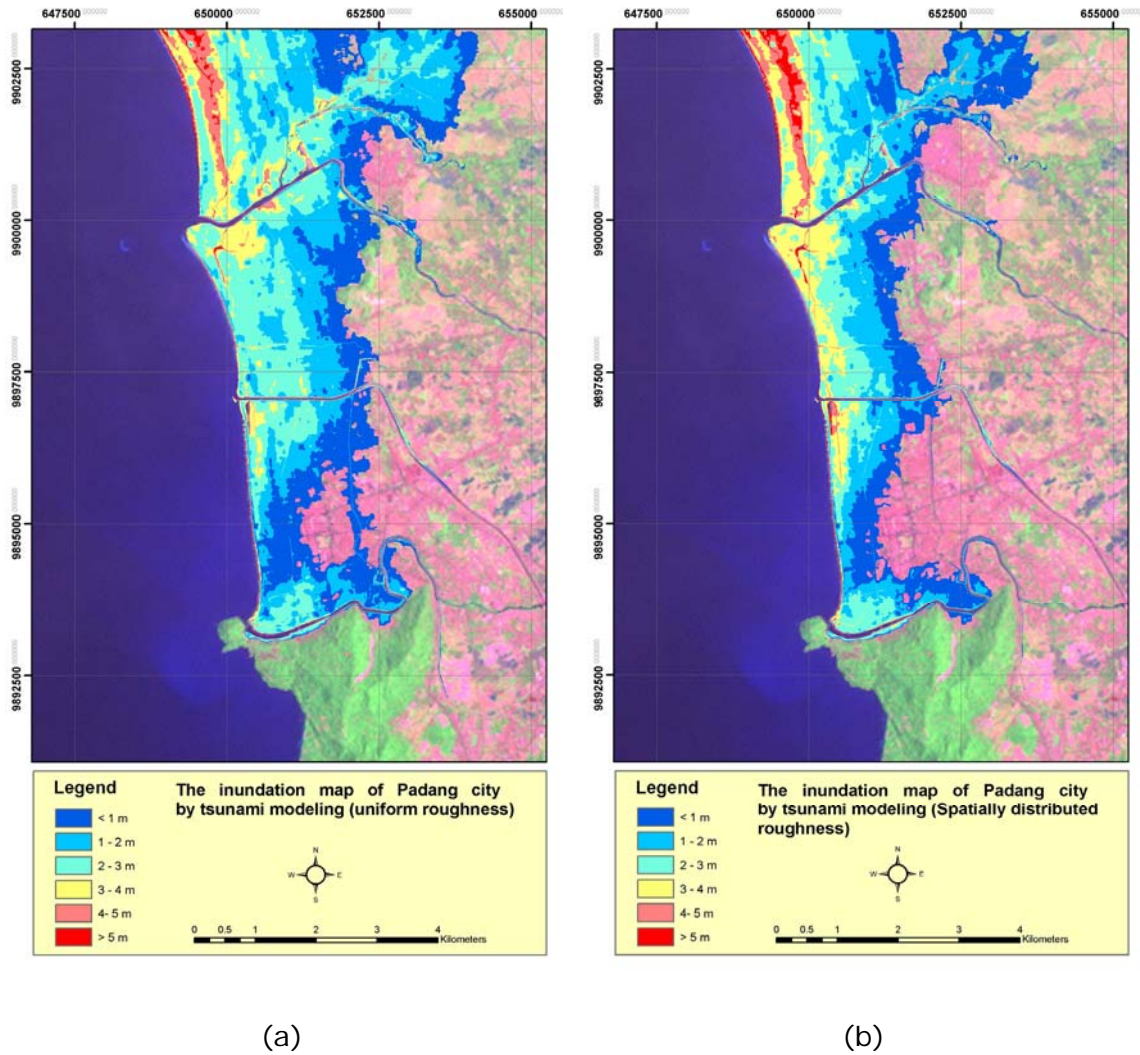


Figure 7.1 Tsunami inundation maps for Padang District based on (a) uniform roughness and (b) spatially distributed roughness

Figure 7.1 clearly shows the differences in the tsunami inundation area. Generally, it can also be stated that for areas with low surface roughness the tsunami flow depth is high and vice versa. For near coastal areas in Padang, the tsunami flow depth modeled based on spatially distributed roughness is higher than for uniform roughness conditions. On the other hand, for inland areas, the roughness values are indicating higher surface resistance, and consequently lower flow depth values and smaller inundated area.

Knowledge on spatially distributed roughness values is an important step towards a deeper consideration of surface characteristics and their impacts on tsunami inundation. Providing spatial distributed roughness information through field surveys is time consuming and costly. The remote sensing approach proposed and applied here provides up-to-date and spatially distributed spatially distributed roughness conditions.

7.3. Human Exposure

Improved spatial resolution of population distribution representation is of great importance for people exposure assessment. In Figure 7.2 and 7.3, the hazard map and the human exposure map for the Denpasar using the results of the population modeling approach are shown. The results show that, for example, the hazard zone is relatively large in the south-eastern part of Denpasar. But, this area is mainly covered by swamp areas with low population density and therefore leads to a low exposure. On the west coast of Denpasar, the situation is different. Here, the hazard zone covers only a small but highly populated area. Consequently, in the western part of Denpasar, there are many more people exposed to tsunami hazards.

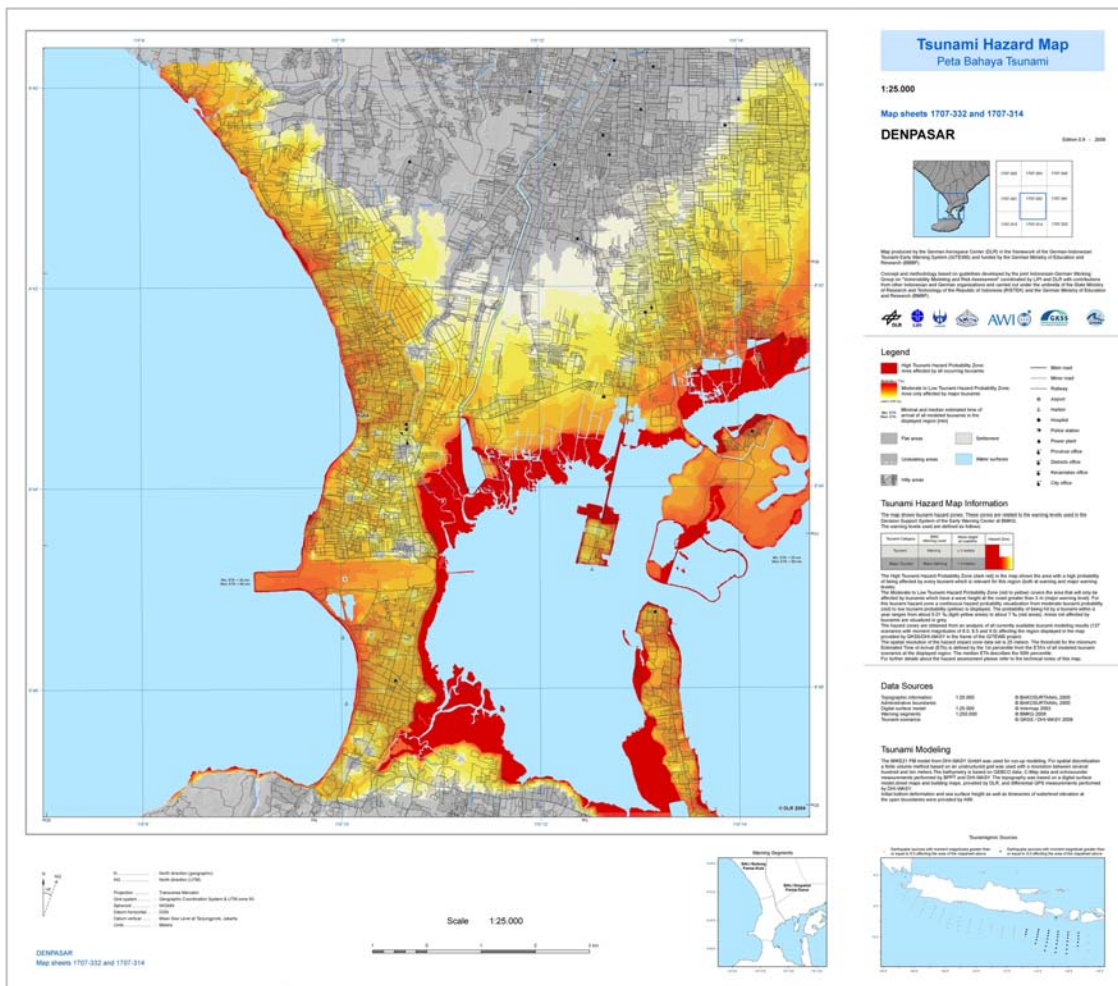


Figure 7.2 Tsunami hazard maps for Kuta, Bali

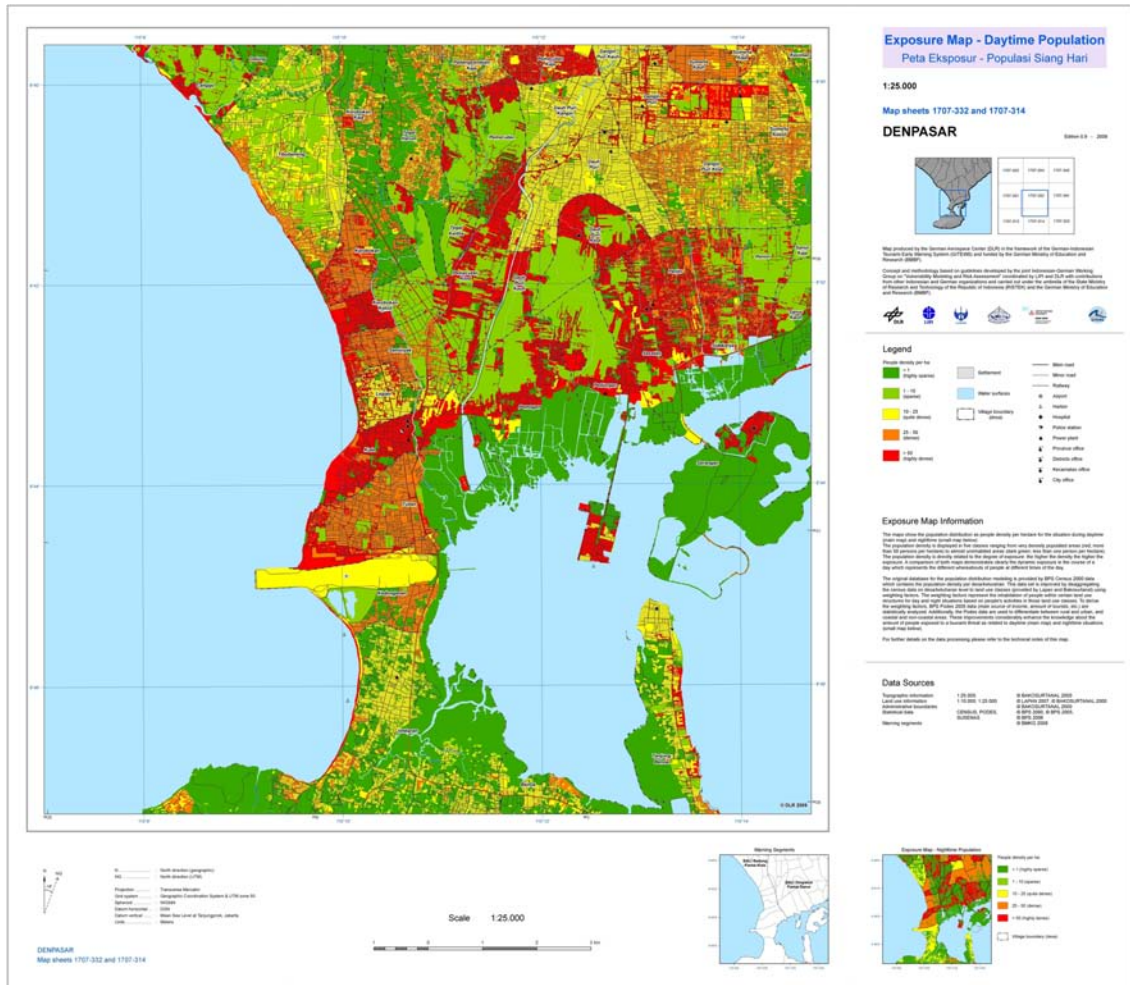


Figure 7.3 Human exposure maps of Kuta, Bali for day-time population

The importance of population distribution analysis is also shown by Kongko *et al.* (2009). The implementation of the developed method for population modeling by day- and night-time in several scenarios of tsunami inundation reveals significant differences in the number of people exposed. Figure 7.4 shows the differences between the numbers of exposed people during the day- and night-time.

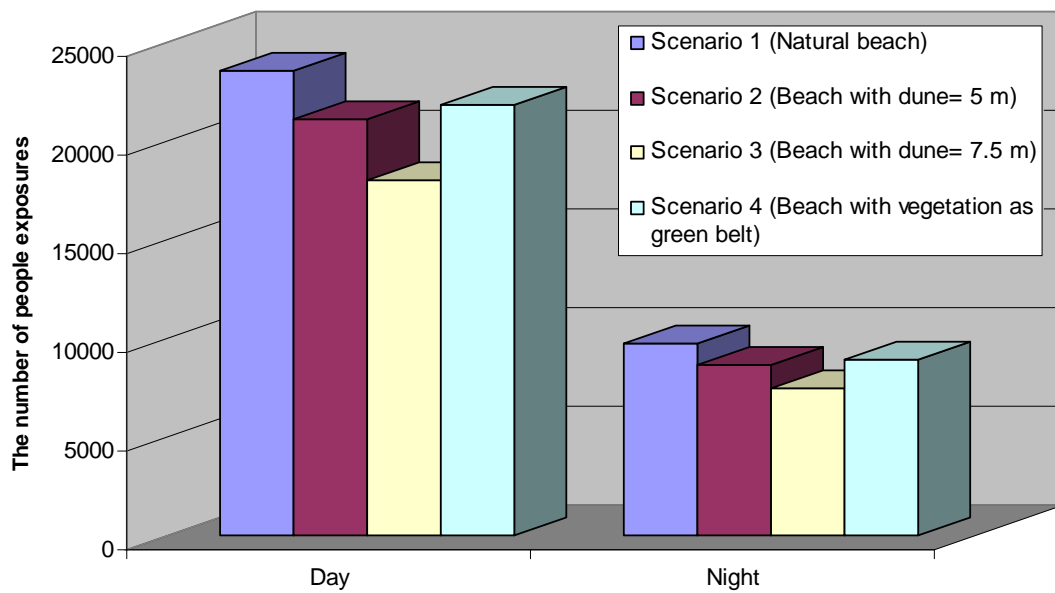


Figure 7.4 Number of exposed people in the hazard zone and changes during the day- and night-time

This result shows the importance of population distribution research for dynamic human exposure assessment. The population distributions during day- and night-time provide enhanced information for tsunami hazard and human exposure assessment.

7.4. Evacuation Planning

The combination of the population distribution with the hazard zone is an important input for evacuation planning. Post *et al.* (2009) implemented the population distribution to estimate the casualties and displaced people by using human immediate response capability. Figure 7.5 shows the estimation of potential casualties on the evacuation time map. The circle on the map shows the number of people who are able to evacuate within 20 and 80 minutes and those who need more than 80 minutes.

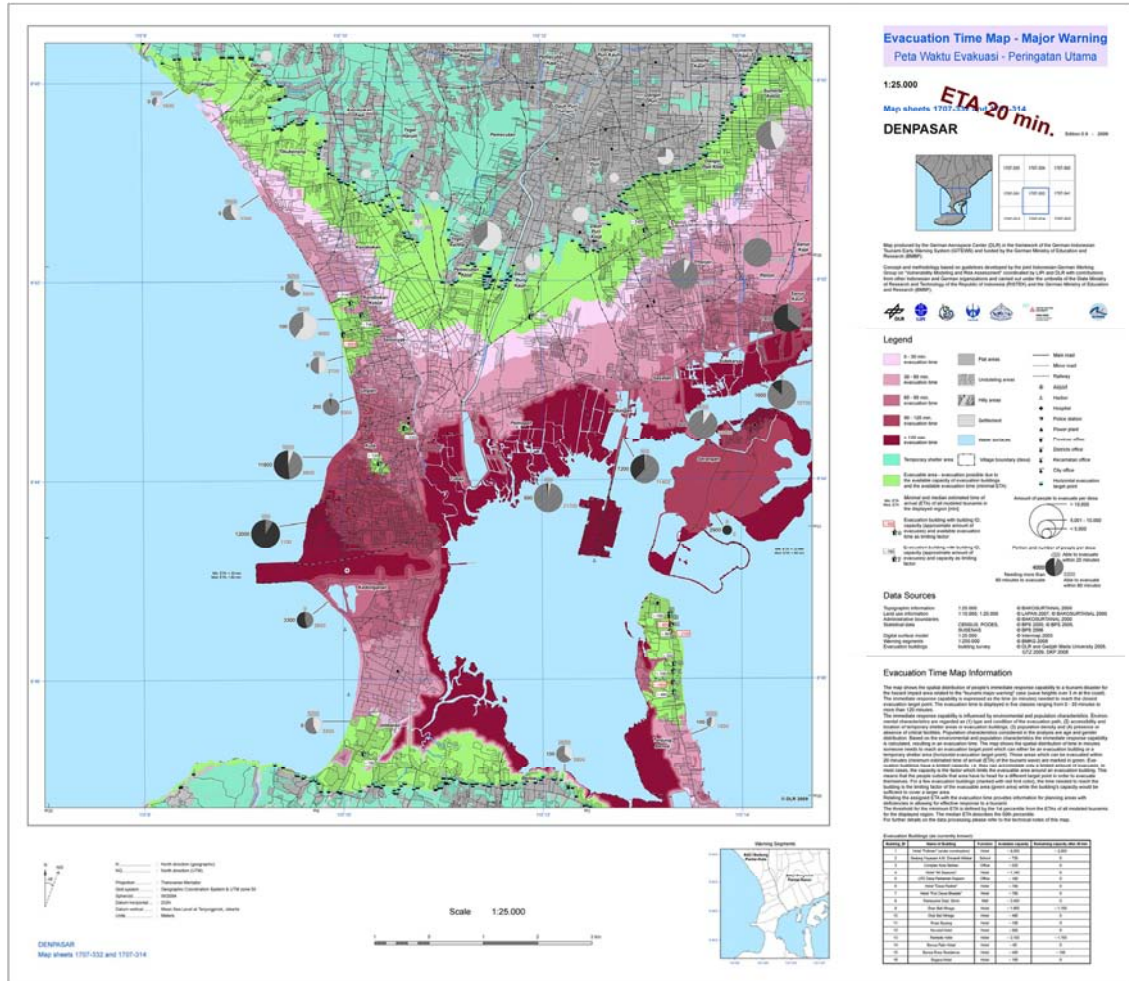


Figure 7.5 Evacuation time map for major warning tsunami in Bali (Source: Strunz *et al.*, 2010)

Another important application of the population distribution modeling is to provide input for the planning of tsunami evacuation buildings and shelters. To map the areas that need additional evacuation shelters, the population distribution map is an important input.

Figure 7.6 shows the planning map with recommendations on where to build tsunami evacuation buildings as shelters. It shows the areas with very high demand (red) or high demand (orange) for additional vertical evacuation buildings.

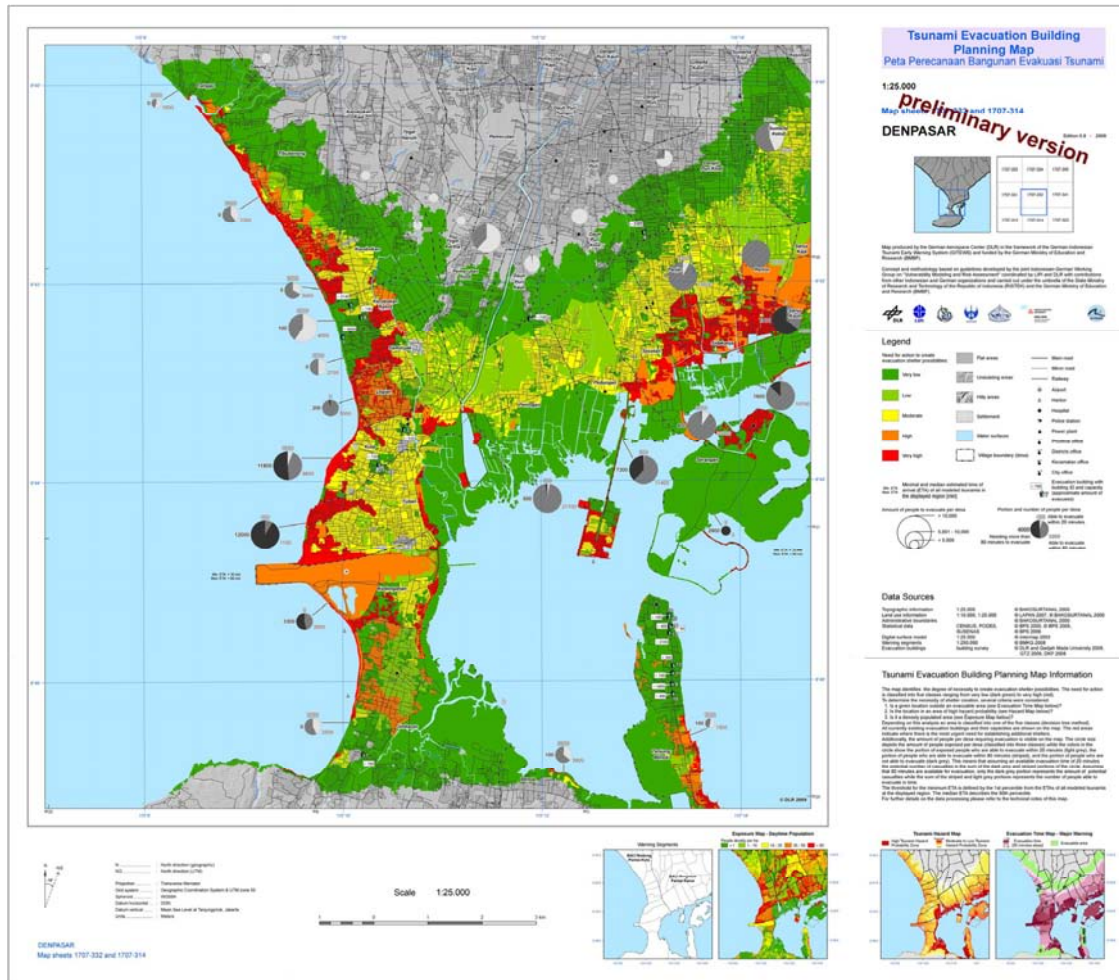


Figure 7.6 A Planning map recommendations for the need of tsunami evacuation buildings (Source: Strunz *et al.*, 2010)

7.5. Integration into the Decision Support System

The Decision Support System (DSS) for Tsunami Early Warning (Raape *et al.*, 2008) forecasts tsunami arrival time and expected wave heights at coast for defined warning segments along the coasts of Indonesia. The wave height at coast is also taken as an indicator for expected tsunami intensity. Wave heights at the coast above 3 m lead to a “major tsunami warning”, whereas wave heights between 0.5 and 3 m lead to a “tsunami warning”; and wave heights below 0.5 m, to a tsunami advisory (Raape *et al.*, 2008). These thresholds are defined by the national Indonesian Tsunami Early Warning Centre. Hazard zones are calculated for both cases by categorizing the modeled tsunami scenarios according to their wave heights at the coast.

By displaying both hazard zones of the “warning Level” with the hazard zone of the “major warning level” the potentially hazard impact areas dependent on the respective warning level may be determined. By overlaying the derived hazard zones with the detailed population distribution, the exposed people in case of a tsunami at a certain warning level within a certain warning segment can be estimated. These numbers are implemented in the DSS.

Figure 7.7 shows how these parameters are integrated in the graphical user interface of the decision support system. For each of the warning levels and warning segments, the number of people endangered by the expected tsunami is shown.

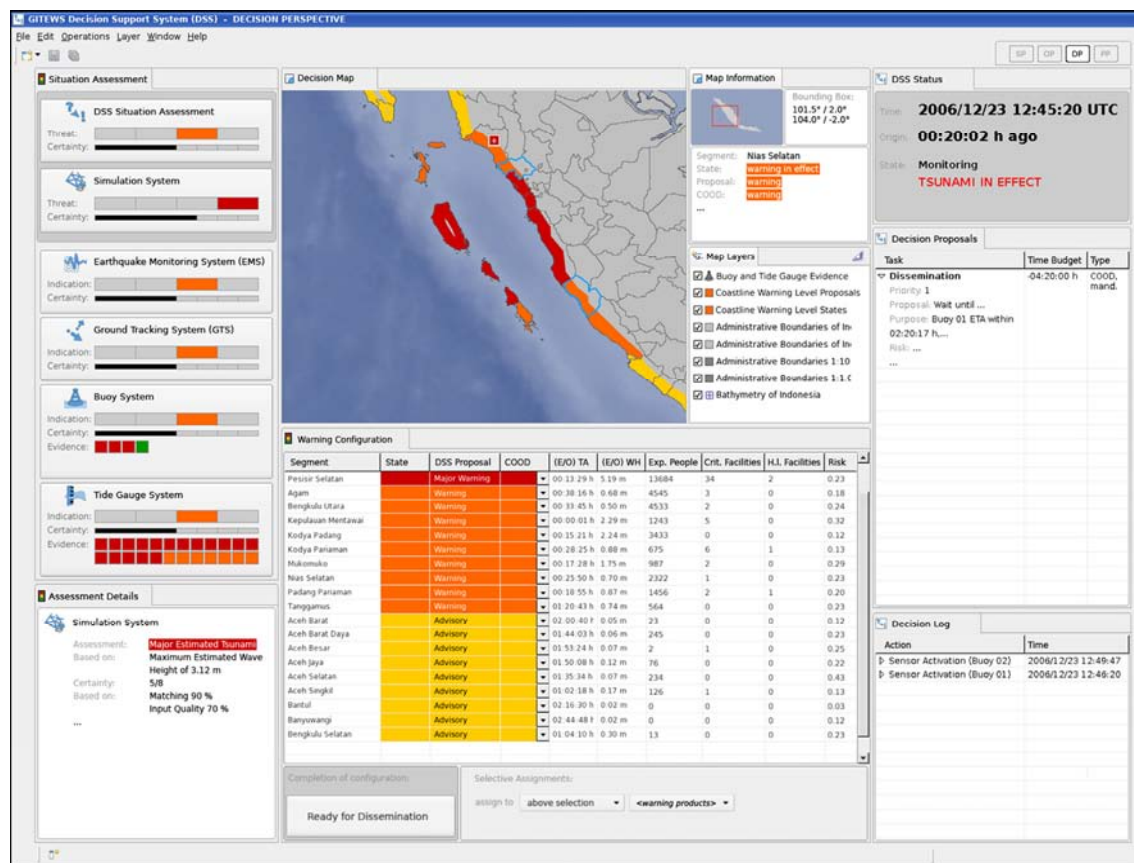


Figure 7.7 Integration of key parameters of risk assessment into the decision support system (Source: Raape *et al.*, 2008)

7.6. Conclusions

From the explanations above, it can be concluded that the results of surface roughness determination and population distribution provide an essential basis for several applications. For tsunami modeling, the spatially distributed surface roughness is an important parameter to more precisely reflect flow resistance properties of the land surface, and hence improve the results of tsunami inundation modeling. The population distribution is a crucial input for human exposure assessment during the day- and night-time, evacuation planning and for estimating potential casualties, and the shelter evacuation building planning. Finally, key parameters derived from population distribution modeling have been implemented in the decision support system for tsunami early warning.

CHAPTER 8: CONCLUSIONS

This research has successfully developed a concept on remote sensing classification for providing spatially distributed surface roughness, a concept for spatial and temporal population distribution modeling, and an improved concept of settlement classification based on speckle divergence and neighborhood analysis of SAR data. These concepts have successfully contributed to tsunami risk and vulnerability assessments. The applications of the research results include tsunami inundation modeling, human exposure analysis, evacuation planning, and integration into the decision support system of the Tsunami Early Warning Center in Jakarta.

It can be concluded that the research questions in chapter 1.2 have been answered and that the research objectives are achieved. Further details of these achievements are concluded below.

8.1. Surface Roughness Determination

The importance of coastal vegetation in terms of tsunami impact reduction is unambiguous. But the amount of reduction depends not only on the mere presence of vegetation. The surface roughness of vegetation varies according to its characteristics, namely the density of vegetation and neighborhood conditions of the different land uses. Thus, it is essential to identify appropriate surface roughness classes and quantify those in order to integrate them into tsunami inundation modeling and hazard assessment.

Analyses of various published surface roughness classes and coefficients made it possible to elaborate the values and variances of the roughness coefficients in each class. This process solved the current problem of coefficient inconsistencies. Furthermore, the variances were applied for uncertainty analyses of tsunami inundation modeling results.

The research developed a method to derive surface roughness classes from remote sensing data including density and neighborhood classification. This classification process is essential for deriving spatially explicit information about surface roughness conditions along the coast. The developed classification procedure consists of two steps — main land use classification and density and neighborhood analyses.

The decision tree method was used to determine relevant and appropriate image variables as input for the main land use classification. It was revealed that texture is a crucial variable to increase the classification accuracy, especially in terms of separating residential areas and bare land.

Standard remote sensing classification techniques such as supervised maximum likelihood are not capable to differentiate adequately and reliably mixed classes. Thus, the specific method of density and neighborhood classification was developed and applied. This method makes it possible to the density and neighboring objects or pixels using the determined a kernel. This two-stage classification procedure then provides spatially distributed surface roughness conditions which were supplied with the previously derived roughness coefficients.

The thus derived spatial distribution of surface roughness was integrated in tsunami inundation modeling. The influence of surface roughness on the inundation results was demonstrated. By varying the roughness coefficients at several locations, the sensitivity of tsunami modeling to surface roughness conditions was quantified.

This research brings a new dimension of knowledge for determining surface roughness conditions and coefficients. Moreover, the sensitivity of the models to surface roughness conditions was investigated, which is important for developing tsunami hazard maps. Such maps are vital for further vulnerability and risk analyses such as the assessment of population exposure or evacuation planning.

8.2. Population Distribution Modeling

For the purposes of risk assessment, especially tsunami risk assessment, knowledge about the distribution of the population is essential. This research successfully developed a concept to significantly improve population distribution data based on available census and other statistical data.

The developed method combines census data with land use data as ancillary data. The underlying idea is to disaggregate people into an administrative unit to more spatially explicit units in order to derive a more realistic spatial distribution. To consolidate the population distribution modeling, further statistical data were used to determine weighting factors for the disaggregation to each land use class.

The inclusion of socio-economic data to determine weightings allows to consider both spatial and temporal changes of the population distribution. The analyses

of the socio-economic data showed that weighting factors must be specific for every village. A generalized weighting factor must be used in case of insufficient or missing data. Missing data is a frequent problem, especially in developing countries like Indonesia. Therefore, a multi-scale disaggregation approach was developed and successfully applied in order to derive village-specific weightings, where possible, and generalized weightings, where necessary.

It was also demonstrated that commuters are an important component to consider in the modeling. Disregarding them decreases the accuracy of the resulting population distribution. As a result of the limited availability of data that captures commuters in Indonesia, it was suggested that infrastructure data could be used in future research to generate information on commuter behavior.

Reference population data derived from survey data was used to analyze the error of the population distribution. Other available population distribution models showed a higher error than this newly developed approach.

The developed population distribution model results in highly accurate data sets, which are of great use for risk and vulnerability analyses. The results were implemented in the field of tsunami risk research, but detailed information about population distribution is also valuable for many other natural hazards.

8.3. Settlement Classification

A method of settlement classification using radar data has been further developed using TerraSAR-X imagery. The improvements include the concept of neighborhood analysis. The settlement areas in the study area were mapped with satisfying overall accuracy. Limitations occur in rural areas where settlement structures are interspersed with vegetation. This results in relatively high omission and commission errors.

The method of settlement classification has been successfully transferred to the Padang District. Here, the same problem regarding rural settlements occurred, also resulting in high omission and commission errors in rural areas. Thus, future research on settlement classification using radar imagery must focus on improving the methodology for the classification of rural settlements.

8.4. Applications of the Research

Accuracy, benefit and usefulness of risk and vulnerability analyses always depend on the availability of spatially and temporally adequate and appropriate data. Remote sensing and GIS technology play an important role in providing, compiling and processing such data.

The findings of the present research provide a valuable contribution to tsunami risk and vulnerability assessments in the framework of the GITEWS project. The results were successfully used for tsunami inundation modeling and hazard mapping, human exposure analysis and evacuation planning efforts. Furthermore, the results were integrated in the decision support system of the Tsunami Early Warning Center in Jakarta.

REFERENCES

- Acqua, F.D., Gamba, P., and Lisini, G. (2003). Improvements to Urban Area Characterization Using Multitemporal and Multiangle SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (9), 1996-2004.
- Acqua, F.D., and Gamba, P. (2003). Texture-Based Characterization of Urban Environments on Satellite SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (1), 153-159.
- Acrement, G.J. and Schneider, V. R. (1989). Guide for Selecting Manning's Roughness Coefficients for Natural Channel and Floodplains. Technical Report WSP2339, USGS.
- Adams, B.J., Huyck, C.K., Mansouri, B., Eguchi, R.T., and Shinozuka, M. (2004). Imagery for Post-Earthquake Damage Assessment: The 2003 Boumerdes (Algeria) and Bam (Iran) Earthquakes. Research Progress and Accomplishments 2003-2004. Multidisciplinary Center for Earthquake Engineering Research. University at Buffalo, State University of New York.
- Ahola, T., Virrantaus, K., Krisp, J.M. and Hunter, G.J. (2008). A Spatio-temporal Population Model to Support Risk Assessment for Emergency Response decision-making. *International Journal of Geographical Information Science - IJGIS*, 21 (8), 935 – 953.
- Annaka, T., Satake, K., Sakakiyama, T., Yanagisawa, K., Shuto, N. (2007). Logic-tree Approach for Probabilistic Tsunami Hazard Analysis and its Applications to the Japanese Coasts. *Pure Appl. Geophys*, 164, 577–592.
- Badan Pusat Statistik Indonesia (BPS). (2005). Publikasi Elektronik Kabupaten Cilacap Tahun 2005. Badan Pusat Statistik Kabupaten Cilacap, Indonesia.
- Balk, D. L., Deichmann, U., Yetman, G., Pozzi, F., Hay, S.I., and Nelson, A. (2005). Determining Global Population Distribution: Methods, Applications and Data. CIESIN (Center for International Earth Science Information Network), USA.
- Belluco, E., Camuffo, M., Ferrari, S., Modenese, L., Silvestri, S., Marani, A., and Marani, M. (2006). Mapping salt-marsh vegetation by multispectral and

- hyperspectral remote sensing. *Remote Sensing of Environment*, 105, 54–67.
- Belward, A.S., Stibig, H.J., Eva, H., Rembold, F., Bucha, T., Hartley, A., Beuchle, R., Khudhairy, D., Michielon, M., and Mollicone, D. (2007). Mapping Severe Damage to Land Cover Following the 2004 Indian Ocean Tsunami using Moderate Spatial Resolution Satellite Imagery, *International Journal of Remote Sensing*, 28 (13), 2977 – 2994.
- Berryman, K. (2006). Review of Tsunami Hazard and Risk in New Zealand. Confidential. Institute of Geological & Nuclear Sciences, New Zealand.
- Bielecka, E. (2005). A Dasymetric Population Density Map of Poland. Institute of Geodesy and Cartography, University of Warsaw, Poland.
- Birkmann, J. (2006). Measuring Vulnerability to Hazards of Natural Origin – Towards Disaster- Resilient Societies, UNU press, Tokyo.
- Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P. (2009). Object-oriented Image Processing in an Integrated GIS/remote Sensing Environment and Perspectives for Environmental Applications. In: Cremers, A., Greve, K. (Eds.), *Environmental Information for Planning, Politics and the Public*, 2. Metropolis Verlag, Marburg, 555-570.
- Briggs, D.J. , Gulliver, J., Fecht, D., and Vienneau, D.M. (2007). Dasymetric Modelling of Small-area Population Distribution using Land Cover and Light Emissions Data. *Remote Sensing of Environment*, 108, 451–466.
- Burbidge, D., Cummins, P.R., Mleczo, R., and Thio, H.K. (2008). A Probabilistic Tsunami Hazard Assessment for Western Australia. *Pure Appl. Geophys.*, 165, 2059–2088.
- Brisbane City Council. (2003). *Natural Channel Design Guidelines*. Grant Witheridge, Catchment & Creeks Pty Ltd. ISBN 187609. http://www.brisbane.qld.gov.au/BCC:BASE::pc=PC_1859.
- Chang, S.E., Adams, B.J., Alder, J., Berke, J.R., Chuenpagdee, R., Ghosh, S., and Wabnitz, C. (2006). Coastal Ecosystems and Tsunami Protection after the December 2004 Indian Ocean Tsunami. *Earthquake Spectra*, 22 (S3), S863–S887.
- Chatenoux B. and Peduzzi P. (2007). Impacts from the 2004 Indian Ocean Tsunami: Analysing the Potential Protecting Role of Environmental Features. *Natural Hazards*, 40, 289–304.

- Chen, K. (2002). An Approach to Linking Remotely Sensed Data and Areal Census Data. *International Journal of Remote Sensing*, 23 (1), 37 — 48.
- Chen, J., and Wang, R. (2007). A Pairwise Decision Tree Framework for Hyperspectral Classification. *International Journal of Remote Sensing*, 28 (12), 2821-2830.
- Chiang, S., Tachikawa, Y., Kojima, T., and Takara, K. (2004). A New Hydrologic Response Function Physically Derived from DEM and Remote Sensing Image. Proc. of the 2nd Asia Pacific Association of Hydrology and Water Resources (APHW) Conference, July 5-8, 2004, Singapore, 2, 268-275.
- Chow, V.T. (1959). *Open-channel Hydraulics*. McGraw-Hill Book Company, Inc., 109–119.
- Comber, A., Anthony, S., and Proctor, C. (2008). The Creation of a National Agricultural Land Use Dataset: Combining Pycnophylactic Interpolation with Dasymetric Mapping Techniques. *Transactions in GIS*, 12(6), 775–791.
- Congalton, R. and Green, K. (1999). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC/Lewis Press, Boca Raton, FL. 137 p.
- Corbane, C., Faure, J.F., Baghdadi, N., Villeneuve, N., and Petit, M. (2008). Rapid Urban Mapping Using SAR/Optical Imagery Synergy. *Sensors*, 8, 7125-7143.
- Cowan, WL. (1956). Estimating hydraulic roughness coefficients. *Agricultural Engineering*, 37(7), 473–475.
- Danielsen, F., Sorensen, M.K., Olwig, M.F., Selvam, V., Parish, F., Burgess, N.D., Hiraishi, T., Karunakaran, V.M., Rasmussen, M.S., Hansen, L.B., Quarto, A., and Suryadiputra, N. (2005). The Asian Tsunami: A Protective Role for Coastal Vegetation. *Science*, 310, 643.
- Dao, M. H., and Tkalich, P. (2007). Tsunami Propagation Modelling – A Sensitivity Study. *Natural Hazards Earth Syst. Sci.*, 7, 741–754.
- Dekker, R.J. (2003). Texture Analysis and Classification of ERS SAR Images for Map Updating of Urban Areas in the Netherlands. *IEEE Transactions on Geoscience And Remote Sensing*, 41 (9), 1950-1958.

- Dobson, J. E., Bright, E. A., Coleman, P. R., Durfee, R. C., Worley, B. A. (2000). LandScan: A Global Population Database for Estimating Populations at Risk. *Photogrammetric Engineering and Remote Sensing*, 66, 849-857.
- Dutta, D., Alam, J., Umeda, K., Hayashi, M. & Hironaka, S. (2007). A two-Dimensional Hydrodynamic Model for Flood Inundation Simulation: A Case Study in the Lower Mekong River Basin. *Hydrological Processes*, 21, 1223-1237.
- Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R. and Davis, C. W. (1997). Relation between Satellite Observed Visible-near Infrared Emissions, Population, Economic Activity and Electric Power Consumption. *International Journal of Remote Sensing*, 18 (6), 1373-1379.
- Epstein, J., Payne, K. M., and Kramer, E. (2002). Techniques for Mapping Suburban Sprawl. *Photogrammetric Engineering & Remote Sensing*, 63 (9), 913-918.
- Esch, T., Thiel, M., Schenk, A., Roth, A., Müller, A., and Dech, S. (2010). Delineation of Urban Footprints from TerraSAR-X Data by Analyzing Speckle Characteristics and Intensity Information. *IEEE Transactions on Geoscience and Remote Sensing*, 48 (2), 905-916.
- Fisher, P.F. and Langford, M. (1996). Modeling Sensitivity to Accuracy in Classified Imagery - A study of Areal Interpolation and Dasymetric Mapping. *Professional Geographer*, 48 (3), 299-309.
- Friesecke, F. (2006). Geodetic Applications for a More Cost-Effective and Sustainable Disaster Risk Management. GSDI-9 Conference Proceedings, 6-10 November 2006, Santiago, Chile.
- Gallego, J. and Peedell, S. (2001). Using CORINE Land Cover to map population density. Towards Agri-environmental indicators, Topic report 6/2001, 92-103. http://reports.eea.eu.int/topic_report_2001_06/en, Copenhagen.
- Gamba, P. and Acqua, F.D. (2007). Fusion of Radar and Optical Data for Identification of Human Settlements. In Weng, Q. (Editor): *Remote Sensing of Impervious Surface*. Taylor and Francis Groups, LLC.
- Gayer, G., Leschka, S., Nöhren, I., and Mir, J. (2008). About the Importance of a Detailed Roughness Description in Tsunami Inundation Modelling (appendix II). GITEWS Report 5326.

- Geist, E.L. and Parsons, T. (2006). Probabilistic Analysis of Tsunami Hazards. *Natural Hazards*, 37, 277–314.
- Harada, K. and Kawata, Y. (2004). Study on the Effect of Coastal Forest to Tsunami Reduction. *Kyoto Daigaku Bosai Kenkyujo nenpo*, 47, 273-279.
- Hasan, N. and Akamatsu, N. (2004). A New Approach for Contrast Enhancement Using Sigmoid Function. *The International Arab Journal of Information Technology*. 1 (2), 221-225.
- Hills, J.G. and Mader, C.L. (1987). Tsunami Produced by the Impacts of Small Asteroids. *Ann NY Acad Sci*, 822, 381-394.
- Hiraishi, T. (2005). Greenbelt Technique for Tsunami Disaster Reduction. APEC-EqTAP Seminar on Earthquake and Tsunami Disaster Reduction. September, 27-28, Jakarta, Indonesia.
- Hofstee, P. and Islam, M. (2004). Disaggregation of Census Districts: Better Population Information for Urban Risk Management. *Proceedings of the 25th Asian Conference on Remote Sensing*, 1-2, 1206– 1211.
- Holloway, S., Schumacher, J., and Redmond, R. (1997). *Population & Place: Dasymetric Mapping using Arc/Info*. Cartographic Design Using ArcView and Arc/Info, Missoula: University of Montana, Wildlife Spatial Analysis Lab.
- Holt, J.B., Lo, C.P., and Hodler, T.W. (2004). Dasymetric Estimation of Population Density Areal Interpolation of Census Data. *Cartography and Geographic Information Science*, 31 (2), 101-121.
- Huete, A. R., (1988). A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25, 295-309.
- Hung, M. and Ridd, M.K. (2002). A Subpixel Classifier for Urban Land-cover Mapping Based on a Maximum-likelihood Approach and Expert System Rules. *Photogrammetric Engineering and Remote Sensing*, 68, 1173–1180.
- Imamura, F., Yalciner, A. and Ozyurt, G. (2006). *Tsunami Modelling Manual*. UNESCO IOC International Training Course on Tsunami Numerical Modelling.
- Imamura, F. (2009). Tsunami Modeling: Calculating Inundation and Hazard Maps. *The Sea*. Volume 15, edited by Eddie N. Bernard and Allan R. Robinson. The President and Fellows of Harvard College, USA.

- Indonesian National Coordinating Agency for Disaster Management (BAKORNAS BPB). (2006). Laporan Perkembangan Penanganan Bencana Gempa Bumi dan Tsunami di Jawa Barat, Jawa Tengah dan DI Yogyakarta (The Progress Report of Disaster Management on Earthquake and Tsunami in the West Java, Central Java and DI Yogyakarta. 1 August 2006.
- ISDR. (2004). *Living with Risk: A global review of disaster reduction initiatives* UNISDR, Geneva.
- ITDB/WLD. (2007). *Integrated Tsunami Database for the World Ocean*. CD-ROM, Version 6.52 of December 31, 2007, Tsunami laboratory, ICMMG SD RAS, Novosibirsk, Russia.
- Iverson, L.R. and Prasad, A.M. (2007). Using Landscape Analysis to Assess and Model Tsunami Damage in Aceh Province, Sumatra. *Landsc. Ecol.*, 22, 323–331.
- Järvelä, J. (2005). Effect of Submerged Flexible Vegetation on Flow Structure and Resistance. *Journal of Hydrology*, 307 (1–4), 233–241.
- Judex, M., Thamm, H.P., and Menz, G. (2006). Improving Land-Cover Classification with a Knowledge Based Approach and Ancillary Data. *Proceedings of the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover*.
- Kathiresan, K. and Rajendran, N. (2005). Coastal Mangrove Forests Mitigated Tsunami. *Estuarine, Coastal and Shelf Science*, 65, 601-606.
- Khomarudin, M. R., Strunz, G., Ludwig, R., Zosseder, K., Post, J., Kongko, W., and Pranowo, W. S. (2010). Hazard Analysis and Estimation of People Exposure as Contribution to Tsunami Risk Assessment in the West Coast of Sumatra, the South Coast of Java and Bali, *Zeitschrift für Geomorphologie*, 54, Suppl. 3, 337-356.
- Kongko, W. (2005). Mangrove as a Tsunami Reduction and Its Application. *APEC-EqTAP Seminar on Earthquake and Tsunami Disaster Reduction*. September, 27-28, Jakarta, Indonesia.
- Kongko, W., Khomarudin, M.R., and Schlurmann, T. (2009). Tsunami Model of Cilacap-Indonesia: Inundation and Its Effectiveness of Mitigation Measures. 2009 AGU Fall Meeting, 14-18 December, San Francisco, California, USA.
- Koshimura, S., Oie, T., Yanasigawa, H., and Imamura, F. (2009). Developing Fragility Functions for Tsunami Damage Estimation Using Numerical Model

- and Post-Tsunami Data from Banda Aceh, Indonesia. *Coastal Engineering Journal*, 51 (3), 243-273.
- Lam, N. S. (1983). *Spatial Interpolation Methods: A Review*. *The American Cartographer*, 10(2), 129-149.
- Langford, M., Maguire, D. J., and Unwin, D. J. (1991). The Areal Interpolation Problem: Estimating Population using Remote Sensing in a GIS Framework, In: Masser, I. and Blakemore, M. (eds.), *Handling Geographical Information: Methodology and Potential Applications*, Longman, London, 55-77.
- Langford, M. and Unwin, D.J. (1994). Generating and Mapping Population-Density Surfaces within a Geographical Information System. *Cartographic Journal*, 31(1), 21-26.
- Latief, H. and Hadi, S. (2007). The Role of Forests and Trees in Protecting Coastal Areas Against Tsunamis. In: *Regional Technical Workshop "Coastal protection in the aftermath of the Indian Ocean tsunami: What Role for Forests and Trees?"* Khao Lak, Thailand, 28 - 31 August 2006.
- Latief, H., Yushahnonta, P., Rahayu, H.P., Gsuman, A.R., and Sunendar, H. (2007). A Contribution of Pre-calculated Tsunami Data Base and Risk Assessment as Last Mile-Evacuation for National TEWS. In: *Workshop on Risk Modeling and Vulnerability Assessment*, 31 July – 3 August, Bandung, Indonesia.
- Leschka, S., Pedersen, C., and Larsen, O. (2009). On the requirements for data and methods in tsunami inundation modeling – Roughness map and uncertainties. *Proc. Of the 3. South China Sea Tsunami Workshop*, 3-5 Nov 2009, Penang/Malaysia.
- Li, J. and Zhao, M. (2003). Detecting Urban Land-Use and Land-Cover Changes in Mississauga Using Landsat TM Images. *Journal of Environmental Informatics*, 2 (1), 38-47.
- Li, G. and Weng, Q. (2005). Using Landsat ETM_ Imagery to Measure Population Density in Indianapolis, Indiana, USA. *Photogrammetric Engineering & Remote Sensing*, 71 (8), 947–958.
- Liu, C., Frazier, P., and Kumar, L. (2007). Comparative Assessment of the Measures of Thematic Classification Accuracy. *Remote Sensing of Environment*, 107, 606–616.

- Liu, K., Li, X., Shi, X., and Wang, S. (2008). Monitoring Mangrove Forest Changes Using Remote Sensing and GIS Data with Decision-Tree Learning. *WETLANDS*, 28 (2), 336–346.
- Liu, X. (2004). Dasymeric Mapping with Image Texture. ASPRS annual conference proceedings, Denver, Colorado.
- Liu, X., Clarke, K., and Herold, M. (2006). Population Density and Image Texture : A Comparison Study. *Photogrammetric Engineering & Remote Sensing*, 72 (2), 187-196.
- Lo, C.P. (2001). Modeling the Population of China Using DMSP Operational Linescan System Nighttime Data. *Photogrammetric Engineering & Remote Sensing*, 67 (9), 1037-1047.
- Lu, D. and Weng, Q. (2007). A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *International Journal of Remote Sensing*, 28 (5), 823 – 870.
- Lu, D., Weng, Q., and Li, G. (2006). Residential Population Estimation using a Remote Sensing Derived Impervious Surface Approach. *International Journal of Remote Sensing*, 27 (16), 3553-3570.
- Martin, D. (2006). Grid Models of Population: Temporal Comparison by Fixing the Geography. Paper Presented at ESRC Research Methods Festival, University of Oxford, 17-20 July.
- Martin, D. and Bracken, I. (1991). Techniques for Modeling Population-Related Raster Databases. *Environment and Planning A*, 23(7), 1069-1075.
- Mason, D.C., Bates, P.D., and Amico, D. (2009). Calibration of Uncertain Flood Inundation Models using Remotely Sensed Water Levels. *Journal of Hydrology*, 368, 224–236.
- McAdoo, B.G., Richardson, N., and Borrero, J. (2007). Inundation Distances and Run-up Measurements from ASTER, QuickBird and SRTM Data, Aceh Coast, Indonesia. *International Journal of Remote Sensing*, 28 (13-14), 2961 – 2975.
- Melesse, A.M., Graham, W.D., and Jordan, J.D. (2003). Spatially Distributed Watershed Mapping and Modeling: GIS-Based Storm Runoff Response and Hydrograph Analysis: Part 2. *Journal of Spatial Hydrology*. 3 (2), 1-28.
- Mesev, V. (1998). The Use of Census Data in Urban Image Classification. *Photogrammetric Engineering & Remote Sensing*, 64 (5), 431-438.

- Mennis, J. (2003). Generating Surface Model of Population Using Dasymeric Mapping. *The Professional Geographer*, 55 (1), 31-42.
- Min, L.A., Ming, L.C., and Jian, L.Z. (2002). Modeling Middle urban Population Density with Remote Sensing Imagery. *Symposium on Geospatial Theory, Processing, and Application*, Ottawa 2002.
- Morin, J., Leclerc, M., Secretan, Y., Boudreau, P. (2000). Integrated Two-Dimensional Macrophytes—Hydrodynamic Modelling. *Journal of Hydraulic Research*, 38(3), 163–172.
- Mubareka, S., Ehrlich, D., Bonn, F., and Kayitakire, F. (2008). Settlement Location and Population Density Estimation in Rugged Terrain using Information Derived from Landsat ETM and SRTM Data. *International Journal of Remote Sensing*, 29 (8), 2339 – 2357.
- Mueller, M., Segl, K., Heiden, U., and Kaufmann, H. (2006). Potential of High-Resolution Satellite Data in the Context for Vulnerability of Buildings. *Natural Hazards*, 38, 247-258.
- Murashima, Y., Takeuchi, H., Imamura, F., Koshimura, S., Fujiwara, K., and Suzuki, T. (2008). Study on Topographic Model Using Lidar for Tsunami Simulation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 37 (B8). Beijing.
- Myers, E.P. and Baptista, A.M. (2001). Analysis of Factors Influencing Simulations of the 1993 Hokkaido Nansei-Oki and 1964 Alaska Tsunamis, *Natural Hazards*, 23(1), 1–28.
- Naesset, E. (1996). Conditional tau Coefficient for Assessment of Producer's Accuracy of Classified Remotely Sensed Data. *ISPRS Journal of Photogrammetry & Remote Sensing*, 51, 91-98.
- Olwig, M.F., Sorensen, M.K., Rasmussen, M.S., Danielsen, F., Selvam, V., Hansen, L.B., Nyborg, L., Vestergaard, K.B., Parish, F. and Karunagaran, V.M. (2007). Using Remote Sensing to Assess the Protective Role of Coastal Woody Vegetation Against Tsunami Waves. *International Journal of Remote Sensing*, 28 (13-14), 3153 - 3169.
- Parish, F., and Yiew, C.T. (2005). The Role of Mangrove in Protecting Coastlines from Tsunamis. *APEC-EqTAP Seminar on Earthquake and Tsunami Disaster Reduction*. September, 27-28, Jakarta, Indonesia.
- Pellizzeri, T.M., Gamba, P., Lombardo, O., and Acqua, F.D. (2003) Multitemporal/ Multiband SAR Classification of Urban Areas Using Spatial

- Analysis: Statistical Versus Neural Kernel-Based Approach. *IEEE Transactions On Geoscience And Remote Sensing*, 41 (10), 2338-2353.
- Post, J., Mück, M., Steinmetz, T., Riedlinger, T., Strunz, G., Mehl, H., Dech, S., Birkmann, J., Gebert, N., Anwar, H., and Harjono, H. (2008). Tsunami Risk Assessment for Local Communities in Indonesia to Provide Information for Early Warning and Disaster Management. In: *Proc. International Conference on Tsunami Warning (ICTW)*. Bali, Indonesia.
- Post, J., Wegscheider, S., Mück, M., Zosseder, K., Kiefl, R., Steinmetz, T., Strunz, G. (2009). Assessment of human immediate response capability related to tsunami threats in Indonesia at a sub-national scale. *Natural Hazards and Earth System Sciences*, 9, 1075–1086.
- Power, W., Downes, G., and Stirling, M. (2007). Estimation of Tsunami Hazard in New Zealand due to South American Earthquakes. *Pure Appl. Geophys.*, 164, 547–564.
- Pozzi, F. and Small, C. (2001). Exploratory Analysis of Suburban Land Cover and Population Density in the U.S.A. *IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas*. Paper 35. Rome, Italy 8-9 November, 2001.
- Raape, U., Fauzi, Riedlinger, T., Tessmann, S., Wnuk, M., Hunold, M., Kukofka, T., Strobl, C., Mikusch, E., Dech, S. (2008). A Newly Developed Decision Support System for Improved Tsunami Early Warning in Indonesia. In: *Proc. International Conference on Tsunami Warning (ICTW)*, 12-14 November 2008, Bali, Indonesia.
- Reibel, M. and Agrawal, A. (2007). Areal Interpolation of Population Counts using Pre-Classified Land Cover Data. *Population Research and Policy Review*, 26 (619).
- RES DAM. (2000). Final Report of Helsinki University of Technology the Use of Physical Models in Dam-Break Flood Analysis.
- Richard, J.A. and Jia, X. (2006). *Remote Sensing Digital Image Analysis an Introduction*. 4th edition. Springer-Verlag Berlin Heidelberg.
- Righetti, M. and Armanini, A. (2002). Flow Resistance in Open Channel Flows with Sparsely Distributed Bushes. *Journal of Hydrology*, 269, 55–64.
- Schneiderbauer, S. (2007). Risk and Vulnerability to Natural Disasters - from Broad View to Focused Perspective. PHD Thesis, University of Berlin, 69-84.

- Schneiderbauer, S., and Ehrlich, D. (2007). EO Data Supported Population Density Estimation at Fine Resolution, Test Case Rural Zimbabwe - in: G. Zeug, M. Pesaresi - Global Monitoring for Security and Stability (GMOSS) - Integrated Scientific and Technological Research Supporting Security Aspects of the European Union - EUR 23033 EN, ISBN 978-92-79-07584-1 (2007) - JRC41424.
- Schumann, G., Matgen, P., Hoffmann, L., Hostache, R., Pappenberger, F., and Pfister, L. (2007). Deriving Distributed Roughness Values from Satellite Radar Data for Flood Inundation Modeling. *Journal of Hydrology*, 344, 96–111.
- Sellin, R., Bryant, T., and Loveless, J. (2003). An Improved Method for Roughening Floodplains on Physical River Models. *Journal of Hydraulics Research*, 41 (1), 3–14.
- Shofiyati, R., Dimiyati, R.D., Kristijono, A. and Wahyunto. (2005). Tsunami Effect in Nanggroe Aceh Darussalam and North Sumatra Provinces, Indonesia. *Asian Journal of Geoinformatics*, 5.
- Sleeter, R. and Gould, M. (2007). Geographic Information System Software to Remodel Population Data using Dasymetric Mapping Methods. U.S. Geological Survey Techniques and Methods 11-C2, 15 p.
- Sleeter, R. and Wood, N. (2006). Estimating Daytime and Nighttime Population Density for Coastal Communities in Oregon. In: Proc. Urban and Regional Information Systems Association, Annual Conference, Proceedings, Vancouver, BC, September 26-29.
- Stasolla, M., and Gamba, P. (2008). Spatial Indexes for the Extraction of Formal and Informal Human Settlements From High-Resolution SAR Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1 (2), 98-106.
- Stathakis, D., and Kanellopoulus, I. (2008). Global Elevation Ancillary Data for Land-use Classification Using Granular Neural Networks. *Photogrammetric Engineering & Remote Sensing*, 74 (1), 1-9.
- Story, M., and Congalton, R.G. (1986). Accuracy Assessment: A User's Perspective. *Photogrammetric Engineering and Remote Sensing*, 52 (3), 397-399.

- Straatsma, M.W. and Baptist, M.J. (2008). Floodplain Roughness Parameterization using Airborne Laser Scanning and Spectral Remote Sensing. *Remote Sensing of Environment*, 112, 1062–1080.
- Strunz, G., Post, J., Zosseder, K., Wegscheider, S., Mück, M., Riedlinger, T., Mehl, H., Dech, S., Birkmann, J., Gebert, N., Harjono, H., Anwar, H.Z., Sumaryono, Khomarudin, R.M., Muhari, A. (2010). Tsunami risk assessment for coastal areas in Indonesia. *Natural Hazards and Earth System Sciences* (submitted).
- Sugimoto, T., Murakami, H., Kozuki, Y., Nishikawa, K., and Shimada, T. (2003). A Human Damage Prediction Method for Tsunami Disasters Incorporating Evacuation Activities. *Natural Hazards*, 29, 585-600. Netherlands.
- Sutton, P., Robert, D., Elvidge, C. D., and Meij, H. (1997). A Comparison of Nighttime Satellite Imagery and Population Density for the Continental United States. *Photogrammetric Engineering and Remote Sensing*, 64, 1303-1313.
- Sutton, P., Roberts, D., Elvidge, C. and Baugh, K. (2001). Census from Heaven: an Estimate of the Global Human Population using Night-time Satellite Imagery. *International Journal of Remote Sensing*, 22 (16), 3061 – 3076.
- Tanaka, N., Sasaki, Y., Mowjood, M.I.M., Jinadasa, K. B. S. N and Homchuen, S. (2007). Coastal Vegetation Structures and Their Functions in Tsunami Protection: Experience of the Recent Indian Ocean Tsunami. *Landscape Ecol. Eng.*, 3, 33–45.
- Tatem, A.J., Noor, A.M., von Hagen, C., Di Gregorio, A., Hay, S.I. (2007). High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa. *PLoS ONE*, 2(12), e1298.
- Thiel, M., Esch, T., Schenk, A. (2008). Object-oriented Detection of Urban Areas from TerraSAR-X Data. In: *Proceeding of ISPRS 2008 Congress* (37), Part B8, Commission VIII, Beijing, pp. 23-27.
- Tian, Y., Yue, T., Zhu, L., and Clinton, N. (2005). Modeling Population Density using Land Cover Data. *Ecological Modelling*, 189, 72–88.
- Tison, C., Nicolas, J.M., Tupin, F., and Maître, H. (2004). A New Statistical Model for Markovian Classification of Urban Areas in High-Resolution SAR

- Images. *IEEE Transactions on Geoscience and Remote Sensing*, 42 (10), 2046-2057.
- Tobler, W.R. (1969). Satellite Confirmation of Settlement Size Coefficients, *Area*, 1, 30-34.
- Tobler, W. R. (1979). Smooth Pycnophylactic Interpolation for Geographical Regions. *Journal of the American Statistical Association*, 74(367), 519-530.
- Van der Sande, C.J., Jong de, S.M. and Roo de, A.P.J. (2003). A Segmentation and Classification Approach of IKONOS-2 Imagery for Land Cover Mapping to Assist Flood Risk and Flood Damage Assessment. *International Journal of Applied Earth Observation and Geoinformation*, 4, 217-229.
- Vermaat, J.E. and Thampanya, U. (2006). Mangroves Mitigate Tsunami Damage: A Further Response. *Estuarine, Coastal and Shelf Science*, 69, 1-3.
- Weng, Q. (2001). A Remote Sensing–GIS Evaluation of Urban Expansion and its Impact on Surface Temperature in the Zhujiang Delta, China. *int. j. remote sensing*, 22 (10), 1999–2014.
- Willige, B.T. (2008). Tsunami Risk Site Detection in Greece Based on Remote Sensing and GIS Methods. *Science of Tsunami Hazards*, 24 (1), 35-48.
- Wilson, C. A. M. E. and Horritt, M. S. (2002). Measuring the Flow Resistance of Submerged Grass. *Hydrological Processes Hydrol. Process.* 16, 2589–2598.
- Wilson, C. A. M. E. (2007). Flow Resistance Models for Flexible Submerged Vegetation. *Journal of Hydrology*, 342, 213– 222.
- Wright, J.K. (1936). A Method of Mapping Densities of Population with Cape Code as an Example. *Geographical Review* 26.
- Wu, S., Qiu, X., and Wang, Li. (2005). Population Estimation Methods in GIS and Remote Sensing: A Review. *GIScience & Remote Sensing*, 42 (1), 80-96.
- Xiao, X., Boles, s., Froking, S., Li, C., Babu, J.Y., Salas, W., and Moore, B. (2006). Mapping Paddy Rice Agriculture in South and Southeast Asia using Multi-temporal MODIS Images. *Remote Sensing of Environment*, 100, 95 – 113.

- Xu, H. (2007). A New Index for Delineating Built-up Land Features in Satellite Imagery. *International Journal of Remote Sensing*, 29 (14), 4269-4276.
- Yuan, Y., Smith, R.M., and Limp, W.F. (1997). Remodeling Census Population with Spatial Information from Landsat TM Imagery. *Comput., Environ. and Urban Systems*, 21 (3/4), 245-258.
- Zha, Y., Gao, J., and Ni, S. (2003). Use of Normalized Difference Built-up Index in Automatically Mapping Urban Areas from TM imagery. *International Journal of Remote Sensing*, 24 (3), 583 — 594.
- Zhang, Q., Wang, J., Peng, X., Gong, P., and Shi, P. (2002). Urban Built-up Land Change Detection with Road Density and Spectral Information from Multi-temporal Landsat TM data. *International Journal of Remote Sensing*, 23 (15), 3057 — 3078.
- Zhuo, L., Ichinose, T., Zheng, J., Chen, J., Shi, P. J., Li, X. (2009). Modelling the Population Density of China at the Pixel Level Based on DMSP/OLS Non-radiance-calibrated Night-time Light Images. *International Journal of Remote Sensing*, 30 (4), 1003 — 1018.

CURRICULUM VITAE

Personal

Name	M. Rokhis Khomarudin
Date of Birth	22.07.1974
Place of Birth	Pekalongan, Indonesia
Nationality	Indonesian
Sex	Male
Degree	Master of Science
Current Address	Römerstr. 59 82205 Gilching
Email	ayah_ale@yahoo.com Rokhis.Khomarudin@dlr.de
Office Address	Jl. LAPAN 70 Pekayon Ps. Rebo Jakarta Timur 13710

Formal Educations

2007-2010	Ph.D-studies, Ludwig-Maximilians-Universität (LMU) Munich – Germany
2002-2005	Master of Science in Agroclimatology Bogor Agricultural University
1993-1998	Bachelor of Science in Agrometeorology Bogor Agricultural University
1990 - 1993	Senior High School SMAN 1 Pekalongan, Central Java, Indonesia
1987 - 1990	Junior High School SMPN Wonopringgo, Central Java, Indonesia
1981 - 1987	Elementary School SDN Kedungpatangewu, Central Java, Indonesia

Training Courses

1996	National Training on Greenhouse Gases Inventory <i>"Asean Least Cost Green House Gases Abatement Strategy (ALGAS)"</i> State Ministry for Environment- Environment Research Center Bogor Agricultural
------	---

University

1996	Training on System Analysis and Modeling Computer for Agrometeorology. Department of Geophysics and Meteorology - Bogor Agricultural University
1999	Training on Remote Sensing and GIS for Land Suitability for Agriculture. Colaboration LAPAN with BPPT.
2003	Training on Remote Sensing and GIS for Disaster Mitigation. AIT Thailand
2007	Risk and Vulnerability, Bonn, Germany
2009	Remote Sensing Techniques, Freiberg, Germany
2009	Peer Reviewed Publishing, DLR Oberpfaffenhofen, Germany

Current Appointment

1999-present	Researcher on The Indonesian National Institute of Aeronautics and Space (LAPAN)
--------------	--

Professional Organizations

1999-now	Member of Indonesian Agricultural Meteorology Society
1999-now	Member of Indonesian Remote Sensing Society
2009	Member of American Geoscience Union
2009 – now	Member of European Geoscience Union