Tropical forest mapping
using polarimetric and interferometric SAR data

A case study in Indonesia

Kemal Unggul Prakoso

Promotor: Prof.dr.ir. R.A. Feddes
Hoogleraar in de Bodemnatuurkunde, Ecohydrologie en
het Grondwaterbeheer, Wageningen Universiteit

Co-promotor: Dr.ir. D.H. Hoekman
Universitair Hoofddocent, leerstoelgroep Bodemnatuurkunde,
Ecohydrologie en het Grondwaterbeheer, Wageningen Universiteit

Samenstelling promotiecommissie:
Dr.ir. Boen M. Purnama M.Sc, Indonesian Ministry of Forestry
Prof.ir. P. Hoogeboom, Technische Universiteit Delft
Prof.dr. F.J.J. Bongers, Wageningen Universiteit
Dr. Y. Hussin, International Institute for Geo-Information Science
and Earth Observation (ITC), Enschede.

Dit onderzoek is uitgevoerd binnen de onderzoekschool WIMEK-SENSE.

Tropical forest mapping using polarimetric and interferometric SAR data

A case study in Indonesia

Kemal Unggul Prakoso

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de Rector Magnificus
van Wageningen Universiteit
Prof.dr. M.J. Kropff
in het openbaar te verdedigen
op dinsdag 3 januari 2006
des namiddags te twee uur in de Aula

ISBN: 90-8504-314-X

Kemal Unggul Prakoso
E-mail: Kemal@SarVision.co.id
Kemal_up@yahoo.com

Tropical forest mapping using polarimetric and interferometric SAR data. A case study in Indonesia / Prakoso, K. U. Doctoral thesis, Wageningen University, Wageningen, The Netherlands - with references - with summaries in English and Dutch.

Tropenbos-Kalimantan Series

The Tropenbos-Kalimantan Series presents the results of studies and research activities related to sustainable use and conservation of forest resources in Indonesia. The multi-disciplinary MOF-Tropenbos Kalimantan programme operates within the framework of the international programme of Tropenbos International. Executing agency is the Forestry Research Institute, Samarinda, governed by the Forestry Research and Development Agency (FORDA) of the Ministry of Forestry.

Ministry of Forestry of
The Republic of Indonesia

Tropenbos International

The International MOF-Tropenbos
Kalimantan Programme

Wageningen University
The Netherlands

NUFFIC, The Netherlands

The Gibbon Foundation

The Borneo Orangutan Survival Foundation (BOS)

Kemal Unggul Prakoso
E-mail: Kemal@SarVision.co.id
Kemal_up@yahoo.com

Tropical forest mapping using polarimetric and interferometric SAR data. A case study in Indonesia / Prakoso, K. U. Doctoral thesis, Wageningen University, Wageningen, The Netherlands - with references - with summaries in English and Dutch.

ISSN: 1566-6522
ISBN-10: 90-5113-081-3 (Tropenbos version)
ISBN-13: 978-90-5113-081-2 (Tropenbos version)
ISBN: 90-8504-314-X (thesis version)

Cover: AirSAR/TopSAR Image, Sungai Wain, East-Kalimantan, Indonesia study area with three different images. Total Power image (left), C-band DEM (Middle) and classified map (right).

Printed by Drukkerij Ponsen en Looijen B.V., Wageningen, the Netherlands

(**QS Al-'Alaq:1**) *Read in the Name of your Lord who created*

To my father H. Herry W. Aloewie
who could not see this work has been completed

Abstract

Prakoso, K. U., 2006. *Tropical forest mapping using polarimetric and interferometric SAR data. A case study in Indonesia*, Doctoral thesis Wageningen University, Wageningen, The Netherlands. Tropenbos-Kalimantan Series 10, 125p.

A study was made of the potential of the combined use of C-, L- and P-band polarimetric and C- and L-band interferometric airborne SAR data for tropical forest mapping. These data were collected with the NASA/JPL AirSAR during the PacRim-2 2000 campaign in Indonesia. The Sungai Wain forest reserve and its surrounding area near Balikpapan city, located in the province of East-Kalimantan, Indonesia, was chosen as study site. This site covers an area of 10 x 60 km^2, which consists of a wide variety and complex mosaic of vegetation and land cover types. Extensive and detailed ground data measurements were made.

The approach presented in this thesis includes new elements such as (1) slope correction, using InSAR, (2) mapping, using a new *reversible transform* technique and (3) Iterated Conditional Modes (ICM), using prior knowledge such as height and texture. The C-band InSAR DEM allowed for a correction of the disturbing effects of relief on the backscatter level. The new method based on a reversible transform of the covariance matrix was introduced to describe the full polarimetric information of land cover type target properties, allowing for the development of simple and robust classifiers. The ICM approach was extended using additional information such as 3-D textural information derived from the InSAR DEM. This approach was demonstrated using a substantial ground truth data set of land cover observations; the result was then validated using a large independent data set with a different legend structure. The results show that most land cover types can be accurately mapped. The land cover classification result reached 88.9% accuracy for a commonly used legend, while for the independent data set, using a 'radar' legend with more classes, the result improved to 93.8%. Validation of the results by inter-comparison provided large consistency. *The best results of land cover type classification are obtained for the C- and P-band and for the C-, L- and P-band fully polarimetric combinations, for which the additional use of relief correction and texture had no noticeable effects.* Biomass data were collected for a large number of forests transects and several non-forest plots. The linear relationships between biomass and the radar responses in C-, L- and P-band were not strong. It is evident that this result is caused by the complexity of vegetation, frequent forest fires and the effect of radar saturation at a certain biomass level.

This study has provided the first experiences with this kind of new technology and methodology in Indonesia. It may offer a substantial contribution to the developments of similar approaches for tropical forest regions in general, especially in those areas where it is difficult to obtain data from optical sensors.

Key words: polarimetry, interferometry, classification, biomass estimation, tropical forest.

Acknowledgements

While working at PT. Mapindo Parama, an Indonesian surveying company facilitated by the Ministry of Forestry (MoF) and the Association of Indonesian Concession Holders (APHI), where conducted aerial photography surveys and mapping, I gained a valuable amount of experience. The work, which consisted of the production of vegetation maps and Digital Elevation Models (DEM's) and supporting optical and radar remote sensing image interpretation and field surveys for all over the country, gave me countless opportunities to assess the capability and the prospects of airborne radar for tropical rain forest application.

In 1992, while I was working on a pilot project with the STAR-1 radar for a forest concession holder in East Kalimantan, I was introduced to Radar Remote Sensing Technology in Intera information technologies Inc, Calgary, Alberta, Canada. Ever since then, my interest in this particular field has grown more intensely by the day. Supported by Mr. Budiono Kartohadiprodjo and Mr. Herman Hidayat, my work responsibilities enabled me to develop and learn more of the technical aspects and data processing of radar data. This work involved the use of several X- and C-band interferometric airborne radar systems such as STAR-3i, Do-SAR and the AeroSensing SAR.

Dr. D.H. Hoekman, the co-ordinator of the 1996 ESA-MoF Indonesian Radar Experiment (INDREX) campaign, kindly invited me at that time to act as a researcher to assess the operational implementation of Do-SAR high resolution SAR interferometry for forest assessment, monitoring and management. The campaign was executed under the auspices of the Indonesian Ministry of Forestry (MoF) and the European Space Agency (ESA). A special fund for stimulating and continuing the use of radar application in Indonesia came from the Tropenbos Foundation, which enabled me to participate in the research program: 'Remote sensing monitoring system for sustainable forest management and land cover in Indonesia' in the Netherlands. After joining Wageningen University in the Netherlands in 1999, where I enrolled in the M.Sc study at the Geo-information Science Department, I continued my research, focusing on the use of radar for forests, and wrote a M.Sc thesis on interferometric SAR tree mapping in tropical rain forest.

The results of the INDREX campaign led to the possibility to continue the research with the NASA PacRim-2 campaign data. This research was carried out as a doctoral study at the WIMEK-SENSE Research School of Wageningen University for a nominal period of four years, starting in June 2000. The radar data studied in this work were made available by NASA.

Through this work, I would sincerely like to thank these organisations for including my study area in the flight plans for the PacRim-2 radar campaign. I also want to gratefully acknowledge Tropenbos, the Gibbon Foundation and MoF for providing the funding, clearance and logistic field support. Without their kind and very much appreciated co-operation, this study would not have succeeded.

With this opportunity, I would like to express my gratitude to all who have contributed to this work.

First of all I would like to direct my first and special acknowledgements to my promotor, Prof.dr. R.A. Feddes and to Dr. Dirk H. Hoekman, my co-promotor. I very sincerely thank these gentlemen for their inspiring guidance, their valuable support during the whole research period, and the faith they have shown in me. I whole-heartedly appreciate their intellectual thoughts, and their sagacious way of putting these valuable thoughts in simple and straightforward arguments and suggestions.

To the Nuffic Huygens Fellowship Program which supported my work and to Tropenbos, The Netherlands for the publication of this thesis in its series, I offer my heartfelt thanks.

I also wish to thankfully acknowledge my colleagues, particularly the Radar Remote Sensing Working Group at the Department of Water Resources of Wageningen University, who provided me with a productive and pleasant working environment. I am particularly grateful to ir. Martin A.M. Vissers for his valuable assistance and support in software development and technical support. The help of other Indonesian colleagues is also greatly appreciated.

I would also like to express my warmest gratitude to Dr. Willie Smits and the Gibbon Foundation for giving the necessary support for study and my family. Finally, I ask my friends at the Indonesian Student Association Wageningen (PPI - *Perhimpunan Pelajar Indonesia* Wageningen) who so kindly encouraged me and frequently provided me with joyful diversion during my studies, to please accept my wholehearted thanks for their part in the successful completion of the present work.

Kemal Unggul Prakoso.

Contents

List of main symbols

Symbol	Unit	Description
C	-	covariance matrix
$C(..)$	-	circular Gaussian distribution
$\|C_i\|$	-	determinant of C_i
D	-	standardized threshold distance
dbh	cm	diameter at breast height (1.3 m)
H	m	height of tree
$li_{i,c}$	-	likelihood of a pixel i belonging to class c
$\boldsymbol{M}$	-	Stokes scattering operator
mli_{ic}	-	ICM modified likelihood of a pixel i belonging to class c
N	-	number of independent radar looks
$N(..)$	-	normal or Gaussian distribution
n	-	number of channels
$P_\gamma(..)$	-	marginal distribution for intensity
$P_\phi(..)$	-	marginal distribution for polarimetric (HH-VV) phase difference
$P_{\|\rho\|}(..)$	-	marginal distribution for coherence magnitude
P_c	-	relative occurrence of class c
$R_{i,c}$	-	relief factor for pixel i and class c
r^2	-	coefficient of determination
$\boldsymbol{S}$	-	scattering matrix
S_{hh}	-	scattering matrix element for HH polarisation
T_i	-	threshold value
TP	-	total power
$T_{i,c}$	-	texture factor for pixel i and class c
T_{mc}	-	mean of the textural coefficient of variation for class c
T_{vc}	-	variance of the textural coefficient of variation for class c
$U_{i,c}$	-	number of neighbouring pixels of pixel i having class c
$\hat{Y}$	-	relationship total above ground biomass estimation
α	^{o}	angle of slope
γ	$m^2\ m^{-2}$	differential radar cross section per unit projected area
γ_i	$m^2\ m^{-2}$	differential radar cross section per unit projected area for field i
γ_f	$m^2\ m^{-2}$	differential radar cross section per unit slope area
$\Delta\hat{K}$	-	significant difference between two classification results
θ_{gr}	^{o}	grazing angle
θ_i	^{o}	incidence angle
$\hat{K}$	-	Kappa statistic
λ	cm	wavelength
μ	-	average value of texture parameter
μ_i	-	average value of texture parameter for class i
$\|\rho\|$	-	polarimetric coherence magnitude or polarimetric correlation
$\|\rho_i\|$	-	polarimetric coherence magnitude for field i
σ	dB	radar cross section
σ^0	$m^2\ m^{-2}$	differential radar cross section
ϕ	^{o}	polarimetric (HH-VV) phase difference
ϕ_i	^{o}	polarimetric (HH-VV) phase difference for field i
χ	^{o}	polarimetric ellipticity angle
Ψ	^{o}	polarimetric orientation angle

1. Introduction

1.1 Introduction

The present awareness and concern about global environmental changes such as high rates of deforestation has prompted governments and international organisations to develop effective measures for a better management of our environment. Deforestation rates are especially high in the vast tropical rain forest areas. At the same time these areas are acknowledged as being very important in the global hydrological and biochemical cycles. They contain important genetic reservoirs, and are potential permanent sources of production for the benefit of an immense number of people. For these reasons, the government of Indonesia has decided to take strong measures in support of sustainable forest management.

For management of these areas, the availability of up-to-date information is imperative. An almost continuous cloud cover is one of the main obstacles for timely acquisition of information, while the pronounced topography of the terrain and the vastness of the area, fragmented over many islands, create additional problems. Approximately 75% of the Indonesian land surface is covered by forest. Of this, 65 million ha is designated as production forests where only selective logging is allowed. These production forests are used by some concession holders to produce timber. Another 49 million ha is designated as protected forests where no deforestation is allowed. These protected forests consist of areas selected for nature conservation, and forest areas on steep slopes selected to protect watersheds. The remaining 30 million ha, so-called conversion forests, are areas in the process of change into other types of land use, such as forest plantations or agricultural areas (Sumitro, 1993).

The government, in its varied duties such as supporting planning and policy-making, enforcing legislation and supporting land management, requires reliable and up-to-date information. Hence, data are required to verify proper execution of rules developed for sustainable management and to evaluate the (long term) effects of these measures. Current remote sensing techniques for acquiring information, such as aerial, photography and optical satellite data (i.e. from LANDSAT and SPOT) have been used widely for forest survey and monitoring, but image acquisition is severely affected by the persistent cloud cover (Gastellu-Etchegorry, 1988). Advanced airborne and space-borne radar techniques are, however, very promising in this respect.

To acquire data on forest and land cover characteristics, modern radar remote sensing techniques such as interferometric and polarimetric synthetic aperture radar are of increasing importance. Considering current information requirements in Indonesia, it would even seem impossible to work without these tools. Many different types of radar sensors exist and, depending on the observation platform used, each specific radar sensor potentially covers a fraction of the necessary information.

The Indonesian Ministry of Forestry (MoF) considers remote sensing, especially radar remote sensing, as the best practical approach to collect and/or verify such information in such large quantities. It would, for example, be very useful if information on the land cover changes, degree of crown cover opening, biomass estimation and timber road construction could be acquired periodically. These considerations motivated MoF, in co-operation with the Tropenbos Foundation and

Wageningen University, to initiate a research project to explore the applicability of new mapping and biomass estimation tools.

1.1.1 Review on the use of radar in land cover classification and biomass estimation

Recent reviews on the application of radar in forestry show that SAR systems have a good capability in discriminating various types of (tropical) forest cover (Van der Sanden, 1997; Varekamp, 2001; Quiñones, 2002; Sgrenzaroli, 2004). In general classification results are poor if only single-frequency, single-polarisation or single-overpass data are used. In order to improve the radar classification capability, either multi-temporal or multi-frequency data are required. Multi-temporal data, which may be acquired by airborne or satellite systems, are particularly important to separate forest types (Ferrazzoli *et al.*, 1999). Multi-frequency or multi-polarisation data are made available by several types of airborne SAR systems, such as the NASA/JPL AirSAR. These are capable of acquiring interesting classification results, due to characteristics of multi-frequency interferometric and polarimetric features, which can be associated to vegetation structures.

A systematic study on the use of ERS radar data for implementing a land cover change monitoring system at the Guaviare site in Colombia shows that linking land cover change models with multi-temporal ERS SAR observations data can be used to detect changes on vegetation cover types and, to some extent, mapping land cover types. Classification accuracies in the order of 60-70% were obtained for forest, secondary vegetation, pastures and natural grasslands (Bijker, 1997). The study of Van der Sanden (1997) in the tropical forest of Guyana indicated that the combination of high and low frequency radar systems could be used for land cover mapping in a tropical environment. Another important finding was that the AirSAR system with its high resolution can be used in forest management at the local scale and is complementary to satellite radar systems. Hoekman and Quiñones (2000) at the Guaviare test site in the Colombian Amazon, studied the capabilities of AirSAR to discriminate four land cover classes and biomass estimation using a classification technique that involves radiometric and polarimetric information. The results indicate the level of accuracy increases when using a multi-frequency combination, and also clearly indicate how SAR systems can be designed to accurately monitor processes of deforestation, land and forest degradation, and secondary forest re-growth. Another study was executed at the Araracuara test site in the Colombian Amazon, using the same technique but extended with the Iterated Conditional Modes (ICM) approach (Hoekman and Quiñones, 2002) to assess AirSAR's potential for forest structural mapping and forest biophysical characteristics. The accuracy of the final product was high, showing many forest structural types and new spatial details. The latter were not visible on previously made maps on the basis of aerial photography.

In a study over a 5-years period in East-Kalimantan, Indonesia, Kuntz and Siegert (1999) using 13 scenes ERS-1 and -2 SAR, showed differences in vegetation and land use classes. Different images processing techniques were investigated for the monitoring of land use and deforestation, using time series of ERS-1 and -2 SAR images. The results suggested that such time series can be used to monitor and analyse forest conversion in tropical rain forest. This study, which was conducted along the Mahakam river, described how several land cover types such as undisturbed

Dipterocarp forests, heath forests, secondary forests, clear-cuts/shifting cultivation and selective logging can be mapped successfully using texture analysis and time series of observations. A related study (Siegert and Hofmann, 2000) used ERS-2 SAR as a complementary source in addition to NOAA-AVHRR, to evaluate the area affected by fire, and detected two classes of fire damage (severe and medium damage) based on ERS SAR composites.

The use of interferometric data for land cover classification is relatively new. Temporal de-correlation properties were presented by Wegmuller and Werner (1995) using ERS-1 repeat-pass data. By examining a variety of sites, they found that urban areas, agricultural areas, bushes and forests had different correlation characteristics, with urban areas showing the highest correlation and forests the lowest (water shows no correlation). Based on this work Wegmuller and Werner (1997) presented a formal classification scheme based on interferometric correlation (coherence), the backscatter intensity, the backscatter intensity change, and a texture parameter. Classification results for a test site near the city of Bern, Switzerland, yielded an overall classification accuracy figure in the order of 90%.

For biomass measurement, many studies have been conducted to investigate the relation between forest biomass and radar backscatter. Both empirical and theoretical methods have been used. Radar backscatter is found to increase uniformly with increasing biomass until the relationship saturates (i.e., flattens) at a biomass level depending on the radar wavelength. The biomass saturation level is lower at shorter wavelengths than at longer wavelengths, and the longer wavelengths have a better relationship to biomass and other stand parameters than short wavelengths. Le Toan *et al*., (1992) used multi-polarisation L- and P-band airborne radar data, and found that the dynamic range of the radar backscatter corresponded highly with forest growth stages and is maximum at P-band HV polarisation. The analysis of P-band data indicated a good correlation between the radar backscatter intensity and the main forest parameters including trunk biomass, height, age, diameter at breast height (*dbh*), and basal area. Dobson *et al*. (1992) showed an increasing range of backscatter with changing biomass from C- to L- to P-band, as well as higher biomass levels at which backscatter relationships to biomass saturate. Hoekman (1990) found poor relationships between X- and C-band backscatter and volume and other stand parameters.

The use of P-band SAR has been proposed for its supposedly superior capabilities for biomass assessment and deforestation mapping. Most study results to date relate to temperate forests and indicate that the P-band, notably the HV-polarisation, is capable to estimate biomass levels up to 150-200 ton ha^{-1} (Le Toan *et al*., 1992). Cross-polarisation data have been found to be better than like-polarized for estimating biomass (Le Toan *et al*., 1992; Dobson *et al*., 1992; Hussin *et al*., 1992) because cross-polarized backscatter is less sensitive to ground conditions (such as local slopes, roughness, or soil moisture).

By combining P-band data with L- and/or C-band data even higher biomass levels may be reached for certain forest types (Ranson and Sun, 1994; Kasischke *et al*., 1995; Quiñones and Hoekman, 2004).

For tropical forests, far fewer studies have been conducted. Imhoff (1995) studied broadleaf evergreen forests in Hawaii and temperate coniferous forests and indicates saturation levels of biomass, for both types of forests, of 100 ton ha^{-1} in P-band and 40 ton ha^{-1} in L-band. Rignot *et al*., (1995) conducted a study in the Amazon forest of Peru and showed P-band's capability to differentiate biomass classes in excess of 200 ton ha^{-1}. Hoekman *et al*., (1996) in Guaviare, Colombia quantitatively compared AirSAR C-, L- and P-band backscatter measurements with on site biomass measurements. Good correlations were found between the backscatter in L- and P-band and the above ground fresh biomass up to a level of approximately 200 ton ha^{-1}. Hoekman and Quiñones (2000) clearly show that L-HV saturates earlier than P-HV in the Colombian Amazon forest, at possibly 30 ton ha^{-1} for L-band and 150 ton ha^{-1} for P-band. However, the lack of data in the range of 30-150 ton ha^{-1} made it difficult to define the biomass saturation value. This lack of data is rather common in reported data sets when the experiments were conducted in natural forests, often not covering a good range of biomass, unlike the regularly exploited plantation forest.

Biomass measurements using VHF indicate that the backscatter is sensitive to biomass up to very high values, i.e. 700-900 m^3 ha^{-1} (Melon *et al*., 2001). Simulations using the Distorted Born Approximation confirm the observations. The result shows that the trunk-ground interaction is the main mechanism, and, also, that the effect of topography was found to be important and necessary to be accounted for. Procedures to estimate biomass from radar observations over hilly terrain should also consider the effect of topography on backscatter level and microwave interaction (Van Zyl, 1993).

One important conclusion of this review is that most of these studies have used radiometric, interferometric correlation (coherence) and fully polarimetric information separately, and did not combine interferometric and full polarimetric information. Most studies took place in the temperate zone and some in the Amazon region. For the South East Asian tropical rain forests, like in Indonesia, the number of studies executed has been very limited. The complex structure of vegetation and hilly terrain complicates mapping and biomass estimation, which may only be solved using a combination of multi-frequency and multi-mode data.

1.1.2 Forest mapping using radar in Indonesia

Forest mapping in Indonesia using radar started in 1992, when PT. Mapindo Parama and the Ministry of Forestry conducted a pilot study on the application of SAR for forest inventory and management in tropical rain forest areas, especially over mountainous or generally hilly areas. An area of almost 6 million hectares in East and Central Kalimantan was covered by the STAR-1 X-band HH polarisation airborne SAR with a resolution 6 by 6 m and a 23 km swath width (Hidayat, 1993). SAR data collection was carried out in double side-looking mode with a nominal 70% overlap between adjacent swaths and two fixed look directions (east and west). The SAR data were used to generate topographic contours, derived through a stereo radargrammetry technique (Leberl, 1990) and forest cover type information. In addition, helicopter airborne photography was taken to provide the density and crown diameter information. Airborne laser profile measurements provided tree height information and were used to closely register the topographic contours. Several problems with visual interpretation of the image occurred because of the characteristic radar shadow

problem, radar lay-over and radar foreshortening. The results had a high correlation of 80-90% for forest type and a poor correlation of 40-50% for timber volume, and were integrated into a Geographic Information System (GIS) with a scale of 1:25,000. Recommendations were made and used to optimise the operational use of the technologies demonstrated (Thompson *et al*., 1993).

Several areas in Sumatra and Kalimantan were covered by RADARSAT in February 1996, making Indonesia the first client of RADARSAT, which was launched in November 1995. Some 200 scenes of the finest resolution of 10 by 10 m was collected and interpreted. Those images were intended for forest monitoring to identify logging and skidding roads. However, the results were not adequate for such a purpose (Hidayat, 1998).

The INDonesian Radar Experiment (INDREX-96) was an airborne radar campaign carried out by the European Space Agency (ESA) within a larger project co-ordinated by the Indonesian Ministry of Forestry and Wageningen University, to develop a "Remote Sensing Monitoring System for Forest Management and Land Cover Change in Indonesia". The aim was to provide an experimental basis (1) to assess how high-resolution radar images could replace aerial photography for detailed forest monitoring purposes, and (2) to determine optimum SAR system configurations for future space-borne instruments. Attention was focussed on two representative test sites in Central Sumatra and Eastern Kalimantan. In July 1996, the Dornier Do-SAR system (Faller and Meier, 1995) was used to collect high-resolution C-band and X-band data, including polarimetry and interferometry. A large number of ERS SAR images have also been studied within the INDREX framework, to develop and further support ERS satellite data interpretation for tropical forest monitoring. INDREX has provided an important set of airborne and satellite SAR data suitable for addressing many important issues concerning the use of radar data for forest management and land cover change in Indonesia and similar tropical countries.

Comparisons made between aerial photography and high-resolution SAR images have demonstrated clearly how (non-interferometric) 1.5 m resolution images can provide very similar information for detailed forest mapping purposes (Wooding *et al*., 1998). Many of the larger trees, with crown diameters of around 10 m can be seen at this resolution but, in comparison with what that can be seen on a 1:20,000 scale aerial photography, most of the smaller trees are not distinguishable. Only the very largest trees can be seen as separate features at 3 m resolution. Analysis of ERS data has shown that both multi-temporal and interferometric processing techniques can be used to produce forest maps and forest fire damage maps (Grim *et al*., 2000). This effort is being continued to produce more examples and to compare fully the results obtained using the two different techniques. These studies recommend the use of a radar satellite monitoring system to collect up-to-date nation-wide information and, once this system is in place, the use of high resolution airborne InSAR for tree mapping (Hoekman and Varekamp, 2001; Varekamp and Hoekman, 2001) in selected areas of interest. Identification of the latter areas is guided by information obtained by the satellite products, thus making the airborne system operation much more effective.

At the same time period several airborne radar systems with a resolution of 3 by 3 m in interferometric mode have been used to cover areas of protected forest throughout Indonesia. Form line maps and forest cover maps have been derived from the

interferometric images. Three systems which were employed, all in May 1997: (1) The Do-SAR (C-band VV) of Germany which covered areas in Kalimantan and Sumatera (Faller, 1998); (2) The Intermap STAR-3i (X-band VV) of Canada/USA which covered areas in Java and Irian Jaya and (3) The AeroSensing Aes-1 (X-band VV) of Germany which covered areas in Maluku and Irian Jaya. They produced orthorectified images and DEMs, orthorectified mosaics and thematic maps, such as forest cover and form line maps, at the scale 1: 25,000 (Hidayat, 1997).

1.1.3 A new campaign with AirSAR in Indonesia

Successful execution of the ESA INDREX'96 airborne radar campaign was followed by NASA's second 'Pacific Rim' campaign (PacRim-2) executed in South East Asia and Australia in the July-October 2000 period (PacRim-2, 2000). The nature of the sites varied widely, from volcanic sites to rain forests. NASA/JPL funded and carried out the campaign to facilitate fundamental research, among others into the role of soils, vegetation and water in the global carbon cycle.

For Indonesia this was the first campaign with NASA using the AirSAR/TopSAR system, for the AIRSAR Pacific Rim campaign (PacRim-96), which was planned to be carried out under the auspices of NASA in the period November-December 1996, could not be executed in Indonesia due to the difficulty of obtaining permission to process data abroad. This campaign was the first opportunity to evaluate multi-frequency polarimetric and interferometric SAR data over tropical forest test sites and was intended to familiarize researchers with polarimetric and interferometric radar data and to prepare them for the analysis of data. Together with the Indonesian Ministry of Forestry (MoF), Wageningen University (WU) participated in the campaign, in order to conduct research at three tropical forest test sites in East-Kalimantan. C-, L- and P-band fully polarimetric data, as well as C- and L-band interferometric data, were successfully acquired over a total area of 3,700 km^2 in September 2000. This study had two main objectives; the first was to familiarize Indonesian scientists with these advanced SAR techniques, which may play an important role in future monitoring and mapping systems in support of sustainable forest management. The second objective was to develop knowledge and applications, not only in forestry but also in agriculture, coastal zone management, mapping, etc.

The primary research test site was the Sungai Wain area, a 10 x 60 km^2 strip, which comprises the Sungai Wain forest reserve having a wide variety of forests including Dipterocarp forest and several types of swamp forests. The 1998 forest fires induced by El Niño affected a small part of this area. Adjacent areas cover two large stretches of mangrove forests along both sides of the Balikpapan Bay, rice fields, fish-ponds, beaches, the Bukit Bangkirai forest reserve and the Wanariset research station.

The C-, L- and P-band polarimetric (HH, HV, VH and VV polarisation combinations) and interferometric (XTI2) data of the TopSAR/AirSAR data will be used for the investigation in this thesis.

1.2 Objectives

Currently aerial photography is used on a routine basis in Indonesia for monitoring and management of tropical forest areas, and for the production of topographic and thematic maps (Hidayat, 1993). Unfortunately, this effort fails to provide the necessary information in time, because of the cloud cover. Interferometric Synthetic Aperture Radar (InSAR) and Polarimetric Synthetic Aperture Radar (PolSAR) high-resolution data has been acknowledged to be a viable alternative to generate appropriate maps. As a tool for collecting information on tropical rain forests, the data has the potential to facilitate the government in establishing appropriate policies and determining the use of forests as natural resources. However, there is little experience with the use of radar data for this field of application at the moment. Recent advances in interferometric and polarimetric radar technology have stimulated research towards development of robust methods to generate maps providing reliable and up to date information.

The main objective of this study is to evaluate the use of the new generation C- and L-band interferometric (TopSAR) and C-, L- and P-band fully polarimetric (PolSAR) NASA/JPL AirSAR data for tropical forest type mapping and biomass estimation and to study the combined use of radar data and additional knowledge in order to improve the results further.

Polarimetric techniques developed for PolSAR mapping (Hoekman and Quiñones, 2000; Hoekman and Quiñones 2002) using data collected in Colombia during the AirSAR 1993 deployment will be tested on Indonesian PacRim-2 PolSAR data. It is assumed that the same method can be applied with some modifications, and that it can be extended utilising the higher spatial resolution and the additional information from the TopSAR mode. A variety of polarisation combinations will be investigated for forest mapping and biomass estimation application. Classification results can be simulated as a function of the number of independent looks. The fully polarimetric multi-band approach used for classification uses probability density functions (*pdf*) for multi-look samples of a certain class, for intensity, phase difference as well as coherence magnitude (Hoekman and Quiñones, 2000; Hoekman and Quiñones 2002). An alternative method for polarimetric classification, the so-called *reversible transform,* recently introduced and tested for an agricultural area, will also be considered (Hoekman and Vissers, 2003). Particular attention will be given to physical understanding of the radar data and validation of the derived information products. Therefore considerable field survey efforts have been made in the framework of this thesis research.

The availability of these multi-frequency polarimetric SAR data enables quantitative comparison of classification capabilities of all polarisation combinations for three frequencies. Furthermore, ground truth maps are available to facilitate the selection of a training set and a reference map. The results of this study will facilitate the selection of optimum sensor parameters in terms of polarisation, frequency and incidence angle for each application under consideration and thus will contribute to an appropriate definition of an advanced SAR system for land observation, with particular emphasis on forestry application in Indonesia.

The selection of radar frequency and polarisation are two of the most important parameters in synthetic aperture radar (SAR) mission design. Nevertheless, a multi-frequency fully polarimetric SAR system is highly desirable. Limitations posed by allowable satellite payload, data rate, financial budget, required resolution, area of coverage, etc. still prevent multi-frequency fully polarimetric SAR from becoming a reality. For a particular application, it is desirable to select the optimum frequency and combination of linear polarisation channels in case a fully polarimetric SAR system is not possible, and to assess the expected loss in classification and geophysical parameter estimation accuracy.

The methodology introduced in this thesis should have an impact in selecting the combinations of polarisations and frequency of a SAR in various applications. For example, the future TerraSAR L-band is expected to become fully polarimetric and would give a boost to accurate mapping and biomass estimation. The combination with advanced C-band systems such as RADARSAT-2 and ENVISAT ASAR would even increase this capability.

1.3 Structure of the thesis

Both theory and experimental studies are covered. This thesis includes an introduction covering some basic principles of interferometry and polarimetry and its use for tropical forests and biomass estimation mapping. A brief overview on radar research conducted in Indonesian forests is also included.

Chapter 2 describes the methods and techniques used to obtain, process and analyze information from available interferometric and fully polarimetric multi-band AirSAR data. A statistical description of fully polarimetric data in terms of backscatter (γ), polarimetric phase difference (ϕ), polarimetric correlation (ρ) and the effect of speckle is described in order to evaluate land cover types classification and biomass estimation. A method based on Iterated Conditional Modes (ICM) (Besag, 1986) is introduced, and, in addition, texture information derived from InSAR data, as *a priori* information, is elucidated. The Maximum Likelihood (ML) approach and Kappa statistics are used for classification and to compare the results.

The selected study area, field data measurements and radar data from the NASA/JPL AirSAR are the subjects of Chapter 3. A general overview of the study area near Balikpapan, East-Kalimantan, Indonesia, is followed by the procedure used to collect extensive ground reference data and detailed intensive transect measurement descriptions of the forest biomass studied. The last two sections explain the radar campaign that provides the data analysed in the present study, and lists the specifications of sensor systems deployed.

The value of this information for the indication of land cover types and biomass estimation is analysed. Validation related to the algorithm and model trial is described; limitations and advantages are also reviewed. The confusion matrix and statistical approach (Kappa analysis) will be used to evaluate the accuracy of the technique for parameter estimation modelling and forest thematic mapping. Effects on the accuracy after the application of the ICM method are presented.

The experimental results are discussed in Chapter 4. Classification approaches are compared, including (1) the recently introduced polarimetric classification method based on the *reversible transform* algorithm (Hoekman and Vissers, 2003), (2) the correction for relief, and (3) the utility of a new approach for *3-D texture* information derived from InSAR data. The relative importance of different frequency and polarisation combinations and the effect of speckle are also addressed. A direct empirical biomass estimation relationship is presented. In addition, polarimetric coherence analyses of several cover types are made.

Chapter 5 presents a validation of the results with an independent data set. A synthesis of all results is made and the relative importance of the classification approaches and radar parameters is assessed.

Summary and conclusions presents the study background, study area, summary of each chapter, and key conclusions and recommendations of the use of interferometric and polarimetric radar for tropical forests classification and biomass estimation mapping, in the perspective of the research objective.

2. Land cover classification and biomass estimation

This chapter describes the methods and techniques used to obtain, process and analyze information from fully polarimetric and interferometric multi-band AirSAR data. A statistical description of fully polarimetric data in terms of backscatter, polarimetric phase difference, polarimetric correlation and the effect of speckle will be given in order to evaluate land cover type classification and biomass estimation. A new method based on a *reversible transform* of covariance matrix elements will be introduced to describe the full polarimetric target properties. The technique of Iterated Conditional Modes (ICM) as well as a new technique to use *3-D textural features* derived from InSAR data as additional information will be described. Maximum Likelihood and Kappa statistics will be used for classification and comparison of results.

2.1 Introduction

The usefulness of a Digital Elevation Model (DEM) derived from interferometric SAR (InSAR), certain combinations of frequency bands and polarisation, polarimetry and the effect of speckle will be evaluated for tropical forest mapping applications and for biomass estimation. This will be done in several steps. Methods and models used for TopSAR/AirSAR data sets will be described first. The main focus is to improve classification accuracy of land cover and biomass estimation using fully polarimetric SAR (PolSAR) data, with the DEM derived from simultaneously acquired InSAR data taken into consideration. The processing sequence of land cover classification and biomass estimation is illustrated schematically in Figure 2.1.

Through orthorectification of the C-, L- and P-band fully polarimetric images, an accurate geometrical registration to the C-band VV-polarisation interferometric image was achieved. This fusion allows correction of the disturbing effects of relief on the backscatter level in the applied fully polarimetric classification procedure (paragraph 2.2). To retrieve the information from corrected radar images, a database was created for the areas visited during the fieldwork stage. The set of Regions of Interest (ROI's) should contain homogeneous areas. These data give information about the physical condition of the land cover and are used to assess and support the analysis of radar data (paragraph 2.3). Basic statistical descriptions of data will be explained (paragraph 2.3.1). A database of plot averaged Stokes scattering operator matrix elements was created, which forms the basis for the analysis. For all locations of field observation ROI's are delineated in the corrected PolSAR image (paragraph 2.3.2). The evaluation of polarimetric classification performance will be done for two techniques (paragraph 2.4). The first is the polarimetric classification technique introduced in Hoekman and Quiñones (2000), used for AirSAR land cover type mapping and forest biophysical characterization. It is assumed that the same method can be applied with some modification (paragraph 2.4.1). In this case the maximum likelihood (ML) classification of an observation to be classified as a certain class is the product of the joint Gaussian distributions of the backscatter multiplied by the likelihood of the phase difference values and the likelihood of the correlation values (paragraph 2.5). The second technique is based on a new *reversible transform* (Hoekman and Vissers, 2003) of covariance matrix elements into backscatter intensities, which will describe the full polarimetric target properties (paragraph 2.4.2). In this case the maximum

likelihood (ML) classification simply is based on multi-variate Gaussian distributions. Biomass and vegetation height data were collected for a large number of forest transects and several non-forest plots. Empirical relations of C-, L- and P-band backscatter and C- and L-band interferometric coherence and height, will be presented. A method based on Iterated Conditional Modes (ICM) (Besag, 1986) will be used to yield radar-derived maps with a high level of agreement with existing maps, as well as with ground observations (paragraph 2.6). *A new 3-D texture analysis derived from InSAR will be introduced as additional information* for feeding the ICM process (paragraph 2.7). The confusion matrix, the overall classification accuracy and the confusion between pairs of classes can be calculated for each classification or simulated classification. The evaluation of the classification capacities will be done through analysis of simulated classification result, using the contingency tables and Kappa statistics. This is done for simulated classifications performed for a single-channel, combinations of two and three channels and full polarimetric information. It is important to notice that a contingency table will give a measure of the overall classification accuracy and present errors of commission and omission while the Kappa statistic will be used to compare the classification results (paragraph 2.8).

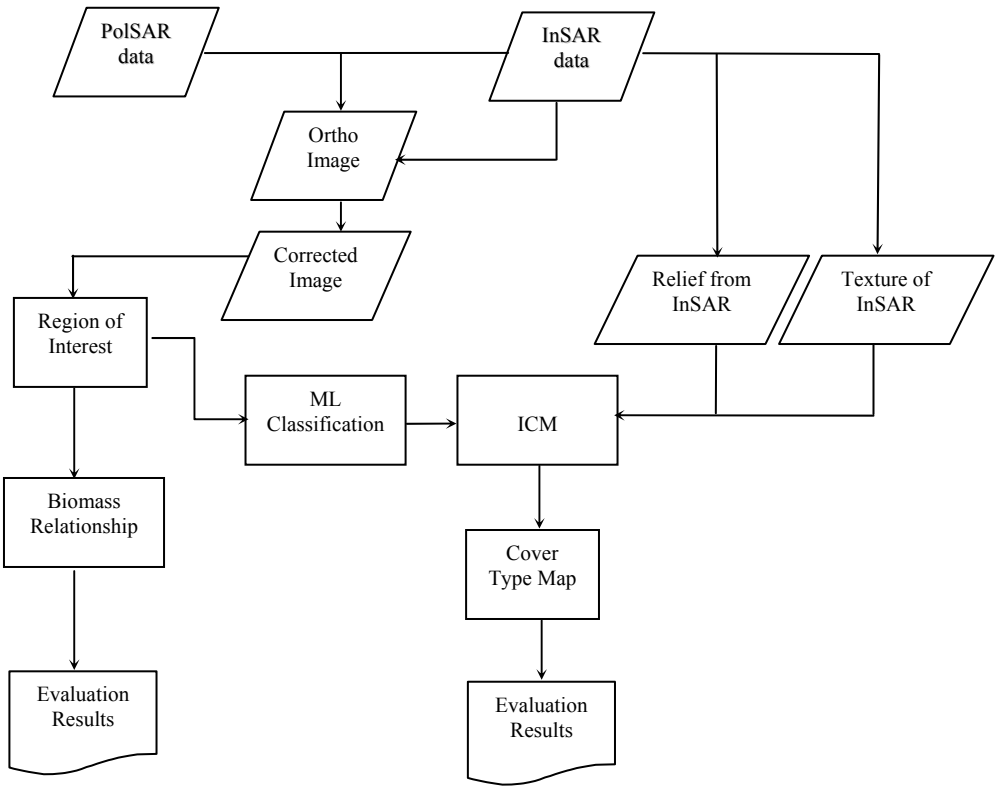


Figure 2.1 Flow diagram of the land cover classification and biomass estimation processing sequence.

2.2 Fusion of InSAR and PolSAR data

2.2.1 The effect of relief and canopy surface geometry

It is well known that airborne Synthetic Aperture Radar (SAR) imaging is based on the principle of range differences between the objects and the sensor for a certain scan line. Because of the imaging mechanism of radar signals, small relief differences can be perceived well, notably at small grazing angles. For an opaque isotropic volume scatterer, the differential radar cross-section, $\gamma = \sigma^o / \cos(\theta_i)$ does not depend on grazing angle, but will depend on the slope of the vegetation surface (Figure 2.2a). Always, and therefore also for the three cases shown in this figure, the ratio between intercepted power and the re-radiated power is the same for every resolution cell. Processing algorithms (which are based on the geometric optics approximation) to compute γ, however, start from the assumption that the terrain is flat, and hence assume the intercepted power is proportional to $\tan(\theta_{gr})$. In fact it is proportional to $\tan(\theta_{gr} + \alpha)$, where α is the angle of slope in range direction which can be derived from InSAR data. The value of γ in the processed image is therefore related to the value of γ_f, for an identical object with the upper surface oriented parallel to the horizontal plane, as the following:

$$\gamma = \gamma_f \frac{\tan(\theta_{gr} + \alpha)}{\tan(\theta_{gr})} \tag{2.1}$$

Figure 2.2b shows this relation in graphical form. The ratio γ / γ_f (or σ^o / σ_f^o) is shown at the dB scale as function of grazing angle and for several slope angles. It can be concluded that, if this mechanism applies, small slopes observed at very small or very large grazing angles have strong effects on the backscatter level. Furthermore, it can be shown that the effects of canopy surface undulations average out for this model in the sense that the (linear) average of γ, for any area within a perimeter located at a horizontal plane and not showing radar shadow, is independent of the degree and location of slopes (Hoekman, 1990).

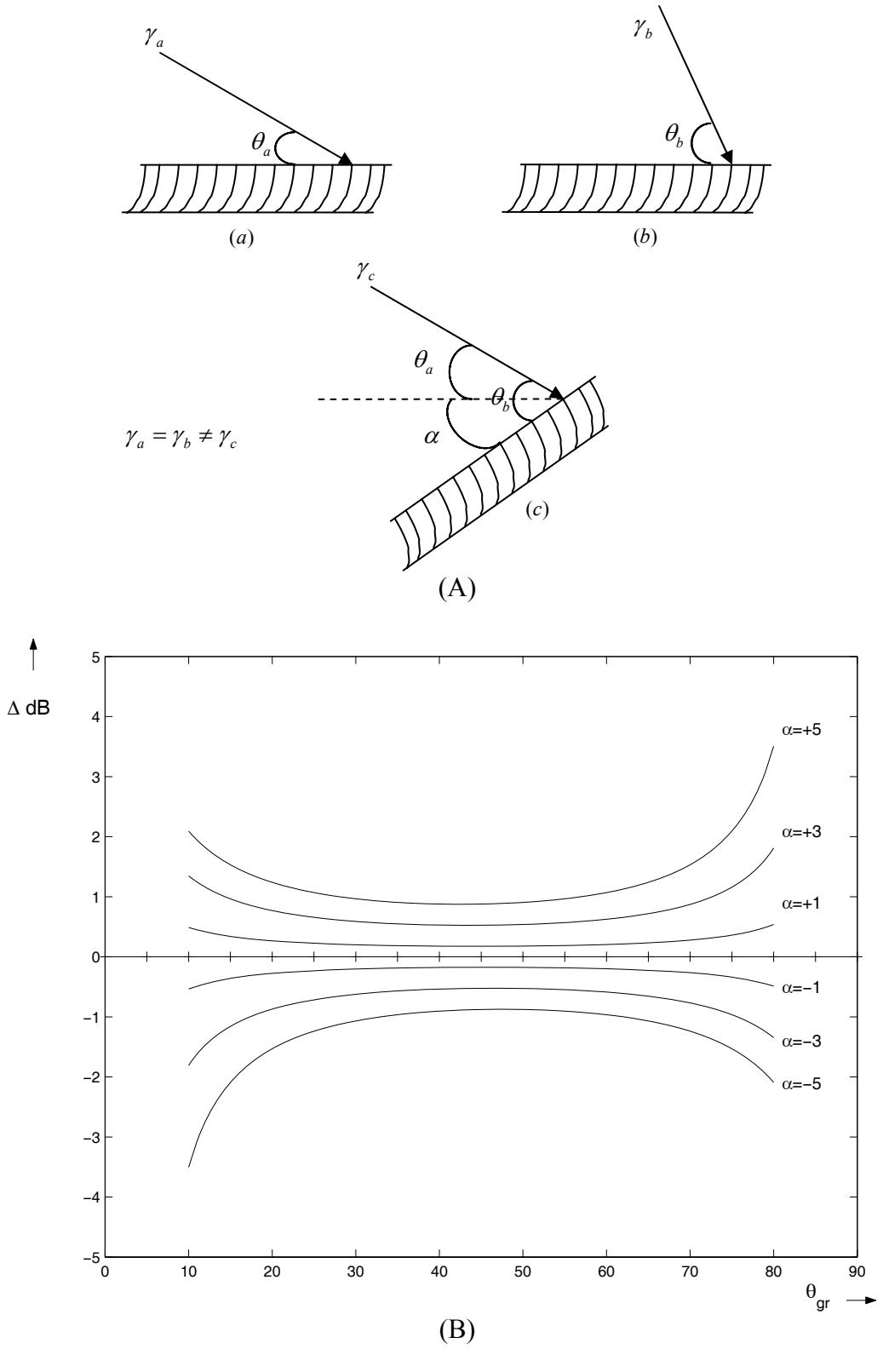


Figure 2.2. A) For an opaque isotropic volume scatterer, γ (the differential radar cross-section, $\gamma = \sigma^o / \cos(\theta_i)$) is independent of grazing angle θ_{gr} (case a and b) and dependent on slope α (case c). B) Effect of canopy undulations if an opaque isotropic volume scatter mechanism applies. Changes in γ level as a function of grazing angle θ_{gr} and angle of slope α are shown (Hoekman, 1990).

This theoretical relationship has been confirmed by experimental observation using ERS-1 SAR and SAREX airborne campaign images of the geomorphology of Guyana's Mabura Hill and Iwokrama areas, which is well perceivable, comprising mountains, plateaux, ridges and dolerite dikes. The actual appearance of these features strongly depends on the radar incidence angle. Since the terrain is completely and densely forested, the theoretical relationship between backscatter modulation and relief as described above is applicable (Hoekman *et al*., 1994). The validation of this relationship in an area in Iwokrama called 'Turtle Mountains' is shown in Figure 2.3. This area was imaged by the ERS-1 (23^{o}), in SAREX track2.2 (66^{o}) and track2.1 (81^{o}). Two large and uniform facing slopes with an angle α of 7^{o} and 17^{o} were used to extract σ^{o} values. The results are shown in table 2.1.

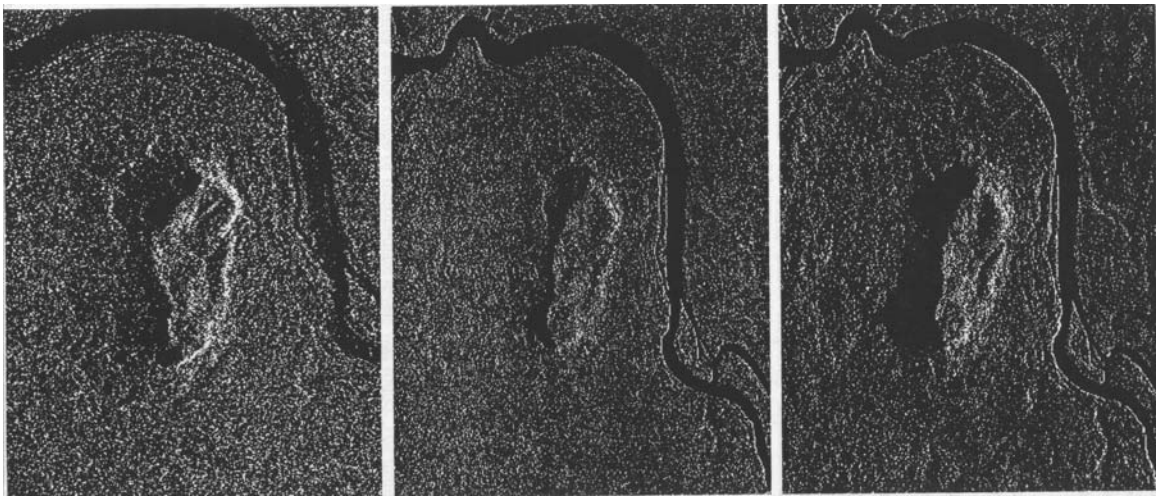


Figure 2.3. Triplet Turtle Mountains showing backscatter modulation by relief at (from left to right) steep (ERS-1), intermediate and grazing incidence angles (Hoekman *et al*., 1994).

Table 2.1. Backscatter level modulation by relief. Theoretical values for dense vegetation cover and experimental values for two uniform slopes and three incidence angles at Turtle Mountains, Iwokrama, Guyana.

	Slope		Theoretical	Experimental
Image	θ_{inc}	α	$\sigma^{o}/\sigma^{o}_{f}$ [dB]	$\sigma^{o}/\sigma^{o}_{f}$ [dB]
ERS-1	23o	17o	4.9	5.4
	23o	7o	3.7	3.5
Track2.2	66o	17o	2.9	3.1
	66o	7o	1.3	1.7
Track2.1	81o	17o	6.1	5.8
	81o	7o	1.7	2.5

These experimental results confirm that backscatter modulation (in forests) through relief is maximal at large and small incidence angles, and is minimal at intermediate incidence angles. The close resemblance of theory and experiment indicates that this

simple methodology can be used for slope correction, and may be useful to quantify slope angles in support of the preparation of Digital Elevation Models. In case a DEM is available, the theory can be applied to decompose the backscatter signal in a part that can be contributed to relief modulation and a part related to cover type. In this way, changes in terrain cover may be recognized more easily.

Alternative theoretical methodologies for the radiometric slope correction with more complex approximations can be found in the literature (Ulander, 1996; Castel *et al*., 2001). These methods are also based on local incidence angle and terrain slope tilt angle, but include the terrain slope aspect angle in addition.

2.2.2 Orthorectification of PolSAR slant range data using the TopSAR DEM

PolSAR and TopSAR data have been acquired in separate passes along the same flight line at nearly identical height and heading (Table 2.2). PolSAR images were delivered as slant range images, while TOPSAR images were delivered as orthorectified images. In case the assumption holds that the two flight lines are nearly identical in heading and height along the complete track, the orthorectification of the PolSAR images can be described in fairly simple geometrical terms when the offset in range and azimuth can be determined. Offset determination can be done using the geometrical description (with these two unknowns) and a number of GCP's. The GCP pairs can be retrieved most easily using (slant range) PolSAR TP composite images and (orthorectified) TopSAR C-VV/L-VV/P-TP composite images.

Table 2.2. Data from PolSAR and TopSAR header files

Near_slant_range PolSAR	9451.33 m, 26.7 deg inc.
Near_slant_range TopSAR	9301.44 m, 24.9 deg inc.
Height PolSAR	8439.9 m
Height TopSAR	8444.6 m
Heading PolSAR	38.8626 deg
Heading TopSAR	38.941 deg

Using the geometry introduced in Figure 2.4 the following two sets of relationships hold:

(1) $srPol^2 = (F_h\text{-}h)^2 + (grTop + \Delta y)^2$ (2.2a)

$srPol = PolPix0 + PolPixNum * PolPixSpac$ (2.2b)

$grTop = TopPix0 + TopPixNum * TopPixSpac$ (2.2c)

(2) $azPol = azTop + \Delta x$ (2.3a)

$azPol = PolLine0 + PolLineNum * PolLineSpac$ (2.3b)

$azTop = TopLine0 + TopLineNum * TopLineSpac$ (2.3c)

where:

F_h	height of flight [m]
h	terrain height, which is the height of the TopSAR pixel located at distance
$grTop + \Delta y$	[m]
$srPol$	slant range PolSAR [m]

azPol	azimuth PolSAR [m]
grTop	ground range TopSAR [m]
azTop	azimuth TopSAR [m]
Δx	offset in azimuth between PolSAR and TopSAR [m]
Δy	offset in (orthorectified) ground range between PolSAR and TopSAR [m]
PolPix0	slant range pixel zero [m]
TopPix0	ground range pixel zero [m]
PolLine0	azimuth distance line zero [m]
TopLine0	azimuth distance line zero [m]
PolPixNum	pixel number PolSAR
TopPixNum	pixel number TopSAR
PolLineNum	line number PolSAR
TopLineNum	line number TopSAR
PolPixSpac	slant range pixel spacing PolSAR [m]
TopPixSpac	ground range pixel spacing TopSAR[m]
PolLineSpac	line spacing PolSAR[m]
TopLineSpac	line spacing TopSAR [m]

(Note: these symbols are only used in this paragraph and are excluded from the list of symbols)

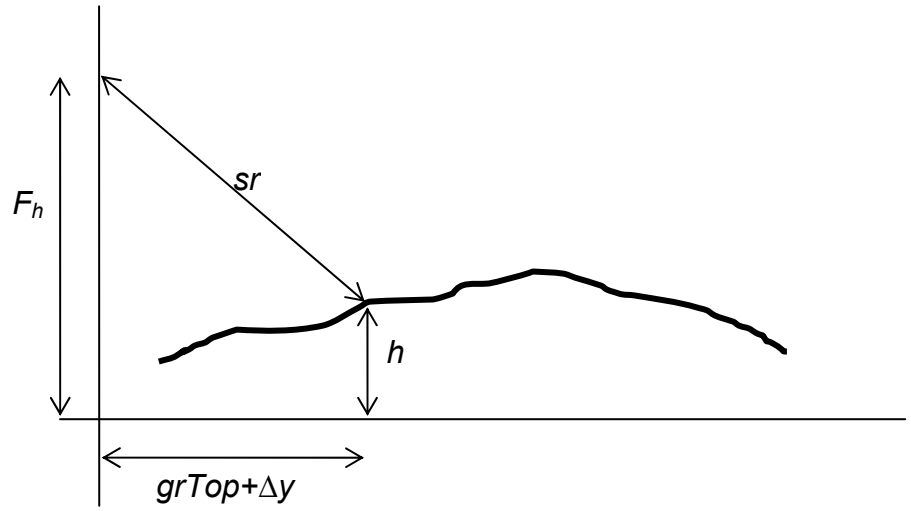


Figure 2.4. Geometry in range direction. Basic relation between slant range image and orthorectified image.

2.2.3 Fusion

After orthorectification of the PolSAR images, the TopSAR DEM can be used to apply the relief correction. Consequently three versions of the PolSAR images resulted which are: (1) the original slant range image, (2) the orthorectified image and (3) the slope corrected (cf. eq. 2.1) orthorectified image.

2.3 Radar data

2.3.1 Statistical description of data

The observation of rural or forest areas is of particular interest in this thesis. These types of targets are known to have fading characteristics. Because of fading, the signal, even for homogeneous areas with an underlying uniform cross section, shows a stochastic fluctuation (also called speckle). Speckle is a salient feature of all radar images and is caused by random variation of constructive and destructive interference from the multiple scattering returns that will occur within each observation point.

The interference of radar echoes clearly obstructs consistent backscatter measurements from single-resolution cells. However, the backscatter fluctuations have a stochastic nature, i.e. behave according to certain probability distributions. For a wide range of distributed land targets, including forests, the backscatter amplitude fluctuations can be described with a theoretical probability distribution usually called the Rayleigh distribution. In practice, many radar systems do not detect the amplitude but the square of the amplitude (or power) of the return signal that can be shown to have an exponential distribution (Hoekman, 1991). Both distributions are single-parameter distributions. The Rayleigh distribution follows directly from the mean amplitude, the exponential distribution from the mean backscattered power. To enable object characterisation using radar remote sensing, it is therefore important that the radar measurements are accurate estimates of either mean amplitude or mean power.

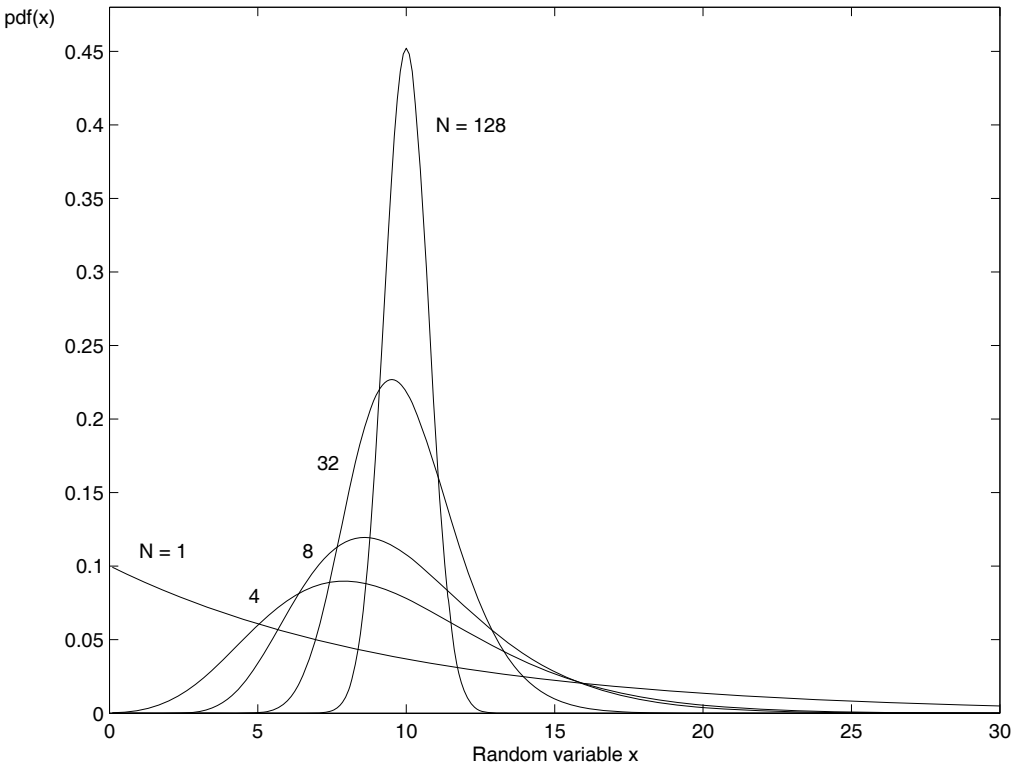


Figure 2.5. Probability density functions (*pdf*) for speckled power have a Gamma distribution: for one look (N=1) the exponential distribution follows, for a large number of looks (N>>100) the *pdf* may be approximated by a Gaussian distribution.

The degrees of freedom or number of independent "looks" involved in the averaging process also affects the shape of the amplitude or power distribution. Figure 2.5 illustrates the distribution of backscattered power for 1, 4, 8, 32 and 128 look radar data. For N=1, the *pdf* is exponential and for $N \to \infty$ it becomes a Gaussian *pdf*. It can be seen that the distributions become higher and narrower as the number of looks increases. This implies that the fluctuations in the measurements of backscattered power are reduced as a result of look averaging. Look averaging improves the radiometric properties, i.e. the radiometric resolution of radar images. However, look averaging may also cause deterioration of the geometrical properties of radar images, and reduce spatial resolution in azimuth direction.

2.3.2 Approach for extraction samples from training areas

Using visual interpretation of total power C-, L- and P-band, composite images of regions of interest (ROI's) were delineated for areas visited during the fieldwork stage. The radiometric attributes of ROI's from the radar image quantify the backscatter (γ). Radar backscatter analysis, however, is usually based not on the backscatter values of individual pixels but on the mean backscatter values for series of pixels, i.e. for image regions. In practice, the accuracy of radar measurements is often improved through linear averaging of measurements from resolution cells that adjoin in the azimuth direction. A measurement from a single-resolution cell is often denoted as a look. This explains why the image resulting from such averaging process is often referred to as a multiple look image. Multiple look images, when compared to single-look images, show fewer fluctuations in backscatter, or in other words, are less "grainy". Relevant averaging, however, requires that the looks are statistically independent. The mean backscatter value for ROI's is computed by averaging the values of all pixels within its boundaries. Region averaged backscatter values are less susceptible to the effects of speckle than backscatter value for pixels. The accuracy of mean amplitude or power estimations will increase with an increase in the number of averaged looks.

It is important to mention that the minimum size of the selected areas was arbitrarily set at 100 pixels. Introducing a lower limit with respect to the number of pixels that has to be averaged can regulate the effect of speckle. In this study it was assumed that the speckle induced standard deviation of the mean backscatter values should be 0.2 dB or less. Based on Hoekman, (1991) this criterion can only be met through the averaging of 500 or more independent backscatter measurements (looks).

This assumption was evaluated using the standard deviation of radar backscatter within the ROI's created. The lowest values were found for relatively homogeneous areas with cover types such as: swamp, *alang-alang* and shrubs. The lower boundaries of standard deviation found were: 1.7-1.8 dB for C-band, 2.2-2.3 dB for L-band and 3 dB for P-band. These figures indicate that the number of independent looks N is 6-7 for C-band, 4-5 for L-band and 2-3 for P-band. This is considerably lower than the specifications of NASA/JPL, which indicate a number of 9 independent looks for all frequency bands (Table 3.4). To compute the number of independent looks for an area, such as an ROI's of 100 pixels, the spatial correlation in azimuth and range direction has to be taken into account. Analysis showed that the effective number of independent looks reduces further to roughly 4.7, 2.8 and 1.5 looks, respectively,

depending also a little bit on shape and orientation of the ROI's. Thus, for an area larger than 100 pixels, the lower values of accuracy follow as a standard deviation of 0.20 dB, 0.26 dB and 0.36 dB, respectively. Since most ROI's are much larger than 100 pixels, the averaged backscatter values are sufficiently accurate (less than ≈ 0.2 dB), with a possible exception for P-band values in small areas (less than ≈ 0.4 dB).

2.3.3 Extraction of polarimetric attributes.

Polarimetric attributes related to the polarisation transformation properties of the observed object can be computed from the polarimetric radar data only. Hence, their extraction was restricted to regions of interest within the AirSAR images. The following data are extracted for each of the three frequency bands: number of pixels, incidence angle, and the elements of the scattering matrix $\boldsymbol{S}$, from which averaged elements of the Stokes scattering operator $\boldsymbol{M}$ can be derived (Table 2.3). Afterwards the averaged γ (HH, VV and HV), averaged complex correlation ρ_{hhvv} and polarimetric phase difference ϕ_{hhvv} can be calculated.

Table 2.3. Elements of the Stokes scattering operator $\boldsymbol{M}$ for backscatter (i.e. $S_{hv} = S_{vh}$) and for an ensemble averaged observation (expressed in elements of the scattering matrix $\boldsymbol{S}$). This matrix is 4x4 real and symmetric (e.g. $M_{32} = M_{23}$).

$$M_{11} = \frac{1}{4}\left[\left\langle S_{hh}\cdot S_{hh}^{*}\right\rangle + \left\langle S_{vv}\cdot S_{vv}^{*}\right\rangle + 2\left\langle S_{hv}\cdot S_{hv}^{*}\right\rangle\right]$$

$$M_{12} = \frac{1}{4}\left[\left\langle S_{hh}\cdot S_{hh}^{*}\right\rangle - \left\langle S_{vv}\cdot S_{vv}^{*}\right\rangle\right]$$

$$M_{13} = \frac{1}{2}\,\mathrm{Re}\left[\left\langle S_{hh}\cdot S_{hv}^{*}\right\rangle\right] + \frac{1}{2}\,\mathrm{Re}\left[\left\langle S_{hv}\cdot S_{vv}^{*}\right\rangle\right]$$

$$M_{14} = -\frac{1}{2}\,\mathrm{Im}\left[\left\langle S_{hh}\cdot S_{hv}^{*}\right\rangle\right] - \frac{1}{2}\,\mathrm{Im}\left[\left\langle S_{hv}\cdot S_{vv}^{*}\right\rangle\right]$$

$$M_{22} = \frac{1}{4}\left[\left\langle S_{hh}\cdot S_{hh}^{*}\right\rangle + \left\langle S_{vv}\cdot S_{vv}^{*}\right\rangle - 2\left\langle S_{hv}\cdot S_{hv}^{*}\right\rangle\right]$$

$$M_{23} = \frac{1}{2}\,\mathrm{Re}\left[\left\langle S_{hh}\cdot S_{hv}^{*}\right\rangle\right] - \frac{1}{2}\,\mathrm{Re}\left[\left\langle S_{hv}\cdot S_{vv}^{*}\right\rangle\right]$$

$$M_{24} = -\frac{1}{2}\,\mathrm{Im}\left[\left\langle S_{hh}\cdot S_{hv}^{*}\right\rangle\right] + \frac{1}{2}\,\mathrm{Im}\left[\left\langle S_{hv}\cdot S_{vv}^{*}\right\rangle\right]$$

$$M_{33} = \frac{1}{2}\left\langle S_{hv}\cdot S_{hv}^{*}\right\rangle + \frac{1}{2}\,\mathrm{Re}\left[\left\langle S_{hh}\cdot S_{vv}^{*}\right\rangle\right]$$

$$M_{34} = -\frac{1}{2}\,\mathrm{Im}\left\langle S_{hh}\cdot S_{vv}^{*}\right\rangle$$

$$M_{44} = \frac{1}{2}\left\langle S_{hv}\cdot S_{hv}^{*}\right\rangle - \frac{1}{2}\,\mathrm{Re}\left[\left\langle S_{hh}\cdot S_{vv}^{*}\right\rangle\right]$$

The pixels in the polarimetric AirSAR data sets represent the Stokes scattering operator $\boldsymbol{M}$. Hence, two steps were needed to compute the region averaged backscatter values from the AirSAR data sets. To begin with, the $\boldsymbol{M}$ values for all pixels within a specific region of interest were averaged linearly, in order to obtain region averaged $\boldsymbol{M}$ values. Subsequently, region averaged backscatter values were

computed through wave synthesis. The wave synthesis process allows the computation of the backscatter for any combination of the received and transmitted polarisations. Selected combinations for this study include HH, VH, VV, RR, LR and LL polarisation. The HV and RL polarisation combinations were not included, because according to the reciprocity relation these combinations are identical to VH and LR, respectively. The set of backscatter values for the described polarisation combinations was complemented with so-called total backscattered power (TP). Table 2.4 shows the orientation angle ψ and ellipticity angle χ for the polarisation combinations, as well as the equations for computing the corresponding σ^o values from the region averaged stokes scattering operator elements.

Table 2.4. Definition of polarisation combinations in term of ψ and χ; equations for computations of σ^o from the region averaged elements of Stokes scattering operator M. The backscatter values for the shown polarisation combinations were computed for C-, L- and P-band.

Polarisation combination	Polarisation parameters				Computation of σ^o from M
	Received wave		Transmitted wave		
	ψ	χ	ψ	χ	
HH	0°	0°	0°	0°	$M_{11}+2.M_{12}+M_{22}$
VH	90°	0°	0°	0°	$M_{11}-M_{22}$
VV	90°	0°	0°	0°	$M_{11}+2.M_{12}+M_{22}$
RR	*	-45°	*	-45°	$M_{11}+2.M_{12}+M_{22}$
LR	*	45°	*	-45°	$M_{11}+2.M_{12}+M_{22}$
LL	*	45°	*	45°	$M_{11}+2.M_{12}+M_{22}$
TP	-	-	-	-	$M_{11}+2.M_{12}+M_{22}$

- * indicates 'not defined'
- - indicates 'not applicable'

Calculation of γ is based on the following relation:

$$\gamma = \sigma^o / \cos(\theta_i) \qquad [m^2 / m^2] \tag{2.4}$$

with σ^o the differential radar cross section and γ the radar cross section per unit projected area.

The complex correlation ρ between the HH- (i.e. horizontal linear receive and horizontal linear transmit) and VV-returns can be computed from elements of the Stokes scattering operator as:

$$\rho = |\rho| \exp(i\phi) = \frac{\langle S_{hh} S_{vv}^* \rangle}{\sqrt{\langle S_{hh} S_{hh}^* \rangle \langle S_{vv} S_{vv}^* \rangle}} \tag{2.5}$$

For each region of interest the C-, L- and P-band polarisation phase difference ϕ of HH and VV (PPD) and the corresponding standard deviation (PPD) were computed. The PPD was computed from spatially averaged C-, L- and P-band Stokes scattering operator according to:

$$PPD = \tan^{-1}\left(\frac{-2M_{34}}{M_{33} - M_{44}}\right) \tag{2.6}$$

It is assumed that the objects display azimuthal symmetry and, consequently, only the correlation and phase difference distributions for HH-VV polarisation are important; these can be ignored for HH-HV and HV-VV polarisation (Nghiem *et al*., 1992).

2.4. Evaluation of classification capacities

2.4.1 Evaluation based on fully polarimetric AirSAR images

To evaluate the classification capacities of the attributes estimated from the AirSAR image, a data base was created for C-, L- and P-band polarimetric data, containing ten different elements of the stokes scattering operator (***M***), the class number and the incident angle for each polygon. The polarimetric classification technique introduced in the work of Hoekman and Quiñones (2000) is exploited to assess AirSAR's potential for forest cover type mapping and biomass classification. The classification results can be simulated as a function of the number of independent looks, a fully polarimetric multi-band approach using probability density functions (*pdf*) for multi-look samples of a certain class, for intensity, phase difference as well as coherence magnitude.

For a homogeneous area *i*, characterized by a spatially uniform differential cross section, phase difference and correlation, and a Gaussian probability density function (*pdf*) for the complex electric field vector as measured by both antennae, multi-look *pdf*'s of the observation can be described by the number of looks *N* (per pixel) and the underlying values for backscatter γ_i, phase difference ϕ_i and correlation $|\rho_i|$. The theoretical *pdf* for multi-look backscatter intensity is the well-known gamma function:

$$P_\gamma\left(\gamma \mid \gamma_i\right) = \frac{1}{\Gamma(N)}\left(\frac{N}{\gamma_i}\right)^N \gamma^{N-1} e^{N\gamma/\gamma i} \tag{2.7}$$

For phase difference and correlation marginal distributions are given by Tough *et al*., (1995):

$$P_{\phi}(\phi \mid \phi_i, |\rho_i|) = \frac{\left(1-|\rho_i|^2\right)^N}{2\pi}\left\{\frac{(2N-2)!}{\left[(N-1)!\right]^2 2^{2(N-1)}}\right.$$
$$\left[\frac{(2N-1)\beta\arccos(-\beta)}{\left(1-\beta^2\right)^{N+1/2}}+\frac{1}{\left(1-\beta^2\right)^N}\right]+$$
$$\left.\frac{1}{2(N-1)}\sum_{r=0}^{N-2}\frac{\Gamma(N-1/2)}{\Gamma(N-1/2-r)}\frac{\Gamma(N-1-r)}{\Gamma(N-1)}\frac{1+(2r+1)\beta^2}{\left(1-\beta^2\right)^{r+2}}\right\}$$

(2.8)

where the last term is 0 for N=1 and $\beta = |\rho_i|\cos(\phi-\phi_i)$, and

$$P_{|\rho|}(|\rho| \mid |\rho_i|) = 2|\rho|(N-1)\left(1-|\rho_i|^2\right)^N\left(1-|\rho|^2\right)^{N-2}\,{}_2F_1\left(N,N,1;|\rho|^2|\rho_i|^2\right),$$

(2.9)

where ${}_2F_1(.)$ is the Gaussian hypergeometric function.

For a classification procedure, statistical descriptions are required for pixels belonging to a certain class, rather than belonging to a certain homogenous area, and it is assumed that probability density functions (*pdf*'s) are well described by Gaussian distribution for parameter γ_i (in dB),

$$N(\gamma_i \mid \gamma_c, \sigma_c) = \frac{1}{\sigma_c\sqrt{2\pi}}\exp\left(-\frac{(\gamma_i-\gamma_c)^2}{2\sigma_c^2}\right) \tag{2.10}$$

with γ_c (in dB) as the mean of the mean field values of class c and σ_c as the standard deviation of γ_c and by circular Gaussian distributions for the polarimetric phase differences

$$C(\phi_i \mid \phi_{c,}, |\rho_c|) = \frac{1-|\rho_c|^2}{2\pi\left(1-\beta^2\right)}\left\{1+\frac{\beta\arccos(-\beta)}{\sqrt{\left(1-\beta^2\right)}}\right\} \tag{2.11}$$

with $\beta = |\rho_c| \quad \cos(\phi_i-\phi_c), \quad -\pi < \phi_i \le \pi$, where ϕ_c = the "effective" mean phase difference for class c and $|\rho_c|$ is the effective mean correlation for class c. Gaussian

distributions are "natural" distributions, which follow from applications of Jayne's maximum entropy principle. For a continuous random variable varying over the (0,1) interval, application of this principle results in the Beta function. Hence, phase correlations may be assumed to be properly described by Beta distributions.

$$B\left(|\rho_i||a,b|\right)=\frac{\Gamma\left(a+b\right)}{\Gamma\left(a\right)\Gamma\left(b\right)}|\rho_i|^{a-1}\left(1-|\rho_i|\right)^{b-1} \quad ,0<|\rho_i|<1 \tag{2.12}$$

where a and b are regression parameters.

The Likelihood for an observation vector to be classified as class c is modelled as the product of the joint Gaussian distributions of the backscatter values. When phase difference and correlation values are included, the likelihood will be determined by this first product multiplied by the likelihood of the phase difference values and the likelihood of the correlation values.

The field averaged Stokes scattering element data of the database are used to calculate field averaged values of backscatter, phase differences and correlation. The accuracy of the estimation of field averaged values depends on the total number of independent looks (see also 2.3).

2.4.2 Evaluation based on reversible transform AirSAR image

To evaluate the classification capacities of the attributes from the AirSAR image in an alternative way, a *reversible transform* of the covariance matrix into backscatter intensities was introduced in the work of Hoekman and Vissers, (2003) in order to describe the full polarimetric target properties, and which allows for the development of simple, versatile and robust classifiers.

Sometimes it is mathematically more convenient to use the covariance matrix instead of the Mueller matrix for the description of the fully polarimetric backscatter from a target. Because of the reciprocity theorem (Nghiem *et al*., 1992) it is sufficient to define the 3x3 covariance matrix with only 9 independent real numbers:

$$C=\begin{pmatrix} \left\langle S_{hh}S_{hh}^{*}\right\rangle & \left\langle S_{hh}S_{hv}^{*}\right\rangle & \left\langle S_{hh}S_{vv}^{*}\right\rangle \\ \left\langle S_{hv}S_{hh}^{*}\right\rangle & \left\langle S_{hv}S_{hv}^{*}\right\rangle & \left\langle S_{hv}S_{vv}^{*}\right\rangle \\ \left\langle S_{vv}S_{hh}^{*}\right\rangle & \left\langle S_{vv}S_{hv}^{*}\right\rangle & \left\langle S_{vv}S_{vv}^{*}\right\rangle \end{pmatrix} \tag{2.13}$$

The covariance matrix contains the second-order statistics of the scatter matrix elements and is Hermitian, which means that the matrix is equal to its transposed complex conjugate. These properties are contained in the three real numbers on the diagonal and the six real and imaginary parts of the three complex numbers above the diagonal. Another way of presenting the full polarimetric information content is by using three intensity values, three polarisation phase difference values, and three correlation magnitude values (with ρ as the complex correlation) as

$$C = \frac{1}{4\pi}\begin{pmatrix} \sqrt{\sigma^0_{hh}} & 0 & 0 \\ 0 & \sqrt{\sigma^0_{hv}} & 0 \\ 0 & 0 & \sqrt{\sigma^0_{vv}} \end{pmatrix}\begin{pmatrix} 1 & \rho_{hhhv} & \rho_{hhvv} \\ \rho^*_{hhhv} & 1 & \rho_{hvvv} \\ \rho^*_{hhvv} & \rho^*_{hvvv} & 1 \end{pmatrix}\begin{pmatrix} \sqrt{\sigma^0_{hh}} & 0 & 0 \\ 0 & \sqrt{\sigma^0_{hv}} & 0 \\ 0 & 0 & \sqrt{\sigma^0_{vv}} \end{pmatrix} \tag{2.14}$$

It is also possible to describe the full polarimetric information content with nine intensities, for example as

$$\begin{pmatrix} \langle S_{hh}\cdot S^*_{hh}\rangle \\ \langle S_{vv}\cdot S^*_{vv}\rangle \\ \langle S_{hv}\cdot S^*_{hv}\rangle \\ \mathrm{Re}\left[\langle S_{hh}\cdot S^*_{vv}\rangle\right] \\ \mathrm{Im}\left[\langle S_{hh}\cdot S^*_{vv}\rangle\right] \\ \mathrm{Re}\left[\langle S_{hh}\cdot S^*_{hv}\rangle\right] \\ \mathrm{Im}\left[\langle S_{hh}\cdot S^*_{hv}\rangle\right] \\ \mathrm{Re}\left[\langle S_{hv}\cdot S^*_{vv}\rangle\right] \\ \mathrm{Im}\left[\langle S_{hv}\cdot S^*_{vv}\rangle\right] \end{pmatrix} = B_r \begin{pmatrix} \sigma^0_{hh} \\ \sigma^0_{vv} \\ \sigma^0_{++45} \\ \sigma^0_{--45} \\ \sigma^0_{ll} \\ \sigma^0_{rr} \\ \sigma^0_{h+45} \\ \sigma^0_{hl} \\ \sigma^0_{+45l} \end{pmatrix},$$

$$\text{with } B_r = \frac{1}{4\pi}\left(\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\ \hline 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & 0 & 0 & -2 \\ -\frac{3}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & 1 & 0 & 0 \\ \hline \frac{3}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & -1 & 0 \\ \frac{3}{8} & -\frac{1}{8} & \frac{5}{8} & -\frac{3}{8} & \frac{1}{8} & \frac{1}{8} & -1 & 0 & 0 \\ -\frac{3}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{5}{8} & \frac{3}{8} & 0 & 1 & 0 \end{array}\right) \tag{2.15}$$

where the subscripts *pq* denote the received and transmitted polarisations of the three common polarisation bases: horizontal (*h*), vertical (*v*), left circular (*l*), right circular (*r*), 45º linear (+ or +45) and -45° linear (- or -45).

In the case of azimuthal symmetry (Nghiem *et al*., 1992) the covariance matrix simplifies to

$$C_r = \begin{pmatrix} \langle S_{hh}S_{hh}^* \rangle & 0 & \langle S_{hh}S_{vv}^* \rangle \\ 0 & \langle S_{hv}S_{hv}^* \rangle & 0 \\ \langle S_{vv}S_{hh}^* \rangle & 0 & \langle S_{vv}S_{vv}^* \rangle \end{pmatrix}, \tag{2.16}$$

and $\rho_{hhhv} = \rho_{hvvv} = 0$. When interpreting remote sensing data it is sometimes useful to consider this and only this information, which is contained in 5 independent values of ***C***. From the intensity representation introduced here, it is also possible to find several sets of 5 independent intensity values containing this and only this information. In all cases at least one composite intensity is needed. One (non-redundant) possibility is

$$\begin{pmatrix} \langle S_{hh} \cdot S_{hh}^* \rangle \\ \langle S_{vv} \cdot S_{vv}^* \rangle \\ \langle S_{hv} \cdot S_{hv}^* \rangle \\ \mathrm{Re}\left[\langle S_{hh} \cdot S_{vv}^* \rangle \right] \\ \mathrm{Im}\left[\langle S_{hh} \cdot S_{vv}^* \rangle \right] \end{pmatrix} = B_r \begin{pmatrix} \sigma_{hh}^0 \\ \sigma_{vv}^0 \\ \sigma_{hv}^0 \\ \sigma_{+-45}^0 \\ \sigma_{+45l}^0 + \sigma_{-45r}^0 \end{pmatrix},$$

$$\text{with } B_r = \frac{1}{4\pi} \left(\begin{array}{ccc:cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hdashline \frac{1}{2} & \frac{1}{2} & 0 & -2 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 & -1 \end{array} \right) \tag{2.17}$$

Ignoring other possibilities, there is one redundant system of particular interest which combines 7 intensity values, including two composite values, namely

$$\{ \sigma_{hh}^0,\ \sigma_{vv}^0,\ \sigma_{hv}^0,\ \sigma_{+-45}^0,\ \sigma_{lr}^0,\ \sigma_{+45l}^0 + \sigma_{-45r}^0,\ \sigma_{+45r}^0 + \sigma_{-45l}^0 \}. \tag{2.18}$$

In this study, four classification models named 3I+, 5I, 7I and 9I will be evaluated:

- (3I+) The first model is identical to paragraph 2.4.1 above. It uses a joint log-normal distribution for the HH-HV- and VV-intensities and independent distributions for the phase difference (circular Gaussian distribution) and the coherence magnitude (Beta distribution).
- (5I) The second model is based on Eq.2.17 and uses a joint log-normal distribution for the 5 intensities.
- (7I) The third model is based on Eq.2.18 and uses a joint log-normal distribution for the 7 intensities.
- (9I) The fourth model is based on Eq.2.15 and uses a joint log-normal distribution for the 9 intensities.

Note that the first three models assume azimuthal symmetry or, alternatively, discard information related to azimuthal asymmetry. The fourth model includes all polarimetric information.

2.5 Maximum Likelihood classification

For classification, the Maximum Likelihood (ML) approach is selected. For each class means, variances and co-variances are calculated on the basis of data derived from the ROI's. Signatures are calculated from a training data set, and then estimates of class membership can be made. Usually, for a given pixel, the classification follows as the most likely class. In more sophisticated approaches, use can be made of additional information, such as class likelihoods of neighbouring pixels (e.g. Markov Random Fields or ICM; see section 2.6) or textural information (see section 2.7).

The operation of a ML classifier is a two-stage process. The first stage is a training stage, where a spectral signature is calculated for each of several *a priori* classes the operator has identified in the image. The second stage, the classification stage, uses these signatures to assess each pixel in the image, and determines the class in which the pixel belongs. The ML classifier used here uses the assumption that the population is Gaussian. Each class is assumed to be normally distributed, and there are several classes in the image. The assumption of normality for each class is typically met in practical situations. In principle, the ML classifier method calculates a probability density function (*pdf*) for each class, and then uses the maximum value of those functions to identify the class to which a pixel belongs. The set of signatures is obtained by calculating the mean vector and covariance matrix for each class. The mean vector describes the location of the curve that represents the class in spectral space, while the covariance matrix describes the orientation of the curve. The parameters are used to calculate a probability density surface for each class.

The classification can be done using all data or a selection of the data. For example, in case only L-band intensities are used with HH, VV and HV polarisation a 3-dimensional feature space is used and a 3-dimensional multi-variate Gaussian distribution applies. In this case 3 so-called radar information channels are used. In general, the number of channels n can vary between 1 and 9 when only one polarimetric frequency band is available, and between 1 and 27 in the case of AirSAR where three frequency bands are available. In this study the utility of many combinations of channels will be evaluated.

The data samples in the training set are related to a number of classes. If there are *m* classes and *n* channels, which for each class ω_i with *i* ranging from 1 to *m* and an *n*-dimensional mean value vector M_i, an *n x n*- dimensional covariance matrix C_i can be computed. Assuming an *n*-dimensional normal distribution is applicable, each observation vector ***X*** to be classified can be assigned to one of the classes ω_i or, in case non of these classes is likely, to a class ω_o.

The likelihood that ***X*** is a number of class ω_i, is

$$P(X \mid \omega_i) = \frac{1}{(2\pi)^{n/2} \mid C_i \mid^{1/2}} \exp(-1/2(X - M_i)^T C_i^{-1}(X - M_i)) \qquad (2.19)$$

with $|C_i|$ = the determinant of C_i

For one of these classes the likelihood value is the highest (i.e. the 'maximum likelihood'). For this class also the logarithmic value of the likelihood is the highest:

$$\ln P(X \mid \omega_i) = -\frac{n}{2}\ln(2\pi) - 1/2\{\ln(\mid C_i \mid) + (X - M_i)TC_i^{-1}(X - M_i)\} \qquad (2.20)$$

The vector ***X*** will be assigned to class ω according to the classification rule:

$$\ln\big(P(X \mid \omega_i)\big) > \ln(P(X \mid \omega_j)) \text{ for all } j \neq i. \qquad (2.21)$$

and

$$\ln\big(P(X \mid \omega_i)\big) > T_i. \qquad (2.22)$$

where $\boldsymbol{T_i}$ is a suitably chosen threshold value. This value can be computed as follows. For the n-dimensional vector ***X*** the standard distance of ***X*** to $\boldsymbol{M_i}$ is

$$d_i(X) = \{\ln(X - M_i)TC_i^{-1}(X - M_i)\}^{1/2}. \qquad (2.23)$$

In case, for example 95% of the observation vectors of class ω_i are not rejected, the standardized threshold distance *D* can be derived from:

$$D^2 \equiv \{\ (X_D - M_i)^T C_i^{-1}(X_D - M_i)\} = X^2{}_{n;0.95} \qquad (2.24)$$

and T_i follows as:

$$T_i = -n/2\ln(2\pi) - 1/2\left\{\ln(\mid C_i \mid) + D^2\right\} \qquad (2.25)$$

For example, for 4 classes and a 95% confidence interval, D^2 is 9.49. In principle the confidence level can be chosen for each class separately. Results can be presented in a

(m+1) * (m)-dimensional contingency table or confusion matrix, percentages of correct and unclassified samples and Kappa statistics. In case the threshold is set to infinity (i.e. the 100% confidence level) all observation vectors will be classified and class ω_o is empty (Hoekman and Vissers, 1995).

For ML classification of fully polarimetric SAR data, several approaches can be utilized. The simplest approach is to classify multi-look pixels. Box averaging or segmentation into homogeneous areas can be applied first, to reduce the effect of speckle. The resulting number of independent looks is usually too small to ignore the effect of speckle on the estimation of the underlying values of γ_i, ϕ_i and $|\rho_i|$.

2.6 Iterated Conditional Modes

In order to improve the classification results, the ML classification using the approaches introduced in section 2.4, could be followed by the technique of Iterated Conditional Modes (ICM) (Besag, 1986). As explained in Hoekman and Quiñones (2002), the Likelihood of a pixel i belonging to class c, $li_{i,c}$, is based on the (multi-frequency) radar signal properties in terms of intensities, phases and coherences. The classification of a pixel is the selection of the class for which $li_{i,c}$ is the highest (the Maximum Likelihood or ML solution). The Likelihood li used for the ML method has to be multiplied with the prior probabilities of each of the neighbouring pixels.

In this approach, the eight surrounding pixels form the neighbourhood. The following notation for the neighbourhood of pixel i is adopted:

1	2	3
4	*i*	5
6	7	8

Depending on its relative position, a neighbouring pixel has a position j, $j \in \{n \mid 1 \le n \le 8\}$. The class has a value x, $x \in \{n \mid 1 \le n \le 14\}$. For every pixel i, we compute for each class x the li and assign the class value for which li is maximal (out of these 14 classes; see Table 3.1) to the pixel i.

In the ICM method the Likelihood $li_{i,c}$ is modified to $mli_{i,c}$ by multiplication with a conditional probability $\exp\left(\beta u_{i,c}\right)$, where $u_{i,c}$ is the current number of neighbours of pixel i having class c, and β is a parameter determining the relative importance of neighbourhood information. The logarithmic version of the modified likelihood $mli_{i,c}$ for ICM-cycle n is denoted as

$$\ln\left(mli_{i,c,n}\right) = \ln\left(li_{i,c}\right) + \beta u_{i,c,n-1} \tag{2.26}$$

For appropriately chosen values of β, the number of cycles and the relaxation scheme, usually determined by trial-and-error, this approach is found to yield major improvements for the classification results. Moreover, taking the dominance of certain cover types into account (i.e. the Bayes criterion) the overall accuracy can increase

further. This method can be enhanced by including additional information to the previous (logarithmic version of the) modified likelihood. For example in Hoekman and Quiñones (2002), the scheme proposed had the following form:

$$\ln\left(mli_{i,c,n}\right)=\ln\left(li_{i,c}\right)+\beta_1 u_{i,c,n-1}+\beta_2 \ln\left(P_c\right)+\beta_3 \ln\left(R_{i,c}\right)-\beta_4 T_{i,c} \qquad (2.27)$$

with

$$T_{i,c}=Min\left\{\frac{\left(t_i-Tm_c\right)^2}{2\,Tv_c},\beta_{\mathrm{T}}\right\}$$

where P_c is the relative occurrence of class *c*,

$R_{i,c}$ is the relief factor for pixel *i* and class c,

$T_{i,c}$ is the texture factor for pixel *i* and class c,

$\beta_2, \beta_3, \beta_4$ are factors defining the relative influence of *prior* information,

β_{T} is a factor defining a threshold for the influence of texture information,

t_i is the (logarithmic version of) the coefficient of variation (CV)

Tm_c, Tv_c are the mean and variance of the CV for class *c*.

In this study, the relief factor could be obtained directly from the DEM derived from the TopSAR data, instead of using geomorphological maps as was done in Hoekman and Quiñones (2002). The textural features give more information as compared to Hoekman and Quiñones (2002), because of the higher resolution of the AirSAR. Moreover, because of the availability of near simultaneously collected TopSAR data, this texture technique can be further developed into 3 dimensions.

2.7 Texture derived from InSAR

3-D textural information in the InSAR images, calculated using statistical measures, can be used as additional information in the ICM process. The classical 2-D image texture may be seen as a spatial pattern arising from a deterministic or random repetition of local sub-patterns or primitives (i.e. pixels), with or without a preferred direction (Hoekman, 1990). Textural of height variation from the vegetation is very common in the tropical rain forest, and may be seen as such a pattern. It can be utilized to discriminate regions of interest or to delineate objects in an image. In radar images, (natural) forest types can be differentiated if sizes of resolution cells are comparable or smaller than sizes of major structural (or architectural) canopy components. Since textural phenomena can be linked to spatial properties, quantification of texture might become a useful tool in the characterization of forest architecture (Oldeman, 1990).

In this study, textural phenomena of height variation derived from InSAR are described with statistical texture measures computed as first order statistics (i.e. standard deviation or SD) and as second order statistics derived from the elements of

the grey level co-occurrence (GLCO) matrix. In this case the elements of the GLCO matrix represent height variation (grey level) second-order statistics of pixel pairs contained in a certain image region or spatial window. This region contains pixels within a moving window (a kernel), and the textural measures are calculated for the centre pixel of this window. The second-order measures describe statistical dependences between two pixels with a set lag to a certain direction inside the kernel. The result depends on this lag or displacement length $|d|$ and the displacement direction α (Haralick, 1986; Hoekman, 1990).

Two of the most commonly used second-order textural operators were selected, namely GLCO contrast (GLCO-CONT) and GLCO correlation (GLCO-COR), and calculated for a number of displacement vectors. Since directional differences were not considered to be of interest, all values corresponding to a certain displacement length were averaged. Displacement lengths of 1, 5, and 10 pixels with window sizes of 7x7, 15x15 and 21x21 pixels were chosen, thus totalling the number of GLCO features to be investigated as 18. The two measures are defined as follows:

Contrast (GLCO-CONT):

$$\sum_{i=1}^{Ng}\sum_{j=1}^{Ng} P(i,j)(i-j)^2 \tag{2.28}$$

Correlation (GLCO-COR):

$$\sum_{i=1}^{Ng}\sum_{j=1}^{Ng} P(i,j)\frac{(i-m_x)(j-m_y)}{s_x s_y} \tag{2.29}$$

where $P(i,j)$ are the probability values of the GLCO matrix, N_g stands for the number of grey levels in the digitized image and m_x, m_y, s_x and s_y stand for the mean values and standard deviations of the row and column positions of the counts in the GLCO matrix, respectively.

2.8 Error matrix

The evaluation of the classification capacities was done through the analysis of the classification result, specifically the error matrix (contingency tables) and Kappa statistics for simulations performed for a single-channel, combination of two and three channels and full polarimetric information. It is important to notice that an error matrix gives a measure of the overall classification accuracy and presents errors of commission and omission, while the use of the Kappa statistic enables comparison between classification results.

The results of a cross-classification between the original class map and the classified class maps were tabulated using the error matrix. By counting cross-classification errors by class, the accuracy of each classification can be assessed. An error matrix is a very effective way to represent map accuracy; the individual accuracies of each

category are plainly described along with both the errors of inclusion (commission errors) and errors of exclusion (omission errors) present in the classification. A commission error is simply defined as including an area into a category when it does not belong to the category. An omission error is excluding that area from the category in which it truly does belong. Every error is an omission from the correct category and a commission to a wrong category.

Contingency tables show the number of correctly and incorrectly classified data points. An example of an error matrix is given in Table 2.5. The columns represent the actual land cover type as verified in the field (ground reference), whereas the rows indicate the land cover type as assigned by the classifier. Contingency tables clearly present errors of omission and errors of commission, while the error matrix can be used to compute other accuracy measures, such as overall accuracy, producer's accuracy, and user's accuracy. Correctly classified data points are located on the major diagonal of the table. The ratio of the number of correctly classified data points and the total number of data points represents the actual agreement between the rows and columns of the table, and may be used as a simple measure of the overall classification accuracy. However, in evaluating classification results it is important not only to note the proportion of correctly classified data points, but also to assess the nature of the errors of omission and commission on a class-by-class basis. When comparing classification results for different data sets, 'chance agreement' also has to be taken into account. Chance agreement in contingency tables results from the fact that any classifier will by chance assign data points to the correct class. It hinders direct comparison of classification results for different data sets, as it is a function of the row and column totals (Congalton, 1999).

Table 2.5. Example of a contingency table, which will be used to evaluate the Gaussian maximum-likelihood classification results.

	Ground truth "True"				
Classification result	D	C	AG	SB	Row Total
D	**65**	4	22	24	115
C	6	**81**	5	8	100
AG	0	11	**85**	19	115
SB	4	7	3	**90**	104
Column Total	75	103	115	141	**424**

Land Cover Categories
D = deciduous
C = conifer
AG = agriculture
SB = shrub

Total classification Accuracy = 100% *(65+81+85+90)/434 = 321/434 = 74%

A further analysis of the classification accuracy is the Kappa $\hat{K}$ (KHAT) statistic with the help of contingency tables, which describes the classification in relation to random classification. With the use of the $\hat{K}$ statistic it is possible to evaluate classification results in contingency tables while taking into account errors of omission and commission, and compensating for the effects of chance agreement. This measure of agreement is based on the difference between the actual agreement in the error matrix (i.e., the agreement between the remotely sensed classification and the reference data as indicated by the major diagonal) and the chance agreement,

which is indicated by the row and column totals (i.e., marginal). The statistic is calculated by:

$$\hat{K} = \frac{N\sum_{i=1}^{l} x_{ii} - \sum_{i=1}^{l} x_{i+} \cdot x_{+i}}{N^2 - \sum_{i=1}^{l} x_{i+} \cdot x_{+i}} \tag{2.30}$$

where l is the number of rows (columns) in the contingence table, x_{ii} the number of data points in row i and column i, x_{i+} the total of row i, x_{+i} the total of column i and N the total number of data points. The maximum value of $\hat{K}$ is 1. $\hat{K}$ and its variance may be used to compute confidence intervals for $\hat{K}$ and thus to construct a hypothesis test for significant difference between the $\hat{K}$'s for different contingence tables (Bishop *et al*., 1984). The test statistic for significant difference between two $\hat{K}$'s (significant difference between two classification results) is given by

$$\Delta\hat{K} = \frac{\left|\hat{K}_1 - \hat{K}_2\right|}{\sqrt{\hat{\sigma}^2_{\infty}\left[\hat{K}_1\right] + \hat{\sigma}^2_{\infty}\left[\hat{K}_2\right]}} \tag{2.31}$$

In this study all tests for significant difference between classification results were carried out at 95 % confidence level, at which two classification results may be considered significantly different when $\Delta\hat{K} > 1.96$ (Benson and DeGloria, 1985). A classification as good as the random assignment should produce a $\hat{K}$ of 0. A perfect classification would be a $\hat{K}$ of 1 (Lillesand and Kiefer, 1994).

3. Description of the study area and radar data

This chapter describes an overview of the study area and data sets that were used in this study. The Sungai Wain and its surrounding region, covering an area of 10 x 60 km^2 which comprises a wide variety of land cover types, located in the province of East-Kalimantan, (Borneo, Indonesia), was chosen as study area. The C-, L- and P-band polarimetric and C- and L-band interferometric data of the AirSAR/TopSAR NASA/JPL were collected in the framework of the PacRim-2 campaign, and was supported with extensive and detailed ground data measurements.

3.1 General study site characteristics

The study area was chosen in a region with tropical rain forest, featuring a large variability of land cover and topographic conditions, located in the surroundings of Balikpapan city, East-Kalimantan Province, Indonesia. Situated around central co-ordinates $0^o83'$ latitude south and $116^o76'$ longitude east (Figure 3.1), the study area is characterised by a complex mosaic of vegetation and land cover types. Tropical lowland evergreen and semi-evergreen rain forest dominate the natural vegetation with *Dipterocarpaceae* as the dominant family (Soerianegara and Lemmens, 1993). Large areas are covered with a variety of primary forests, including a Dipterocarp forest with emergent trees exceeding 30 m in height and 30 cm in diameter. More than 50% of the forest was affected by the El Niño forest fires of 1997/1998 (Priadjati, 2002). Secondary forest vegetation was formed partly as the result of deforestation and/or burning, caused by natural factors as well as by local farmers. The human population is concentrated along the main roads. Human influence largely affected terrain conditions and changed the primary vegetation partially or totally (Sidiyasa, 2001). Another important species to mention is *Imperata cylindrica* (*alang-alang*), which presently covers vast areas throughout the whole study site. In the southern section, a part of the Balikpapan Bay and several transmigration areas are included. Two large stretches of mangrove forests lie along both sides of the bay. The rest of the area consists of plantations, rice fields, wastelands, shrimp ponds and beaches. Slopes are between 8% and 30% and the elevation is between 50-150 m above sea level (Bremen *et al*., 1990).

3.2 Ground data collection

3.2.1 Description of the land cover classes determined by fieldwork

To conform to the objectives of the study, ground data were collected, focusing on the most important variables considering the radar data application. Additional data were collected later, when the analysis of the radar data set indicated the need to do so. Four types of information were collected: extensive observation, intensive observation, road tracking and collection of existing maps/data (Hoekman *et al*., 2000).

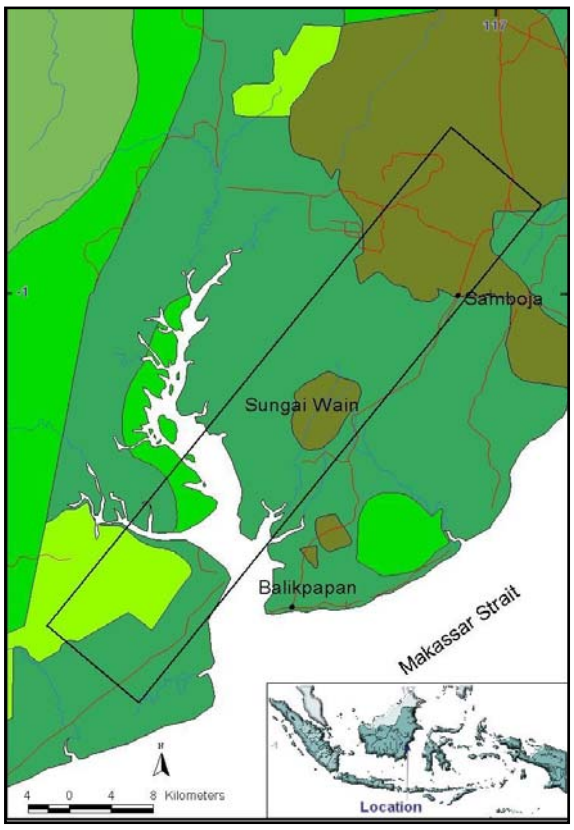


Figure 3.1. Location of the study area at *East-Kalimantan (Borneo),* Republic of Indonesia. (Derived from: 1996 Forest land use by consensus map of the Ministry of Forestry).

The fieldwork data sets contain several selected areas. Global Positioning System (GPS) measurements and several photographs were also acquired to support analysis for accuracy assessment and validation of classification models. During the fieldwork, the identification and the description of vegetation and land cover types were made at 142 locations. With the help of the topographic map and terrain knowledge, a total of 386 regions of interest (ROI's) were delineated and 14 land cover classes were distinguished. Those classes were grouped according to their natural composition, structure and function in three levels of division.

Land cover types defined in this study are divided in three main groups: non-forest, forest and water. At the second level, the forest class is subdivided into primary forest un-burnt and burnt, secondary forest and mangrove. In the same way, non-forest is subdivided into potential and real agricultural areas and industrial production areas. The potential and real agricultural areas are subdivided once again on the third level, and consist of: rice (*padi*) field, *alang-alang*, shrubs, mixed (transition from *alang-alang* to shrubs), shrimp ponds, urban area and bare soil. Industrial production areas are subdivided into oil palm plantations and rubber plantations. The study focuses on a total of fourteen classes (classes C1 through C14) at three different levels. The number of training areas per class is indicated in Table 3.1. To be able to study classification potential in a systematic way, and to enable proper validation, another independent training data set was collected by another researcher (Rodriguez, 2002; Vargas, 2002) in September 2001 (Table 3.1).

Table 3.1. Land covers type classification system for region of interest (ROI's) and validation. Bold characters indicate classes determined by fieldwork and used in this study. *Remark *): Class C4 was added later, after the evaluation made in section 4.4. Class C4 originates from a redefinition of classes C2 and C3 by adding a transition class.*

Level I	**Level II**	**Level III** Training area	Number of ROI's	Independent training area for validation	Number of ROI's
Non Forest	Potential and real agricul-tural areas	Rice (*Padi*) field (C1)	12	Swamp	10
		Alang-alang (C 2)	23	*Alang-alang*	12
		Shrubs (C3)	32	Shrubs	18
				Shrubs-trunks	10
		Mixed *) (C4)	11	Mixed	17
		Shrimp ponds (C5)	32	Shrimp ponds	11
		Urban area (C6)	11	Man-made structures	13
	Industrial production areas	Bare soil (C7)	33	Dead Mangrove	10
		Oil palm plantation (C8)	22	Coconut plantation	10
		Rubber plantation (C9)	29	Rubber plantation	11
	Mangrove (C10)		41	Mangrove	10
Forest	Secondary forest (C11)		38	Secondary forest	13
				Secondary forest-wet	12
				Secondary forest-O	10
				Secondary forest-Y-T	10
				Secondary forest-young	22
	Primary forest un-burnt (C12)		52	Primary forest un-burnt	11
	Primary forest burnt (C13)		19	Primary forest burnt	10
Water (C14)			31		-
Total			**386**		**220**

3.2.2 Land cover class description

The following descriptions of land cover type are based on visual information obtained from the field data collection and literature study.

Alang-alang (C2), the *alang-alang* (*Imperata cylindrica)* grass reproduces prolifically using seeds or a vegetative mechanism, and probably this is the reason why this grass is one of the few classes that can be found all over the study area. This type of grass is a result of the continuous disturbance of forested areas that have degraded to such a stage that it is difficult to recover. The area occupied by *alang-alang* increases continuously because of certain activities or events such as forest cutting, shifting cultivation, incorrect soil management and abandoned agricultural areas after two or three harvests of crops and after forest fires (Temmes, 1992). Often, it is associated

with shrubs. Those activities or events play an important role in the persistence of this *alang-alang* plant covering the land, because it collectively impedes natural regeneration or succession. This class has the ability to thrive on infertile soil, and has a high growth rate and biomass production. All these factors make this species a strong competitor with other plants for water, light and nutrients (Priadjati, 2002; Tolkamp *et al*., 2001).

Shrubs (C3), this class is characterised by very dynamic vegetation with an average height between 1 to 3 m and the presence of some pioneer trees. Most of the time, this class is associated with *alang-alang* and/or agricultural areas. This type of vegetation appears as a result of the succession or natural recovery of severely degraded areas that in many cases were probably reduced to the presence of only grasses. Some standing trunks are often present. These are mainly remnants of forest trees that were present before major fire events.

Mixed (transition from alang-alang to shrubs) (C4), under this label are those heterogeneous areas that present a chaotic mix of plants that make it impossible to define a dominant type of vegetation. Usually they resulted from severe intervention and continuous interruption of the succession or natural recovery of the area, sometimes associated with *alang-alang* and/or shrubs. Usually the mixed class is composed of grasses, shrubs, small trees, and maintained gardens (pepper, banana, etc.) and abandoned gardens.

Water, Shrimp ponds and Rice (Padi) field (C5), in the study area, the following were found: open surface water present in the Balikpapan Bay; shrimp ponds; rice (*padi*) fields; rivers; and an artificial lake. Ponds are created mainly inside a mangrove forest and are the cause of destruction of this important type of vegetation. These ponds are used to produce shrimps or fish. The conversion of mangrove forests to ponds is a common practice, not only in this area, but also in general all over Indonesia (Zuhair *et al*., 2001).

Urban area (C6), this class represents the settlements along the main roads, Balikpapan Bay harbours, industrial oil palm production areas, and transmigration areas in the south part of the study area

Bare soil (C7), this class, which is characterised as having no vegetation, includes main roads, abandoned land, and logging trails.

Oil palm plantations (C8), also mainly found in the south part of the study area within zones designated for oil palm plantation as a mono-crop. The areas designated for coconut plantations were in general fewer and smaller in size compared to those areas dedicated to rubber and oil palm production. They were usually located adjacent to the coast, and maintained as big gardens near a village.

Rubber plantation (C9), these easily identifiable zones are mainly found in the south part of the study area, and are designated for rubber plantation as a mono-crop. Rubber plantations in different stages of development were found covering large areas, with trees ranging from 15 to 25 m in height. The ground in areas occupied by rubber was completely covered by a plant belonging to the *Leguminaceae* family.

Mangrove forest (C10), this class is found along the Balikpapan Bay and occurs along the river corridors with seawater influence. Mangrove forest occupies the coastline and continues between 200-500 m into the main land. In the study area *Rhizopora sp.* is the most dominant, building a homogeneous canopy of 5-15 m in height. This type of vegetation constitutes the base of an important ecosystem that has been demonstrated to have a high productivity and an essential role in supplying organic materials to marine ecosystems (Mougin *et al*., 1999).

Secondary forest (C11), the secondary forests are areas initially covered by primary forest, but afterwards have been continuously degraded by selective cutting, intensive logging or fires during the last decades. The vegetation is characterised by a community of trees with average height between 4 up to 20 m and has standing trunks as remnants of forest fires, as well as some isolated high trees that survive.

Primary forest (C12), the type of vegetation included in this class is the natural vegetation that remains in the region as the original lowland forest. It still has its structural and functional pristine characteristics: high density, tall trees with a large diameter, large number of species and several layers. Its presence is a consequence of a combination of factors like topography, geology, climate, as well as a low or absent influence of humans in the area. This vegetation was found in the protected areas known as Sungai Wain, Bukit Bangkirai and Wartono Kadri. This kind of forest is typically three layered, with an upper layer of 30 to 40 m in height (Sidiyasa, 2001).

Burnt primary forest (C13), during the past fire events, not only secondary vegetation was burnt away, but also primary forest. Its main characteristic is the presence of big living trees, remains of the original vegetation, with scorched standing trunks and the presence of light demanding species such as *Macaranga spp.* Another characteristic of burnt primary forest is that many primary forest plant families are suppressed, the most important one being *Dipterocarpaceae*, which is a very important family of this kind of forest in this region of the world (Keßler, 2001). Species of this family disappear after fire because they burn readily due to their relatively thin bark, high content of flammable oleo-resins, and because of their inability to sprout.

3.2.3 Land cover class description of the independent validation set

The following descriptions of land cover type are valid for the independent data set, which has been collected by Rodriguez (2002) in September 2001, using a different legend.

Alang-alang (equals C2), an identical definition was used

Shrubs and shrubs-trunks (2 classes), these classes are characterized by very dynamic vegetation with an average height between 1 to 3 m and the presence of some pioneer trees. This type of vegetation appears as a result of the succession or natural recovery of severely degraded areas that in many cases were probably reduced to the presence of only grasses. The difference between these two classes is based on the presence or absence of standing trunks, which mainly are remnants of fire events that made both classes easy to distinguish.

Coconut plantation and rubber plantation (2 classes), located mainly in the south part of the study area. This area has zones dedicated to maintain coconut palms and to plant and produce rubber as a mono-crop. The areas dedicated to coconut were in general fewer and smaller in comparison with those dedicated to rubber production. Usually they were adjacent to the coast and maintained as big gardens near to a village. Rubber plantations (*Leguminaceae* family) were found covering big areas and in different stage of development, with trees ranging from 6 to 9 m in height. Note that the rubber in the training set features heights in the range of 15 to 25 m.

Primary forest (equals C12), basically the same definition is used. In Rodriguez (2002), additional information is given for Sungai Wain. This forest is located only 15 km north of Balikpapan city, along the Balikpapan – Samarinda road. According to Sidiyasa (2001), the Sungai Wain forest has a density of 535 trees ha^{-1}, with a basal area of 23.5 m^2 ha^{-1}. The composition of a plot, with an area of 3.6 ha, showed the presence of 385 tree species belonging to 143 genera and 49 families. The most important families are *Euphorbiaceae, Lauraceae, Myristicaceae, Myrtaceae, Dipterocarpaceae, Fagaceae, Leguminoseae, Burseraceae, Annonaceae* and *Anacardiaceae* (Sidiyasa, 2001). This kind of forest is typically three-layered, with an upper layer of 30 to 40 m height.

Burnt primary forest (equals C13), basically the same definition is used.

Secondary forest and secondary forest, wet, (2 classes), labelled as secondary forest, are those areas that were initially covered by primary forest but that have been continuously degraded by selective cutting, intensive logging or fires during the last decades with vegetation characterized by a community of trees with average height between 10 and 20 m. Wet secondary forest may form a separate category. It also has been subject to intervention, but differs from the former secondary forest category by being associated to wet soils, usually present in the north part of the study area in V-shaped narrow valleys.

Secondary forest, young, and secondary forest, young-trunks (2 classes), secondary forest, young, includes those areas that were severely degraded, and, as a consequence of the succession, have recovered. High shrubs colonise them and trees with an average height of around 4 to 9 m. A variation of this class is "secondary forest, young-trunks*"*, which shares all the characteristics mentioned above, but has standing trunks as remnants of fire events and some isolated high trees that survive.

It is noted that the latter 4 classes, in fact, comprise the class C11 of the training set in section 3.2.1.

Secondary forest old, this kind of forest was defined taking into account the height of the trees; it was located in areas protected from forest fires during the last years. The average height of the trees was of at least 20 m with emergent trees reaching a height of 30 to 35 m. It was found that almost all early colonizers such as *Macaranga* had disappeared and the tree species composition was typically that of a secondary forest.

Mangrove forest and dead mangrove forest (2 classes), these two classes were found along Balikpapan Bay and the estuaries of the rivers that flow into it. The mangrove

forest is occupying the coastline and is growing into the mainland as far as 200 to 500 m. The main genus is *Rhizophora* and the height of the trees is about 10 to 15 m.

Mangrove forests provides human communities with food resources and other products like firewood, timber, charcoal. This forest is being destroyed to create shrimp ponds, and as a part of this transformation process there is a stage in which areas with dead mangrove trees are found in the study area, covering considerable areas which were classified as dead mangrove.

Swamps, swamps were defined as those zones located in low parts of the study area, usually adjacent to rivers with flooded soil. Those areas were found usually covered by rice crops or grasses and in some cases there are a few small shrubs. The vegetation in general does not exceed 1 m height.

Mixed (equals C4), basically the same definition is used.

Water or pond (equals C5), basically the same definition is used.

3.2.4 Intensive measurement of field data

The intensive observations for the purpose of biomass measurement were made for 9 primary forests un-burnt, 9 primary forests burnt, 4,000 m^2 each and 7 secondary forests of 8,000 m^2. In addition, vegetation characteristics were collected for 15 plots of *alang-alang* with varying degrees of bush invasion. The method applied to estimate forest biomass differs from the method used for *alang-alang.*

Based on literature (Brown *et al*., 1989) the relationship between the total above ground biomass estimation, the diameter at breast height *dbh* (or *D*) (m) and tree top height *H* (m) is as follows:

$$\hat{Y} = \exp\left\{-3.3012 + 0.9439 \ln\left(D^2 H\right)\right\} \qquad [\text{kg tree}^{-1}] \qquad (3.1)$$

With this allometric relationship it was possible to calculate the total above ground biomass estimation for the tropical forest types studied, using the collected ground data. A conversion was made to calculate the total above ground biomass estimate in ton ha^{-1} from the tree biomass of individual trees present.

The intensive observations to measure *dbh* and tree top height for the purpose of biomass measurement were all made in so-called super-transects of 8,000 m^2 area. For the primary forest, the area was divided into a burnt and un-burnt part (each 50%), since all primary forest transects were located at the sharp transition between undisturbed and fire-affected areas. Each super-transect was divided into 8 transects (Figure 3.2). The origin (*x, y*) = (0, 0) was chosen in the middle. Along the negative *x*-axis transects a, b, c and d were situated, and along the positive *x*-axis transects e, f, g and h. For all 9 super-transects in the Sungai Wain forest reservation, the middle was situated on the border between un-burnt and burnt primary forest. The heights of the terrain and the trees in a super-transect were all relative to the height at location (*x, y*) = (0, 10), which was arbitrarily set to $z = 0$. Transect drawings were made for four

transects in once, combining the left (a, b, c, d) and the right (e, f, g, h) transects into two separate drawings of the super-transect.

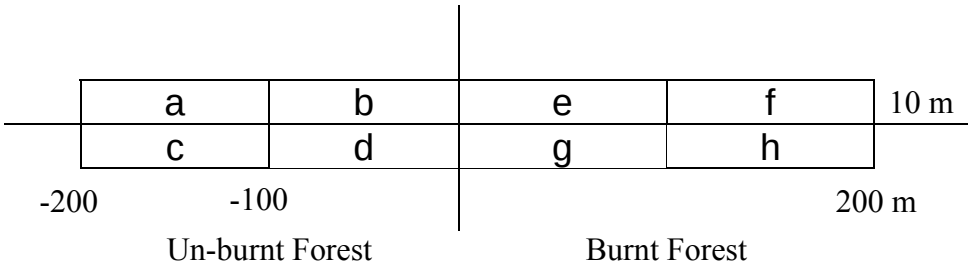
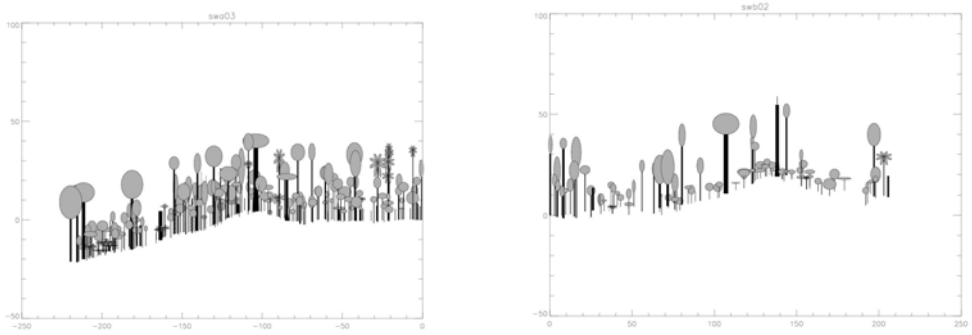


Figure 3.2. Name convention for transects within super-transects in study area, each block 10 x 100 m.

All trees with a diameter at breast height (*dbh*) of 10 cm or more (if buttresses were present, the diameter was measured about 30 cm above the buttress) were measured in each transect, including the *dbh*, tree height, height to the first living branch and crown size. All trees in the subplots were numbered permanently using aluminum tags. Wooden poles were placed in grids of 20 meters, for making elevation corrections. Terrain height of these grid points was measured relative to one of the grid points with an accuracy +/- 0.5 m.

Tree stem position was measured using tape, compass and clinometer, and calculated with a basic triangular method. The azimuth angle of a tree was measured from two reference points within the chosen strip line from the grid. Tree stem height (on the ground) was derived from interpolation of the height from the poles in the grid. Electronic distance measurements were made in four directions, to determine the tree crown projection. Tree heights were measured relatively to the stem base using a clinometer. Object related errors could occur in dense stands, where tree top height or the stem base was not clearly visible because of the presence of a dense canopy or shrubs, respectively. Another error could occur in circumstances when the tree was leaning, either towards or away from the observer, and also because of bad visibility, faulty instrument operation, and incorrect techniques in taking readings.



(a) (b)

Figure 3.3. Example of transects drawing (a) un-burnt and (b) burnt primary forest.

The diameters of the trees were measured at 1.3 m height (*dbh*) using diameter tape. In this case errors could occur when the tape was not placed exactly around the measurement plane, perpendicular to the stem. This caused an overestimation of the *dbh*, which is in order of 0.5% for a tilting angle of 5%. Another error could occur when cross sections were not circular, resulting in an overestimation of the basal area and biomass estimation. Figure 3.3 shows a sample result of the transect drawing. More details of ground data collection can be found in (Hoekman *et al*., 2001).

The biomass of the *alang-alang* was estimated by cutting and weighing all vegetation within some small sample areas within these plots. This sampling procedure was executed by taking a random location within a 3 x 1 m^2 grid area. First the square meter with the highest biomass was chosen (visually); the biomass was then cut and weighed (kg m^{-2}) using pruning-shears, scissors, plastic bags and steelyard. Random sampling was repeated twice, to measure the square with the medium biomass, and finally the square with the lowest biomass (Table 3.2).

Biophysical characteristics in each transect using the above measurement procedures are summarised in table 3.3, namely: averaged dominant height [m] (i.e. from 10 x 10 m^2 blocks), biomass [ton ha^{-1}], basal area [m^2 ha^{-1}], density of trees [ha^{-1}] and number of species.

Table 3.2. *Alang-alang* biomass (Hoekman *et al*., 2001).

Cover Type	Soil Condition	Terrain	Weight (kg m^{-2})			Height (cm)
			High	Med	Low	
Alang-alang	dry	flat	1.50	1.00	0.50	80
Alang-alang	dry	flat	1.50	1.50	1.00	90
Alang-alang	dry	flat	2.00	1.75	1.50	150
Alang-alang/Shrub	wet	flat	7.00	5.00	3.00	250
Alang-alang/Shrub	wet	flat	2.50	1.50	1.00	100
Alang-alang	dry	flat	3.00	2.00	1.50	150
Alang-alang	dry	flat	3.00	2.00	2.00	90
Alang-alang/Mixed	wet/flood	flat	5.50	5.50	5.00	130
Alang-alang/Shrub	wet/flood	flat	4.50	2.00	1.75	120
Alang-alang/Mixed	dry	flat	3.00	2.00	2.00	90
Alang-alang	dry	flat	2.00	2.00	1.30	85
Alang-alang	dry	flat	2.50	2.50	2.00	75
Alang-alang	dry	flat	2.50	2.50	2.10	75
Alang-alang	dry	flat	1.50	1.30	1.20	70
Alang-alang/Pepper	dry	flat	3.00	2.75	1.30	95

Table 3.3. Biophysical characteristics of different forest vegetation types (Hoekman *et al*., 2001).

	Height upper canopy (m)	Biomass (tons ha^{-1})	Basal area (m^2 ha^{-1})	Density of trees (ha^{-1})	Number of species
Primary forest					
SWA01	23.93	114.78	19.65	582.50	112
SWA03	25.68	141.88	21.70	545.00	96
SWA05	26.88	141.08	19.80	510.00	92
SWA07	24.70	199.60	28.78	525.00	107
SWA09	25.88	112.03	18.10	527.50	88
SWA11	20.60	100.05	17.20	547.50	106
SWA13	21.73	130.63	21.38	642.50	109
SWA15	28.80	121.65	17.35	702.50	109
SWA17	30.75	141.03	19.90	845.00	109
Primary forest (burnt)					
SWB02	17.50	84.23	11.58	612.50	63
SWB04	20.75	113.93	17.05	372.50	76
SWB06	18.53	100.50	16.13	367.50	61
SWB08	9.68	48.63	8.65	215.00	45
SWB10	18.33	115.73	17.05	287.50	53
SWB12	12.73	79.78	12.80	297.50	63
SWB14	10.10	68.03	9.93	185.00	26
SWB16	15.28	52.10	7.13	205.00	56
SWB18	8.70	31.58	4.23	135.00	31
Secondary forest					
WAN01	10.76	25.16	3.71	72.50	38
WAN02	18.65	72.39	12.33	187.50	71
WAN03	24.80	153.18	27.09	547.50	103
WAN04	16.63	75.69	16.39	232.50	83
WAN05	17.14	150.59	28.23	413.75	118
WAN06	1.95	132.53	19.83	400.00	85
WAN07	17.23	119.64	20.25	432.50	114

3.3 NASA/JPL Airborne SAR data

The radar data used in this study comes from the second generation AIRSAR system (AIRSAR, 2000). For investigation of the C-, L- and P-band fully polarimetric (HH, HV, VH and VV polarisation combinations) and interferometric C- and L-band data the AirSAR/TopSAR acquisition of September 14, 2000 was used, covering a 10-km wide and 60-km long stretch. The training and validation set samples consist of 606 delineated areas of at least 100 pixels in a 29^o-61^o range of incidence angles. The field-averaged Stokes scattering element data of the database was used to calculate field-averaged values for backscatter, phase difference and correlation. A database of plot averaged Stokes scattering operator matrix elements was created, which forms

the basis for the analysis. The most important technical specifications of the NASA/JPL AirSAR system and a brief description of nominal sensor parameters and image characteristics are given in table 3.4. A more detailed description of the system can be found in (AIRSAR, 2000). In the header file of the raw data details of the flight and radar data acquisition can be found. All data were processed on the JPL frame processor, which includes absolute radiometric calibration.

Note that the C- and L-band interferometric TopSAR data seems to have a multi-path error (Imel, 2002). If a multi-path signal is present, possibly from signals returning from the scattering scene, which bounce off of wing or engine before arriving at the receiving antenna, then some fraction of the signal will be present and will give height variations with range not corresponding to the scattering topography. This effect was quantified by taking the first few kilometres of the scene, making a horizontal slice through the digital elevation model (DEM), subtracting the mean elevation, and averaging this operation along-track. The multi-path height errors here vary from close to zero to as much as +/-3 meters. The data quality is still being studied at Wageningen University (see also section 3.4) and re-processing by NASA may be necessary to take full advantage of this radar data set.

3.4 Evaluation of DEM data quality and consequences for fusion

When applying the technique introduced in section 2.2.2, a problem with the TopSAR DEM dataset became apparent. It is clear that at least one of the two DEMs contains large errors. When the L-band DEM is subtracted from the C-band DEM, fringes with a height difference of plus and minus 20 m appear (Figure 3.4). It can also be noted from Figure 3.4 that patterns in the hills of Sungai Wain show up, which may be an effect of differences in penetration depth. It seems impossible to warp the PolSAR image accurately onto the TopSAR image. The proposed algorithm may be too rigid to make proper warping since it assumes correct height, and only a displacement in range and azimuth has to be fitted (the flying heights are within a few m only). In fact we are faced with a significant height error which slowly varies in a fringe like pattern. This problem could be solved within a more sophisticated approach, such as the use of local correlation techniques; however this is out of the scope of this research.

Two types of results have been obtained so far.

(1) The first uses a filtered DEM using a circular low-pass filter with a radius of 17 pixels. It produces sharp orthorectified images. The match with TopSAR is very poor (local deviations are up to 10 pixels), which may be a result of the combined effect of DEM errors and the low-pass filtering (which reduces hill top heights and increases valley heights). The errors seem to be equally severe when using the C-band DEM or the L-band DEM.

(2) The best result is obtained when no DEM filtering is applied. The warped PolSAR image is not sharp anymore; however the spatial errors seem to be limited to 3-4 pixels.

Therefore it may be concluded that (a) the analysis is still well possible at the level of polygons and (b) the warped PolSAR image (i.e. version 2) can still be linked to the (filtered L-band) DEM to achieve a slope correction. This correction may not be optimal. However, it could increase the classification accuracy considerably. This issue will be studied in Chapters 4 and 5.

Table 3.4. Some relevant AirSAR image specifications.

Parameter	PolSAR			InSAR	
Full polarimetry	C-band	L-band	P-band		
Interferometry				C-band	L-band
Central frequency (GHz)	5.26	1.25	0.44	5.26	1.25
Wavelength (cm)	5.7	24	68	5.7	24
Slant range sample spacing (m)	3.3	3.3	3.3	3.3	3.3
Azimuth range sample spacing (m)	4.6	4.6	4.6	4.6	4.6
Range pixel spacing (m)	3.3	3.3	3.3	5	5
Azimuth pixel spacing (m)	4.6	4.6	4.6	5	5
Height resolution for low-relief terrain (m)				1-3	1-3
Height resolution for high-relief terrain (m)				3-5	3-5
Approx. DEM resolution (m)				10 x 10	10 x 10
Platform	DC-8 aircraft				
Radar altitude (m)	8439.9				
Incidence angle range	24.9^{o} to 61.7^{o}				
Processor version	6.10				
Bandwidth	40 MHz				
Approx. SR pixel size (m)	4 x 3				
Approx. SR resolution (m)	5 x 5				
Independent looks per pixel	9				
NE sigma-nought	-45 dB				
Absolute calibration	< 3dB				
Relative calibration between channels	< 1.5 dB				
Relative polarisation calibration within channel	< 0.5 dB				

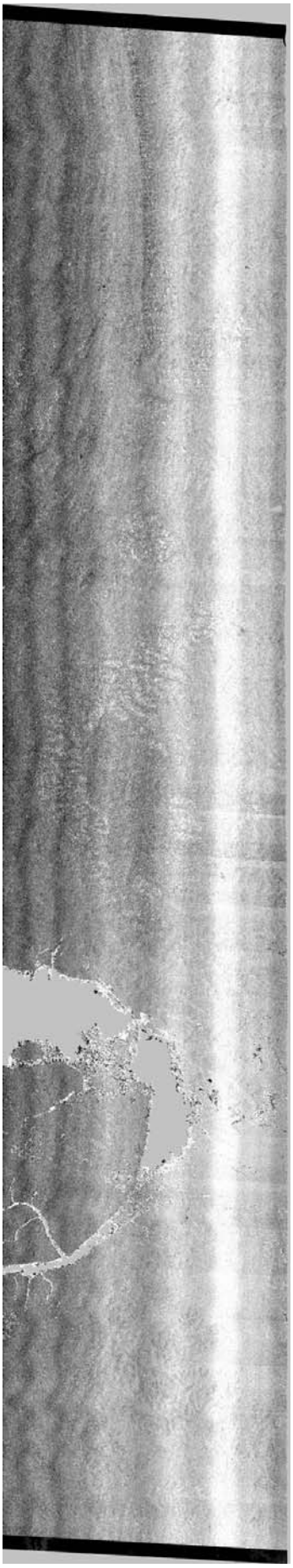


Figure 3.4. *C-band DEM height minus L-band DEM height*. Fringes with a height difference of plus and minus 20 m appear, which is supposed to be a result of multi-path fading for at least one of the frequency bands. It can also be noted that patterns in the hills of Sungai Wain show up, which may be an effect of differences in penetration depth.

4. Biophysical parameter retrieval and land cover type classification

4.1 Introduction

The theory and approaches discussed in Chapter 2 will be applied using high resolution radar data acquired by the AirSAR/TopSAR airborne SAR system. The field data discussed in Chapter 3 will be used for training and validation.

The main objective of this chapter is to compare different approaches of biophysical parameter retrieval and land cover classification. Main differences between approaches relate to options to include relief correction, to include 3-D textural information, or to exclude a number of frequency bands.

Total Power information is used to obtain clues/indications regarding the significance of the relief correction (Chapter 4.2).

Subsequently fully polarimetric classification simulations will be performed to establish the relative importance of certain frequency band and polarisation combinations, and the effect of the number of independent radar looks (Chapter 4.3).

A first classification of the full multi-frequency will be executed to evaluate the appropriateness of the legend introduced in Table 3.1 (Chapter 4.4).

3-D textural information, derived from the high resolution InSAR DEM, will be evaluated for its utility to produce additional information (Chapter 4.5).

Finally, the impact of tuning parameters of the Iterated Conditional Modes (ICM) methodology, the inclusion of relief correction and 3-D textural information and the exclusion of frequency bands will be evaluated (Chapter 4.6).

In addition, empirical relationships between biomass and backscatter characteristics will be studied (Chapter 4.7).

To deepen the understanding of the physical interpretation of the results, the relation between several forest structure characteristics and multi-frequency complex coherence signatures as introduced by Hoekman and Quiñones (2002) will be studied (Chapter 4.8).

4.2 Relief correction

A multi-band C-, L- and P-band PolSAR image is shown in Figure 4.1a. This combination gives excellent possibilities for visual discrimination. The radar platform was moving from the bottom to the top of the image and was looking towards the left side; the radar incidence angle varies from 24^{o}-71^{o} from the right to the left of the image. The interferometric SAR derived DEM is shown in Figure 4.1b as a grey scale image. The pixel spacing is 5 m by 5 m, and the image shown has 2385 x 12211

pixels. The DEM derived from C-band InSAR (Figure 4.1b) enables compensation for the relief effect of the co-registered C-, L- and P-band PolSAR data. Slope angles in range direction were computed from the DEM. Equation 2.1 was then applied to obtain the relief corrected image in Figure 4.1c. A new multi- band composite was created, and the result shows that the illumination effect of sloping ground or canopy undulation, which can confuse the target identification for classification, has been reduced, and in general the image has become more uniform and homogeneous (Figure 4.1c). The corrected PolSAR image is used as a base image to retrieve biophysical parameters, and for classification.

The Total Power (TP) is the sum of backscatter in all polarisations. Within the research area, the total power of the land cover type can be interpreted visually from the BGR (Blue, Green, Red) colour of the pixels; Blue represents the C-band; Green represents the L-band; and Red represents the P-band.

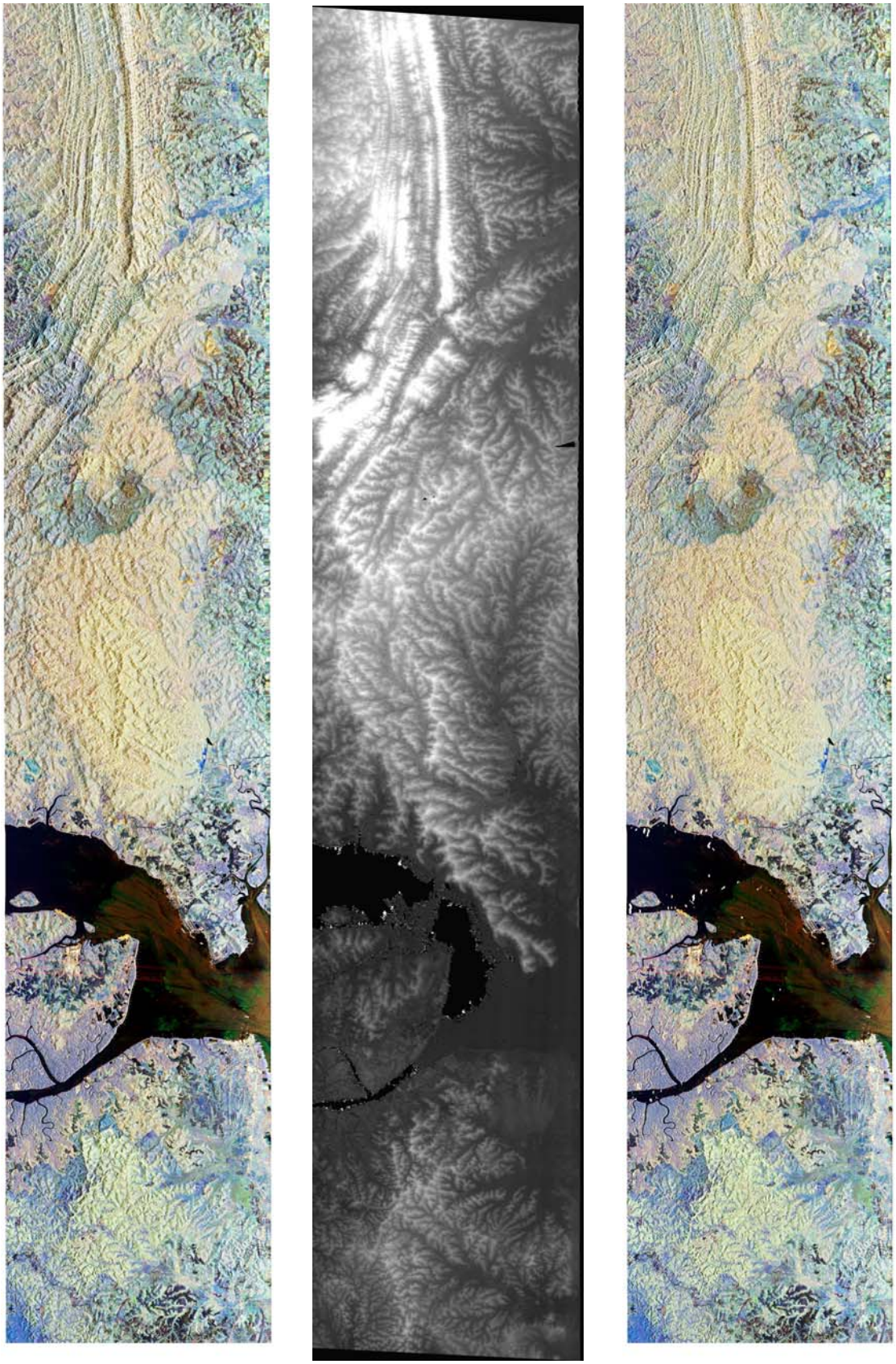
In the total power C-, L- and P-band image (Figure 4.1c) the bare soil shows up dark; *alang-alang*, grasslands (as well as the transition of *alang-alang* into shrubs) and shrubs in blue tones. The *alang-alang* (*Imperata cylindrica*) covers vast areas throughout the whole study site. The main factor for this variation seems to be the degree shrub invasion in *alang-alang* fields. The higher the shrub density, the higher the backscatter level. Some degraded grasslands can be recognised by their heterogeneous appearance. Mangrove appears as dark purple. The oil palm plantation appears as a mixture of yellow and red pixels. A rubber plantation appears as light green with a rough texture. Water, *padi* (rice) fields and shrimp ponds appear in very dark tones, in contrast to urban areas, which appear in very bright tones. Primary forest appears in yellow bright tones. Burnt primary forests appear as a mixture of yellow and red bright tones, with the red colour being slightly darker than the red of the oil palm plantation. Secondary forest appears in mostly yellow and red pixels with a few blue pixels.

Though the spatial resolution is only around 5 m, the availability of full polarimetry in three bands yields an unprecedented view of individual tropical rain forest trees. The image indicates that many groups of trees can be recognised. The trees that are most clearly observed by the radar are the large trees of the upper canopy, which are all in a mature stage. It seems that, since radar is sensitive to structure and the mature trees have a fairly well defined structure, species-related information may be obtained. Large dead standing trees, many of which remained after the 1998 El Niño fires, are clearly recognisable as bright yellow dots, because of high L- and P-band backscatter levels and very low C-band backscatter level (Figure 4.2).

The theoretical relationship between backscatter modulation and relief was tested quantitatively and qualitatively. For this purpose, several land cover types were selected: mangrove, oil palm plantation, primary forest and burnt primary forest. Qualitative comparison of radar total power (TP), before and after slope correction, for a selection of four land cover types, were made in C-, L- and P-band. The backscatter ranges for these four land cover types are shown in Figure 4.2. These results show that for the 41 mangrove samples the TP of C-band has a smaller range than for L- and P-band. However, the differences before and after slope correction are small, because of the flat topography and a homogeneous canopy. For oil palm plantations, the C-, L- and P-band backscatter ranges are smaller than for mangroves.

However, after slope correction, these ranges become even smaller, because this land cover type occurs also on undulating terrain. For the primary forest and burnt primary forest cover classes, which appear on undulating terrain, the backscatter range is large in all frequency bands, and this range diminishes considerably after slope correction. This general behaviour qualitatively confirms the assumption that relief induced modulation can be reduced after slope correction.

Quantitative comparisons of radar intensity or total power (TP) extracted from the AirSAR image before and after slope correction in C-, L- and P-band areas of mangrove, oil palm, primary forest and burnt primary forest were made and are shown in Table 4.1. The average incidence angle of the plots range from 28.6^{o} to 61.5^{o}. It can be seen that for mangrove, the TP ranges in C-, L- and P-band before and after slope correction show very small differences, and that the same applies to the standard deviations. As was mentioned before, this is caused by the flat topography and homogeneous canopy. For oil palm plantations, the C-band backscatter has a larger range than mangrove, because of a more irregular canopy architecture and terrain undulation. After slope correction this difference almost disappears. In general, for all classes on non-flat terrain, the range of TP backscatter and its standard deviation becomes smaller (i.e. more homogenous) after slope correction.



(a) (b) (c)

Figure. 4.1 (a) The C-, L- and P-band orthorectified total power composite image of the research area. The image size is 2385 x 12211 pixels. (b) The C-band VV interferometric SAR generated DEM. The pixel spacing is 5 m x 5 m. (c) The C-, L- and P-band relief corrected total power composite image.

(a)

(b)

(c)

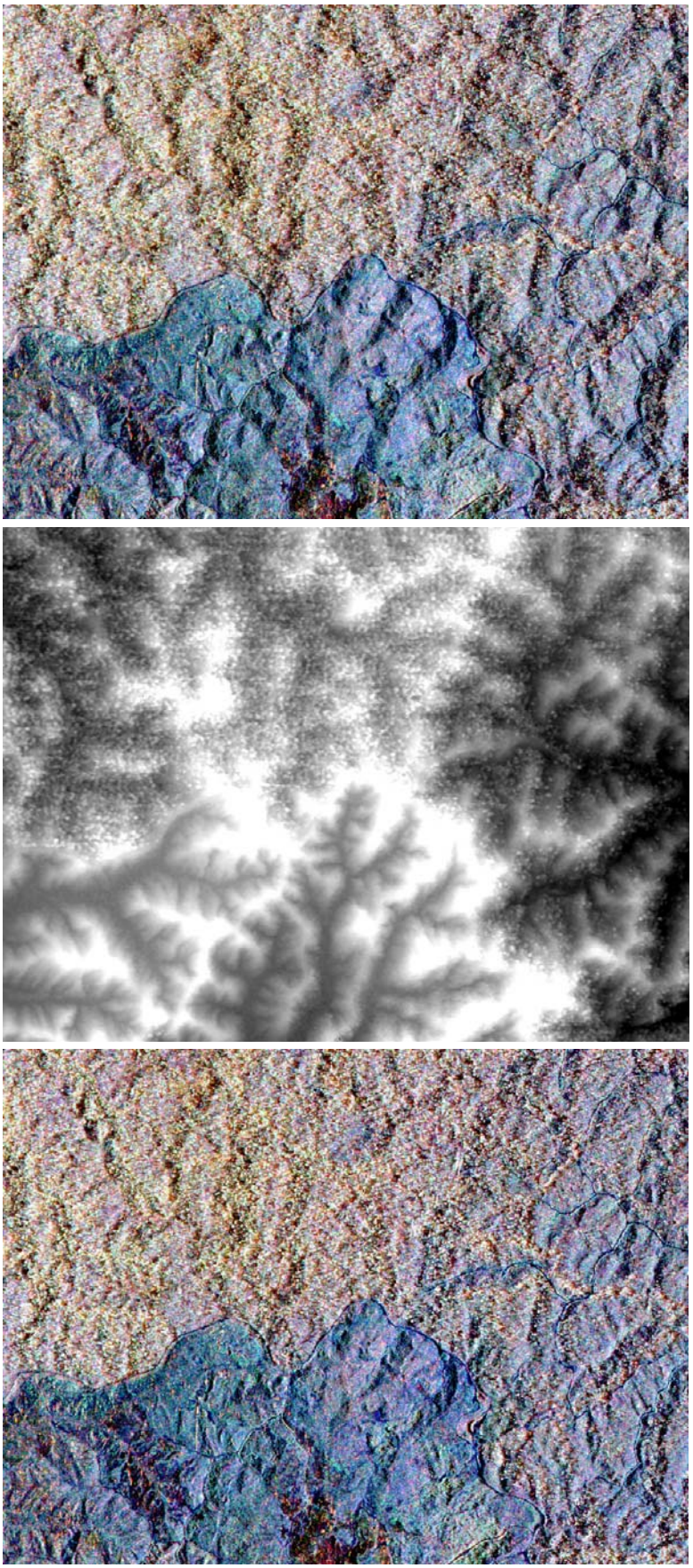
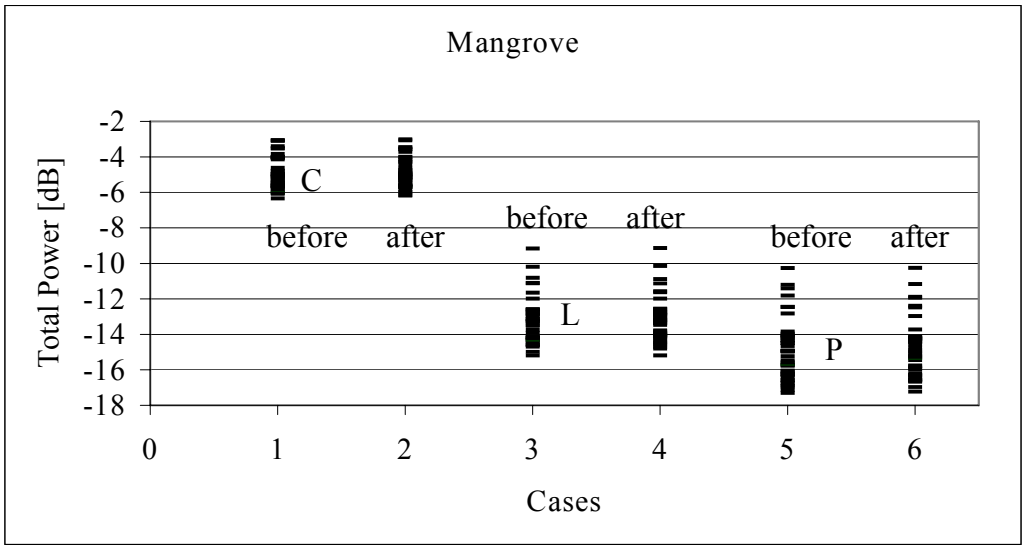
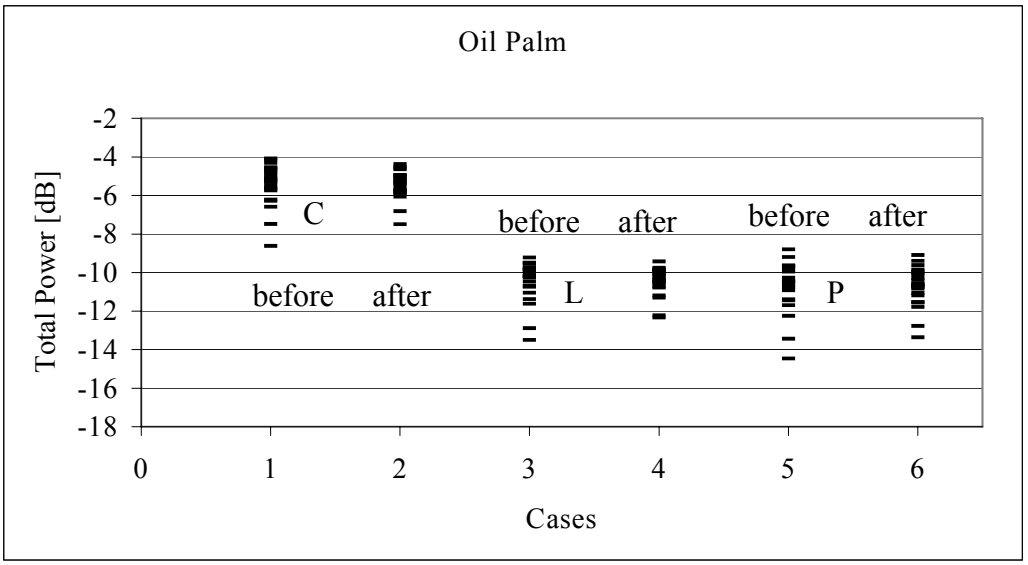


Figure 4.2 Detail (part of the study area) of (a) C-, L- and P-band orthorectified total power composite image of the research area. (b) The C-band VV interferometric SAR generated DEM. The pixel spacing is 5 m x 5 m. (c) The C-, L- and P-band relief corrected total power composite image.



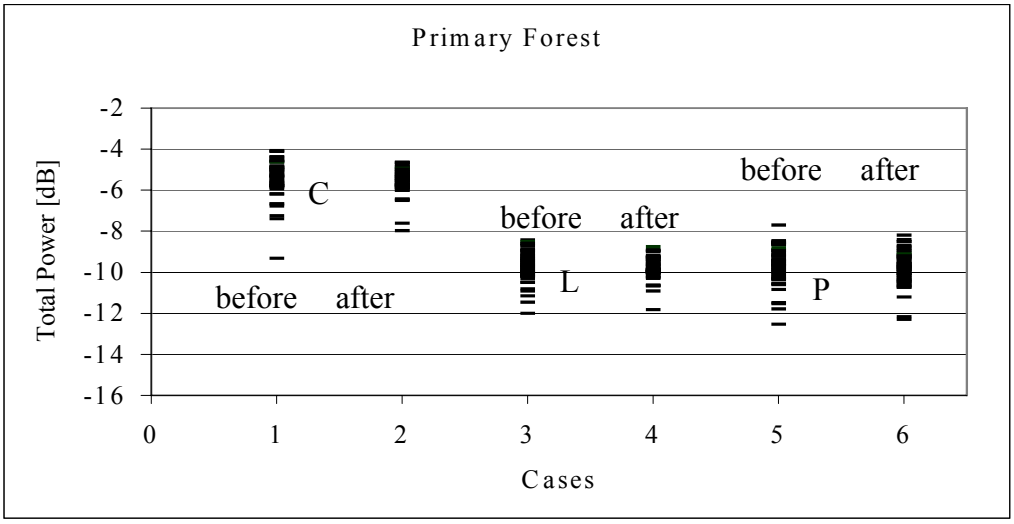
(a)

Figure 4.3a. Radar intensity (TP) extracted from the orthorectified AirSAR image *before and after slope correction* in C-, L- and P-band areas in (a) Mangrove (41 samples or polygons). Cases: (1) C-band before correction, (2) C-band corrected, (3) L-band before correction, (4) L-band corrected, (5) P-band before correction and (6) P-band corrected.



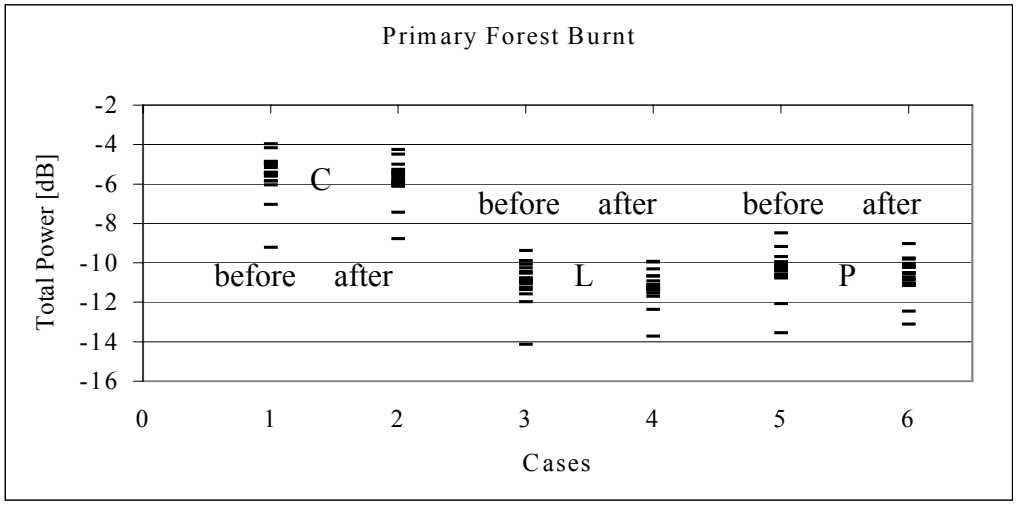
(b)

Figure 4.3b. Radar intensity (TP) extracted from the orthorectified AirSAR image *before and after slope correction* in C-, L- and P-band areas in (b) Oil palm (22 samples or polygons). Cases: (1) C-band before correction, (2) C-band corrected, (3) L-band before correction, (4) L-band corrected, (5) P-band before correction and (6) P-band corrected.



(c)

Figure 4.3c. Radar intensity (TP) extracted from the orthorectified AirSAR image *before and after slope correction* in C-, L- and P-band areas in (c) Primary Forest (52 samples or polygons). Cases: (1) C-band before correction, (2) C-band corrected, (3) L-band before correction, (4) L-band corrected, (5) P-band before correction and (6) P-band corrected.



(d)

Figure 4.3d. Radar intensity (TP) extracted from the orthorectified AirSAR image *before and after slope correction* in C-, L- and P-band areas in (d) Primary Forest Burnt (19 samples or polygons). Cases: (1) C-band before correction, (2) C-band corrected, (3) L-band before correction, (4) L-band corrected, (5) P-band before correction and (6) P-band corrected.

Table 4.1 Range of radar intensity (TP) and standard deviation (SD) *before and after slope correction* for C-, L- and P-band areas in Mangrove (41 samples of polygons), Oil palm (22 samples), (c) Primary Forest (52 samples) and (d) Primary Forest Burnt (19 samples).

Classes	Band	TP range (dB)		SD		Inc. Angle	
		Before	After	Before	After	Min	Max
Mangrove	C	3.2881	3.1625	0.9337	0.8594		
Mangrove	L	6.0292	6.0561	1.3401	1.2855	28.56	61.53
Mangrove	P	7.0266	6.9716	1.7292	1.6610		
Oil palm	C	4.5434	3.1173	1.0968	0.7410		
Oil palm	L	4.2845	2.9179	1.0814	0.7499	46.67	52.52
Oil palm	P	5.6694	4.2696	1.3380	1.0290		
Primary Forest	C	5.2478	3.3235	0.9406	0.6441		
Primary Forest	L	3.5782	3.0696	0.7493	0.5474	43.72	56
Primary Forest	P	4.8262	4.1111	0.9257	0.8396		
Primary Forest Burnt	C	5.2485	4.5236	1.1192	0.9828		
Primary Forest Burnt	L	4.7445	3.7777	0.9758	0.8397	45.03	57.01
Primary Forest Burnt	P	5.0615	4.0980	1.0362	0.9232		

Another point of concern needs to be mentioned. For C-band the analysis has been done on modified data. The original data showed an artefact (dark band) on C-VV, which appeared as an abnormally dark strip, and which heavily contaminated significant features (Vargas, 2002; Hoekman *et al*., 2001). In order to retain geophysical information across this strip, a *radiometric balancing* algorithm was used to adjust the brightness relative to the neighbouring and the C-HH intensity values. It should be noted that this step does not affect polarimetric phase difference and coherence.

4.2.1 Conclusions

The main conclusions derived from the relief correction analysis using total power information above are:

- Multi-band total power composite images of C-, L- and P-band PolSAR allow visual interpretation of land cover. Because of the high resolution they also *allow for assessment of forest structures, large emergent trees, and sometimes individual trees, especially in areas which have fire damage.*
- The *relief correction reduces illumination effects from the terrain slope and canopy undulation*, and produces images that *visually appear more uniform and homogenous* in character.
- The quantitative comparison of the total power values of several land cover types, before and after relief correction, confirms that:
    - Mangrove shows no difference because of the flat topography and its homogenous canopy.
    - Oil palm shows a slight difference because of the slight undulation.
    - Primary forest burnt and un-burnt, which are located on undulating terrain and have an undulating canopy, reveal significant differences, showing the significant accomplishment of relief correction.

4.3 Land cover type classification simulation

The simulated classification results for the 13 land cover types of Table 3.1 using several single and multi-channel configurations are shown in Table 4.2. The level of confidence is 95%, which means that roughly 5% of the data are classified as the class "unknown" and results far in excess of 95% are unlikely. In this table the accuracy is expressed as the percentage of well-classified pixels.

Using the 7I model at the '1 dB' level (i.e. 20-look data), *and the slope correction*, the classification simulation of a single-channel showed a very low overall classification accuracy, ranging from 27.3% for C-VV to 36.1% for L-HH. In general the C-band channel combinations obtained the lowest accuracy, while the P-band channel combinations obtained the highest accuracy; however, they are not significantly different. Combinations of channels from two frequency bands showed an increase in the overall classification accuracy when compared with the results for only one frequency band. The values ranged from 31.9% for the combination of C-HV and C-VV to 64.9% for the combination C-pol and P-pol. Nevertheless, these values can be considered low, which could indicate that the information provided for carrying out the classification was insufficient to successfully separate the samples in each of the respective classes observed during the field work. Consequently, a simulation was performed for the same level of speckle but utilizing more information, in this case a three-channel combination, i.e. a simulation using full polarimetry and three bands (C-, L- and P-band). The overall classification accuracy after slope correction obtained was 71.3%. The fully polarimetric overall classification result for the combination of C-, L- and P-band was good (71.3%), C- and P-band (64.9%) performed slightly better than the C- and L-band (60.5%). In general, the results of 64-look data (0.5 dB level) are better than that of 20-look data, and after slope correction the classification results are slightly improved.

It is interesting to compare the performance of the different fully polarimetric classification models 7I and 3I+ at the 0.5 dB and 1 dB speckle level, as shown in Table 4.3. In general, the 7I system is better, especially for the C-, L- and P-band fully polarimetric combination. A slightly lower result is obtained using the 3I+ model with dual band combinations. Results for single-band polarisation combinations using the 3I+ model are slightly better than for the 7I model, especially at the 0.5 dB speckle level. It may be noted, as explained in Chapter 2, that in the 3I+ model only the co-variances between the three intensity values are taken into account (and phase information is treated independently), while for the model 7I all co-variances (between 7 intensities) are taken into account. Apparently, for this simulation, the advantage of having a more sophisticated model, i.e. the 7I model, is especially evident when many information channels are used. As expected, the results of 64-look data (0.5 dB level) are better than that of 20-look data.

Table 4.2. Comparison of classification simulation results *before and after slope correction* for the 7I algorithm for single and multiple channels. Overall Maximum Likelihood (ML) classification accuracy (expressed in percentages) at the 95% level of confidence. Single and multi-channel fully polarimetric classification results are shown for speckle levels of 1 and 0.5 dB, (i.e. N=20, N=64, respectively), together with the statistic $\hat{K}$ and its large sample variance $\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$. Note: 'int' is HH+VV+HV polarisation; 'pol' is full polarimetry.

Channel	N=20	Non-corrected		N=20	Corrected		N=64	Non-corrected		N=64	Corrected	
	Result	$\hat{K}$	$\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$	Result	$\hat{K}$	$\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$	Result	$\hat{K}$	$\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$	Result	$\hat{K}$	$\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$
C-,L-,P-pol	70.6	0.6818	0.0006	71.3	0.6894	0.0006	79.2	0.7756	0.0004	85.3	0.8392	0.0003
C-,L-,P-int	63.3	0.6032	0.0006	64.2	0.6132	0.0006	71.6	0.6927	0.0006	72.2	0.6997	0.0005
C-pol,P-pol	64.0	0.6113	0.0006	64.9	0.6201	0.0006	70.7	0.6835	0.0006	72.0	0.6981	0.0005
L-pol,P-pol	62.2	0.5907	0.0007	64.3	0.6136	0.0006	70.9	0.6858	0.0006	72.7	0.7042	0.0005
C-pol,L-pol	60.0	0.5672	0.0007	60.5	0.5715	0.0007	67.9	0.6528	0.0006	67.7	0.6509	0.0006
C-int,L-pol	57.9	0.5448	0.0007	57.5	0.5412	0.0007	65.7	0.6298	0.0006	65.7	0.6294	0.0006
P-pol	52.8	0.4881	0.0007	54.2	0.5032	0.0007	58.9	0.5550	0.0007	60.4	0.5712	0.0007
P-int	49.2	0.4473	0.0007	51.0	0.4665	0.0007	53.5	0.4937	0.0007	56.0	0.5213	0.0007
L-pol	49.0	0.4470	0.0007	50.0	0.4592	0.0007	55.5	0.5185	0.0007	57.9	0.5443	0.0007
L-int	45.3	0.4057	0.0007	45.7	0.4110	0.0007	49.8	0.4551	0.0007	51.1	0.4690	0.0007
C-pol	38.7	0.3365	0.0007	39.6	0.3465	0.0007	43.3	0.3875	0.0007	44.6	0.4000	0.0007
L-HH	36.1	0.3006	0.0006	36.2	0.3037	0.0006	38.7	0.3299	0.0007	39.3	0.3383	0.0006
C-int	35.0	0.2951	0.0006	35.2	0.2991	0.0006	39.3	0.3421	0.0007	40.1	0.3513	0.0007
C-HH	29.4	0.2295	0.0006	32.3	0.2557	0.0006	29.8	0.2388	0.0006	35.0	0.2839	0.0007
C-VV	27.3	0.2112	0.0006	31.9	0.2567	0.0006	28.6	0.2304	0.0006	31.4	0.2522	0.0006
C-HH-HV	32.6	0.2683	0.0006	32.7	0.2714	0.0006	36.9	0.3152	0.0007	37.2	0.3197	0.0007
C-VV-HV	30.3	0.2530	0.0006	31.4	0.2608	0.0006	34.7	0.2989	0.0006	36.2	0.3110	0.0006
C-VV-HH	30.7	0.2476	0.0006	32.5	0.2628	0.0006	33.2	0.2751	0.0006	35.5	0.2953	0.0007
L-VV	33.6	0.2741	0.0006	36.0	0.3023	0.0006	37.0	0.3103	0.0006	37.1	0.3133	0.0006
L-HH-HV	42.1	0.3698	0.0007	41.7	0.3672	0.0007	45.8	0.4106	0.0007	46.2	0.4155	0.0007
L-VV-HV	42.1	0.3699	0.0007	41.8	0.3689	0.0007	45.9	0.4118	0.0007	46.1	0.4153	0.0007
L-VV-HH	40.8	0.3543	0.0007	41.4	0.3629	0.0007	46.1	0.4120	0.0007	47.4	0.4271	0.0007

The confusion matrix results for simulated classification at the 1 dB level using the 7I model for the C-, L- and P-band fully polarimetric combination are shown in Table 4.4. The overall accuracy is 71.3% (see Table 4.2). The water, *padi* (rice) field and shrimp pond classes have an accuracy of more than 81.6%. These classes could collectively be considered as water. Urban area has a 93.6% accuracy; bare soil has an 87.2% accuracy. *Alang-alang* has a 55.0% accuracy; this class is mainly confused with shrub and mangrove. Shrubs have a 49.4% accuracy; this class is mainly confused with *alang-alang*, secondary forest and mangrove. Oil palm has a 48.4% accuracy; this class is mainly confused with secondary forest, primary forest, primary forest burnt and rubber. Rubber has a 78.2% accuracy; this class is mainly confused with oil palm, secondary forest and primary forest, and is less confused with shrub, *alang-alang* and burnt primary forest. Mangrove has 85.1% accuracy and is slightly confused with *alang-alang* and shrub. Secondary forest, primary forest and burnt primary forest have an accuracy of 51.7%, 69.8% and 68.3%, respectively. Secondary forest is mainly confused with oil palm, primary forest, and burnt primary forest, and is less confused with rubber, *alang-alang* and shrub. Primary forest is mainly confused with oil palm, primary forest burnt and secondary forest. Primary forest burnt, is mainly confused with oil palm, primary forest and secondary forest.

Table 4.3. Comparison of classification simulation *after slope correction* for the 3I+ and 7I algorithm for single and multiple channels. Overall Maximum Likelihood (ML) classification accuracy (expressed in percentages) at a 95% level of confidence. Single and multi-channel fully polarimetric classification results are shown for speckle levels of 1 and 0.5 dB, (i.e. N=20, N=64, respectively), together with the statistic $\hat{K}$ and its large sample variance $\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$.

Channel	N=20	3I+		N=64	3I+		N=64	7I	
	Result	$\hat{K}$	$\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$	Result	$\hat{K}$	$\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$	Result	$\hat{K}$	$\hat{\sigma}^2_{\infty}\left[\hat{K}\right]$
C-,L-,P-pol	66.9	0.6404	0.0006	74.7	0.7256	0.0005	85.3	0.8392	0.0003
C-,L-,P-int	64.6	0.6156	0.0006	72.1	0.6982	0.0006	72.2	0.6997	0.0005
C-pol,P-pol	62.2	0.5878	0.0007	69.1	0.6644	0.0006	72.0	0.6981	0.0005
L-pol,P-pol	61.7	0.5820	0.0007	69.1	0.6645	0.0006	72.7	0.7042	0.0005
C-pol,L-pol	57.9	0.5403	0.0007	66.2	0.6321	0.0006	67.7	0.6509	0.0006
C-int,L-pol	58.3	0.5444	0.0007	65.9	0.6282	0.0006	65.7	0.6294	0.0006
P-pol	54.2	0.4987	0.0007	60.2	0.5658	0.0007	60.4	0.5712	0.0007
P-int	51.9	0.4724	0.0007	57.0	0.5290	0.0007	56.0	0.5213	0.0007
L-pol	50.9	0.4599	0.0007	57.8	0.5379	0.0007	57.9	0.5443	0.0007
L-int	47.7	0.4244	0.0007	52.7	0.4810	0.0007	51.1	0.4690	0.0007
C-pol	41.2	0.3508	0.0007	46.0	0.4047	0.0007	44.6	0.4000	0.0007
L-HH	38.8	0.3193	0.0007	41.4	0.3493	0.0007	39.3	0.3383	0.0006
C-int	38.7	0.3216	0.0007	43.0	0.3697	0.0007	40.1	0.3513	0.0007
C-HH	34.8	0.2713	0.0007	37.2	0.2985	0.0007	35.0	0.2839	0.0007
C-VV	32.8	0.2481	0.0007	35.2	0.2754	0.0007	31.4	0.2522	0.0006
C-HH-HV	36.7	0.2974	0.0007	40.8	0.3434	0.0007	37.2	0.3197	0.0007
C-VV-HV	36.0	0.2899	0.0007	40.6	0.3423	0.0007	36.2	0.3110	0.0006
C-VV-HH	35.9	0.2870	0.0007	38.5	0.3168	0.0007	35.5	0.2953	0.0007
L-VV	37.3	0.3013	0.0006	41.2	0.3463	0.0007	37.1	0.3133	0.0006
L-HH-HV	44.1	0.3840	0.0007	48.1	0.4296	0.0007	46.2	0.4155	0.0007
L-VV-HV	45.7	0.4006	0.0007	49.7	0.4459	0.0007	46.1	0.4153	0.0007
L-VV-HH	45.9	0.4024	0.0007	48.6	0.4333	0.0007	47.4	0.4271	0.0007

Table 4.4. Confusion matrix for the C-, L- and P-band fully polarimetric combination using a 95% confidence interval for classification using the 20 looks simulated data set.

	1	2	3	4	5	6	7	8	9	10	11	12	13	%	
Unknown	2.7	0.8	2.1	3	2.5	0.7	1.7	1.1	1.6	2.7	2	3.1	1.2		
1	**27.4**	0	0	0	0.1	0	0	0	0	0	0	0	0	99.6	Water
2	0	**10.8**	0.4	0	1.4	0	0.4	0	0	0	0	0	0	83.1	Padi field
3	0	0.1	**18.7**	6.9	0.1	0	0.5	0	0.1	1.7	0.8	0	0	64.7	Alang-alang
4	0	0	6	**15.9**	0	0	0.6	0.1	0.3	1.6	0.8	0.2	0.2	61.9	Shrub
5	0.9	0.1	0.1	0	**26.1**	0	1	0	0	0	0	0	0	92.6	Shrimp pond
6	0	0	0	0	0	**10.2**	0	0	0	0	0	0	0	100	Urban area
7	0	0.2	1.9	0.1	1.8	0	**28.7**	0	0	0	0	0	0	87.8	Bare soil
8	0	0	0	0.5	0	0	0	**10.7**	2.4	0	5.3	5.2	2.6	40.1	Oil palm
9	0	0	0.1	0.6	0	0	0	2.2	**22.6**	0	1.5	1.6	0.1	78.7	Rubber
10	0	0	3.9	1.9	0	0	0	0	0.2	**34.9**	0.1	0	0	85.1	Mangrove
11	0	0	0.7	2.9	0	0	0	3	0.8	0.1	**19.7**	2.5	0.8	64.6	S. forest
12	0	0	0	0	0	0	0	2.7	0.8	0	3.4	**36.3**	1.1	81.9	P. forest
13	0	0	0.1	0.4	0	0	0	2.3	0.1	0	4.5	3.1	**12.9**	55.1	P. forest burnt
%	88.4	90	55	49.4	81.6	93.6	87.2	48.4	78.2	85.1	51.7	69.8	68.3		

In summary, it may be concluded that the classes which feature low classification accuracy are *alang-alang*, shrub and oil palm. The failure to distinguish between the forest types is considerable, with the exception of mangrove.

The level of inability to distinguish between *alang-alang* and shrub is very high. This may be a result of the fact that these classes have similar vegetation types which grow very quickly in uncovered areas.

4.3.1 Conclusions

From the above simulation of land cover type classification *before and after relief correction*, which had, as a purpose, finding optimal band combinations, with the number of looks and models used, and utilizing the confusion matrix to analyse the result and accuracy of each land cover class, the conclusions are the following:

- At the 1.0 dB level (i.e. 20-look data), a single-channel shows a very low overall classification accuracy. When compared to a single-channel a combination of 2 channels shows an increase of the overall accuracy.
- The combination of *3 fully polarimetric bands shows optimal results*. In general, results at the 0.5 dB level (64-look data) are better than at the 1.0 dB level (20-look data).
- The *best result is obtained from the simulation result which uses the 7I model* at the 0.5 dB speckle level.
- *The land cover classes alang-alang, shrub and oil palm yield the lowest accuracy. Other land cover classes yield satisfying results. The failure to distinguish between the forest types is considerable, with the exception of mangrove.*

4.4 Land cover type mapping

From the simulated classification table above, it can be concluded that it is possible to choose the best band combination to make a land cover map. In this case the C-, L- and P-band fully polarimetric combination yields the best result, as could be expected. It is necessary to mention here that simulated classification is merely an approach for pre-evaluation, such as assessing the performance of the different models and band-polarisation combinations, the effect of speckle and the choice of classes. In real images we are dealing with more complex situations, related to factors such as incidence angle dependence, presence of additional not-well-known classes and small features (houses, roads) and disturbing effects of texture or relief. The texture issue will be dealt with in section 4.5. Incidence angle effects and choice of classes will be discussed next.

The distributions used for classification are based on training areas within the 30º-60º incidence angle range of the relief corrected image. As a consequence, training areas of classes which have an incidence angle exclusively outside this range will not be taken into account (e.g. shrimp pond). It needs to be remarked here that *also* for evaluation purposes, only those validation areas inside this incidence angle range are taken into consideration. The effect of incidence angle dependence is ignored. To mitigate possible significant effects of incidence angle (Hoekman and Quiñones, 2000) the backscatter parameter γ is used instead of σ^o ($\gamma = \sigma^o / \cos(\theta_i)$).

Following the evaluation of the confusion matrix (Table 4.4), the land cover type classes have been adapted slightly. The class urban area, because of its very high backscatter, can be classified with very high precision, and can therefore be masked, and will be excluded from further analysis. Also the class of shrimp ponds is excluded, because it is outside the incidence angle range studied. Since the *alang-alang* and shrub classes are not well-defined because of transition stages, these two classes are intentionally separated into three classes: *alang-alang*, mixed (transition *alang-alang* to shrub) and shrub. Consequently, we end up with 12 classes.

Pre-processing steps have been considered. The first is spatial aggregation of 3x3 pixels or 4x4 pixels. Figure 4.5a shows the land cover map based on 4x4 aggregation step, which gave slightly better results than the 3x3 aggregate, because of the higher level of speckle reduction.

Tables 4.5a and 4.5b, respectively, show the percentage and number of pixels of the classification results. The probabilities of a pixel being correctly classified at the 100% level of confidence have been calculated following the technique in Chapter 2.

The overall classification map accuracy is 63.3%. Water and *padi* (rice) fields have an accuracy of more than 90.1%. Mangrove is classified with an 85.7% accuracy. Bare soil and rubber plantations also obtain good results, 77.5%. *Alang-alang* has a 42.1% accuracy; this class is confused mainly with mixed, bare soil, secondary forest and mangrove and is less confused with burnt primary forest, shrub, oil palm and rubber plantations. Mixed has an accuracy of 65.6%; this class is mainly confused with mangrove, *alang-alang* and burnt primary forest and is less confused with secondary forest, bare soil and rubber plantations. Shrubs have an accuracy of 33.6%; this class

is mainly confused with *alang-alang*, secondary forest, burnt primary forest and mangrove. Bare soil has an accuracy of 77.6%; this class is confused mainly with *alang-alang*, *padi* fields and mixed and is less confused with mangrove, secondary forest and burnt primary forest. Oil palm has 50.3% accuracy; this class is mainly confused with burnt primary forest, primary forest and rubber. Rubber has 77.6% accuracy; this class is mainly confused with oil palm plantations, secondary forests and primary forests, and is less confused with mixed, shrub, *alang-alang* and burnt primary forests. Mangrove has 85.7% accuracy and is a little bit confused with mixed, *alang-alang* and shrubs. Secondary forest, primary forest and burnt primary forest have 34.8%, 58.7% and 55.5% accuracy, respectively. Secondary forests are mainly confused with burnt primary forests, oil palm plantations, primary forests and rubber and are less confused with mixed, *alang-alang* and shrub. Primary forests are mainly confused with oil palm, burnt primary forests, secondary forests and rubber. Burnt primary forests are mainly confused with oil palm, primary forests and secondary forests, and are less confused with rubber and mixed. For more details see Tables 4.5a and 4.5b.

It may be concluded that the classes which feature the most confusion are *alang-alang*, shrubs and oil palm. This result corresponds entirely with the classification simulation presented earlier. *Alang-alang* and shrubs are highly subject to inaccuracies in differentiation. The same applies for plantation forests (oil palm) and secondary forests. The diversity, i.e. the wide range of physical structure within the secondary forest, primary forest, and *alang-alang* classes may also affect the classification potential. In general, the differences between forest (primary, mangrove, etc) and non-forest are well distinguishable, but it is especially difficult to differentiate between natural forests (except mangrove) and forest plantations (oil palm and rubber).

Table 4.5a. Confusion matrix for the C-, L- and P-band fully polarimetric combination of classification results in Figure. 4.5a. Results are expressed *in percentages*.

	WATE	PADI	ALAN	MIXE	SHRU	BARE	PALM	RUBB	MANG	SECO	PRIM	BURN
WATE	**96.5**	2.5	0.0	0.0	0.3	0.3	0.0	0.0	0.0	0.0	0.0	0.0
PADI	2.0	**90.1**	0.3	0.0	0.3	4.7	0.0	0.0	0.0	0.0	0.0	0.1
ALAN	0.5	0.5	**42.1**	7.6	14.8	11.1	0.8	0.6	3.4	2.0	0.3	0.8
MIXE	0.0	1.5	15.2	**65.6**	17.6	3.7	0.3	0.8	5.1	2.3	0.2	1.1
SHRU	0.4	0.0	4.5	1.7	**33.6**	1.4	0.3	0.6	2.7	1.6	0.8	0.7
BARE	0.6	5.5	10.3	3.4	0.9	**77.6**	0.0	0.0	0.0	0.0	0.0	0.0
PALM	0.0	0.0	3.8	0.0	4.0	0.0	**50.3**	10.8	0.3	16.9	14.3	18.3
RUBB	0.0	0.0	3.1	2.5	4.6	0.0	11.3	**77.6**	0.6	8.8	5.5	3.2
MANG	0.0	0.0	7.6	11.8	7.1	0.8	0.0	0.6	**85.7**	0.9	0.1	0.2
SECO	0.0	0.0	8.6	3.4	8.3	0.2	9.0	4.1	1.1	**34.8**	7.6	7.9
PRIM	0.0	0.0	0.0	0.0	0.6	0.0	13.3	3.6	0.0	14.0	**58.7**	12.3
BURN	0.0	0.0	4.5	4.2	7.7	0.2	15.0	1.5	1.3	18.8	12.5	**55.5**

Table 4.5b. Idem. Results are expressed *in pixels*.

	WATE	PADI	ALAN	MIXE	SHRU	BARE	PALM	RUBB	MANG	SECO	PRIM	BURN
WATE	**959**	5	0	0	1	2	0	0	0	0	0	0
PADI	20	**181**	1	0	1	28	0	0	0	0	0	1
ALAN	5	1	**122**	9	48	66	3	3	24	14	9	17
MIXE	0	3	44	**78**	57	22	1	4	36	16	6	23
SHRU	4	0	13	2	**109**	8	1	3	19	11	22	14
BARE	6	11	30	4	3	**460**	0	0	0	0	1	0
PALM	0	0	11	0	13	0	**201**	58	2	119	398	389
RUBB	0	0	9	3	15	0	45	**415**	4	62	153	69
MANG	0	0	22	14	23	5	0	3	**611**	6	2	5
SECO	0	0	25	4	27	1	36	22	8	**246**	213	169
PRIM	0	0	0	0	2	0	53	19	0	99	**1637**	262
BURN	0	0	13	5	25	1	60	8	9	133	348	**1182**

Since the results show the existence of confusion between plantation forests and natural forests, additional information, such as texture of the canopy, may be valuable in improving the classification results. This can be done by considering the differences of textural features (see Section 2.7) in the classification process. Texture from radar images may be described as being smooth, rough, fine, coarse, random, grainy, granulated, etc. In the section (Section 4.5) the use of 3-D texture information derived from the C-band DEM and texture information derived from C-band total power image will be studied.

The creation of a classified image, as well as the evaluation of a classification result, is, in general, not a very straightforward task. This may be particularly true for the complex structure of tropical land cover types. Some additional points of consideration are: the occurrence of many other types of forests and plantations, the absence of well-defined boundaries, the existence of transitions between forest types, and the presence of mixed vegetation (such as secondary forests). The existence of transition stages, for example from *alang-alang* to shrub, is often found, and randomly located in the study area. Yet, only 12 classes have been defined so far on the basis of carefully selected training areas. This may be on the low side (see also Chapter 5).

4.4.1 Conclusions

From the results of land cover type mapping using the C-, L- and P-band fully polarimetric combination above, the following may be concluded:

- The selection of training areas was limited to the 30^o–60^o incidence angle range, the effect of incidence angle dependence was ignored, and the backscatter parameter gamma was used because it is less incidence angle dependent. Classification was conducted at the 100% level of confidence. Pre-processing entailed spatial aggregation of 4x4 pixels.
- The definition of land cover classes was slightly adapted. The urban area class was excluded from further analysis, however is very easy to distinguish. *Alang-alang* and shrub were difficult to distinguish and were separated into 3 classes,

namely: *alang-alang*, mixed (transition *alang-alang* to shrub) and shrub. Consequentially a legend with 12 classes resulted.

- *Overall accuracy reached 63.3%.* Water and *padi* (rice) fields had a relatively high accuracy, followed by mangrove, bare soil and rubber plantation. Primary forest, burnt and un-burnt, had a medium accuracy level, while *alang-alang*, mixed and shrubs had low accuracy.

4.5 Image classification enhancement

4.5.1 3-D Texture

The results above show that the possibilities to discriminate the different forest types and forest plantations are very limited. For this reason, the use of texture as additional information obtained from the radar backscatter images (C-band total power) and height images (C-band InSAR DEM) may be useful. The potential of the extracted texture attributes for additional classification information to differentiate between plantation forests (rubber and oil palm) and natural forests (secondary forest, primary forest and burnt primary forest) types studied was evaluated on the basis of a class separability measure. This measure represents the statistical distance between class pairs and is an indirect and *a priori* estimate of the probability of correct classification. The index for class pair (i, j) is given by:

$$\text{Index} = \left| \frac{\mu_i - \mu_j}{sd_i + sd_j} \right|$$

where μ_i is the average value for class i, μ_j is the average value for class j; sd_i and sd_j are the standard deviations for class i and j.

This study aims to rank the various attributes and/or attribute combinations according to their ability to discriminate successfully the five forest cover classes studied, i.e. according to their classification potential. Arbitrarily a separability value in excess of 1.2 is considered to be useful. Table 4.6 shows that both the Standard Deviation (SD) and GLCO-COR TP features can not sufficiently discriminate the 10 class pairs, while the GLCO-CONT[d1-w21] DEM feature can only distinguish seven out of 10 class pairs and fails to differentiate between secondary and primary forest. However, *it can make a distinction between plantation forests and natural forests*, which may have been expected, because of their different canopy architecture. All natural forests classes, i.e. secondary, primary forest and burnt primary forest, could not be correctly distinguished. In contrast, plantation forests with a smooth upper canopy can be told apart.

Though it is apparent that the classification capacity of SD and GLCO-COR textural attributes are less pronounced, for GLCO-CONT the classes are texturally more distinct. Texture might be used as an additional important source of information for identifying tropical land cover types in high frequency and high-resolution radar images. The combined use of two or more GLCO attributes might result in improved textural descriptions, and therefore might enhance the chances of classifying results correctly. The results of data sets with highly ranked attributes and/or attribute

combinations will be used later for classification enhancements. These numbers can serve as an additional *prior* in the extended ICM method.

Because of its ability to describe the dimensions of textural sub-patterns, GLCO-CONT can be said to be more sensitive to canopy architecture than SD. The values in Table 4.6 show the changes in the GLCO attributes as a function of displacement length, which reflects physical properties of the observed canopy. The results indicate that only the GLCO-CONT[d1] and -CONT[d5] textural features are sufficiently sensitive and stable to identify the forest types in a relatively big (21x21) image window.

To provide better assessment of the potential of texture as additional information for classification, a qualitative analysis using scatter plots (Figure 4.4) of the performance of textural features based on the GLCO approach was made. The feature combinations which are able to discriminate the forest types and plantations are found to be GLCO-CONT[d1-w21] from the C-band TP image and GLCO-CONT[d1-w21] and GLCO-CONT[d5-w21] from the C-band DEM.

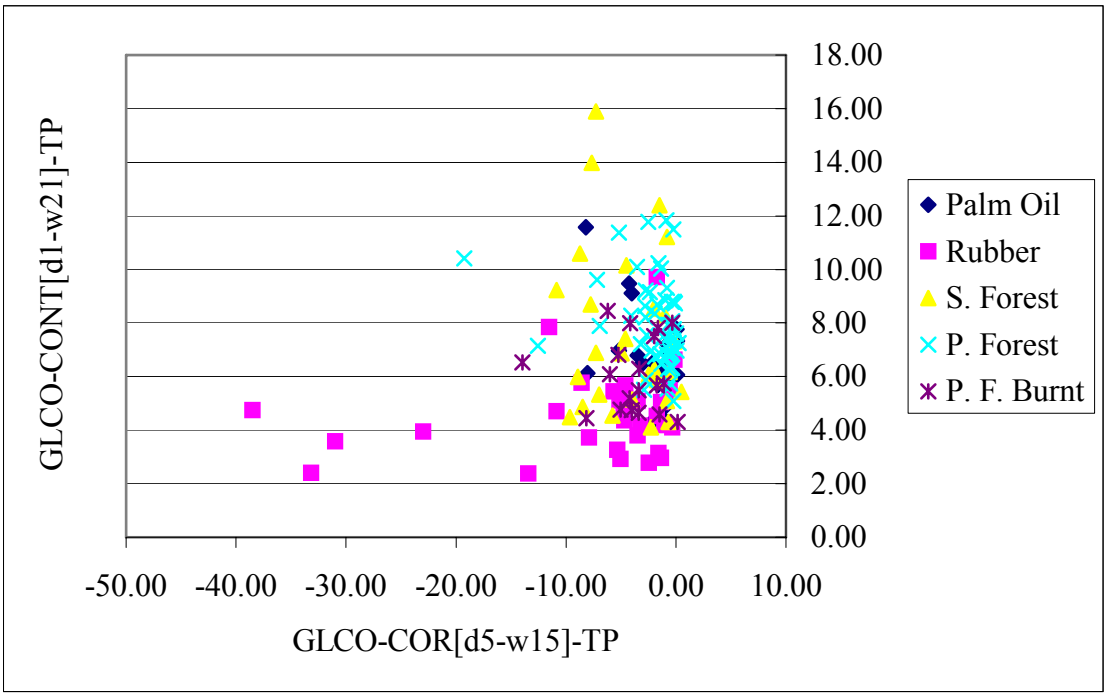
Table 4.6. *Texture separability* indices for a selection of land cover type pairs. Each column gives the result of the texture index between classes: (1) oil palm (2) rubber (3) secondary forest (4) primary forest and (5) burnt primary forest. The higher the values, the better the class separation. *The numbers in the shaded boxes indicate the best results*.

	Class Pairs									
Standard Deviation	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
C-TP-W07	0.751	0.046	0.371	0.033	0.798	1.084	0.741	0.327	0.081	0.412
C-TP-W15	0.651	0.138	0.606	0.085	0.763	1.107	0.593	0.480	0.227	0.698
C-TP-W21	0.585	0.154	0.704	0.182	0.699	1.103	0.471	0.551	0.342	0.912
C-DEM-W07	0.226	0.596	0.646	0.456	0.765	0.865	0.672	0.069	0.219	0.180
C-DEM-W15	0.037	1.084	0.715	0.495	1.012	0.650	0.438	0.417	0.626	0.222
C-DEM-W21	0.208	1.423	0.755	0.521	1.166	0.513	0.282	0.699	0.962	0.257
C-band TP	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
CONT[d1-W7]	0.748	0.019	0.305	0.062	0.767	1.004	0.716	0.288	0.082	0.373
CONT[d1-W15]	1.018	0.037	0.528	0.250	1.049	1.470	0.864	0.593	0.227	0.802
CONT[d1-W21]	1.226	0.108	0.694	0.403	1.201	1.795	0.865	0.836	0.320	1.076
CONT[d5-W7]	0.506	0.033	0.260	0.150	0.591	0.765	0.437	0.249	0.204	0.443
CONT[d5-W15]	0.652	0.062	0.498	0.278	0.730	1.052	0.476	0.459	0.361	0.806
CONT[d5-W21]	0.671	0.094	0.696	0.415	0.725	1.144	0.397	0.592	0.499	1.096
C-band TP	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
COR[d1-W7]	0.374	0.827	0.020	0.171	0.468	0.381	0.212	0.829	0.677	0.183
COR[d1-W15]	0.584	1.001	0.022	0.273	0.658	0.593	0.366	1.004	0.893	0.289
COR[d1-W21]	0.528	0.904	0.053	0.350	0.608	0.560	0.248	0.917	0.771	0.395
COR[d5-W7]	0.379	0.973	0.002	0.127	0.618	0.384	0.259	0.979	0.865	0.131
COR[d5-W15]	0.543	1.072	0.016	0.268	0.746	0.545	0.316	1.071	0.954	0.275
COR[d5-W21]	0.469	1.035	0.073	0.339	0.743	0.518	0.184	1.057	0.882	0.402

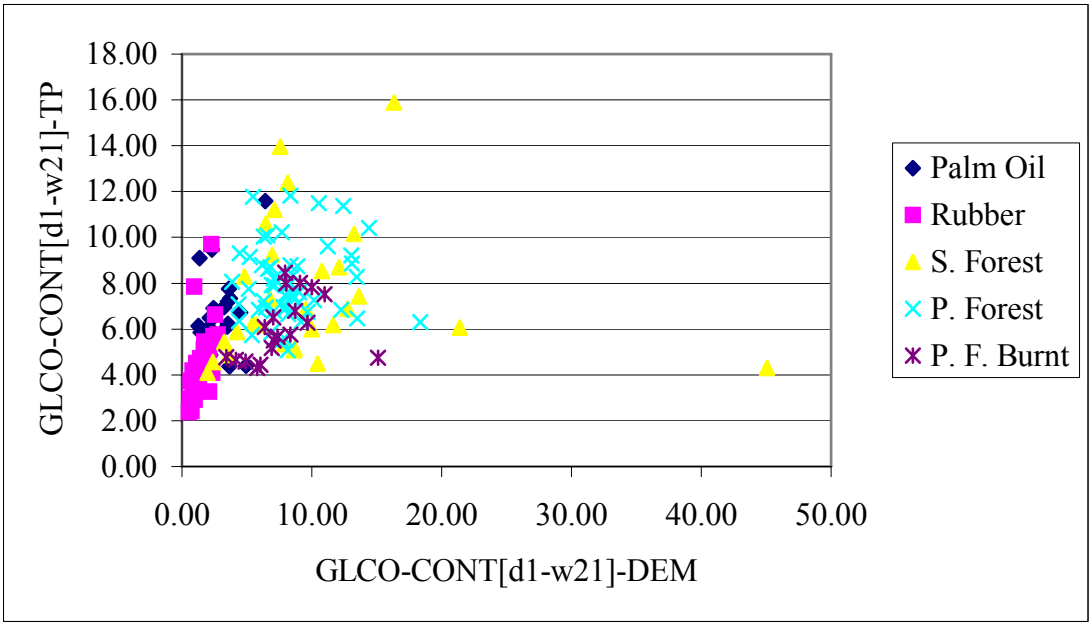
DEM	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
CONT[d1-W7]	0.647	0.509	0.747	0.487	0.572	1.096	0.742	0.314	0.330	0.087
CONT[d1-W15]	1.073	1.060	1.523	1.017	1.182	2.104	1.440	0.624	0.638	0.132
CONT[d1-W21]	1.356	1.386	2.200	1.488	1.531	2.937	2.013	0.808	0.809	0.140
CONT[d5-W7]	0.257	0.543	0.454	0.271	0.704	0.658	0.489	0.155	0.311	0.179
CONT[d5-W15]	0.407	0.978	0.845	0.576	1.182	1.132	0.896	0.311	0.538	0.276
CONT[d5-W21]	0.505	1.425	1.265	0.934	1.606	1.532	1.253	0.446	0.731	0.349

DEM	1-2	1-3	1-4	1-5	2-3	2-4	2-5	3-4	3-5	4-5
COR[d1-W7]	0.242	0.027	0.211	0.056	0.246	0.347	0.268	0.151	0.024	0.137
COR[d1-W15]	0.326	0.248	0.067	0.077	0.548	0.395	0.408	0.195	0.192	0.009
COR[d1-W21]	0.242	0.345	0.087	0.045	0.546	0.174	0.219	0.456	0.433	0.047
COR[d5-W7]	0.396	0.084	0.312	0.187	0.435	0.529	0.483	0.228	0.101	0.137
COR[d5-W15]	0.469	0.209	0.096	0.094	0.604	0.538	0.541	0.120	0.125	0.003
COR[d5-W21]	0.367	0.278	0.084	0.053	0.577	0.312	0.355	0.377	0.372	0.040

From the scatter plots of TP and DEM textural characteristics (Figure 4.4), it is clear that these are insufficient to make clear differentiations for all forest classes studied. The strong difference of canopy architecture between natural forest and forest plantations does not show up very clearly. Figure 4.4a shows that only rubber and secondary forest can be distinguished, while other classes cannot easily be discriminated. Oil palm seems to be confused with primary and burnt primary forest. Figure 4.4b shows the same result. The natural forest type areas, characterised by complex topography and rough canopy, can not be correctly differentiated from the forest plantation areas, which are characterised by flat to gentle topography and regular canopy architecture. Transitions between these areas are sometimes very gradual. Forests with rough to very rough canopy structures (i.e. secondary forest, primary forest and burnt primary forest) appeared to be the most difficult to be differentiated with the GLCO-COR approach. In this case, the GLCO-CONT parameters seemed to offer the best possibilities. The results indicate that the combination of total power and InSAR 3-D textural properties might be suitable for use as additional information to improve the discrimination of the natural forest types and forest plantations.



(a)



(b)

Figure 4.4. (a) Scatter plot of GLCO-COR[d5-w15]-TP vs. GLCO-CONT[d1-w21]-TP and (b) GLCO-CONT[d1-w21]-DEM vs. GLCO-CONT[d1-w21]-TP textural features for the main forest types and forest plantations**.**

4.5.2 Conclusions

The results obtained from image texture analysis using C-band total power and C-band InSAR height images can be summarised as follows:

- The 3 textural features Standard Deviation, GLCO-COR and GLCO-CONT were evaluated on the basis of a class separability index for 5 forest cover classes (oil palm, rubber, secondary forest, burnt and un-burnt primary forest).
- The GLCO-CONT textural feature, both in total power and height images, is clearly superior to the other two in distinguishing types of canopy architecture.
- Qualitative analysis using scatter plots showed that textural features based on GLCO-CONT[d1-w21] can *clearly contribute to the differentiation of the forest classes under study*.

4.6. Classification enhancement using neighbouring pixel information

To obtain reasonable classification results the use of additional information, such as 3-D textural features, seems to be promising. However, because of its limited ability to distinguish canopy architecture other additional features, such as height information derived from InSAR, should be considered. For example, it can be noted that mangrove occurs at low altitudes only. To utilise such as potential sources of information, the Iterated Conditional Modes (ICM) method (Besag, 1986) is suitable. It will be explored whether a new approach of ICM (Hoekman and Quiñones, 2002) combined with several types of *prior* information can enhance the classification result.

The use of ICM facilitates image reclassifying by considering the neighbourhood cells to enable an increase in the overall classification accuracy. Table 4.7 shows the ICM result with several band combinations, number of iterations, and additional information to be used (i.e. texture and DEM). The first column (Slope-corr) shows the use of slope correction. Y means that the slope correction is used, and N means it is not used. In the second column (Set), Kem_ag4 means that the training data set (the 386 areas of Table 3.1) is used for classification (training and validation) while Zul_ag4 means that the independent data set (the 220 areas of Table 3.1) is used for classification (training and validation). Note that Chapter 5 discusses the situation where the training data set and the independent data set are used for mutual classification performance evaluation. The column 'Bands' shows the combinations of bands being used, the column ' Number iteration' shows the number of ICM iterations, the column 'β1' shows the relative importance of neighbouring pixels, the column 'β8' shows the relative importance of the textural features, the column 'DEM_THR' shows the terrain altitude threshold value in meters. The columns DEM and Texture show switches: if the value is 1, the information is used, and the value is 0 if it is not used. The column 'End result' shows the final overall classification results after the iteration process is finished.

From Table 4.7 it is shown that the β1 values being used are 0.4, 1, 3, 6, 8 and 10, the β8 values being used are 1, 2, 3, 4 and 5, while the number of iteration used are 4, 5, 8, 10, 20 and 50. Because the number of combinations of all these values is enormous, and computing time for each classification is considerable because of the large image size, it is only possible to evaluate a number of combinations.

The classification results show that for the C-, L- and P-band combination with $\beta 1$=10, iterations=20 and with the use of *a priori* knowledge (i.e. texture and DEM), the classification has an 88.0% accuracy (case a). For the C-, L- and P-band combination with $\beta 1$=10, iterations=50, and without the use of *a priori* knowledge, the classification is even better with an 88.9% accuracy (case b). The latter result is shown in Figure 4.5b. For the C-, and L-band combination with $\beta 1$=10, iterations=20, and with the use of *a priori* knowledge, the classification has an 81.4% accuracy (case c), and the result is shown in Figure 4.6a. For the C-, and P-band combination with $\beta 1$=10, iterations=20 with the use of *a priori* knowledge, the classification has an 82.2% accuracy (case d); the result is shown in Figure 4.6b. For the L- and P-band combination with $\beta 1$=10, iterations=20 with the use of *a priori* knowledge, the classification has an 83.1% accuracy (case e), and the result is shown in Figure 4.6c. However, it is clear that for $\beta 1$=10 and iterations less than 20, the classification result is less optimal.

Table 4.7. Overall classification results for training/validation areas of the extended ICM approach for several bands, iteration values and *a priori* knowledge combinations. All combinations are fully polarimetric and the 7I model is chosen. *) For case j see Chapter 5.

Slope-corr	Set	Run	Bands	Number iterations	$\beta 1$	$\beta 8$	DEM	Texture	DEM_THR	End result	Case
Y	Kem_ag4	ICM_run01	CLP	4	0.4	0	0	0	0	70.6	
Y	Kem_ag4	ICM_run02	CLP	8	1	0	0	0	0	77.6	
Y	Kem_ag4	ICM_run03	CLP	50	10	0	0	0	0	88.9	b
Y	Kem_ag4	ICM_run04	CLP	10	8	2	0	1	0	85.6	
Y	Kem_ag4	ICM_run05	CLP	5	1	1	1	1	50	77.0	
Y	Kem_ag4	Run_f01	CLP	20	10	1	0	0	50	87.6	
Y	Kem_ag4	Run_f02	CLP	20	10	1	1	0	50	87.8	h
Y	Kem_ag4	Run_f03	CLP	20	10	1	1	1	50	88.0	a
Y	Kem_ag4	Run_01	CL	20	10	1	0	0	50	81.6	
Y	Kem_ag4	Run_02	CL	20	10	1	1	0	50	81.8	i
Y	Kem_ag4	Run_03	CL	20	10	1	1	1	50	81.4	c
Y	Kem_ag4	Run_04	CL	20	10	4	1	1	50	79.9	
Y	Kem_ag4	Run_05	CLP	10	3	2	0	0	50	83.7	
Y	Kem_ag4	Run_06	CLP	10	6	1	0	0	50	85.4	
Y	Kem_ag4	Run_07	CLP	10	6	2	1	0	50	85.6	
Y	Kem_ag4	Run_08	CLP	10	6	1	1	1	50	85.6	
Y	Kem_ag4	Run_09	CLP	10	6	2	1	1	50	85.4	
Y	Kem_ag4	Run_10	CLP	10	6	3	1	1	50	84.7	
Y	Kem_ag4	Run_11	CLP	10	6	4	1	1	50	84.0	
Y	Kem_ag4	Run_12	CLP	10	6	5	1	1	50	82.9	
Y	Kem_ag4	Run_13	CP	20	10	0	1	0	50	82.2	d
Y	Kem_ag4	Run_13a	LP	20	10	0	1	0	50	83.1	e
Y	Zul_ag4	Run_14	CLP	20	10	0	1	0	50	93.8	j*)
N	Kem_ag4	Run_15	CL	20	10	0	1	0	50	82.3	
N	Kem_ag4	Run_16	CLP	20	10	0	1	0	50	87.8	
N	Kem_ag4	Run_17	CP	20	10	0	1	0	50	82.5	f
N	Kem_ag4	Run_18	LP	20	10	0	1	0	50	82.0	g

When comparing cases d and f for the CP combination, or comparing cases e and g for the LP combination, it is clear that the effect of the additional use of slope correction is not significant. Similarly, when comparing cases h and a for the CLP combination, or comparing cases i and c for the CL combination, it is clear that the effect of the additional use of texture is not significant. *Apparently, when using the full polarimetric information of 2 or 3 bands, the information content is already that high that additional information hardly adds anything.*

It may be concluded that the approach chosen is relatively satisfying. Nevertheless, it is not clear how the values of the influence factors $\beta 1$, $\beta 8$ (section 2.6), their relaxation schemes, and other factors (i.e. texture and DEM) can be optimised. However, it can be shown that this approach yields major improvements in classification.

It can be concluded that this simple and computationally fast ICM technique yields significantly better results (Figure 4.5b) compared to the previous ML results (Figure 4.5a), although explicit application of DEM and texture may not provide major improvements.

Table 4.8a presents the confusion matrix in percentages of land cover type classification for the third case of Table 4.7, which is the best result achieved for the 'Kem_ag4' data set. Figure 4.5b is the resulting image after completion of 50 cycles of the extended ICM method. The classes are coded with the colour legend described in Figure 4.5c. In general, satisfying results are obtained. The overall accuracy was 88.9%. The classification was highly accurate for water, *padi* fields, mangrove, mixed, bare soil, rubber, oil palm, rubber plantations, primary forests and burnt primary forests. For these classes, the classification accuracy varied from 86.5% for oil palm to 98.8% for water. The classification was less successful for *alang-alang* (46.7%), shrubs (34.0%) and secondary forest (73.2%). *Alang-alang* and shrubs were often confused with each other. Secondary forests were often classified as burnt primary forest, primary forest and oil palm.

Table 4.8a. Confusion matrix for classification results (*in percentages*) for the 12 classes after applying 50 cycles of the extended ICM approach. Note this is the third case of Table 4.7.

	WATE	PADI	ALAN	MIXE	SHRU	BARE	PALM	RUBB	MANG	SECO	PRIM	BURN
WATE	**98.8**	1.1	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PADI	0.3	**95.6**	0.0	0.7	0.5	2.2	0.0	0.0	0.0	0.0	0.0	0.0
ALAN	0.2	1.1	**46.7**	5.0	11.0	8.4	0.1	0.2	0.2	0.1	0.0	0.1
MIXE	0.0	1.7	26.6	**87.5**	31.9	1.4	0.0	0.1	0.9	0.2	0.0	0.1
SHRU	0.4	0.0	1.9	0.3	**34.0**	0.2	0.0	0.0	0.4	0.1	0.1	0.2
BARE	0.4	0.6	7.2	0.3	0.1	**87.6**	0.0	0.0	0.1	0.0	0.0	0.0
PALM	0.0	0.0	0.6	0.6	3.7	0.1	**86.5**	1.5	0.0	6.5	3.5	1.3
RUBB	0.0	0.0	0.5	0.2	3.5	0.0	1.2	**96.2**	0.1	0.9	0.4	0.2
MANG	0.0	0.0	4.3	0.6	3.3	0.1	0.0	0.2	**98.3**	0.0	0.0	0.0
SECO	0.0	0.0	8.9	4.5	3.6	0.0	3.8	1.0	0.1	**73.2**	0.7	0.5
PRIM	0.0	0.0	0.0	0.0	0.0	0.0	6.3	0.7	0.0	7.8	**91.8**	2.8
BURN	0.0	0.0	3.3	0.3	8.1	0.0	2.0	0.1	0.1	11.1	3.4	**94.8**

Table 4.8b. Idem. Results are expressed *in pixels*.

	WATE	PADI	ALAN	MIXE	SHRU	BARE	PALM	RUBB	MANG	SECO	PRIM	BURN
WATE	**50064**	108	0	0	51	5	0	0	0	0	0	0
PADI	152	**9802**	6	43	74	656	0	0	0	0	0	20
ALAN	91	112	**6903**	302	1811	2529	15	42	63	52	61	80
MIXE	17	171	3932	**5313**	5277	410	6	30	310	59	33	91
SHRU	194	0	286	15	**5615**	50	5	6	130	47	161	161
BARE	176	58	1058	19	9	**26505**	0	0	32	0	14	0
PALM	0	0	94	39	618	39	**17640**	413	5	2343	5027	1433
RUBB	0	0	66	13	578	0	248	**26243**	18	324	564	260
MANG	0	0	635	37	546	39	0	45	**35741**	16	10	12
SECO	0	0	1322	270	595	3	780	283	42	**26367**	1002	576
PRIM	0	0	0	0	5	0	1291	187	0	2794	**130528**	3023
BURN	0	0	488	18	1345	7	415	36	22	4004	4839	**103025**

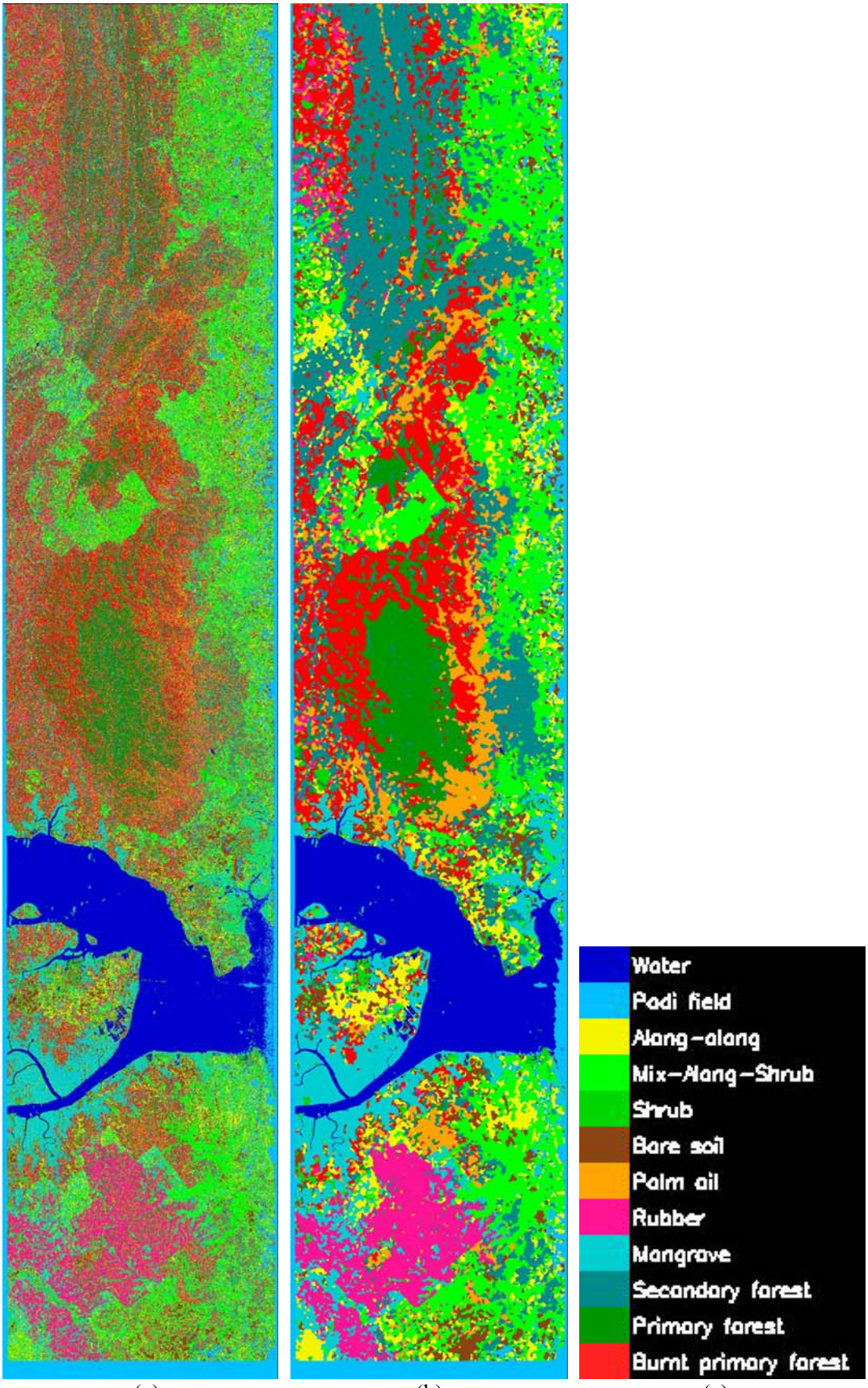
It can be observed from Figure 4.5b that a few areas in the area just above the bay and below the primary forest region (the Sungai Wain protected forest) are labelled as oil palm, but in fact are primary forest, according to a field survey. This area of primary forest differs from other primary forest areas in the sense that it is very swampy. The oil palm area actually only occurs in the southern area, just below the bay. Primary forest areas were correctly labelled.

Using dual band combinations of fully polarimetric SAR data, the classification results after ICM are shown in Figure 4.6. The classes are coded with colours described in Figure 4.5c. Visual interpretation of the C- and L-band combination (Figure 4.6.a) shows that burnt primary forest is difficult to separate from the secondary forest, while the primary forest is clearly separable. In the southern side, just below the bay, the burnt primary forest appears in the map while it is actually a part of the nearby oil palm plantation, and the rubber plantation appears in the map to be larger than it is. The classes *alang-alang* and mixed can be clearly separated, although they are often mixed up with each other. For the C- and P-band combination (Figure 4.6b) primary forests and secondary forests can also be clearly differentiated. In the northern part, rubber plantations and oil palm appear in the map, although they are actually secondary forest. The classes *alang-alang*, mixed and secondary forest can be well differentiated from other classes, although they are mutually hard to differentiate. In the southern part, burnt primary forest appears in the map, although the area is actually a part of the oil palm plantations. For the L- and P-band combination (Figure 4.6c) the classification result for the primary forest is slightly inferior compared to the two other combinations mentioned previously. Secondary forests can be clearly differentiated. In the Sungai Wain protected forest area and along its eastern and north-eastern side, oil palm appears on the map, although the area is actually burnt primary forest in the Sungai Wain area and secondary forest in the north-eastern part. In the southern part below the bay the oil palm and rubber plantations can be well differentiated. The classes *alang-alang*, shrub and mixed are difficult to separate. *In general using dual-band combinations the classification result of the C- and P-band combination provides the most realistic outcome.*

4.6.1 Conclusions

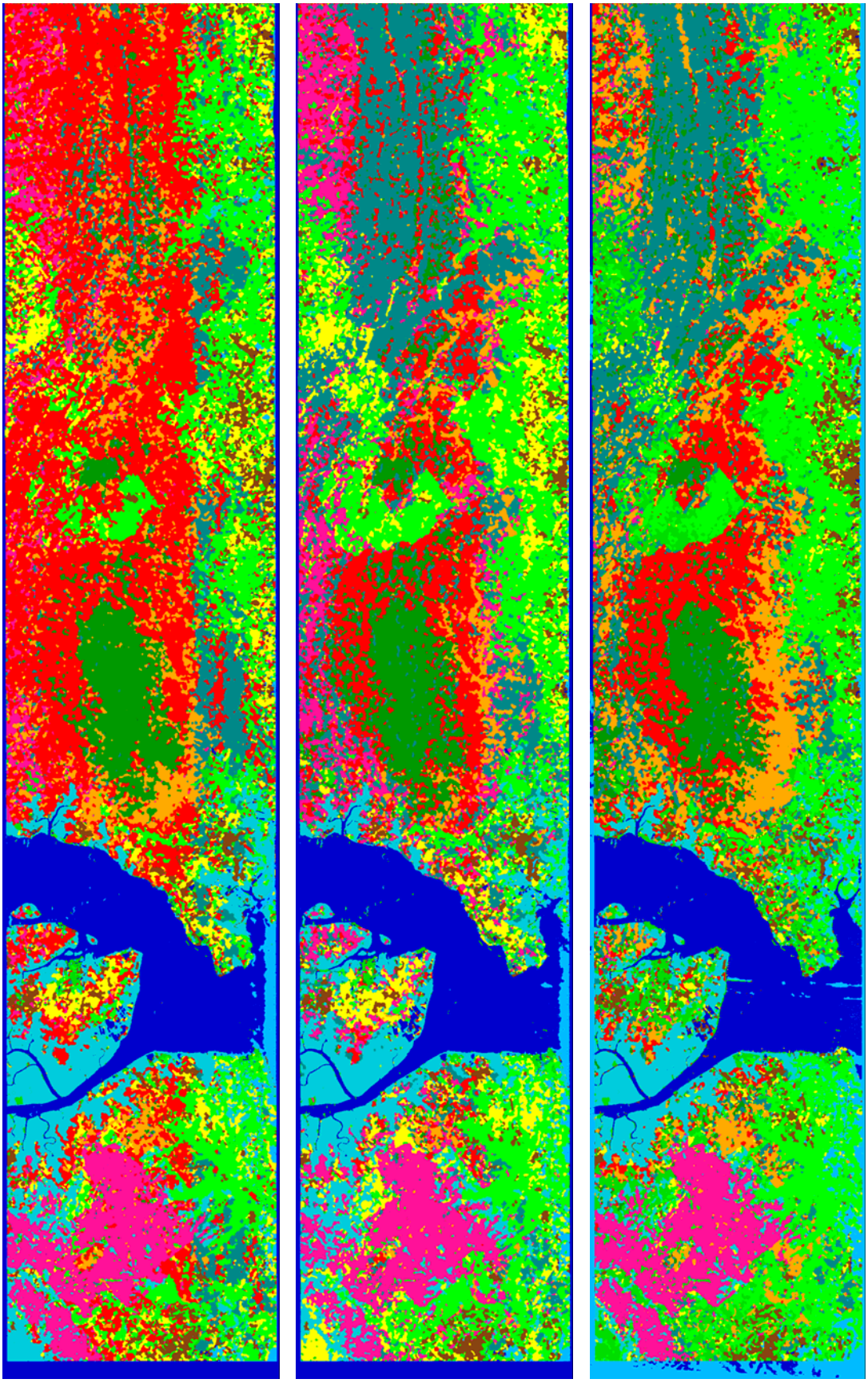
From the result of classification enhancement using neighbouring pixel information (ICM), DEM, and 3-D textural information as additional information, the following can be concluded:

- ICM, by taking into consideration the class of neighbouring pixels and additional knowledge from the DEM (height and texture), increases the overall classification accuracy considerably. In section 4.4 the best simulated result was 85.3% for 64-look data and 71.3% for 20-look data; here, *after applying ICM the accuracy increases to 88.9% for a 4x4 pixel aggregate*. From section 2.3.2 it can be concluded this corresponds to roughly 75-, 45- and 24–look data for C-, L- and P-band, respectively.
- Classification results showed that the use of the C-, L- and P-band combination with $\beta 1=10$ and 50 iterations, without *a priori* knowledge (i.e. texture and DEM), yielded the best results. The use of *a priori* knowledge (i.e. texture and DEM) did not always yielded significant increases in the overall classification results.
- *Apparently, when using the full polarimetric information of 2 or 3 bands, the information content is already that high that additional information hardly adds anything.*
- For water, *padi* fields, mangrove, mixed, bare soil, rubber, oil palm, rubber plantations, primary forests, and burnt primary forests, the classification accuracy varies between 86.5% for oil palm and 98.8% for water. For *alang-alang*, shrubs, and secondary forests, the classification was less successful.
- When using dual-band combinations: *C- and L-band showed that burnt primary forest was difficult to separate from secondary forest*. Using the C- and P-band combination, primary forest and secondary forest can be well differentiated while for the L- and P-band combination the classification result for the primary forests is slightly inferior compared to the two other combinations mentioned previously. *In general using dual-band combinations the classification result of the C- and P-band combination provides the most realistic outcome.*



(a) (b) (c)

Figure 4.5. The study area showing (from left to right): (a) ML-classification and (b) ICM (50)-classification and (c) legend for colour-codes used in the map.



(a) (b) (c)

Figure 4.6. Classification result from dual-band combinations after ICM (a) for C- and L-band (b) for C- and P-band and (c) for L- and P-band. Legend of colour-codes used in the map can be seen in Figure 4.5c.

4.7 Biomass estimation

An allometric equation (Eq. 3.1) valid for the East-Kalimantan province was applied to *estimate the (total above-ground) biomass of the plots of primary forest, burnt primary forest and secondary forest*, using trunk diameter and tree top height (Brown *et al*., 1989). Since many dead trees with standing trunks were present, and these could have an important effect on the radar backscatter, this estimation was done twice: with and without dead standing trees. It is noted here that dead fallen trees were never included. The potential of biomass class mapping was studied by evaluating the backscatter for all nine transects of primary forest and burnt primary forest, and for seven transects of secondary forest. A clear empirical relationship between backscatter and biomass would allow the development of a straightforward direct approach in biomass mapping. In addition, vegetation characteristics were collected for 15 plots of *alang-alang* with varying degrees of bush invasion. In these fields, the above-ground fresh biomass was found to vary over the range of 112-199 ton ha^{-1} (1 ton = 1,000 kg; 1 ha = 10,000 m^2) for primary forest, 31-115 ton/ha for burnt primary forest and 25-153 ton ha^{-1} for secondary forest (see Chapter 3 for more details). The correlation coefficients r^2 are calculated from linear regression between gamma (γ_i) values (in dB) and the logarithm of biomass.

The main results are summarised in Tables 4.9 and 4.10. Since slope effects are significant, these results are shown for backscatter values *before and after slope correction*. As mentioned above, results are also shown for biomass with and without the inclusion of dead standing trees. Therefore four cases are considered. The backscatter range for each radar band-polarisation combination reflects the sensitivity of radar bands and polarisation to biomass. This backscatter range is commonly referred to as dynamic range. The dynamic range, and hence the sensitivity to biomass, is shown to increase with an increase in wavelength. This phenomenon is well known, and may be explained by the relationship between wavelength and the scattering behaviour. The available data show that L-HV, P-VV and P-HH have the largest dynamic range (Table 4.10, after slope correction).

It is shown that the coefficient of correlation (r^2) is smaller for biomass excluding dead standing trees compared to biomass including dead tress. The correlation coefficients obtained before and after slope correction are slightly different; after slope correction a small increase in the correlation coefficient is observed. The maximum value of r^2 is 0.68, which is for P-HV with a range/SEE ratio of 5.53. For C-VV the r^2 is 0.34 with a range/SEE ratio of 5.09 and for L-VV r^2 is 0.62 with a range/SEE ratio 0.34. All other r^2 values are between 0.29 and 0.60 with range/SEE ratios between 4.63 and 6.17. It seems that the r^2 values of the cross-polarised backscatter in the long wavelengths (L-HV and P-HV) are higher than those for the short wavelengths C-HV).

In case dead standing trees are excluded, the correlation coefficients are much lower in general than in case dead standing trees are included. The maximum value of r^2 is only 0.60 and results for P-HV. All other values are lower than 0.57. The lowest correlation coefficients are measured at C-band with a maximum value of 0.34 from C-VV. In general, the correlation coefficients are higher for L- and P-band. HH and HV polarisation do not always yield better results than VV polarisation. The maximum values of r^2 with dead standing trees included are 0.62 and 0.68, which are

found for L-VV and P-HV respectively. In addition, the range/SEE ratio values for the cases which exclude dead standing trees are lower than for the cases which include dead standing trees. Apparently, the impact of the presence of dead standing trees provides an undesirable effect in the estimation of actual biomass vegetation in burnt and intact primary forest, because the dead trunk remnants in such forest types will increase the amount of backscatter measured, causing the increase of biomass estimation calculated by the empirical relation. This approach, i.e. using backscatter-log10 biomass relationships, leads to moderate results for this site.

Scatter plots in Figure 4.7 show the backscattering coefficient as a function of the estimated log10 biomass for C-HH, L-VV, P-HH and P-HV, respectively. For C-HH (Figure 4.7a) the correlation and the range/SEE ratio are very low. No clear relationship can be observed. For L-VV (Figure 4.7b) the correlation and the range/SEE ratio are moderate (see Table 4.10). The secondary forest biomass value seems to have a wide range (from medium to high). The heterogeneous vegetation of this class features an irregular upper canopy due to the clumped occurrence of dominant canopy trees and the presence of emergent pioneer trees. For P-HH (Figure 4.7c) the ratio of range/SEE is higher: 6.17 and the correlation is 0.55. In this case there is a slight relation with biomass; however there is also a clear dispersion between samples from different forest types, caused by factors such as the presence of many dead trunks in the burnt primary forest. The theoretical concept of threshold or saturation level is the maximum biomass level for which the radar backscatter is still sensitive. For the P-HH result it points to a biomass level around 100 ton ha^{-1}. In the case of P-HV (Figure 4.7d), there is a slight relation with biomass, however there is also no clear dispersion between secondary forest and burnt primary forest. The ratio is 5.53 and the correlation is 0.68. Although less dispersion of the data is observed, the saturation is reached very early, around 2.0 (or 100 ton ha^{-1}).

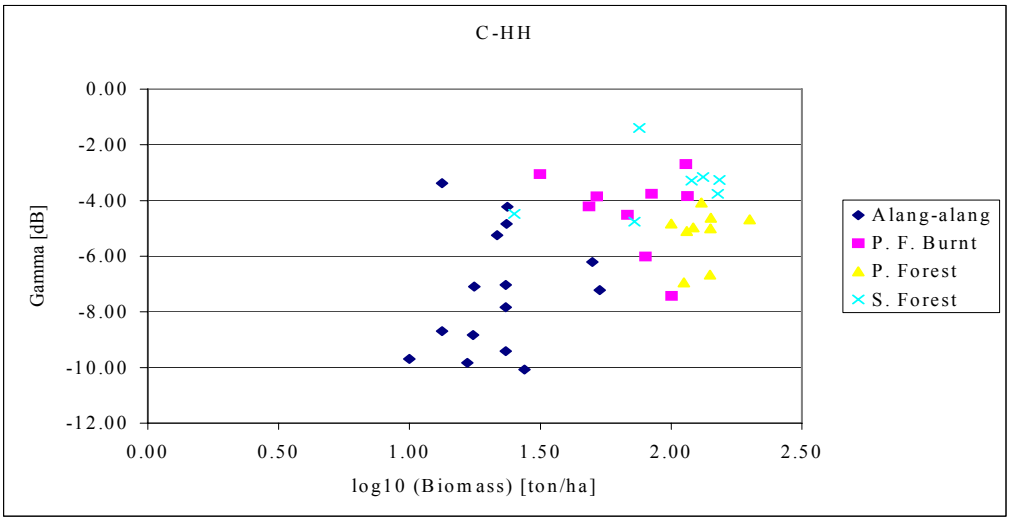
Table 4.9. Relationship between backscatter [dB] and log10 biomass *without dead standing trees* in ton ha^{-1} for several frequency and polarisation combinations. The correlation coefficient r^2, standard error of estimate (SEE), the total range of the experimental data and the ratio of range and SEE are shown *before and after slope correction*.

Channel	Before					After				
	r	r^2	SEE [dB]	range [dB]	Range/ SEE [dB]	*r*	r^2	SEE [dB]	range [dB]	range/ SEE [dB]
C-HH	0.50	0.25	2.22	9.41	4.24	0.53	0.28	1.99	8.68	4.37
C-VV	0.55	0.30	2.35	11.02	4.69	0.58	0.33	2.14	10.26	4.80
C-HV	0.53	0.28	2.11	9.25	4.39	0.55	0.31	1.90	8.51	4.49
L-HH	0.72	0.52	2.60	12.35	4.74	0.75	0.56	2.41	11.53	4.78
L-VV	0.73	0.53	2.53	11.26	4.45	0.75	0.57	2.35	10.55	4.50
L-HV	0.71	0.50	3.08	15.21	4.93	0.72	0.52	2.93	14.50	4.94
P-HH	0.68	0.46	3.54	17.58	4.97	0.70	0.49	3.32	16.76	5.05
P-VV	0.71	0.51	3.07	13.98	4.55	0.73	0.54	2.88	13.52	4.69
P-HV	0.76	0.58	3.83	16.19	4.22	0.77	0.60	3.65	15.88	4.35

Table 4.10. Relationship between backscatter [dB] and log10 biomass *including dead standing trees* in ton ha^{-1} for several frequency and polarisation combinations. The correlation coefficient r^2, standard error of estimate (SEE), the total range of the experimental data and the ratio of range and SEE are shown *before and after slope correction*.

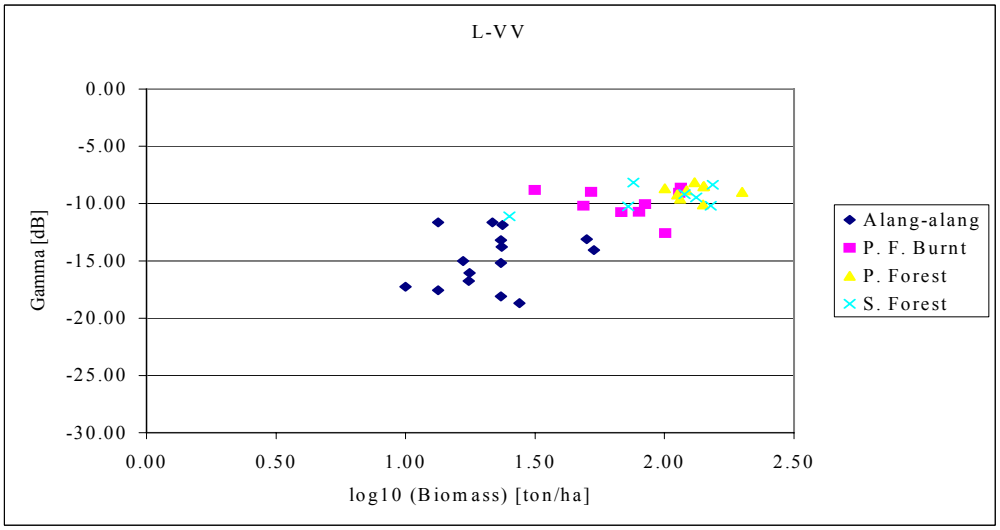
Channel	Before					After				
	r	r^2	SEE [dB]	range [dB]	range/ SEE [dB]	*r*	r^2	SEE [dB]	range [dB]	range/ SEE [dB]
C-HH	0.51	0.26	2.19	9.41	4.29	0.54	0.29	1.88	8.68	4.63
C-VV	0.56	0.31	2.32	11.02	4.74	0.58	0.34	2.01	10.26	5.09
C-HV	0.54	0.29	2.08	9.25	4.46	0.56	0.31	1.80	8.51	4.73
L-HH	0.74	0.54	2.53	12.35	4.88	0.76	0.57	2.06	11.53	5.60
L-VV	0.77	0.59	2.38	11.26	4.73	0.78	0.62	1.98	10.55	5.34
L-HV	0.75	0.56	2.91	15.21	5.22	0.76	0.57	2.55	14.50	5.68
P-HH	0.72	0.52	3.35	17.58	5.24	0.74	0.55	2.72	16.76	6.17
P-VV	0.76	0.57	2.87	13.98	4.88	0.78	0.60	2.40	13.52	5.64
P-HV	0.81	0.66	3.48	16.19	4.65	0.82	0.68	2.87	15.88	5.53

In general backscatter in L- and P-band increases with increasing biomass until it saturates at a certain threshold level. L- and P-band have the capability to discriminate the secondary forest, primary forest and burnt primary forest, as long as the biomass levels of the secondary forests are lower than those of primary forest and burnt primary forest. Biomass estimation is complicated further by the fact that the backscatter in L- and P-band is not only a function of biomass, but also of forest composition and condition. The results, obtained for this site, may indicate the general trend observed in this type of forest conditions (primary forest and burnt primary forest). The response of the radar to burnt areas will be complex and more unpredictable, depending on the burning intensity, nature of the forest and underlying ground surface. Ground fires, for example, may cause no difference in radar response, but dead standing trees might have an impact upon radar response. Boundaries of burn scars are not regular in tropical areas and burning gradations frequently occur.



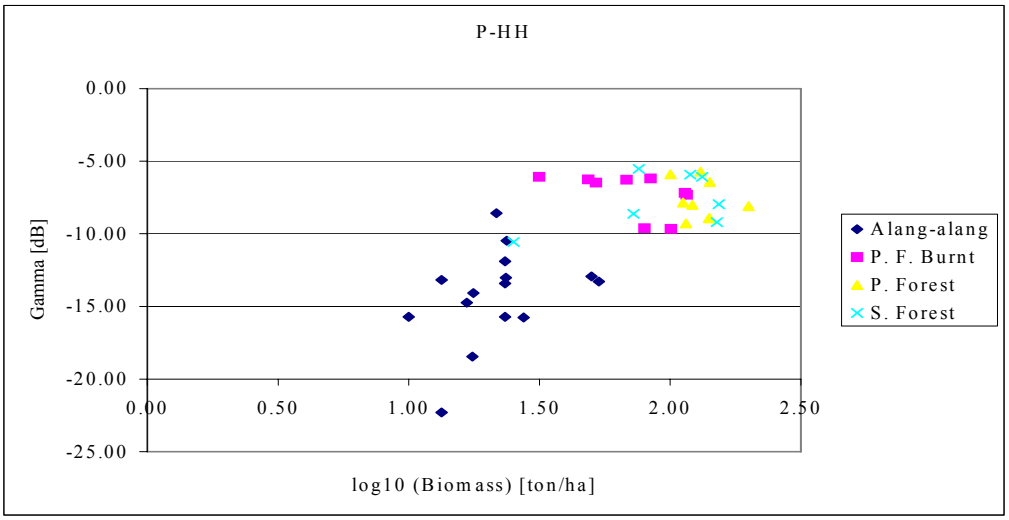
(a)

Figure 4.7a. Scatter plots for the relation between the estimated of fresh weight above-ground biomass for *alang-alang* and other forest cover types including dead standing trees and the corresponding intensity value expressed in gamma [dB] *after slope correction*, for C-HH.



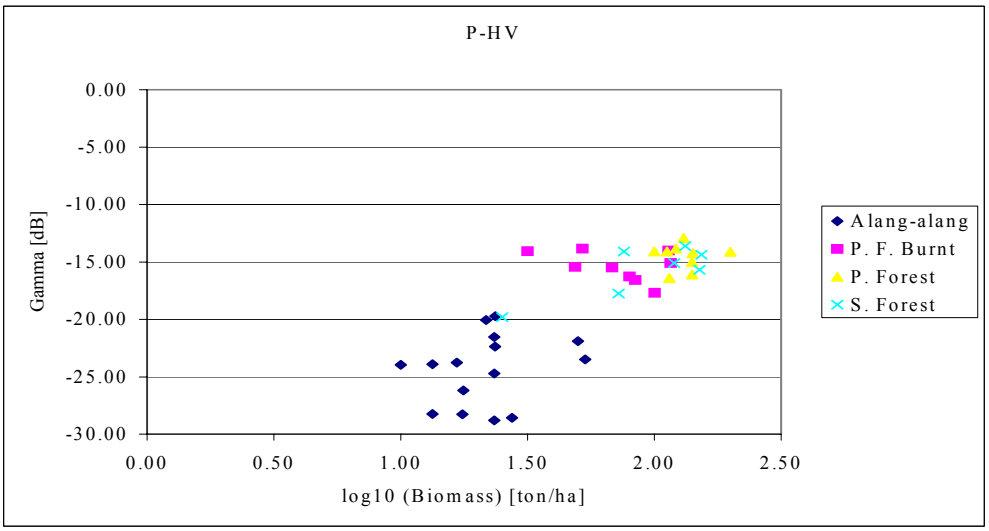
(b)

Figure 4.7b. Scatter plots for the relation between the estimated fresh weight above-ground biomass for *alang-alang* and other forest cover types including dead standing trees and the corresponding intensity value expressed in gamma [dB] *after slope correction*, for L-VV.



(c)

Figure 4.7c. Scatter plots for the relation between the estimated fresh weight above-ground biomass for *alang-alang* and other forest cover types including dead standing trees and the corresponding intensity value expressed in gamma [dB] *after slope correction*, for P-HH.



(d)

Figure 4.7d. Scatter plots for the relation between the estimated fresh weight above-ground biomass for *alang-alang* and other forest cover types including dead standing trees and the corresponding intensity value expressed in gamma [dB] *after slope correction*, for P-HV.

Taking into account the unsatisfying result of having (direct) empirical relationships between backscatter and biomass which strongly depend on forest conditions (primary forest and burnt primary forest) and due to the early saturation of the backscatter intensity level, which will also largely effect the accuracy of the biomass map, the possibility of using indirect relationships using structural forest maps, like proposed in (Quiñones, 2002) should be considered as an alternative solution.

4.7.1 Conclusions

Conclusions from the analysis of biomass estimation using empirical relationships between backscatter and biomass for *alang-alang* fields, secondary forests, primary forests and burnt primary forests, including as well as excluding dead standing trees, are the following:

- The coefficient of correlation is smaller for biomass excluding dead standing trees compared to biomass *including dead standing trees*.
- The coefficient of correlation before and after slope (relief) correction is slightly different. After slope correction, the coefficient shows a small increase.
- P-HV shows a maximum correlation result which is 0.68.
- In general, the backscatter in L- and P-band increases with the increase of biomass until a certain saturation limit is reached, which is around 100 ton ha^{-1}
- *In case burnt forests would be excluded, the relationships are much more pronounced*, and the saturation limit may reach 200 ton ha^{-1}.

4.8 Physical interpretation of signal characteristics

To understand the physical aspects of the effect of some forest structural characteristics (such as canopy structure, trunk or terrain condition) on the radar return, the multi-frequency complex coherence signature of field plots with different structural characteristics was studied, using a physical model introduced in (Hoekman and Quiñones, 2002). Several specific cases, such as individual trees, mangrove, primary forest and burnt primary forest are considered in order to elucidate the direct effect of physical structure on the radar return signal.

Changes in intensity are expected to occur as described in the literature, i.e. higher backscatter values where more scatterers of a certain size (in relation to wavelength) occur. To understand the physical interaction between the forest and radar waves, the scattering mechanisms can be simplified into a model composed of three contributions (Ulaby *et al*., 1986). These are: (1) the direct backscattering from the vegetation layer, (2) the direct backscattering from the ground attenuated by the vegetation cover and (3) the backscattering originating from ground-trunk interaction attenuated by the vegetation cover. These are sometimes referred to as the diffuse (volume) term, the single-bounce (odd) term and the double-bounce (even) term, respectively. For diffuse scattering the polarimetric coherence is low and the polarisation phase difference (PPD) can be anything in the range between 0^o to 360^o. For single-bounce scattering the coherence is high and the PPD is around 0^o. For double-bounce scattering the coherence is high and the PPD is around 180^o (Van Zyl, 1989). The dominance of a certain scattering mechanism might, in some cases, be related to certain vegetation structures. In Hoekman and Quiñones, 2002, a model is introduced in which values of complex coherence are described as the combined effect of such scattering mechanisms. Pure scattering mechanisms as described in (Van Zyl, 1989) can be located in the complex coherence plane as points or, when including the effect of speckle, as small areas (Hoekman and Quiñones, 2002). For the complex structure of the tropical forest, pure scattering mechanisms are not expected to occur. But it is

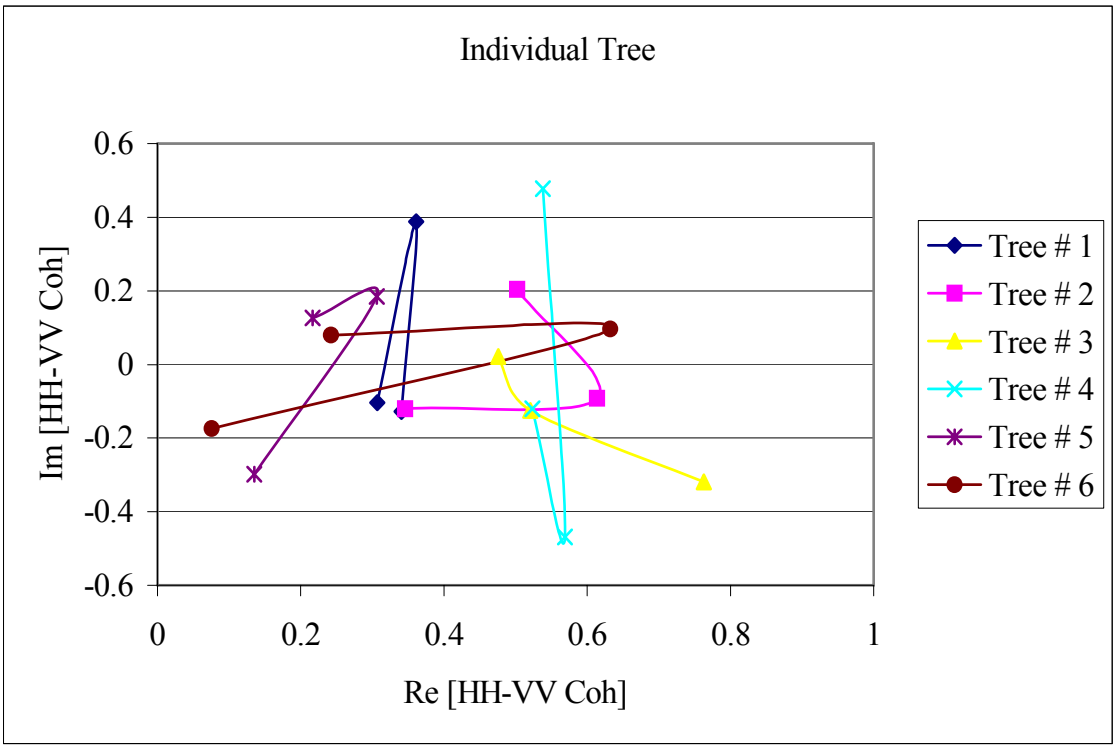
expected that some patterns in the signatures can be related to specific forest structural characteristics.

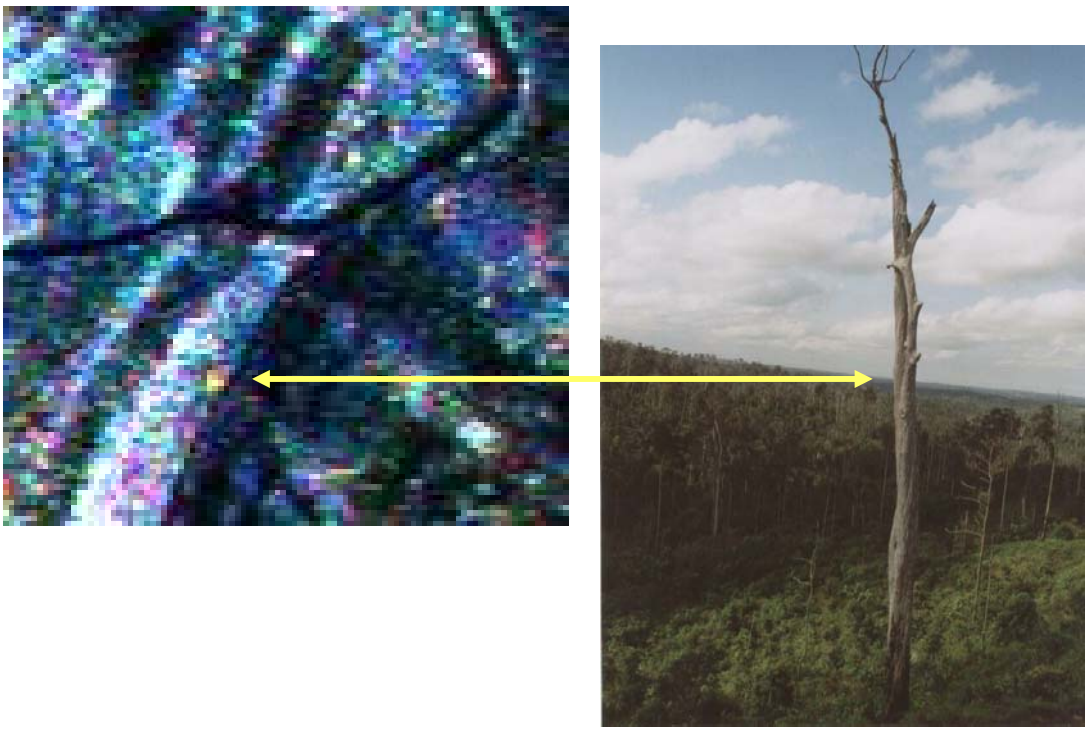
Figure 4.8 shows the frequency dependence of the complex coherence of some characteristic features, which may relate strongly to the forest structure as illustrated in the photograph of actual field observation. The signatures of four different specific cases of forest structures are presented in Figure 4.8a-d and the complex coherence value of the individual tree, mangrove, primary forest and burnt primary forest are shown. The curves connect the C-band coherence (in these examples almost always the right-most point) with the L-band (in the middle) and P-band value (at the other, i.e. the left-end, of the curve). Except for the curves of two mangrove plots (in Figure 4.8b) all curves lie in the first and fourth quadrant.

The first case (Figure 4.8a) shows individual dead standing trees with standing trunks, but without or with very few remaining branches. The second case is a mangrove forest which features a high density and a very closed canopy over partly flooded terrain. The third is a primary forest with a high density and with a closed canopy over undulating terrain. The fourth case is burnt primary forest with medium density and an open canopy and many dead standing tree trunks remaining after fire events, over undulated terrain.

The multi-frequency complex coherence curves of dead standing trees (Figure 4.8a) do not show a distinct pattern. Since all curves lie completely in the first and fourth quadrant, it is clear that the double-bounce mechanism, or the trunk-ground interaction, is not dominant. A more detailed analysis for tree number 6 is even more revealing (see detail of radar image and photograph). The tree is observed as a bright yellow dot on the left side of a small road. The actual position of the tree is at the right side of the road, i.e. more to the far range of the radar image. Considering the radar observation geometry it can be shown that the scattering centre is roughly in the middle of the trunk, because of its positional shift or parallax towards the radar. These two facts: parallax shift and absence of double-bounce characteristics strongly support the hypothesis that a *direct trunk scattering* mechanism is observed. Moreover, since the tree is located in an area with steep slopes, and is devoid of any substantial flat area between radar and tree, a double-bounce return is highly unlikely. The yellow colour is indicative for strong scattering in L- and P-band (conventional AirSAR colour coding is used). Thus, it may be concluded that *dead standing trunks contribute significantly to the overall forest backscatter in L- and P-band*.

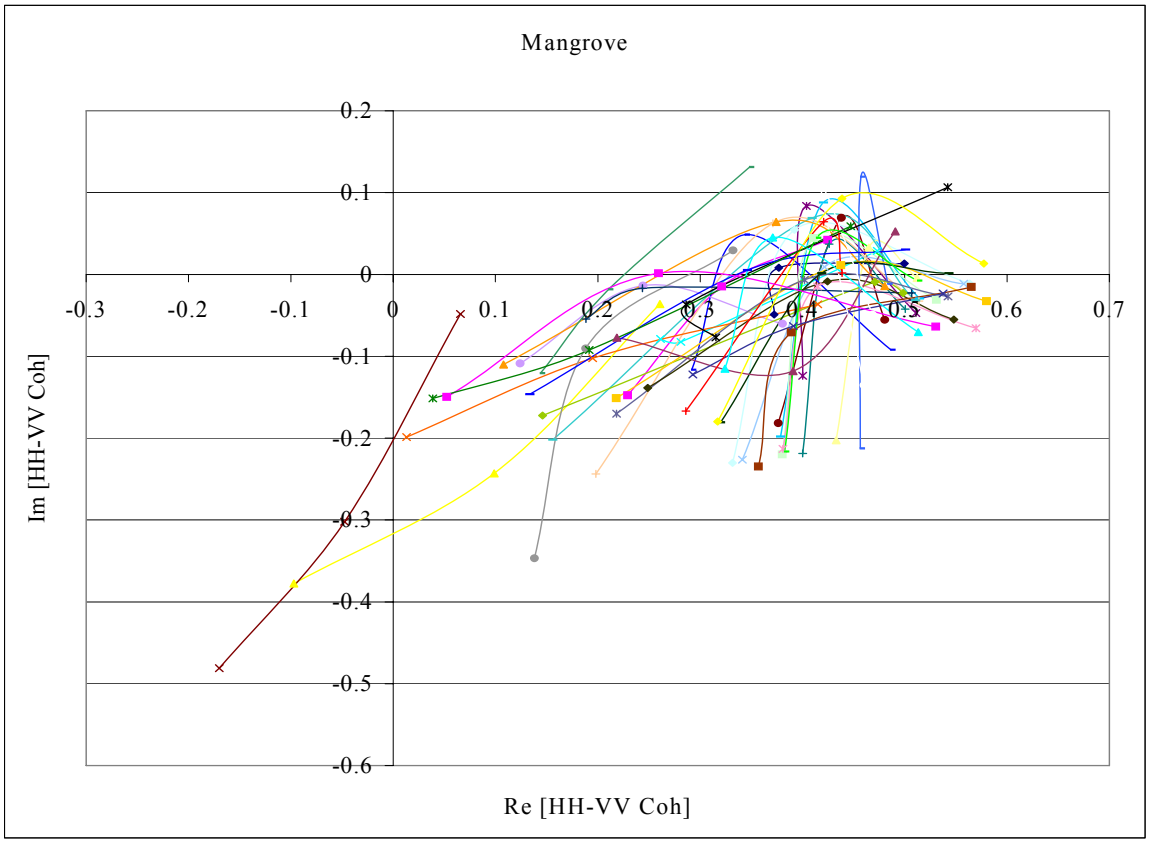
For all other forest types (Figure 4.8b, c and d) there is a decrease in the correlation from C-band through L-band towards P-band. This indicates a gradual shift from surface scattering (i.e. a top of canopy observation) to volume scattering or increasing penetration with wavelength. Since a strong double-bounce contribution would lead to a high correlation in combination with a phase around 180 degrees (i.e. in the second or third quadrant), it may be concluded that large trunk ground interaction contributions are absent. Physically this may be associated with a combination of thick vegetation with high attenuation and/or a forest floor which is sloping. The only exceptions are two curves of mangrove plots which feature such behaviour for the P-band in a lesser extent. This may be caused by a relatively large fraction of not-dense canopy over a flooded terrain.





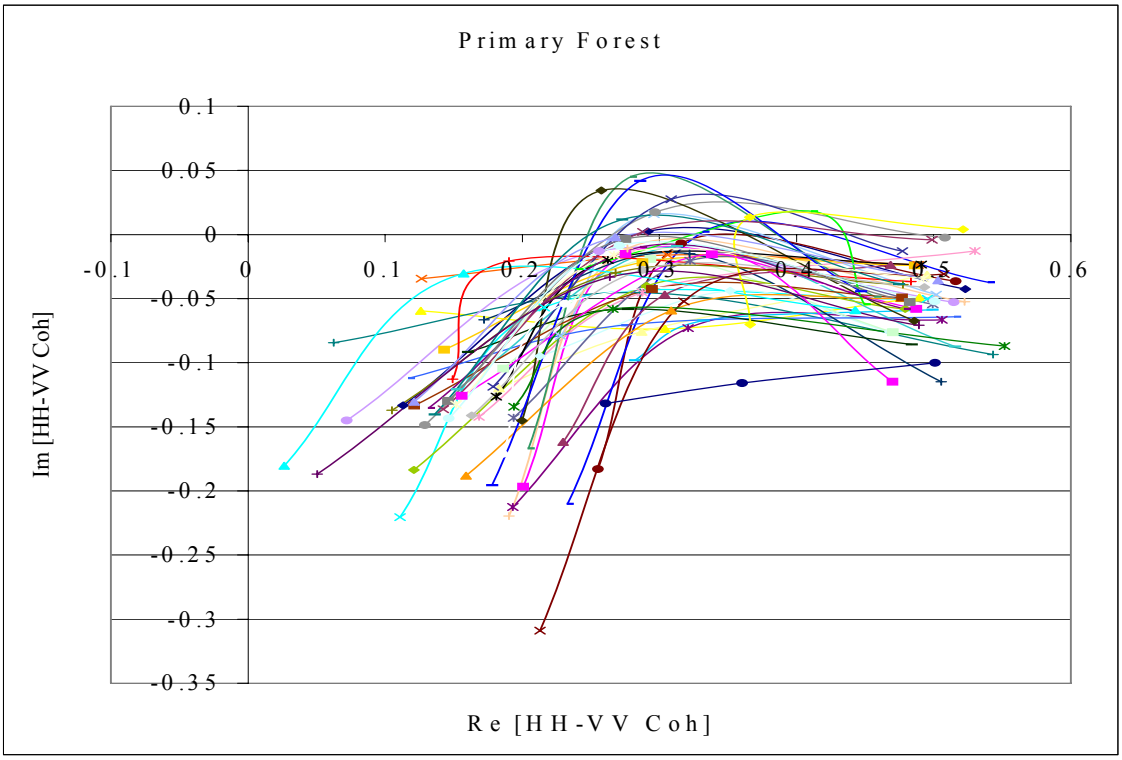
(a)

Figure 4.8a. Multi-frequency complex coherence curves for *individual dead standing trees*, with an example image fragment and a photograph from the field for tree # 6. The curves connect C-band complex coherence (in these examples mostly the right-most point), with the L-band and the P-band value (the latter mostly at the left-end of the curve).



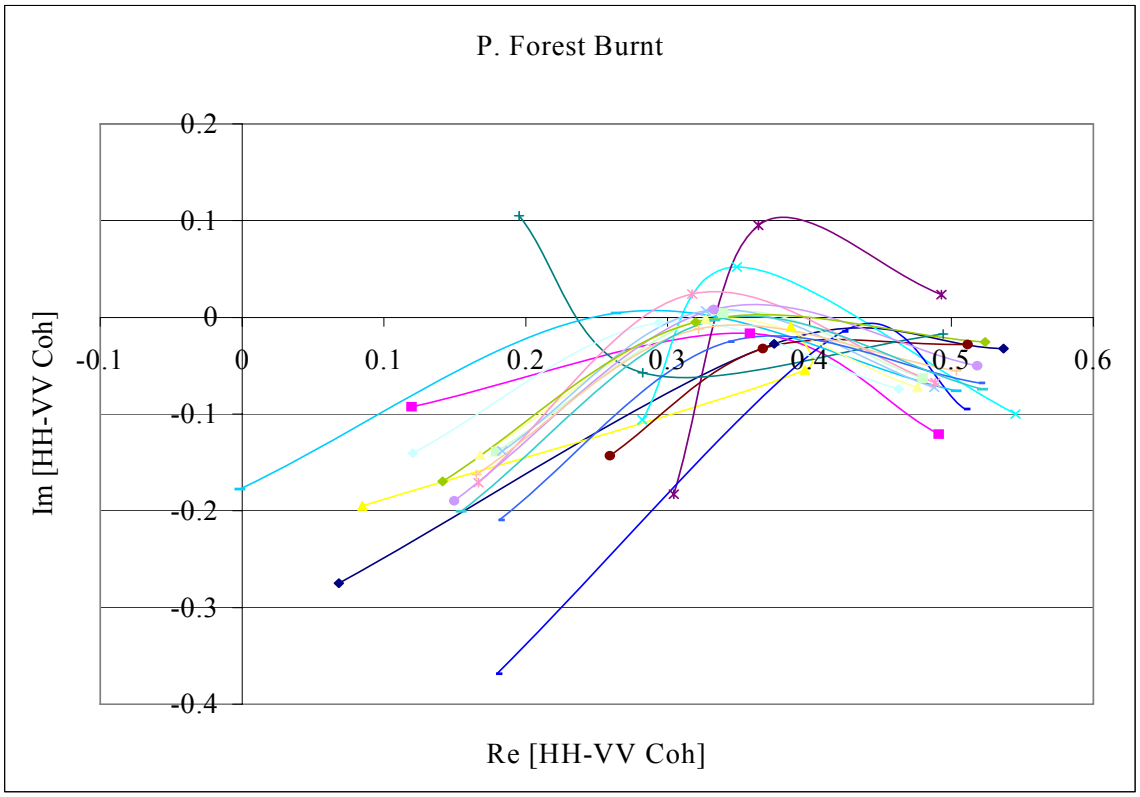
(b)

Figure 4.8b. Multi-frequency complex coherence curves for *mangrove*, with characteristic photograph. The curves connect C-band complex coherence (in these examples generally the right-most point), with the L-band and the P-band value (the latter generally at the left-end of the curve).



(c)

Figure 4.8c. Multi-frequency complex coherence curves for primary forest, with photograph. The curves connect C-band complex coherence (in these examples always the right-most point), with the L-band and the P-band value (the latter always at the left-end of the curve).



(d)

Figure 4.8d. Multi-frequency complex coherence curves for (d) primary forest burnt, with photograph. The curves connect C-band complex coherence (in these examples always the right-most point), with the L-band and the P-band value (the latter always at the left-end of the curve).

4.8.1 Conclusions

Conclusions of the physical interpretation of signal characteristics of several examples of land cover types as mentioned are the following:

- The approach using interpretation of multi-frequency complex coherence can provide insight in physical backscatter mechanisms taking place for every forest structural type.
- *Isolated dead standing trees give a significant contribution in L- and P-band through direct (single-bounce) scattering from the trunk,* and not through the double-bounce trunk-ground interaction mechanism.
- *Double bounce scattering is rare*. Some indications were found it occurs in flooded parts of relatively open canopy mangrove areas in P-band.
- In general, for all forests canopies, single bounce dominates in C-band (high coherence) and diffuse scattering dominates in L- and P-band.

5. Comparison and validation of land cover type classification with independent data sets

5.1 Introduction

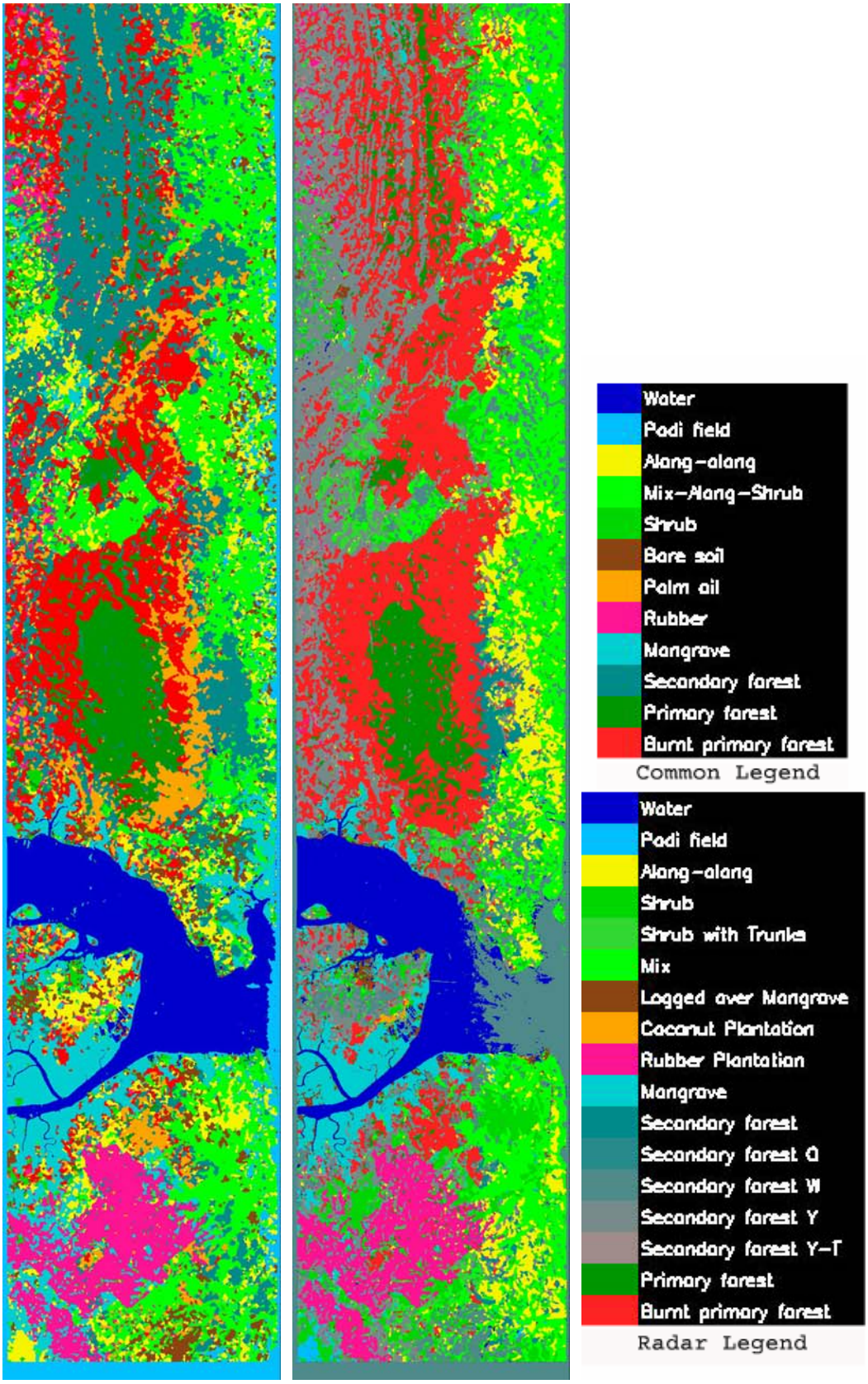
This chapter presents an evaluation and validation of the results of land cover type classification obtained in Chapter 4, using an independent data set acquired by another researcher (Rodriguez, 2002), --- see also section 3.2. The validation will be both quantitative (ML and ICM approaches) and qualitative (visual comparison).

In Chapter 4 the accuracy assessment was based on all available ground truth areas in the 30°-60° incidence angle range. However, the training of the classifier was based on the same ground truth areas. In principle the available ground truth areas could have been split into separate groups of training and validation data, which may have lead to a slightly poorer result. However, since the aim of this study was to assess relative comparison between utilities of factors such as different approaches or certain frequency band combinations, the accurate assessment of the actual accuracy is not of primary importance. Fortunately, the presence of a completely independent data set allows the assessment of a good validation. Of course, both sets can mutually act as training and validation sets.

It is very difficult to make a good quantitative assessment, because the legend of the independent data set is different. The legend used in Chapter 4 of this study uses class *definitions commonly used in Indonesia*. The legend of the independent data set is the *'radar' legend based on physical structures the radar is expected to be able to observe* and, therefore, may have a more general application. It should be noted, however, that there are no real one-to-one or one-to-several correspondences between the *independent data set or 'radar' legend set* and the other *'original' data set or 'common' legend set* (see Figure 5.1c). It should be noted that one year elapsed between the acquisitions of both sets of data. For this study the data were collected during the campaign in September 2000. *The independent data set was collected 12 months later*, i.e. during the same season, and was guided by the availability of the radar data. Of course, the field plots were selected with care and fields that experienced recent change (i.e. during the previous year) were excluded.

5.2 Land cover types of independent data set

The independent data set consists of *18 classes* and was shown in Table 3.1. To make it more comparable with the changes to the original data set made in Chapter 4, the class "water" was added and the class "man-made structure" removed, and the class "shrimp ponds" was added to the class "water". The class "swamp" in the independent data set is actually abandoned *padi* field, and is re-labelled as *padi* field. Therefore, 17 classes were used. In case this data set is used for training as well as for classification using the 7I model and ICM technique, then an overall map accuracy of 93.8% results (Table 4.7; Case j). Table 5.1a and 5.1b show the confusion matrix in percentages and in number of pixels, repectively. Figure 5.1b shows the resulting image after completion of 20 ICM cycles. The classes are coded with the colour legend described in Figure 5.1c.



(a) (b) (c)

Figure 5.1. (a) The classification result for the original data set and the *common legend*, (b) the classification result for the independent data set and the *radar legend* and (c) the legends.

In general, *satisfying results were obtained*. The classification was highly accurate (more than 95%) for water, swamp, shrubs-trunks, logged (or dead) mangrove, rubber plantation, mangrove, secondary forest (wet), secondary forest Y-T (young with trunks), primary forest and burnt primary forest. For the classes *alang-alang*, shrubs, mixed, coconut plantations, secondary forest (old) and secondary forest (young), the classification accuracy varied from 79.3% for mixed to 94.7% for secondary forest (young). The classification was less successful for the secondary forest only (34.4%). The class mixed was often confused with shrubs, *alang-alang* and shrubs-trunks. Secondary forest was often classified as primary forest and burnt primary forest. In this independent database, the classes which yielded the greatest confusion were secondary forest (34.4%) and, to a much lesser extent, mixed (79.3%).

Table 5.1a. Confusion matrix for the C-, L- and P-band fully polarimetric combination, using the independent set for training as well as for validation. The map is shown in Fig. 5.1b. Results are expressed *in percentages*.

	WATE	SWAM	ALAN	SHRU	SHRT	MIXE	LOGG	COCO	RUBB	MANG	SECF	SECO	SECW	SECY	SECT	PRIM	BURN
WATE	**99.0**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SWAM	0.0	**98.7**	0.1	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ALAN	0.0	0.0	**86.5**	4.7	0.3	5.5	0.0	0.0	0.1	0.0	1.0	0.0	0.0	0.1	1.5	0.0	0.0
SHRU	0.0	0.4	0.4	**81.0**	0.1	8.1	0.0	0.0	0.2	0.1	0.3	0.0	0.3	0.1	0.0	0.0	0.0
SHRT	0.0	0.0	1.1	5.5	**98.8**	5.0	0.1	0.0	0.0	0.1	0.2	0.0	0.0	0.2	0.1	0.0	0.0
MIXE	0.1	0.8	1.5	3.4	0.3	**79.3**	0.0	12.2	0.0	0.0	1.3	0.0	2.1	0.4	0.3	0.0	0.2
LOGG	0.1	0.0	0.0	0.0	0.0	0.0	**95.3**	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.4	0.0
COCO	0.0	0.0	0.0	0.0	0.0	0.0	0.3	**86.0**	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
RUBB	0.0	0.0	6.4	0.2	0.4	0.2	0.0	0.0	**96.4**	0.0	0.9	0.2	0.0	2.3	1.0	0.2	0.4
MANG	0.0	0.1	0.6	0.8	0.2	0.2	0.8	0.1	1.4	**99.3**	0.0	0.0	0.1	0.3	0.4	0.0	0.0
SECF	0.0	0.0	0.0	0.0	0.0	0.1	0.4	0.3	0.0	0.0	**34.4**	0.0	0.0	0.1	0.0	0.0	0.0
SECO	0.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	0.0	1.6	**83.5**	0.0	0.0	0.1	0.9	0.3
SECW	0.8	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**97.6**	0.1	0.0	0.0	0.0
SECY	0.0	0.0	2.3	3.7	0.0	1.3	0.0	0.0	1.5	0.2	0.6	0.4	0.0	**94.7**	0.4	0.2	0.4
SECT	0.0	0.0	1.1	0.1	0.1	0.3	0.2	0.0	0.2	0.0	0.3	0.0	0.0	0.8	**95.7**	0.1	0.4
PRIM	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	44.1	4.9	0.0	0.0	0.0	**95.7**	0.8
BURN	0.0	0.0	0.0	0.0	0.0	0.0	0.9	1.4	0.1	0.1	14.4	11.0	0.0	1.0	0.5	2.5	**97.6**

Table 5.1b. Idem Results are expressed *in pixels*.

	WATE	SWAM	ALAN	SHRU	SHRT	MIXE	LOGG	COCO	RUBB	MANG	SECF	SECO	SECW	SECY	SECT	PRIM	BURN
WATE	**50266**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SWAM	0	**1824**	2	20	0	1	0	0	0	0	0	0	0	0	0	0	0
ALAN	14	0	**2944**	265	5	318	0	0	7	2	31	0	0	4	25	0	1
SHRU	0	7	14	**4524**	1	471	0	0	8	7	8	0	8	7	0	0	0
SHRT	0	0	36	306	**1763**	294	2	0	2	10	5	0	0	11	2	0	0
MIXE	26	15	51	192	5	**4629**	0	179	0	0	39	0	63	21	5	2	7
LOGG	71	0	0	0	0	0	**2501**	0	0	1	25	0	0	0	0	19	1
COCO	0	0	0	0	0	0	9	**1264**	0	1	3	0	0	0	0	0	0
RUBB	0	0	218	10	7	14	1	0	**4656**	3	28	2	0	128	17	12	17
MANG	0	2	19	47	3	14	22	2	66	**7006**	0	0	2	15	6	0	0
SECF	0	0	1	0	0	7	10	4	0	2	**1032**	0	0	3	0	0	0
SECO	0	0	0	0	0	0	47	0	2	0	47	**1105**	0	0	1	50	15
SECW	380	0	0	12	0	0	0	0	0	0	0	0	**2930**	3	0	0	0
SECY	0	0	79	205	0	74	0	0	72	15	19	5	0	**5190**	7	12	18
SECT	0	0	38	4	1	16	4	0	11	3	10	0	0	42	**1608**	6	17
PRIM	0	0	0	0	0	0	6	0	2	2	1325	65	0	0	0	**5205**	34
BURN	0	0	0	1	0	0	23	21	4	4	431	146	0	57	9	133	**4447**

5.3 Land cover type validation with independent data set

The results obtained in Chapter 4 will now be evaluated and compared with the independent data set. Figures 5.1a and 5.1b show the maps of the classification for the original and the independent data sets, respectively. Table 5.2a shows the confusion matrix for the classification result using the original data set for training and the field data within the sample plots (polygons) of the independent data set. Table 5.2b shows the confusion matrix for the reverse case.

5.3.1 Forward approach

To quantify the performance of the classification results, the confusion matrix was used. The independent database polygons were used as absolute truth, and afterwards the misclassifications were counted in those regions. It is noted that the relationship is not one-to-one; classes with no clear corresponding classes in the other data set will be included in the class considered as the most similar one (see table 3.1 and sections 3.2.2 and 3.2.3 for description of each class). The results are shown as a confusion matrix (Table 5.2a) which shows the number of pixels that were misclassified. The confusion matrix may be able to explain the occurrence of misclassification.

Water has very high classification accuracy, i.e. no confusion occurs between water and other cover types. Rice *(padi)* fields exist only in the original data set, and were validated mainly as water and less frequently as *alang-alang*, mixed and logged mangrove in the areas of the independent data set. This could be a logical consequence of the fact that *padi* fields consist of water and small vegetation, just like *alang-alang*, mixed and logged mangrove.

Alang-alang in the original data set is mainly validated by the class shrubs, and less frequently as coconut plantation and wet secondary forest. The class *mixed* in the classification map is mainly validated as class mixed, *alang-alang* and shrubs and less frequently as shrubs-trunks, secondary forest-W (wet), swamp, logged mangrove, and coconut plantation. The class *shrubs* in the classification map is mainly validated as coconut plantation, logged mangrove and even water, and less frequently as secondary forest-Y-T (young with trunks). The class *bare soil* in the classification map is mainly validated as secondary forest-W (wet), mixed and swamp, and less frequently as shrubs and *alang-alang*.

The *oil palm* class exists only in the original data set and is validated mainly as secondary forest-Y-T (young with trunks) and secondary forest, and less frequently as mixed and *alang-alang*. The validation accuracy for the *rubber plantation* class is high, but is sometimes validated as secondary forest-Y (young), secondary forest-Y-T (young with trunks) and *alang-alang*. The validation accuracy for *mangrove* is also high and only sporadically validated as shrubs, logged mangrove and water.

Secondary forest in the classification map is mainly validated as secondary forest-Y (young) and secondary forest, and less frequently as *alang-alang*, secondary forest-O (old), secondary forest-Y-T (young with trunks) and shrubs. The classification accuracy for the *primary forest* is relatively high and is less frequently validated as secondary forest, secondary forest-O (old) and burnt primary forest. The validation

accuracy for burnt primary forest is also relatively high, and less frequently validated as logged mangrove, secondary forest-Y-T(young with trunks), secondary forest-Y (young) and shrubs-trunks.

It can be concluded that classes which have a high agreement are water, rubber plantation and mangrove. Primary forest and burnt primary forest have a good level of agreement, secondary forest also has a good level of agreement with validation classes such as secondary forest (young), secondary forest-O (old) and secondary forest-Y-T (young with trunks), which all can be considered as secondary forest classes. Classes with lower validation level are: mixed, shrubs, *alang-alang*, bare soils and oil palm. It should be noted that the oil palm class exists only in the original data set.

5.3.2 Reversed approach

For further assessment of the potential of mapping and to test the physical consistency, the process has been reversed. A map was made based on the independent data set and this map is validated using the polygons from the original data set. The confusion matrix of this result is shown in table 5.2b, now with the original data set as the absolute truth.

Water has very high classification accuracy, and is only slightly confused with shrubs and rice (*padi*) fields. *Swamp* exists only in the independent data set and is not coinciding with the basic truth of the original data set (there only 2 pixels of class mixed). *Alang-alang* is mainly validated as secondary forest and shrubs, and less frequently with mangrove and mixed. *Shrubs* has a low agreement, it is mainly validated with mangrove and *alang-alang* and less frequently with bare soils and mixed. *Shrubs-trunks* exists only in the independent data set and is validated mainly as shrubs, and is less frequently validated as *alang-alang* and mixed. The class *mixed* has a low agreement and is mainly validated with mangrove and less frequently with shrubs, *alang-alang* and *padi* field. The *coconut plantation* class exists only in the independent data set and, correctly, is not found in any of the areas of the original data set. The validation accuracy for the *rubber plantation* and *mangrove* classes is very high.

Secondary forest is mainly validated as shrubs and rubber. *Old* s*econdary forest* has a good agreement and is less frequently validated as primary forest. *Wet secondary forest* is mainly validated as bare soils, water and *padi* field. *Young secondary forest* is mainly validated as secondary forest and less frequently as shrubs and *alang-alang*. *Young secondary forest with trunks* is validated as shrubs. However, if classes such as young secondary forest, old secondary forest and young secondary forest with trunks can be considered as one class, i.e. the secondary forest, the validation has good agreement.

Primary forest has a high accuracy and is less frequently validated as secondary forest, oil palm and rubber plantation. The accuracy for *burnt primary forest* is also relatively high and less frequently validated as secondary forest, oil palm and primary forest.

It can be concluded that classes with good agreement between both data sets are water, rubber plantation, mangrove, logged mangrove and primary forest. Burnt primary forest, old secondary forest and young secondary forest also have a good level of agreement. Wet secondary forest is difficult to validate because there is no clear corresponding class. Classes with a low validation level are swamp (abandoned *padi* field), *alang-alang*, shrubs, shrubs-trunks and mixed. It should be noted that the coconut plantation class exists only in the independent data set.

5.3.3 Visual approach

For *qualitative evaluation*, a visual comparison between the maps shown in Figures 5.1a and 5.1b has been made. The legend of both Figures is shown in Figure 5.1c. According to the author's terrain knowledge, as a general rule, the classification result in Figure 5.1a seems more general than the one of Figure 5.1b. The classification results show a good agreement in size and location for the classes water, mangrove, rubber plantation, secondary forest, primary forest and burnt primary forest. For the classes *alang-alang*, mixed, shrubs and bare soils, the agreement is poor.

The main error in the 'common' legend map derived from the original data set appears to be a primary forest area located at the northern side of the Balikpapan Bay (at the eastern side of the Sungai Wain protected forest area). According to the common legend map, it is misclassified as oil palm. In the 'radar' legend map it is correctly classified as a burnt primary forest area. This misclassification may be caused by the absence of suitable legend units in the common legend map.

5.3.4 Conclusions

From the evaluation results above, it can be concluded that the radar legend and common legend maps have an acceptable consistency level. Water, rubber plantation and mangrove have a high validation level. Primary forest and burnt primary forest have a good level of validation, while the secondary forest also has a moderate validation level. Classes with a low validation level are mixed, shrubs, *alang-alang* and bare soils. This can be explained by the fact that these classes have a large variation of vegetation structure with a rapid vegetation transformation, and can be found on dry as well as on wetland, making them difficult to be clearly discriminated.

It can be noted that the map resulting from the independent data set yields significantly better classification results. This may be explained by the fact that the variation in physical structures are much better captured in the 'radar-based' legend. A similar conclusion was drawn by Hoekman and Quiñones (2002). Moreover, it should be noted that the fact that a larger number of classes generally gives a lower classification result may not apply to these data sets. In Quiñones and Hoekman (2004) theoretical simulation results clearly reveal that multi-frequency polarimetric data sets have an enormous information content, and that well chosen classes, even when there are several hundreds of classes, can still be correctly separated.

Since the classification approach utilizes the supervised classification method, the subjectivity of choice and spread of ROI's may influence the classification results.

Accurate field observation and measurement play a crucial role in analysis and validation. Local variation in physical characterisation of the land cover types related to factors such as variation in forest structure and different conditions may also have important effects. In addition, the effect of a strong ICM process can be positive or negative. In this process, certain cells that initially are classified incorrectly can receive a new and correct class when neighbouring cells of the same class are correctly classified. However, when the relaxation parameters have values which are too high, isolated small areas may disappear. In order to achieve more optimal results such effects may need to be studied in more detail, i.e. maintaining good overall results without losing the fine spatial structures. It is noted that a large proportion of training area (ROI's) are located in such fine-structured regions.

Though an extensive good quality data set was available, *the importance of repeating the experiment in other areas of the Indonesian forests, and in different conditions, should be emphasized*. Different land cover situations, such as levels of fragmentation, soil moisture variations or relief factors may have significant effects on the results. *Such a wider scope study will provide further examination of how robust the algorithm employed in this study is*, to enable operational applications in the future. It is interesting to note that such an experiment has been executed in November 2004 in the framework of the ESA-MOF INDREX-2 campaign.

Table 5.2a. Confusion matrix for land cover type classification. The columns present the result of the classification *in pixels* and the rows represent data from the independent database used for validation. The result is obtained after applying 50 cycles of the extended ICM approach.

	WATE	PADI	ALAN	MIXE	SHRU	BARE	PALM	RUBB	MANG	SECO	PRIM	BURN	
WATE	2803	14	0	0	11	3	0	0	7	0	0	0	Water
SWAM	0	0	5	34	0	49	0	0	0	0	0	0	Swamp
ALAN	0	6	7	118	0	12	8	10	0	19	0	0	Alang alang
SHRU	0	0	170	118	0	17	0	0	57	6	0	0	Shrubs
SHRT	0	0	5	65	0	0	0	0	0	3	0	12	Shrubs-T
MIXE	0	6	3	260	0	58	13	1	0	3	0	5	Mixed
LOGG	0	6	9	19	11	1	0	0	22	0	0	57	Logged Mangrove
COCO	0	0	37	19	13	0	1	0	0	0	0	0	Coconut plantation
RUBB	0	0	0	0	0	0	0	240	8	1	0	0	Rubber plantation
MANG	0	0	0	0	0	0	0	0	401	0	0	0	Mangrove
SECF	0	0	0	2	0	0	26	3	0	65	55	1	Secondary forest
SECO	0	0	0	0	0	0	0	0	0	14	47	2	Secondary forest-O
SECW	0	0	31	22	0	90	0	0	0	0	0	0	Secondary forest-W
SECY	0	0	9	6	4	1	0	14	0	248	0	16	Secondary forest-Y
SECT	0	0	0	0	0	0	37	9	0	12	0	22	Secondary forest-Y-T
PRIM	0	0	0	0	0	0	0	0	0	0	254	5	Primary forest
BURN	0	0	0	0	0	0	0	0	0	2	22	193	Burnt primary forest

Table 5.2b. Confusion matrix for the validation of land cover type classification using the reversed approach. The columns present the result of the classification of the independent data set *in pixels* and the rows represent data from the original data set. The result has been obtained after applying 20 cycles of the extended ICM approach.

	WATE	SWAM	ALAN	SHRU	SHRT	MIXE	LOGG	COCO	RUBB	MANG	SECF	SECO	SECW	SECY	SECT	PRIM	BURN	
WATE	1002	0	0	0	0	2	0	0	0	0	1	0	414	3	0	0	0	Water
PADI	1	0	0	2	0	11	0	0	0	0	0	0	187	0	0	0	0	Padi field
ALAN	0	0	3	27	17	56	0	0	0	6	0	0	97	70	0	0	14	Alang-alang
MIXE	0	2	13	8	12	60	0	0	0	0	0	0	17	7	0	0	0	Shrub
SHRU	2	0	55	36	36	59	1	0	0	0	3	0	45	98	1	0	15	Shrimp ponds
BARE	0	0	1	13	2	11	0	0	0	0	0	0	563	3	0	0	0	Bare soil
PALM	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	30	361	Oil palm
RUBB	0	0	0	2	1	0	3	0	497	0	2	0	0	10	0	20	0	Rubber
MANG	8	0	27	51	0	170	16	0	1	699	0	0	3	0	0	0	0	Mangrove
SECO	0	0	65	0	0	1	1	0	0	0	0	25	0	229	0	46	339	Secondary forest
PRIM	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	2570	211	Primary forest
BURN	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2130	Burnt primary forest

5.4 Evaluation of land cover type classification results

To evaluate the *relative importance of the steps leading to the overall or final land cover type classification results*, it is useful to distinguish between the *choice of radar system parameters* and the *choice of image processing procedures*. Important system parameters are polarisation, frequency band(s) and radiometric resolution (or number of looks). Image processing procedures can be divided into three main steps: the pre-processing steps (such as relief correction), the processing steps (such as legend choice, the full-polarimetry model, ML or texture), and post-processing steps (such as ICM).

5.4.1 Effect of the radar parameters

Tables 4.2, 4.3 and 4.7 showed that the overall classification result is affected by many factors, including polarisation combination, combination of bands being used, and also the number of looks (radiometric resolution).

Interaction of the radar with the land cover might result in a change of polarisation. First, microwave interaction may cause the polarisation plane of the scattered waves to be different from that of the incident wave. Second, the microwaves' interaction with the land cover may cause depolarisation. Due to the depolarisation, the scattered waves become partially polarised. Changes in the polarisation and the depolarisation of the radar waves, and the combination of both, are often object specific, and, therefore, an important source of information for differentiating objects in radar images. Use of all polarisation factors will result in better classification results. Therefore, the results for the so-called fully polarimetric cases are superior

The wavelength λ is of primary importance in the interaction of microwaves with forest vegetation, as it affects the penetrating capacity of the microwaves and the spatial distribution of the scattered power. For instance, for C-band the penetration is low, and the backscatter from the crown layer dominates; for L- and P-band the penetration is higher and contributors to backscatter are the trunks and the trunk-soil interaction, respectively. Combining the information of these three bands may yield information over a range of forest elements, in relation to the size of the wavelength and forest scatterers. This implies that particular wavelengths are more suitable for certain applications. Therefore, multi-band radar provides complementary information sources of the forests. Hence, multi-band combinations are more suitable in differentiating objects (classification) of the forest than single-band. From the results in this thesis, the combination of 3 bands is superior to the combination of 2 bands.

The accuracy of the estimation of field averaged values depends on the total number of independent looks N. The number of looks or radiometric resolution will affect the object detection. Detection becomes more difficult when speckle dominates the intensity images. By increasing the number of looks from N=20 to N=64, the classification result shows a considerable increase.

5.4.2 Relative importance of factors in the classification procedure

Effect of relief correction

Illumination effects caused by relief have been corrected in the AirSAR fully polarimetric data, which visually gives a much more homogenous perception (Figure 4.1). Relief correction is a pre-processing step, an important step that must be executed before the classification process. However, for classification purposes, as shown in Table 4.7, cases d/e and cases f/g show that the effect of the relief correction is not significant for the overall classification results (in these cases for dual fully polarimetric band combinations). It can be explained as follows.

Re-writing equation 2.14:

$$C = \frac{1}{4\pi} \begin{pmatrix} \sqrt{\sigma^0_{hh}} & 0 & 0 \\ 0 & \sqrt{\sigma^0_{hv}} & 0 \\ 0 & 0 & \sqrt{\sigma^0_{vv}} \end{pmatrix} \begin{pmatrix} 1 & \rho_{hhhv} & \rho_{hhvv} \\ \rho^*_{hhhv} & 1 & \rho_{hvvv} \\ \rho^*_{hhvv} & \rho^*_{hvvv} & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\sigma^0_{hh}} & 0 & 0 \\ 0 & \sqrt{\sigma^0_{hv}} & 0 \\ 0 & 0 & \sqrt{\sigma^0_{vv}} \end{pmatrix} \tag{5.1}$$

by separating the factor σ^o_{hh}, the equation becomes:

$$C = \frac{1}{4\pi} \sigma^0_{hh} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\sigma^0_{hv}/\sigma^0_{hh}} & 0 \\ 0 & 0 & \sqrt{\sigma^0_{vv}/\sigma^0_{hh}} \end{pmatrix} \begin{pmatrix} 1 & \rho_{hhhv} & \rho_{hhvv} \\ \rho^*_{hhhv} & 1 & \rho_{hvvv} \\ \rho^*_{hhvv} & \rho^*_{hvvv} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\sigma^0_{hv}/\sigma^0_{hh}} & 0 \\ 0 & 0 & \sqrt{\sigma^0_{vv}/\sigma^0_{hh}} \end{pmatrix} \tag{5.2}$$

Written in this way, fully polarimetric radar target properties contain five values: two polarisation ratios, three complex coherences (3 x 2 numbers) *and only one backscatter value*. The complex coherence numbers represent the difference in the phase angle for the HH and VV polarized signal, the two polarisation ratio numbers represents correlation magnitude for the HH and VV polarized signal and both of them are not affected by backscatter intensity. Since only this single backscatter value is modulated by relief, and all other values are not, relief correction only affects 1 out of 9 numbers of the polarimetric backscatter signal properties. *The effect of relief correction on the full polarimetric classification can be small in the case that the other 8 numbers provide sufficient information.* The runs mentioned above prove that this is indeed the case. Of course, in case the full polarimetric information is not used, *then the effect of relief correction can be large*, such as visually in the intensity image (Figure 4.1), or in the relationship between biomass and backscatter intensity (Section 4.7), or in the classification of single polarisation data

Effect of using texture

Since 3-D texture derived from C-band InSAR DEM and C-band backscatter intensity differentiates between forest structural types, it can be expected to have a significant influence on the classification result. Nevertheless, this additional source of knowledge, as Table 4.7: cases i/c and cases h/a show, hardly has any effect. The reason may be very similar as the one discussed above for the relief. If we only have the C-band for the classification process, the use of texture might be important, but using *a fully polarimetric combination of C- and L-band or C-, L- and P-band will provide enough information, so that additional information of texture will not be required.*

This evaluation shows that the main factor affecting the classification result is the amount of polarimetric information and the number of frequency bands; the best results will be achieved by using fully polarimetric information. *Definitely, a high number of radar looks will be very helpful, while relief correction and texture or 3-D texture is not significant in the process of improving classification results, provided a sufficient amount of polarimetric information and frequency bands is available*.

Effect of using ICM

Post-processing through ICM is very useful. The likelihood of a pixel was modified by a conditional probability in which the number of neighbours of a certain class is used in the classification. This technique proved to have an important effect on the classification accuracy. Variations in the confusion between classes have to be carefully studied when applying this algorithm, since classes occurring in spatially small areas can disappear under the presence of more extended classes. *The number of ICM cycles applied to a classified image increased the classification accuracy until reaching a maximum,* in which a stable solution was found, although the appropriate selection of parameters used in the neighbourhood operations has to be optimised by trial-and-error.

Effect of legend choice

Two different types of legend have been used to create classifications of the polarimetric radar images. First, a ‘common’ map classification legend (14 classes), where the land cover is divided into different classes based on the usual definitions of land cover in Indonesia. The second legend (the ‘radar’ legend with 18 classes), is based on the identification of land cover types which are expected to give different backscatter signatures. These classes correspond more to different structural types for the forested areas, where the main difference between the 2 legends for example is the sub-division of the ‘secondary forest‘ class into 5 different classes based on observed differences in the forest structure. Initially, all secondary forests types were grouped in one single class. *The results showed that the radar based legend has higher overall accuracy than the common legend*. Some non-forest classes could not be differentiated well with the radar in both legends, and therefore confusion between these classes remained (e.g *alang-alang*, shrubs and mixed). Other forested classes (e.g. primary forest, primary forest burnt) could be differentiated with the radar

legend. *Even though more legend sub-divisions were used, the radar based legend still gave better results*.

5.4.3 Conclusions

Using the classification procedures introduced here the polarimetric AirSAR images allow for the creation of map legends with more details related to forest structure than are currently in use. A combination of at least two bands will yield good classification results, but *combinations of C-, L- and P-band fully polarimetric radar data will be the best*. A good radar based legend should present classes that are possible to be distinguished with a certain level of accuracy. *ICM post-processing cycles improve the result significantly.*

In the future, the further developments of *unsupervised classification* approaches of polarimetric data instead of using a good radar based legend may automatically allow better structural classifications of the images. These new developments will probably bring new insights into the capabilities of radar systems for tropical forest mapping applications, especially in those areas where it is difficult to obtain field data.

Summary and conclusions

Problem statement

The deforestation rate in Indonesia tends to increase sharply. The Indonesian Government, through the Ministry of Forestry, has decided to stop and/or mitigate this process, and to support actions for sustainable forest management. The Indonesian forests contain a large biodiversity and are an important factor of the hydrological and biochemical cycles, indispensable to keep nature in balance.

To monitor and evaluate the condition of these forests, the government needs up-to-date and accurate information, which is currently still based on photographic images obtained through optical technology, such as airborne aerial images, LANDSAT or SPOT. This kind of technology always encounters problems related with cloud cover. Therefore, the technology that is based on microwaves (RADAR) has the potential to solve the problem. Modern radar techniques such as interferometry and polarimetry are currently advancing rapidly, and this progress is expected to offer the best solution to provide up-to-date and accurate information related with land cover alteration, degree of crown cover opening, biomass, and so on.

Currently, significant obstacles in providing continuous and up-to-date information regarding the forests are:

- The forests in Indonesia are almost continuously covered with clouds, haze or smoke, making it difficult to obtain clear images with optical technology.
- There is limited experience in using alternative technologies such as radar.
- Suitable approaches/software to analyse the accomplishment of this radar technology are not yet commonly available.

Even though the assessment/investigation of radar technology applications in Indonesian forests started in 1991, the results had not been yet satisfactory. In the last decade scientists/researchers spent considerable time and effort to develop algorithms able to classify vegetation structure and estimate the biomass. In the case of Indonesian forests, such algorithms had not been tested or used. In consideration of that situation, this research aimed to *evaluate the use of the new generation C- and L-band interferometric (TopSAR) and C-, L- and P-band fully polarimetric (PolSAR) NASA/JPL AirSAR data for tropical forest type mapping and biomass estimation, and to study the combined use of radar data and additional knowledge base in order to improve the results.*

To assess the value of information contained in the data of this sophisticated AirSAR radar system, this study used and analysed available algorithms, and modified these in order to optimise the application. This study combined a DEM, which was derived from InSAR data, and fully polarimetric data, for which the parameters are: spatial variability, radar backscatter intensity, polarisation and phase. These parameters were corrected for relief illumination effects. Predefined regions in the corrected image, where ground truth was available, were used to construct a database (in this thesis termed as '*the original data set'*). A new method based on a *reversible transform* of the polarimetric radar covariance matrix was applied to describe the full polarimetric target properties. 3-D spatial variations of InSAR data were conceived as image

texture, which could be related to the forest canopy roughness. Subsequently, post-processing steps were performed; using information of the neighbouring pixels (ICM) and with the additional information from the texture the accuracy of the classification results could be increased. Maximum likelihood and Kappa statistics were utilised to evaluate the overall classification results. This technique and the theoretical analysis have been elaborated in Chapter 2.

Study area

The research for this study was executed in the area of tropical forest near Sungai Wain, not far from Balikpapan City ($0^0$83'S, $116^0$76'E), in the East Kalimantan Province, Republic of Indonesia. The size of the area covered by the radar was 10x60 km^2, and was characterised by a complex mosaic of vegetation and land cover types, from tropical lowland evergreen and semi-evergreen dominated by *Dipterocarpaceae*, a variation of *primary forest with emergent trees exceeding 30 m*, and burnt primary forest. The secondary forest originated from deforestation, burnt remains and natural regeneration, and also from vegetation alteration caused by local farming activities. In this area, *alang-alang* covers vast areas throughout the whole study site. The mangrove forest covers the southern area which is located nearby the Balikpapan Bay, and the rest consists of plantations, rice fields, transmigration areas, shrimp ponds, and wasteland.

In the year 2000, the study area was imaged by the C-, L-band InSAR and C-, L- and P-band PolSAR data of the AirSAR/TopSAR NASA/JPL in the framework of the PacRim-2 campaign. To support the study objectives, field data were collected through extensive and intensive observation. Another '*independent data set'* was available to validate the results of this study. Detailed description of the study areas and available radar data can be found in the Chapter 3.

Biophysical parameter retrieval and land cover type classification

Chapter 4 discussed the research results with emphasis on the evaluation of radar system characteristics in capturing tropical forest parameters, land cover classification and biomass estimation.

Analyses of the relief correction applied to the multi-band composite of C-, L- and P-band PolSAR data using the InSAR DEM showed that the illumination effects of terrain slope and canopy undulation had been reduced. Images produced generally appeared more uniform and more homogenous and enabled to visual distinction of land cover differences more clearly.

Land cover type classification simulations, before and after relief correction, were required to find optimal band combinations, to assess the effect of the number of radar looks and the polarimetric description models. Confusion matrices were used to analyse the result and accuracy for every individual land cover class. Results showed that the combination of three fully polarimetric bands gave optimal results, and the results obtained from 0.5 dB data (64-look data) were always better than from 1.0 dB data (20-look data). Simulation results using a new reversible transform method (7I model) gave the best results.

Analysis of 3-D textural features derived from C-band InSAR data and 2-D textural attributes derived from C-band total power (TP) data showed that the classification capacity of Standard Deviation and GLCO-COR textural features are less pronounced compared to the GLCO-CONT feature which was more sensitive to canopy architecture, while qualitative analysis using scatter plots showed that textural features based on GLCO-CONT[d1-w21] are able to differentiate forest types and plantation. Scatter plots of TP and DEM textural features could not clearly differentiate all the forest classes studied.

Post-processing using neighbouring pixel information (ICM), DEM information and 3-D textural information as additional information, resulted in a considerable increase in the overall classification accuracy. Classification results showed that the C-, L- and P-band combination with $\beta 1 = 10$ and 50 ICM iterations without *a priori* knowledge showed optimal results. The use of *a priori* knowledge (i.e. texture and relief correction) did not necessarily provide better results. When using dual-band combinations the C- and L-band combination showed that burnt primary forest was difficult to separate from secondary forest; the C- and P-band combination could sufficiently separate primary and secondary forest, while for the L- and P-band combination the classification result for the primary forest was slightly inferior compared to the two bands combinations mentioned previously.

Single-band data were not sufficient to give good classification results; it appears that L-band or P-band alone can not be used for classifying the images. However, combined with other information (e.g. C- or L-band fully polarimetric data) the results could be improved considerably.

Biomass estimation using empirical relationships with backscatter intensity was studied for secondary forest, primary forest and burnt primary forest classes. By including as well as excluding dead standing trees, two cases were considered. It was shown that the coefficient of correlation was smaller for biomass excluding dead standing trees compared to biomass including dead standing trees. Correlation before and after slope (relief) correction were slightly different. A small increase took place after slope correction. P-HV showed maximum correlation results. The L- and P-band backscatter intensity increased with the increase of biomass until a certain saturation limit was reached.

A model describing the multi-frequency complex coherence was used to achieve a better physical interpretation of the radar backscatter signal characteristics. Different forest types showed characteristic multi-frequency complex coherence signatures which could be linked to different structural forest types. A clear example was the single and double-bounce mechanism. Single-bounce (direct trunk scattering) occurred at the trunks in ‘open’ forest, whereas double-bounce only occurred in open flooded mangrove areas. The existence of dead standing trunks caused significant backscatter in L- and P-band. Hence these are an important factor to be considered in the biomass estimation.

Comparison and validation of land cover type classification with the independent data set

In Chapter 5 the classification results have been evaluated by the use of an independent data set acquired by another researcher. The map resulting from the independent data set yields significantly higher classification results. The use of a structural or 'radar' legend increased the accuracy. This 'radar' legend is based on radar's ability to capture forest information, such as the structure of forest/trees, and the sub-division of secondary forest into old secondary forest, secondary forest with young trunks, etc.

Maps with 'common 'legends based on the original data set and maps with the 'radar' legend based on the independent data set show acceptable consistency levels. Water, rubber plantation and mangrove have high accuracy and agreement levels. Primary forest and burnt primary forest also have good levels of accuracy and agreement. The secondary forest has a moderate result. Classes with poor results are mixed, shrubs, *alang-alang* and bare soils.

Qualitative evaluation by visual comparison shows, generally, that the classification results show a good agreement in size and location for the classes of water, mangrove, rubber plantation, secondary forest, primary forest and burnt primary forest. For the classes *alang-alang*, mixed, shrubs and bare soils the agreement is poor.

Analysis of factors affecting classification processes has been made. The overall results depend on many factors such as polarisation or polarisation combinations, single or multi-band and radiometric resolution (or number of looks). *The effect of relief correction on the full polarimetric classification can be small.* Of course, in case the full polarimetric information is not used, *then the effect of relief correction can be large*, such as visually in the intensity image, or in the relationship between biomass and backscatter intensity, or in the classification of single polarisation data. Similarly, *a fully polarimetric combination of C- and L-band or C-, L- and P-band will provide enough information, so that additional information of texture will not be required.* Post-processing through ICM is very useful and was used to produce radar derived classifications. A good radar based legend is one of main key factors to achieve a good classification result.

Key Conclusions and recommendations

A new approach for tropical forest type and biomass mapping using interferometric and multi-band polarimetric SAR has been introduced and evaluated. The PacRim-2 experiment has provided the first experiences with this kind of new technology and methodology in Indonesia. The approach presented in this thesis includes *new elements such as (1) slope correction using InSAR, (2) mapping using a new reversible transform technique and (3) ICM using prior knowledge* (e.g. terrain height or image texture) including spatial aggregation. The new *reversible transform* of the covariance matrix was used to describe the full polarimetric target properties in an alternative way, allowing simpler statistical descriptions. It has been shown that this transform yields versatile and robust classification approaches. Comparisons of results for the various classification methods have been given, using several combinations of frequency bands. Classification results could be simulated for certain combinations of

frequency bands and polarisation as a function of speckle level. In addition biomass estimations before and after relief correction by using a digital elevation model (DEM) derived from InSAR data have been compared.
The problem caused by the presence of topographic slopes in the interpretation of the AirSAR images has been solved. Backscatter quality was improved by utilising the information on the topography. The radar cross section modulation by slopes could be compensated for by radiometric slope correction, using slopes derived from across track interferometry or from a DEM.

The use of the new *reversible transform* technique for simulation studies was proven to be useful; most classes could be recognized and distinguished. Such results cannot be related easily to image classification results because of the presence of texture and relief. Commonly used image processing techniques to mitigate or to capitalise on these effects appear to be of limited use for several reasons, which were indicated in Chapter 4. The transform to an 'intensities-only' system is very supportive as it allowed for the development of the 7I system, which outperforms slightly the system 3I+, as introduced in Hoekman and Quiñones (2000). Though the improvement is sometimes not more than a few percent, this can be relatively large in case of high values of accuracy.

In general, the difference of forest types in tropical regions are by far the most difficult to identify. This is partly due to the fact that tropical rain forests do not generally consist of a collection of homogeneous, well-defined forest types with distinct boundaries. In fact, these forests are characterised by a high variability in species composition and architecture, with the presence of transitional forest types as well as transitions between forest types. Image 3-D texture derived from a DEM might offer a possibility for distinguishing natural forest and plantation forest. However, the appraisals in Table 4.6 and previous discussion lead to the general conclusion that textural patterns in high resolution of C-band DEM and total power backscatter radar images are only of limited *additional* value in the case multi-band fully polarimetric data are available.

A technique based on iterated conditional modes (ICM) appeared to be useful. The classification result reached 88.9%. An independent data set with 17 classes yielded even a much higher accuracy of 93.8%. Validation of the results and its reverse process provide sufficient consistency. Natural forest types (primary forest, burnt primary forest and secondary forest) were successfully discriminated, as well as forest plantation types (oil palm and rubber). *Alang-alang*, mixed and shrub cannot be easily discriminated because they have similar vegetation and grow quickly everywhere. The vigorous growth may also have caused discrepancies between conditions at the time of field observation and the time of radar observation. By visual interpretation, the area which appears to be the most misclassified is the primary forest located at the northern side of Balikpapan Bay (eastern side of the Sungai Wain protected forest area). According to the classification map it is classified as oil palm while it is actually a burnt primary forest area, as shows up correctly in the 'radar' legend map. The misclassification may occur because the primary forest at that location may have a different vegetation structure than the primary forest in the middle of the protected forest area. This may be due to a low degree of fire damage, consequently yielding radar signal characteristics more similar to oil palm or burnt primary forest in the ground reference sets.

The use of direct empirical relationships for biomass estimation in tropical forests based on C-band backscatter shows poor results. For L- and P-band, moderate results are achieved. The HH and HV polarisations do not always provide better results than VV polarisation. The results are difficult to interpret because of the complexity of vegetation, frequent forest fires, relief modulation on the backscatter level and the effect of radar saturation at a certain biomass level. The possibility of using an indirect approach, in which forest type classification is a first step, as proposed in Quiñones (2002), was considered. In this case, a biomass map will show a number of physical vegetation structure classes; and biomass levels can be associated to each cover class. Thus an estimation of biomass levels beyond the saturation level may be obtained.

An analysis of physical interpretation of signal characteristics data appeared to be useful to increase the understanding of specific effects of forest structure. These results clearly show that the different sets of multi-frequency complex coherence numbers can be recognised and linked to the different structural types, based on the physical interaction with the waves. It was shown that forests with similar structures generate similar radar signatures. Changes in intensity are expected to occur as described in the literature, i.e. higher backscatter values where more scatterers of a certain size (in relation to wavelength) occur. To understand the physical interaction between the forest and radar waves, the scattering mechanisms can be simplified into a model composed of three contributions (Ulaby *et al*., 1986). These are: (1) the direct backscattering from the vegetation layer, (2) the direct backscattering from the ground attenuated by the vegetation cover and (3) the backscattering originating from ground-trunk interaction attenuated by the vegetation cover. These are sometimes referred to as the diffuse (volume) term, the single-bounce (odd) term and the double-bounce (even) term, respectively.

The main conclusions could be summarised as follows. *Isolated dead standing trees give a significant contribution in L- and P-band through direct (single-bounce) scattering from the trunk,* and not through the double-bounce trunk-ground interaction mechanism. *Double bounce scattering is rare*. Some indications were found it occurs in flooded parts of relatively open canopy mangrove areas in P-band. In general, for all forests canopies, single bounce dominates in C-band (high coherence) and diffuse scattering dominates in L- and P-band. These finding can be used as a model for a direct (unsupervised) physical approach to classification in less known areas, and may contribute to classification potential on inaccessible areas with very little additional field information.

The results presented in this thesis give insight into the difficulty of optimising the utility of interferometry and C-, L- and P-band polarimetry in tropical rain forest areas, and the accuracy that can be obtained. In the near future some capabilities that are currently exclusive to airborne SAR may be available from spaceborne SAR and could measure global biomass to increase the accuracy of climate change models. Japan's Advanced Land Observing Satellite (ALOS) PALSAR and possibly the TerraSAR-X or the combination of advanced C-band systems such as RADARSAT-2, ENVISAT with Shuttle Radar Topography Mission (SRTM) data may appear very useful for land cover type mapping as well as biomass estimation. In addition the combination of two techniques: interferometry and polarimetry or polarimetric interferometry (Pol-InSAR) can be used to assess forest height and three-dimensional forest structural views.

Samenvatting en conclusies

Probleem omschrijving

De ontbossingsnelheid in Indonesië heeft de neiging sterk toe te nemen. De Indonesische regering, middels het Ministerie van Bosbeheer, heeft besloten dit proces te verminderen of te stoppen, en om maatregelen voor duurzaam bosbeheer te ondersteunen. De Indonesische bossen hebben een grote biodiversiteit en zijn een belangrijke factor in hydrologische en biochemische cycli, die onontbeerlijk zijn om de natuur in evenwicht te houden.

Om de conditie van deze bossen te monitoren en te evalueren heeft de regering informatie nodig die nauwkeurig en up-to-date is. Deze informatie is op het moment gebaseerd op fotografische beelden verkregen door optische technologie zoals luchtfoto's, LANDSAT of SPOT. Deze technieken geven vaak problemen als gevolg van bewolking. Daarom heeft de technologie die gebaseerd is op microgolven (RADAR) de potentie om dit probleem op te lossen. De vooruitgang van moderne radartechnieken zoals interferometrie en polarimetrie gaat snel en verwacht wordt dat deze vooruitgang de beste oplossing biedt om nauwkeurige en up-to-date informatie te geven op het gebied van verandering van land cover, hoeveelheid opening in het kronendak, biomassa, enz..

Belangrijke obstakels voor het verstrekken van continue en up-to-date informatie met betrekking tot de bossen zijn momenteel:

- De bossen in Indonesië zijn bijna voortdurend bedekt met wolken, mist of rook, wat het moeilijk maakt om met optische technologie duidelijke beelden te verkrijgen.
- Er is onvoldoende ervaring in het gebruik van alternatieve technologieën zoals radar.
- Geschikte toepassingen/software om de kracht van deze radar technologie te analyseren is nog niet algemeen beschikbaar.

Hoewel de toepassing van radar technologie in Indonesische bossen sinds 1991 wordt beoordeelt/onderzocht zijn de resultaten nog niet bevredigend. De laatste tien jaar hebben wetenschappers/onderzoekers een aanzienlijke hoeveelheid tijd en moeite gestopt in het ontwikkelen van algoritmes om de vegetatie structuur te classificeren en de hoeveelheid biomassa te schatten. In het geval van de Indonesische bossen zijn deze algoritmes niet getest of gebruikt. Dit onderzoek heeft als doel, deze situatie in ogenschouw nemend, om *het gebruik van de nieuwe generatie C- en L-band interferometrisch (TopSAR) en C-, L- en P-band volledig polarimetrisch (PolSAR) NASA/JPL AirSAR data voor het in kaart brengen van tropische bostypes en biomassa schattingen te evalueren, en het gecombineerde gebruik van radar data en toegevoegde kennis om de resultaten te verbeteren te bestuderen.*

Om de waarde van de informatie in de data van het verfijnde AirSAR radar systeem te beoordelen werden in deze studie beschikbare algoritmes gebruikt, geanalyseerd en aangepast om de toepassing ervan te optimaliseren. Deze studie combineerde een DEM, afgeleid van InSAR data, en volledig polarimetrische data, met de volgende parameters: ruimtelijke variabiliteit, intensiteit radar backscatter, polarisatie en fase.

Deze parameters werden gecorrigeerd voor de hellingseffecten op licht. Van te voren vastgestelde regio's, waarvan de situatie in het gecorrigeerde beeld in het veld bekend was, werden gebruikt om een database te construeren (in deze thesis genoemd '*the original data set'*). Een nieuwe methode gebaseerd op een *reversible transform* van de polarimetrische radar covariantie matrix werd toegepast om de volledige polarimetrische eigenschappen te beschrijven. 3-D Ruimtelijke variaties van InSAR data werden ontvangen als beeldtextuur, welke gerelateerd kon worden aan de ruigheid van het kronendak van het bos. Vervolgens werden naverwerkingsstappen uitgevoerd; gebruikmakend van informatie van naburige pixels (ICM) en met additionele informatie van de textuur kon de nauwkeurigheid van de classificatieresultaten verbeterd worden. Maximum likelihood en Kappa statistiek werden gebruikt om de overall classificatie resultaten te evalueren. Deze techniek en de theoretische analyse zijn uitgewerkt in Hoofdstuk 2.

Studiegebied

Het veldwerk voor deze studie werd uitgevoerd in een tropisch bosgebied vlakbij Sungai Wain, wat niet ver is van Balikpapan City ($0^0$83'Z, $116^0$76'O), in de provincie Oost-Kalimantan in Indonesië. Het gebied wat door de radar bestreken werd was 10x60 km^2, en werd gekarakteriseerd door een complex mozaïek van vegetaties en land cover types, van tropisch altijdgroen laagland en semi-altijdgroen bos gedomineerd door *Dipterocarpaceae*, een verscheidenheid aan *primair bos met sommige bomen uitschietend boven de 30 m*, en verbrand primair bos. Het secundaire bos is voortgekomen uit ontbossing, overblijfselen van branden en natuurlijke regeneratie, maar ook door vegetatieverandering door lokale agrarische activiteiten. A*lang-alang* beslaat grote gebieden door het gehele studiegebied. Het mangrove bos beslaat het zuidelijke gebied wat vlakbij de Baai van Balikpapan ligt, de rest bestaat uit plantages, rijstvelden, overgangsgebieden, garnalen vijvers en wildernis.

In 2000 werd het studiegebied binnen het kader van de PacRim-2 campagne in beeld gebracht door de C-, L- en P-band PolSAR data van de AirSAR/TopSAR NASA/JPL. Voor het ondersteunen van de studieobjecten werden velddata verzameld door zowel extensieve als intensieve observaties. Een tweede '*onafhankelijke data set'* was voorhanden om de resultaten van deze studie te valideren. Een gedetailleerde omschrijving van de studiegebieden en de aanwezige radar data kan gevonden worden in Hoofdstuk 3.

Biofysische parameter inwinning en landbedekkings classificatie

Hoofdstuk 4 behandelt de onderzoeks resultaten met nadruk op de evaluatie van de radar karakteristieken bij het vastleggen van de parameters van het tropische bos, land cover type classificatie en biomassa schatting.

Analyses van de hellingcorrecties toegepast op de multi-band composiet van C-, L- en P-band PolSAR data met gebruik van de InSAR DEM laat zien dat de effecten van terreinhelling en golvingen in het kronendak op de invallende straling gereduceerd worden. Geproduceerde beelden bleken over het algemeen meer uniform en homogeen te zijn en maakten visueel onderscheid tussen landbedekkingen duidelijker.

Land cover type classificatie simulaties, voor en na hellingcorrectie, waren noodzakelijk om een optimale combinatie van banden te vinden, om het effect van de hoeveelheid *radar looks* en de beschrijvende polarimetrische modellen te onderzoeken. Verstrooiing matrices werden gebruikt om het resultaat en de nauwkeurigheid van elke individuele land cover klasse te analyseren. De resultaten lieten zien dat de combinatie van 3 volledig polarimetrische banden een optimaal resultaat gaf, en de resultaten verkregen van 0.5 dB data (64-look data) waren altijd beter dan van 1.0 data (20-look data). Simulatieresultaten, gebruikmakend van een nieuwe *reversible transform* methode (7I model) gaf de beste resultaten.

Analyse van 3-D textuur eigenschappen afgeleid uit de C-band InSAR data en 2-D textuur attributen afgeleid uit de C-band *total power* (TP) data lieten zien dat de classificatie capaciteit van Standaard Deviatie en GLCO-COR texturele karakteristieken minder uitgesproken zijn in vergelijking met de GLCO-CONT karakteristieken die gevoeliger was voor kronendak architectuur, terwijl een kwalitatieve analyse met behulp van scatter plots liet zien dat textuurkarakteristieken gebaseerd op GLCO-CONT(d1-w21) in staat zijn om onderscheid te maken in bostypes en plantage. Scatter plots van TP en DEM textuur eigenschappen konden geen duidelijk onderscheid maken in alle bestudeerde bostypes.

Naverwerking met behulp van informatie van naburige pixels (ICM), DEM informatie en 3-D textuur informatie als toegevoegde informatie, resulteerde in een aanzienlijke toename van de overall classificatie nauwkeurigheid. Classificatie resultaten lieten zien dat de C-, L- en P-band combinatie met $\beta 1 = 10$ en 50 ICM herhalingen zonder *a priori* kennis (bijv. textuur en hellingcorrectie) niet noodzakelijk betere resultaten opleverden. Wanneer een duo-band combinatie werd gebruikt, lieten de C- en L-band combinatie zien dat verbrand primair bos moeilijk te onderscheiden was van secundair bos; de C- en P-band combinatie kon voldoende onderscheid maken tussen primair en secundair bos, terwijl bij de L- en P-band combinatie het classificatie resultaat voor primair bos iets minder was vergeleken met de twee eerder genoemde band combinaties.

Data van een enkele band was niet voldoende om goede classificatie resultaten te verkrijgen; het lijkt erop dat de L-band of P-band alleen niet gebruikt kan worden om de beelden te classificeren. Hoewel een combinatie met andere informatie (bijv. C- of L-band volledige polarimetrische data) de resultaten aanzienlijk kon verbeteren.

Biomassaschatting met gebruik van empirische relaties met backscatter intensiteit is bestudeerd voor secundair bos, primair bos en verbrande primaire bostypes. Door het wel of niet uitsluiten van dode staande bomen werden twee gevallen in acht genomen. Aangetoond is dat de correlatiecoëfficiënt kleiner was voor biomassa zonder dode staande bomen in vergelijking met biomassa met dode staande bomen. Correlatie voor en na hellingcorrectie verschilden erg weinig. Een kleine toename vond plaats na hellingcorrectie. P-HV liet maximale correlatie resultaten zien. De L- en P-band backscatter intensiteit nam toe met een toename van biomassa tot een zekere verzadigingslimiet werd bereikt.

Een model welke de multi-frequente complexe coherentie beschreef werd gebruikt om een betere fysische interpretatie van de karakteristieken van de signalen van de radar backscatter te krijgen. Verschillende bostypes lieten een karakteristieke signatuur van de multi-frequente complexe coherentie zien die in verband konden worden gebracht

met de verschillende structurele bostypes. Een duidelijk voorbeeld was het enkele en dubbele reflectie mechanisme. Enkele reflectie (directe reflectie van de stam) vond plaats op stammen in 'open' bos, terwijl dubbele reflectie alleen voorkwam in open ondergelopen mangrove gebieden. De aanwezigheid van dode staande stammen zorgde voor een significante backscatter in de L- en P-band. Vandaar dat dit een belangrijke factor is die meegenomen moet worden bij de schatting van de biomassa.

Vergelijking en validatie van land cover type classificatie met de onafhankelijke dataset

In Hoofdstuk 5 zijn de classificatieresultaten geëvalueerd met behulp van een, door een andere onderzoeker verzamelde, onafhankelijke dataset. De onafhankelijke dataset levert een kaart met significant hogere classificatie resultaten op. Het gebruik van een structurele of 'radar' legenda deed de nauwkeurigheid toenemen. Deze 'radar' legenda is gebaseerd op de mogelijkheid van radar om informatie over bossen vast te leggen zoals de structuur van het bos/de bomen, het onderverdelen van secundair bos in oud secundair bos en secundair bos met jonge stammen, enz..

Kaarten met 'normale' legenda's gebaseerd op de originele dataset en kaarten met de 'radar' legenda's gebaseerd op de onafhankelijke dataset laten acceptabele niveau's van eenduidigheid zien. Water, rubber plantages en mangrove hebben een hoge nauwkeurigheid. Primair bos en verbrand primair bos hebben ook een goede nauwkeurigheid. Het secundaire bos heeft een matige nauwkeurigheid. Klassen met slechte resultaten bestaan uit een gemengde vegetatie, struikgewas, *alang-alang* en kale bodems.

Een kwalitatieve evaluatie door een visuele vergelijking laat zien, dat over het algemeen, de classificatieresultaten een goede overeenstemming in grote en plaats hebben voor de klassen water, mangrove, rubberplantage, secundair bos, primair bos en verbrand primair bos. Bij de klassen *alang-alang*, gemengde vegetatie, struikgewas en kale bodems is deze overeenstemming slecht.

Er is een analyse van de factoren die classificatieprocessen beïnvloeden gemaakt. De overall resultaten hangen van veel factoren af zoals de polarisatie of polarisatie combinaties, enkele of meerdere banden, radiometrische resolutie (aantal looks). Het effect van hellingcorrectie op de volledig polarimetrische classificatie kan klein zijn. Natuurlijk, wanneer de volledig polarimetrische informatie niet gebruikt wordt, kan het effect van helling correctie groot zijn, zoals visueel in het intensiteitsbeeld, in de relatie tussen biomassa en backscatter intensiteit, of in de classificatie van enkel gepolariseerde data. Op dezelfde wijze bevat een volledig polarimetrische combinatie van C- en L-band of C-, L- en P-band genoeg informatie, zodat additionele informatie over de textuur niet noodzakelijk is. Nabewerking door middel van ICM is erg behulpzaam en werd gebruikt om classificaties afgeleid van radar te maken. Een goede, op radar gebaseerde legenda, is een van de belangrijkste factoren om een goede classificatie te krijgen.

Belangrijkste conclusies en aanbevelingen

Een nieuwe benadering van tropische bostypes en het in kaart brengen van biomassa met gebruik van interferometrische en multi-band polarimetrische SAR werd geïntroduceerd en geëvalueerd. Het PacRim-2 experiment heeft de eerste ervaringen met deze technologie en methodologie in Indonesië mogelijk gemaakt. De benadering die in deze thesis gepresenteerd wordt bevat *nieuwe elementen zoals (1) hellingcorrectie met behulp van InSAR, (2) cartograferen met behulp van een nieuwe reversible transform techniek en (3) ICM met gebruik van a priori kennis* (bijv. terreinhoogte of beeldtextuur) waaronder ruimtelijke aggregatie. De nieuwe *reversible transform* van de covariantie matrix werd gebruikt om de volledig polarimetrische doeleigenschappen op een alternatieve manier te beschrijven waardoor simpelere statistische beschrijvingen mogelijk werden. Getoond werd dat deze transformatie een veelzijdige en robuuste manier van classificeren is. Vergelijkingen van resultaten van de verschillende classificatie methoden, met verschillende combinaties van frequentiebanden, zijn gepresenteerd. Classificatieresultaten voor sommige combinaties van frequentiebanden en polarisatie konden gesimuleerd worden als functie van ruisniveau. Bovendien werden ook de biomassa schattingen voor en na hellingcorrectie met behulp van een digitaal hoogtemodel (DEM) afgeleid van InSAR data, vergeleken.

Problemen, veroorzaakt door de aanwezigheid van hellingen, bij de interpretatie van de AirSAR beelden zijn opgelost. Backscatter kwaliteit werd verbeterd met behulp van informatie over de topografie. De radar cross section modulatie als gevolg van reliëf kan gecompenseerd worden door de radiometrische hellingcorrectie met behulp van reliëf afgeleid van across track interferometrie of van een DEM.

Het gebruik van de nieuwe *reversible transform* techniek voor simulatie studies is nuttig gebleken; de meeste klassen konden herkend en onderscheiden worden. Zulke resultaten kunnen niet eenvoudig gerelateerd worden aan beeldclassificatie resultaten door de aanwezigheid van textuur en reliëf. Algemeen gebruikte beeldverwerkingstechnieken hebben een beperkt nut om deze effecten te matigen of geheel het hoofd te bieden om verschillende redenen die aangegeven zijn in hoofdstuk 4. De omschakeling naar een systeem met uitsluitend ‘intensiteit’ verdiend ondersteuning omdat het de ontwikkeling van het 7I systeem, welke net beter was dan het 3I+ systeem, zoals geïntroduceerd door Hoekman en Quiñones (2000) mogelijk maakte. Hoewel de verbetering soms niet meer dan een paar procent is kan dit relatief veel zijn in geval bij hoge nauwkeurigheden.

Over het algemeen is het verschil in bostypes in tropische regio’s verreweg het moeilijkst te identificeren. Dit ligt deels aan het feit dat tropische regenwouden over het algemeen niet bestaan uit een samenstelling van homogene, goed te definiëren bostypes met duidelijke grenzen. In feite worden deze bossen gekarakteriseerd door een grote variabiliteit in soortensamenstelling en architectuur, met de aanwezigheid van zowel overgangs bostypes als overgangen tussen bostypes. 3-D beeldtextuur afgeleid van een DEM geeft mogelijk de mogelijkheid om onderscheid te maken tussen natuurlijk bos en plantage bos. Hoewel, de schattingen in tabel 4.6 en voorgaande discussie leiden tot de algemene conclusie dat texturele patronen in de hoge resolutie C-band DEM en de total power backscatter radarbeelden slechts van *aanvullende* waarde zijn wanneer de multi-band volledig polarimetrische data beschikbaar zijn.

Een op iterated conditional modes (ICM) gebaseerde techniek bleek bruikbaar te zijn. Het classificatieresultaat bereikte een waarde van 88.9%. Een onafhankelijke dataset met 17 klassen leverde zelfs een veel hogere nauwkeurigheid van 93.8% op. Controle van deze resultaten en het omgekeerde proces hiervan leverde voldoende eenduidigheid op. Natuurlijke bostypes (primair bos, verbrand primair bos en secundair bos) werden succesvol onderscheiden, net als bosplantage types zoals oliepalmen en rubberplantages. *Alang-alang*, gemengde vegetaties en struikgewas kunnen niet gemakkelijk onderscheiden worden omdat ze een vergelijkbare vegetatie hebben en overal snel groeien. Deze snelle groei kan ook verschillen tussen het tijdstip van veldobservatie en radar observatie veroorzaakt hebben. Bij een visuele interpretatie, lijkt het gebied met de meeste misclassificaties het primaire bos aan de noordkant van de Baai van Balikpapan (oostzijde van het Sungai Wain beschermde bosgebied) te zijn. In overeenstemming met de classificatiekaart is dit gebied geclassificeerd als olie palm terwijl het eigenlijk verbrand primair bosgebied is, zoals correct wordt weergegeven in de kaart met de 'radar' legenda. De misclassificatie kan het gevolg zijn van een verschillende vegetatiestructuur van het primaire bos in die locatie en primair bos in het midden van het beschermde bosgebied. Dit kan veroorzaakt zijn door een kleine hoeveelheid schade door brand, wat consequent leidt tot karakteristieken van de radarsignalen die meer overeenstemming vertonen met oliepalmen of verbrand primair bos bij de referentie data.

Het gebruik van directe empirische relaties gebaseerd op C-band backscatter voor het schatten van biomassa in tropische bossen laat slechte resultaten zien. Bij de L- en P-band worden matige resultaten bereikt. De HH en HV polarisaties leveren niet altijd betere resultaten op dan VV polarisatie. De resultaten zijn moeilijk te interpreteren vanwege de complexiteit van de vegetatie, frequente bosbranden, reliëf modulatie op het backscatter niveau en het effect van radar verzadiging bij een zeker niveau van biomassa. De mogelijkheid van het gebruik van een indirecte benadering, waarbij bostypes classificatie een eerste stap is, zoals voorgesteld door Quiñones (2002), is overwogen. In dit geval zal het een biomassa kaart met een aantal fysische vegetatiestructuur klassen opleveren waarbij biomassa niveau's gerelateerd kunnen worden aan elke cover klasse. Hierdoor kan een schatting van biomassa niveau's voorbij het verzadigingsniveau gemaakt worden.

Een analyse van de fysische interpretatie van data van signaalkarakteristieken bleek nuttig te zijn om specifieke effecten van bosstructuur beter te kunnen begrijpen. Deze resultaten lieten duidelijk zien dat verschillende sets van complexe multi-frequentie coherentie waarden konden worden herkend en gelinked konden worden aan verschillende structuurtypes, gebaseerd op de fysische interactie bij deze golflengtes. Bossen met vergelijkbare structuren leverden vergelijkbare radar signaturen op. Veranderingen in intensiteit worden verwacht voor te komen zoals in de literatuur beschreven, bijv. hogere backscatter waardes wanneer er meer scatterers van een zekere grootte (in relatie met golflengte) zijn. Om de fysische interactie tussen bos en radar golven te begrijpen kunnen de scatter mechanismes versimpeld worden in een model met 3 factoren (Ulaby *et al*., 1986). Deze zijn: (1) de directe backscattering van de vegetatie laag, (2) de directe backscattering van de grond verminderd door de vegetatie bedekking en (3) de backscattering van grond-stam interactie verminderd door de vegetatie bedekking. Deze worden ook wel respectievelijk de diffuse (volume) term, de single-bounce (odd) term en de double-bounce (term) genoemd.

De hoofdconclusies kunnen als volgt samengevat worden. *Geïsoleerde dode staande bomen geven een significante bijdrage in de L- en P-band door middel van directe (single-bounce) scattering van de stam*, en niet door het double-bounce stam-grond interactie mechanisme. *Double bounce scattering is zeldzaam*. Er zijn een paar aanwijzingen gevonden dat het voorkomt in de P-band in overstroomde delen van mangrove gebieden met een relatief open kronendak. Over het algemeen, bij alle kronendaken in bossen, domineert single bounce in C-band (hoge coherentie) en diffuse scattering domineert in de L- en P-band. Deze bevindingen kunnen gebruikt worden als een model voor directe (unsupervised) fysische benadering voor het classificeren van minder onderzochte gebieden, en kan het classificatie potentieel van ontoegangbare gebieden met erg weinig informatie uit het veld verhogen.

De resultaten die in deze thesis gegeven worden geven een inzicht in de moeilijkheden voor het optimaliseren van het gebruik van interferometrie en C-, L- en P-band polarimetrie in tropische regenwouden en de nauwkeurigheid die bereikt kan worden. In de nabije toekomst kunnen sommige mogelijkheden die nu uitsluitend met SAR vanuit de lucht mogelijk zijn ook beschikbaar komen met SAR vanuit de ruimte, hiermee kan de globale biomassa gemeten worden om de nauwkeurigheid van klimaatveranderingsmodellen te vergroten. Japan's Advanced Land Observing Satellite (ALOS) PALSAR en mogelijk de TerraSAR-X of de combinatie van geavanceerde C-band systemen zoals RADARSAT-2, ENVISAT met Shuttle Radar Topography Mission (STRM) data zijn mogelijk erg nuttig bij het in kaart brengen van land cover types en het schatten van de biomassa. Bovendien kan de combinatie van 2 technieken: interferometrie en polarimetrie of polarimetrische interferometrie (Pol-InSAR) gebruikt worden om boshoogte en driedimensionale bosstructuren te schatten.

References

AIRSAR NASA/JPL PACRIM-2 campaign web site, 2000, (http://airsar.jpl.nasa.gov).

Benson, A.S. and SD DeGloria, 1985. Interpretation of Landsat-4 Thematic Mapper and Multispectral Scanner data for forest surveys. *Photogrammetric Engineering and Remote Sensing*, vol.51, no.9, pp.1281-1289.

Besag, J., 1986, On the statistical analysis of dirty pictures, *Journal Royal Statistical Society B*, Vol.48, No.3, pp.259-302.

Bijker, W., 1997, Radar for rain forest. A monitoring system for land cover change in the Colombian Amazon. *Ph.D. Thesis* Wageningen Agricultural University, The Netherlands, 192 p.

Bremen, H. van, M. Iriansyah and W. Andriesse, 1990. *Detailed Soil Survey and Physical Land Evaluation in a Tropical Rain Forest, Indonesia. A Study of Soil and Site Characteristics in Twelve Permanent Plots in East Kalimantan*, The Tropenbos Foundation, Ede, The Netherlands, 188 p.

Brown, S., A.J.R. Gillespie, and A.E. Lugo, 1989, Biomass estimation methods for tropical forests with applications to forest inventory data, *Forest Science*, Vol.35, pp.881-992.

Castel, A. Beaudoin, N. Stach, N. Stussi, T. Le Toan and P. Durand, 2001. Sensitivity of space-borne SAR data to forest parameters over sloping terrain. Theory and experiment. *International journal of remote sensing*, 2001, vol. 22, no. 12, 2351-2376.

Congalton, R. G., 1999. *Assessing the accuracy of remotely sensed data: principles and practices*. Lewis Publishers, 137 p.

Dobson, M.C., F.T. Ulaby, T. Le Toan, A. Beaudoin, E.S. Kasischke, and N. Christensen, 1992. Dependence of Radar Backscatter on Coniferous Forest Biomass, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.30(2):421-415.

Faller, N.P and E.H. Meier, 1995. First result with the airborne single-pass DO-SAR interferometer, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.33, No.5, pp1230-1237.

Faller, N.P., 1998, *Operational topographic mapping in Indonesia with interferometric DO-SAR*, Unpublished.

Ferrazzoli, P., L. Guerriero, and G. Schiavon, 1999, Experimental and model investigation on radar classification capability, *IEEE transactions on Geoscience and remote sensing*, Vol.37. No.2, pp.960-968.

Gastellu-Etchegorry, J.P., 1988, Cloud cover distribution in Indonesia, *International Journal of Remote Sensing*, Vol.9:1267-1276.

Grim, R.J.A, F.M. Seifert, D.H. Hoekman, C. Varekamp, Y.A. Hussin, M.A. Sharifi, M. Weir, Kemal Unggul Prakoso, Ruandha Agung Sugardiman, Muljanto Nugroho, Bambang Suryokusumo, 2000, SIRAMHUTAN Sistem Informasi untuk Manajemen HUTAN, A demonstration project for forestry in Indonesia, *BCRS Publication NRSP-2 Report 00-07*, The Netherlands, 83p.

Hidayat, H. 1993, *Forest inventory and monitoring over production forest areas in Indonesia*, GIS World Symposium, Vancouver, British Columbia, Canada.

Hidayat, H., 1997. Operationalisation using radar in Indonesia. *Proc. Seminar on "Radar technology for sustainable forest management"*, 3 April 1997, Departemen Kehutanan-PT.Mapindo Parama-Tropenbos. Jakarta, Indonesia, pp48-52.

Hidayat, H., 1998. Inventorying and Monitoring forestry in Indonesia. http://www.spot.com/home/appli/FOREST/indofor/welcome.htm.

Hoekman, D.H., 1990. Radar Remote Sensing Data for Applications in Forestry, *Ph.D. Thesis*, Wageningen Agricultural University, The Netherlands, 277 p.

Hoekman, D.H., 1991. Speckle ensemble statistics of logarithmically scaled data. *IEEE Transactions on Geoscience and Remote sensing*, vol.29, no.1, pp.180-182.

Hoekman, D.H, K.U. Prakoso, B. Suryokusumo, M. Nugroho, R.A. Sugardiman., 2000*, Experiment plan for the PacRim 2 campaign in East-Kalimantan, Indonesia*, Version 1, August 2000, Department of Environmental Sciences, Wageningen University.

Hoekman, D.H, K.U. Prakoso, B. Suryokusumo and M.A.M. Vissers., 2001*, PacRim 2 database Sungai Wain*, ESTEC Contract No. 15193/01/NL/DC "PacRim-2 Ground Data Collection in Kalimantan", ESTEC, The Netherlands.

Hoekman, D.H. and C. Varekamp, 2001. Observation of tropical rain forest trees by airborne high resolution interferometric radar, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.39, No.3, pp.584-594.

Hoekman, D.H. and M.A.M. Vissers, 1995, *Manual ML-class*. Wageningen University, unpublished.

Hoekman, D.H., M.A.M. Vissers, R.A. Sugardiman, and J. Vargas, 2001, ENVISAT forest monitoring Indonesia. NRSP-2 project 3.3/DE-08, *NUSP-2 report 02-01*, Wageningen, The Netherlands, 72 pages.

Hoekman, D.H., and M.A.M. Vissers, 2003. A new polarimetric classification approach evaluated for agricultural crops, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.41, No.12, December 2003, pp.2881-2889.

Hoekman, D.H. and M.J. Quiñones, 2000. Land cover type and biomass classification using AirSAR data for evaluation of monitoring scenarios in the Colombian Amazon, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.38, pp.685-696.

Hoekman, D.H. and M.J. Quiñones 2002. Biophysical Forest Type Characterization in the Colombian Amazon by Airborne Polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*. Vol.40(6):1288–1300.

Hoekman, D.H., J.J. van der Sanden and W. Bijker, 1994. Radar remote sensing of tropical rain forests: The SAREX-92 campaign in Guyana and Colombia. First interim report. Delft (BCRS, Netherlands Remote Sensing Board), *BCRS report,* no.94-01, 34 p. + 4 appendices.

Hoekman, D.H., J.J. van der Sanden and W. Bijker, 1996. Radar remote sensing of tropical rain forests: The AIRSAR-93 campaign in Guyana and Colombia. Delft, (Netherlands Remote Sensing Board) *BCRS report* no.96-02, 59 p.

Hussin, Y.A, R.M. Reich, and R.M Hoffer, 1992. Effect of polarisation on radar backscatter in relation to slash pine stand biomass using aircraft and SIR-B data, *Int. Soc. for Photogrammetry and Remote Sensing*, XVII Congress, Vol. XXIX, Commission VII (B7), pp.661-667.

Imel, D.A. 2002, What's wrong with my AirSAR data?, *AirSAR earth science and applications workshop*, 4-6 March 2002.

Imhoff, M.L., 1995. Radar Backscatter and Biomass Saturation; Ramifications for Global Biomass Inventory, *IEEE Transactions on Geoscience and Remote Sensing*, March 1995, No.2.Vol.33, pp.511-518.

Kasischke, E.S, J.M. Melack; and M.C. Dobson, 1995. The Use of Imaging Radars for Ecological Applications – A Review. *Remote Sensing of Environment*, Vol.59:141-156.

Keßler, P. 2001. *The balance between biodiversity conservation and sustainable use of tropical rain forest*. pp 139-145. The Tropenbos Foundation, Wageningen. The Netherlands.

Kuntz, S. and F. Siegert, 1999. Monitoring of deforestation and land use in Indonesia with multi-temporal ERS data. *International Journal of Remote Sensing*, Vol.20, No.14, pp.2835-2853.

Le Toan, T., A. Beaudion, J. Riom, and D. Guyon, 1992. Relating Forest Biomass to SAR Data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.30(2):403-411.

Leberl, F., 1990. *Radargrammetric Image Processing*, Artech House, Inc. Norwood, MA, 595 p.

Melon P., J.M. Martinez, T. Le Toan, L.M.H. Ulander and A. Beaudoin, 2001. On the retrieval of forest stem volume from VHF SAR data: Observation and modelling, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.39, pp.2364-2372, No.11,November 2001.

Mougin, E., C. Proisy, G. Marty, F. Fromard, H. Puig, J. Betoulle and J. P. Rudant, 1999. Multi-frequency and multi-polarisation radar backscattering from mangrove forest. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, 2 part, pp 69-77.

Nghiem, S.V., S.H. Yueh, R. Kwok and F. K. Li, 1992, Symmetry properties in polarimetric remote sensing, *Radio Science*, Vol. 27, pp. 693 - 711.

Priadjati, A. 2002. Dipterocarpaceae: Forest Fires and Forest Recovery. Wageningen (The Tropenbos Foundation), Tropenbos-Kalimantan Series 8, *PhD thesis*, Wageningen University, The Netherlands, 214p.

Quiñones, M.J., 2002. Polarimetric data for tropical forest monitoring. Studies at the Colombian Amazon. *PhD. Thesis*, Wageningen University, The Netherlands, 184 p.

Quiñones, M.J., and D.H. Hoekman, 2004, Exploration of factors limiting biomass estimation by polarimetric radar in tropical forests, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.42, No.1, January 2004, pp.86-104.

Ranson, K. J and G. Sun, 1994, Mapping biomass of a northern forest using multi-frequency SAR data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.32, pp.388-396.

Rignot, E.J, R. Zimmermann, and J.J. van Zyl, 1995, Spaceborne applications of P-band imaging radar for measuring forest biomass, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.33, pp.1162-1169.

Rodriguez, Z. 2002. Capabilities of POLSAR data for land cover classification in a tropical rain forest environment. East Kalimantan, Indonesia. *MSc Thesis*, Wageningen University, The Netherlands.

Sgrenzaroli, M., 2004. Tropical forest mapping at regional scale using the GRFM SAR mosaics over The Amazon in South America. *Ph.D. Thesis* Wageningen University, 25 Feb 2004.

Sidiyasa, K. 2001. Tree Diversity in the Rain Forest of Kalimantan. In: *Proceedings of The Balance Between Biodiversity Conservation and Sustainable Use of Tropical Rain Forest*. The Tropenbos Foundation, Wageningen. The Netherlands.

Siegert, F. and Hofmann A.A., 2000. The 1998 forest fire in East Kalimantan (Indonesia): A quantitative evaluation using high resolution, multi-temporal ERS-2 SAR images and NOAA-AVHRR hotspot data. Remote Sensing of Environment, Vol.71, No.1, pp.65-77.

Soerianegara, I. and Lemmens, R.H.M.J. 1993. Timber trees: major commercial timbers. *PROSEA, Plant Resources of South-East Asia*. No. (5). Wageningen, The Netherlands. Pudoc. 610p.

Sumitro, A., (Dean Faculty of Forestry, Gajah Mada University), 1993, *The Indonesia's preparedness toward target year 2000 for sustainable forest management.*

Temmes, M. 1992. Reforestation operations manual for alang-alang grasslands. *Reforestation and tropical forest management project ATA-267*. Ministry of Forestry Indonesia.

Thompson, M.D., J.B. Mercer, and M.C. Wride, 1993. A decade of commercial radar operations: Intera's STAR-1 and STAR-2 services, *Proceedings of the International Symposium on Operationalization of Remote Sensing*, Enschede, The Netherlands, 19-23 April 1993, Vol.5, pp.191-202.

Tolkamp, G. W, A. Priadjati and R. Effendi. 2001. Towards ans ecology-based strategy for the reforestration of Imperata cylindrica Grasslands in East-Kalimantan. In: *Proceedings of The Balance Between Biodiversity Conservation and Sustainable Use of Tropical Rain Forest*. The Tropenbos Foundation, Wageningen. The Netherlands.

Tough, R.J.A., D. Blacknell and S. Quegan, 1995, A statistical description of polarimetric and interferometric synthetic aperture radar data, *Proc. Royal Society London A*, Vol. 449, pp. 567-589.

Ulaby, F.T., R.K. Moore, and A.K. Fung, 1986, *Microwave Remote Sensing, Volume III: From Theory to Applications*, Athech House.

Ulander, L.M.H., 1996. 'Radiometric slope correction of synthetic-aperture radar images'. *IEEE Transactions on Geoscience and Remote Sensing*, vol.34, no.5, pp. 1115-1122.

Van der Sanden, J.J. 1997. Radar Remote Sensing to Support Tropical Forest Management. *Ph.D. Thesis*, Wageningen Agricultural University, The Netherlands, 330 p.

Van Zyl, J.J., 1989, Unsupervised classification of scattering behaviour using radar polarimetry data, *IEEE Transactions on Geoscience and Remote Sensing*, GE-27, no 1, pp.36-45.

Van Zyl, J.J. 1993. The effect of topography on radar scattering from vegetated areas, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.31, No.1, pp153-160.

Varekamp, C., 2001. Canopy reconstruction from interferometric SAR, *Ph.D. Thesis* Wageningen University, 12 January 2001.

Varekamp, C. and D.H. Hoekman, 2001. Segmentation of high-resolution InSAR data of tropical forests using Fourier parameterised deformable models, *International Journal of Remote Sensing*, Vol.22, No.12, pp.2339-2350.

Vargas, J. 2002. Potential of ENVISAT-ASAR data to support the assessment of land cover in a tropical rain forest environment. East Kalimantan, Indonesia. *MSc Thesis*, Wageningen University, The Netherlands.

Wegmuller, U. and C.L. Werner, 1995. SAR interferometric signatures of forest, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.33, pp.1153-1161.

Wegmuller, U and C.L. Werner, 1997, Retrieval of vegetation parameters with SAR interferometry, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.35, pp.18-24.

Wooding, M.G., A.D. Zmuda, D.H. Hoekman, J.J. de Jong and E.P.W. Attema, 1998. The Indonesian Radar Experiment (INDREX-96), *In: Proceedings of Final results workshop Indonesian radar experiment*, ESTEC, The Netherlands, 1999, ESA SP-489, pp1-9.

Zuhair, M, M. Y. A. Hussin and M. Weir. 2001. Monitoring mangrove forests using remote sensing and GIS. In: *Proceedings of The Balance Between Biodiversity Conservation and Sustainable Use of Tropical Rain Forest*. The Tropenbos Foundation, Wageningen. The Netherlands.

About the author

Kemal Unggul Prakoso received the B.Sc degree in physics from Bandung Institute of Technology, Bandung, Indonesia, in 1991, and the M.Sc. degree in Geo-information Science from Wageningen University, Wageningen, The Netherlands, in 2000. From 1991 to 1998, he worked with PT. Mapindo Parama, Jakarta, an Indonesian Aerial Survey Mapping and Geographic Information System company as a data processing co-ordinator to produce forest inventory maps all over the country. At this company he dedicated his work to several airborne radar systems such as the STAR-1, STAR-3i, Do-SAR and AeroSensing SAR's, focusing on the production of Topographical (Contour) and Vegetation Maps. His first introduction to research in tropical rain forests was in 1996 when he was involved in the preparations for the Indonesian Radar Experiment 1996 (INDREX'96) campaign initiated by the European Space Agency (ESA), Ministry of Forestry, Tropenbos Foundation and Wageningen University at test sites in Sumatera and Kalimantan using the Do-SAR airborne system. In September 2000 he also was involved in the PACRIM-2 campaign in East-Kalimantan using the AirSAR/TopSAR of NASA/JPL. He started his Ph.D. thesis research in 2000 at the Department of Environmental Sciences, Wageningen University, Wageningen, The Netherlands. His main research interest is in the area of radar remote sensing of tropical forest, with emphasis on the development and accuracy assessment of techniques for classification and biomass measurement using interferometric and polarimetric Synthetic Aperture Radar. Since the year 2004, he was involved in setting up and managing PT. SarVision Indonesia, a foreign capital company working in consultancy for nature conservation mapping. In the same year, he acted as the local coordinator for the Indonesian Radar Experiment 2 (INDREX2), a research project initiated by the European Space Agency (ESA) and the Indonesian Ministry of Forestry.

List of publications

[1] **Prakoso, K.U**, B. Suryokusumo, D.H Hoekman, 1999, Radar tree mapping and FIEPLP methodology to support sustainable forest management. International MOFEC-Tropenbos and NOW workshop, Balikpapan Indonesia Dec 6-8 1999.

[2] **Prakoso, K.U**, B. Suryokusumo and D. H. Hoekman. 1999. Radar tree mapping and FIEPLP methodology to support sustainable forest management. In proceeding of Final results workshop Indonesian radar experiment, ESTEC, The Netherlands.

[3] **Prakoso, K.U**, B. Suryokusumo. 2000. Interferometric SAR tree mapping: potentials and limitations in support of sustainable forest management in Indonesia, M.Sc thesis Wageningen University and research Centre, Wageningen, The Netherlands.

[4] Grim, R.J.A, F.M. Seifert, D.H. Hoekman, C. Varekamp, Y.A. Hussin, M.A. Sharifi, M. Weir, **Kemal Unggul Prakoso**, Ruandha Agung Sugardiman, Muljanto Nugroho, Bambang Suryokusumo, August 2000, SIRAMHUTAN Sistem Informasi untuk Manajemen HUTAN, A demonstration project for forestry in Indonesia, BCRS Publication NRSP-2 report 00-07.

[5] Hoekman, D.H, **K.U. Prakoso**, B. Suryokusumo, M. Nugroho, R.A. Sugardiman., 2000, Experiment plan for the PacRim 2 campaign in East-Kalimantan, Indonesia, Version 1, August 2000, Department of Environmental Sciences, Wageningen University.

[6] **Prakoso, K.U**, 2001 Assessment of InSAR tree mapping accuracy in tropical rain forests, 3rd International Symposium on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications, 11th - 14th September 2001 SCEOS, University of Sheffield, UK.

[7] Hoekman, D.H, **K.U. Prakoso**, B. Suryokusumo and M.A.M. Vissers., 2001, PacRim 2 database Sungai Wain, ESTEC Contract No. 15193/01/NL/DC "PacRim-2 Ground Data Collection in Kalimantan", ESTEC, The Netherlands.

[8] **Prakoso, K.U**, 2003 Tropical Forest Mapping using Multi-Band Polarimetric and Interferometric SAR Data, POLinSAR workshop on applications of SAR Polarimetry and Polarimetric Interferometry, 14 -16 January 2003 ESRIN, Frascati, Italy.

[9] Hoekman, D.H., M.J. Quiñones, **K.U. Prakoso** and Gary Smith, 2004, Appendix G: Sungai Wain Test Site Report, 21 pages. In: Smith, *et.al*., Applications of Low-Frequency SAR, Final Project Report, November 2004, ESA-ESTEC contract no. 16115/02/NL/MM, November 2004, 119 pages + 13 Appendices.