

VEGETATION CLASSIFICATION USING SEASONAL  
VARIATIONS OF SCATTEROMETER DATA AT C-BAND AND  
KU-BAND

by  
Gardner Watt

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Science

Department of Electrical and Computer Engineering  
Brigham Young University

August 1999

Copyright © 1999 Gardner Watt

All Rights Reserved

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Gardner Watt

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

\_\_\_\_\_

Date

\_\_\_\_\_

Dr. David Long, Chair

\_\_\_\_\_

Date

\_\_\_\_\_

Dr. David Arnold

\_\_\_\_\_

Date

\_\_\_\_\_

Dr. Mike Jensen

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Gardner Watt in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

---

Date

---

Dr. David Long  
Chair, Graduate Committee

Accepted for the Department

---

Lee Swindlehurst  
Graduate Coordinator

Accepted for the College

---

Douglas M. Chabries  
Dean, College of Engineering and Technology

## ABSTRACT

# VEGETATION CLASSIFICATION USING SEASONAL VARIATIONS OF SCATTEROMETER DATA AT C-BAND AND KU-BAND

Gardner Watt

Department of Electrical and Computer Engineering

Master of Science

Satellite data has played a large part in helping create global weather models and in monitoring global climate and weather. Scatterometers, in particular, have been very useful determining global near surface winds over the ocean surface. However, lately there have arisen other applications of scatterometry besides measuring wind. Specifically, there have been many experiments in classifying vegetation using various types of satellite data. This thesis presents a new algorithm for classifying vegetation regions using the seasonal response of scatterometry data at two different frequencies.

In leading up to the algorithm a number of studies looking at the seasonal response of  $\sigma^o$  are done. The  $\sigma^o$  seasonal response is observed over different vegetation classes at two different frequencies over both North America and South America. The results show that there is a significant difference in seasonal response over different vegetation regions at different frequencies.

The algorithm discussed in this thesis uses the differences in the seasonal response of  $\sigma^o$  to classify vegetation. The algorithm takes a training set from two

different vegetation maps and uses singular value decomposition (SVD) to create a basis set. The results are compared with the vegetation maps and found to perform well over large areas. Misclassification occurs with classes of similar types. Overall, results are consistent with previous work done and could be combined with other methods to improve classification.

# Contents

<b>List of Tables</b>	<b>xi</b>
<b>List of Figures</b>	<b>xviii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Contributions . . . . .	2
1.3 Overview . . . . .	3
<b>2 Background</b>	<b>5</b>
2.1 Overview of Scatterometry . . . . .	5
2.2 Scatterometry in Land and Ice studies . . . . .	6
2.3 Vegetation classification and studies . . . . .	7
2.4 Overview of the ERS-1/2 AMI and NSCAT scatterometers . . . . .	8
2.5 The SIRF Algorithm . . . . .	10
2.6 Singular Value Decomposition . . . . .	10
<b>3 The Seasonal Response of <math>\sigma^o</math></b>	<b>17</b>
3.1 Temporal Response of $\mathcal{A}$ and $\mathcal{B}$ values over Land . . . . .	17
3.2 $\sigma^o$ versus Precipitation and Temperature . . . . .	26
3.3 $\sigma^o$ Variation over Vegetation Regions . . . . .	31
3.4 Seasonal Response of $\sigma^o$ . . . . .	40
<b>4 Vegetation Classification Algorithm</b>	<b>49</b>
4.1 The Steps . . . . .	49
4.2 Selecting a Training Set . . . . .	50

4.3	Finding a Basis Set . . . . .	51
4.4	Classification . . . . .	53
<b>5</b>	<b>Results using the Algorithm</b>	<b>57</b>
5.1	Results over South America . . . . .	57
5.2	Results over North America . . . . .	64
5.3	Accuracy and Inaccuracy of the Classification . . . . .	71
<b>6</b>	<b>Conclusion</b>	<b>79</b>
6.1	Overview . . . . .	79
6.2	Contributions . . . . .	80
6.3	Future Work . . . . .	80
<b>A</b>	<b>Temporal Average Algorithm</b>	<b>83</b>
A.1	Abstract . . . . .	83
A.2	Introduction . . . . .	83
A.3	The Time Index Algorithm . . . . .	88
A.4	Non-temporally Uniform Measurements . . . . .	90
A.5	Conclusion . . . . .	92
	<b>Bibliography</b>	<b>99</b>

## List of Tables

3.1	The 32 different vegetation classes found in the Matthews vegetation map. . . . .	34
3.2	The 16 different vegetation classes from the AVHRR-derived vegetation map. . . . .	35
3.3	New vegetation classes and how they were combined from the Matthews and AVHRR-based maps. . . . .	35
3.4	Vegetation classes over North America . . . . .	38
3.5	Vegetation classes over South America . . . . .	39
5.1	Percentage of NSCAT vegetation classification over South America for Matthew's using $L^2$ Norm. . . . .	58
5.2	Percentage of NSCAT vegetation classification over South America for Matthews using maximum likelihood (ML). . . . .	58
5.3	Percentage of NSCAT vegetation classification over South America for AVHRR using $L^2$ Norm. . . . .	58
5.4	Percentage of NSCAT vegetation classification over South America for AVHRR using maximum likelihood (ML). . . . .	59
5.5	Percentage of ERS-2 vegetation classification over South America for Matthews using $L^2$ Norm . . . . .	61
5.6	Percentage of ERS-2 vegetation classification over South America for Matthews using maximum likelihood. . . . .	61
5.7	Percentage of ERS-2 vegetation classification over South America for AVHRR using $L^2$ Norm . . . . .	62
5.8	Percentage of ERS-2 vegetation classification over South America for AVHRR using maximum likelihood (ML). . . . .	63

5.9	Percentage of NSCAT vegetation classification over North America for Matthews using $L^2$ Norm . . . . .	64
5.10	Percentage of NSCAT vegetation classification over North America for Matthews using maximum likelihood (ML). . . . .	65
5.11	Percentage of NSCAT vegetation classification over North America for AVHRR using $L^2$ Norm . . . . .	66
5.12	Percentage of NSCAT vegetation classification over North America for AVHRR using maximum likelihood (ML). . . . .	67
5.13	Percentage of ERS-2 vegetation classification over North America for Matthews using $L^2$ Norm . . . . .	67
5.14	Percentage of ERS-2 vegetation classification over North America for Matthews using maximum likelihood (ML). . . . .	68
5.15	Percentage of ERS-2 vegetation classification over North America for AVHRR using $L^2$ Norm. . . . .	69
5.16	Percentage of ERS-2 vegetation classification over North America for AVHRR using maximum likelihood (ML). . . . .	70
5.17	Class-Average and overall accuracy coefficients over South America for the given classification parameters used in the classification algorithm	71
5.18	Class-Average and overall Accuracy coefficients over North America for the given classification parameters used in the classification algorithm	71
5.19	New vegetation classes over North America . . . . .	73
5.20	New vegetation classes over South America . . . . .	73
5.21	The overall accuracy, $A(X)$ , class-average, $C(X)$ ,and $J_p(X)$ accuracy coefficient over South America for the reduced vegetation set. . . . .	74
5.22	The overall accuracy, $A(X)$ , class-average, $C(X)$ ,and $J_p(X)$ accuracy coefficient over North America for the reduced vegetation set. . . . .	74
5.23	The $R_p$ inaccuracy coefficients for the given , cost functions over South America for the given classification parameters used in the classification algorithm. . . . .	77

5.24 The  $R_p$  inaccuracy coefficients for the given , cost functions over North America for the given classification parameters used in the classification algorithm. . . . . 77



## List of Figures

2.1	ERS-1 fan beam pattern . . . . .	8
2.2	NSCAT fan beam pattern. . . . .	9
2.3	SIRF A image of North America for NSCAT-V. . . . .	11
2.4	SIRF A image of North America for ERS-2. . . . .	11
2.5	SIRF A image of South America for NSCAT-V. . . . .	12
2.6	SIRF A image of South America for ERS-2. . . . .	12
2.7	SIRF B image of North America for NSCAT-V. . . . .	13
2.8	SIRF B image of North America for ERS-2. . . . .	13
2.9	SIRF B image of South America for NSCAT-V. . . . .	14
2.10	SIRF B image of South America for ERS-2. . . . .	14
2.11	Example of SVD and how it is used to form an orthogonal basis set along lines of highest variance. The basis set (x,y) before SVD (left) and the basis set (u,v) after rotating the basis. . . . .	15
3.1	SIRF image over North America showing the strip taken to create the Hovmoller diagrams. . . . .	17
3.2	Time series strips over a part of the Midwest region of North America for NSCAT-V and the corresponding average values over each slice shown with slice one at the top. . . . .	19
3.3	Time series strips over a part of the Midwest region of North America for ERS-1/2 in 1996 and the corresponding average values over each slice shown with slice one at the top. . . . .	20
3.4	Time series strips of B images over a part of the Midwest region of North America for NSCAT-V and the corresponding average values over each slice shown with slice one at the top. . . . .	21

3.5	Time series strips of B images over a part of the Midwest region of North America for ERS-1/2 in 1996 and the corresponding average values over each slice shown with slice one at the top. . . . .	22
3.6	Normalized histogram plots of the strips at various times throughout the year for NSCAT-V $\mathcal{A}$ values. . . . .	23
3.7	Normalized histogram plots of the strips at various times throughout the year for ERS-1/2 $\mathcal{A}$ values. . . . .	24
3.8	Thresholded temporal images of $\mathcal{A}$ . . . . .	25
3.9	$\mathcal{A}$ values versus time for ERS-1/2 (left) and NSCAT (right) over the small test region. . . . .	26
3.10	$\mathcal{B}$ values versus time for ERS-1/2 (left) and NSCAT (right) over the small test region. . . . .	27
3.11	C-band $\sigma^o$ $\mathcal{A}$ values and $\mathcal{B}$ values vs. temperature. o-1994 *-1995 +-1996. . . . .	28
3.12	Ku-band $\sigma^o$ $\mathcal{A}$ values for both v-pol(o) and h-pol(*) vs. temperature. . . . .	29
3.13	Ku-band v-pol(o) and h-pol(*) vs. C-band $\mathcal{A}$ during the NSCAT mission. . . . .	29
3.14	NSCAT-H $\mathcal{A}$ values vs. NSCAT-V $\mathcal{A}$ values over the study region showing the offset of $\mathcal{A}$ values for horizontal polarization with vertical polarization. . . . .	30
3.15	The Matthews Vegetation map of the world showing all 32 vegetation classes. . . . .	31
3.16	The AVHRR-derived Vegetation map of the world showing its thirteen vegetation classes. . . . .	32
3.17	The vegetation maps over North America showing the new projection and pixel resolution. . . . .	36
3.18	The vegetation maps over South America showing the new projection and pixel resolution. . . . .	37
3.19	The unified vegetation classes for Matthews and AVHRR maps over North America. . . . .	38

3.20	The unified vegetation classes for Matthews and AVHRR maps over South America. . . . .	39
3.21	Seasonal response of $\mathcal{A}$ -values over each vegetation region of North America for NSCAT-V. . . . .	40
3.22	Seasonal response of $\mathcal{A}$ -values over each vegetation region of North America for ERS-2. . . . .	41
3.23	Seasonal response of $\mathcal{B}$ -values over each vegetation region of North America for NSCAT-V. . . . .	42
3.24	Seasonal response of $\mathcal{A}$ -values over each vegetation region of South America for NSCAT-V. . . . .	43
3.25	Seasonal response of $\mathcal{A}$ -values over each vegetation region of South America for ERS-2. . . . .	44
3.26	Seasonal response of $\mathcal{B}$ -values over each vegetation region of South America for NSCAT-V. . . . .	45
3.27	Yearly seasonal response of $\mathcal{A}$ -values over each vegetation region of North America for ERS-1/2. . . . .	46
3.28	Yearly seasonal response of $\mathcal{A}$ -values over each vegetation region of South America for ERS-1/2. . . . .	47
4.1	The log of the principal component values of $X$ , showing that the highest two basis vectors account for most of the variance in the vegetation responses. . . . .	52
4.2	The eigenvectors of $X$ , showing that the highest basis vectors is mainly a DC component while the other vectors account for actual variations in the vegetation responses. . . . .	53
5.1	NSCAT Vegetation Classification over South America using the Matthews vegetation map. The left one is the classification using the $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier. . . . .	59

5.2	NSCAT Vegetation Classification over South America using the AVHRR vegetation map. The left one is the classification using the $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier. . . . .	60
5.3	ERS-2 Vegetation Classification over South America using the Matthews vegetation map. The left one is the classification using the $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier. . . . .	62
5.4	ERS-2 Vegetation Classification over South America using the AVHRR vegetation map. The left one is the classification using the $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier. . . . .	63
5.5	NSCAT Vegetation Classification over North America using the Matthews vegetation map. The top one is the classification using the $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier. . . . .	65
5.6	NSCAT Vegetation Classification over North America using the AVHRR vegetation map. The top one is the classification using the $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier. . . . .	66
5.7	ERS-2 Vegetation Classification over North America using the Matthews vegetation map. The top one is the classification using the $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier. . . . .	68
5.8	ERS-2 Vegetation Classification over North America using the AVHRR vegetation map. The top one is the classification using the $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier. . . . .	69
A.1	(a) Simulated SIR image (b) linear average image (c) filtered image at day 3. . . . .	84

A.2	Simulated ice edge images from (a) day 0 to (m) day 6, in 1/2 day increments. . . . .	85
A.3	Ice edge of the SIR image, linear average image and filtered image at day 3 . . . . .	86
A.4	SIR image and linear average image shifted from the center to an earlier period. . . . .	87
A.5	(a) The simulated SIR $\mathcal{A}$ image (b) Time index image in days (c) Filtered image at day 2.2 (d) Filtered image at day 3. . . . .	89
A.6	(a) Horizontal slice of simulated SIR $\mathcal{A}$ image, filter image at day 3 and shifted average image with center corresponding to day 2.2 (b) Horizontal slice of simulated SIR $\mathcal{A}$ image, filtered image at day 2.2 and shifted linear average image with center at day 2.2 (c) Time index image in days. . . . .	90
A.7	(a) Histogram of the time of the measurements used to create image (b) SIR $\mathcal{A}$ image (c) Horizontal slice of the SIR image, the filtered image at day 1.1, and the shifted linear average image with center at day 1.1 (d) Time index image in days (e) Filtered image at day 1.1 (f) Horizontal slice of the time index image . . . . .	91
A.8	(a) Histogram of the time of the measurements used to create image (b) SIR $\mathcal{A}$ image (c) Horizontal slice of the SIR image, the filtered image at day 3.1, and the shifted linear average image with center at day 3.1 (d) Time index image in days (e) Filtered image at day 3.1 (f) Horizontal slice of the time index image . . . . .	93
A.9	(a) Histogram of the time of the measurements used to create image (b) SIR $\mathcal{A}$ image (c) Horizontal slice of the SIR image, the filtered image at day 4.2, and the shifted linear average image with center at day 4.2 (d) Time index image in days (e) filtered image at day 4.2 (f) Horizontal slice of the time index image . . . . .	94

A.10 (a) Histogram of the time of the measurements used to create image  
(b) SIR  $\mathcal{A}$  image (c) Horizontal slice of the SIR image, the filtered  
image at day 1.5, and the shifted linear average image with center at  
day 1.5 (d) Time index image in days (e) filtered image at day 1.5 (f)  
Horizontal slice of the time index image . . . . . 95

# Chapter 1

## Introduction

Satellite data has played a large part in helping create global weather models and in monitoring global climate and weather. Scatterometers, in particular, have been very useful determining global near surface winds over the ocean surface. However, recently new applications of scatterometry have arisen in addition to measuring wind. One such application includes studies of annual global ice over the Arctic and Antarctic. Another application is studies of regional and global vegetation and soil moisture.

This thesis presents a new algorithm for classifying vegetation regions using the seasonal response of scatterometer data, specifically, the normalized radar cross section,  $\sigma^o$ . An analysis of the seasonal response observed by two different instruments, NSCAT and ERS-1/2, at two different frequencies over various vegetation areas is made and compared to different ‘truth’ vegetation maps. This analysis assumes that the  $\sigma^o$  response is homogeneous over a particular vegetation and does not address the issues of topography, climate, and snow/freezing on vegetation and how they affect the seasonal variation of  $\sigma^o$ . It is shown that this technique gives results consistent with existing vegetation maps.

### 1.1 Motivation

Vegetation is a crucial factor in controlling such things as global oxygen content,  $CO_2$  production, global warming and over-all global weather patterns. Vegetation studies and classification are therefore important in creating global models of

the earth and in helping determine how the environment is being affected by human activities.

The first scatterometer flew on Seasat (SASS) and was put up in 1978. The European Space Agency put ERS-1 up in 1991, followed by ERS-2 in 1995. NASA launched their version (NSCAT) in 1996 and will be following it up with QuickSCAT in the Spring of 1999. Although scatterometers were originally designed to measure near-surface winds over the ocean surface, studies have used the data over the land and polar regions. Scatterometers are useful in studying vegetation because they are insensitive to cloud cover at their operating frequencies. They also obtain global coverage over a reasonably short time and therefore are useful in global studies of seasonal changes over vegetation regions. Although many studies have used scatterometer data in vegetation studies at different frequencies, few studies have tried to incorporate scatterometer data over a variety of frequencies in vegetation classification. An algorithm is needed that can incorporate the seasonal nature of vegetation at multiple frequencies so that vegetation studies and classification can be improved.

## 1.2 Contributions

Several methods have been employed to derive vegetation maps from different types of satellite data. Both trained and untrained algorithms have been used. Most vegetation studies that use satellite data tend to be regional in nature. Further most studies using microwave data have employed SAR or radiometers for vegetation classification. Also, the studies usually consider only one instant in time over a particular area and do not incorporate the seasonal response of the vegetation.

The method developed in this thesis differs from previous algorithms. Since it combines the information over a time range at two different frequencies, it examines the seasonal response of the vegetation over a number of areas and uses that information to classify the vegetation cover. The method, however, does not consider topography, climate and snow/freezing and how they affect the seasonal variation of  $\sigma^0$ . The classification technique discussed requires training the algorithm first, although the algorithm can be modified for use in untrained classifications. The

following contributions are found in this thesis:

1. An assessment of the seasonal response of  $\sigma^o$  over a given time period for the following:
  - (a) Over different vegetation regions
  - (b) At multiple frequencies (C-band and Ku-band)
  - (c) At different polarizations
2. A method of combining two different maps at two different projections and grid resolution for easy comparison.
3. An operational algorithm using the seasonal response of  $\sigma^o$  to classify vegetation.
4. An assessment of the vegetation classification algorithm over two regions: North America and South America at both bands.
5. Comparison of the algorithm to two different vegetation sets (Matthews and AVHRR-derived).

### 1.3 Overview

Chapter two of this thesis presents background information on scatterometry, vegetation classification and the SIRF algorithm. It discusses what scatterometry is and how it relates to vegetation classification. The SIRF algorithm, which creates images from  $\sigma^o$ , is considered.

Chapter three gives some motivation in using the seasonal response of the  $\sigma^o$  values over different vegetation regions as a means of classifying vegetation. A method of combining maps with different projections and resolutions for easy comparison is also discussed.  $\sigma^o$  is compared with temperature and precipitation in a particular area to evaluate any correlation.

Chapter four puts forth the actual steps to the vegetation classification algorithm as well as talk about what Singular Value Decomposition (SVD) is and how it

is used in the algorithm. Chapter five presents results over North and South America and gives a quantitative assessment. The final chapter summarizes the results, contributions and proposes future work areas to be considered.

## Chapter 2

### Background

Since  $\sigma^o$  measurements from scatterometers are to be used in the vegetation classification algorithm put forth in this thesis, a brief explanation and principles of scatterometry is discussed in this chapter. The linear scattering model and how the SIRF algorithm creates images is also discussed.

#### 2.1 Overview of Scatterometry

Scatterometers are radars that measure the backscatter return off of the earth's surface. Like other radars, the power returned is given by the radar equation

$$P_s = \frac{P_t G^2 \lambda^2 A \sigma^o}{(4\pi)^3 r^4} \quad (2.1)$$

where  $P_s$  is the power returned,  $P_t$  is the power transmitted,  $G$  is the gain of the antenna,  $\lambda$  the wavelength,  $\sigma^o$  the normalized radar backscatter coefficient (or the normalized radar cross section),  $A$  the area of the target, and  $r$  the distance to the target.

The key parameter of interest in remote sensing is  $\sigma^o$ , the normalized radar cross-section. This parameter gives an indication of how large a target looks to the radar. It is a function of incidence angle and is sensitive to the surface roughness and the surface's electrical properties.  $\sigma^o$  is a ratio of the amount of power returned over the power transmitted. The radar equation can be rearranged to obtain

$$\sigma^o = \frac{P_s}{X} \quad (2.2)$$

where

$$X = \frac{G^2 \lambda^2 P_t}{(4\pi)^2 r^4} \quad (2.3)$$

is termed the X-factor. The X-factor can be computed from known parameters, thus  $\sigma^o$  can be calculated by measuring the return power and dividing that by  $X$ . Scatterometers actually measure  $P_r$ , the return power, which consists of both signal power,  $P_s$  and noise power,  $P_n$ , i.e.  $P_r = P_s + P_n$ . A separate measurement of the noise power is also taken and subtracted from the return power to estimate the signal power.

Scatterometers were originally designed to measure near-surface winds over the ocean. Wind blowing over the ocean's surface creates waves. The stronger the winds, the rougher the surface and the more backscatter power is returned from areas of ocean with rough waves (corresponding to higher wind speeds). Scatterometers can therefore correlate the return backscatter power to wind speed.

Scatterometer measurements can also be used over ice and land. Different land regions reflect back a different amount of power depending on the characteristics of the land surface. For example, flat smooth areas return less power than areas that are more rough. Also, different vegetation characteristics give different responses. Grassland, for example, has a different  $\sigma^o$  response than does deciduous forest due to the dense canopy of leaves.

## 2.2 Scatterometry in Land and Ice studies

Both ERS-1 and NSCAT scatterometers were designed for measuring near-surface winds. However, these instruments have also been found useful in studies over land and ice. For example, studies have been done over the Boreal forests regions in Canada, and over the Siberian Forests at C-band [1], [2], [3]. It has been shown that during the summer, the variation in the backscatter coefficient over these land areas is dependent on soil and vegetation moisture, while during the winter, soil freezing and snow cover drastically changes the response of  $\sigma^o$  [4]. Scatterometers have also been used in ice studies over Greenland [5] and vegetation studies over the Amazon at

Ku-band [6]. Specifically, vegetation classification with these instruments has lately become more important.

### 2.3 Vegetation classification and studies

Vegetation is a crucial factor in controlling such things as global oxygen,  $CO_2$  production, global warming and over-all global weather patterns. Vegetation studies and classification are therefore important in creating global models of the earth and in helping determine how the environment is being affected by human activities.

Although scatterometers have been used in a variety of vegetation classification studies, they have not been extensively used for vegetation classification. Most vegetation classification studies use radiometers, Synthetic Aperture Radars (SARs), and optical sensors instead of scatterometers. The Advanced Very High Resolution Radiometer (AVHRR), has been used over many parts of the world for different vegetation classification studies [7], [8]. SAR instruments have been used for vegetation classification over many parts of the world including over areas of Michigan [9] as well as areas over central Sumatra [10]. Other radiometers such as Landsat TM and MSS have also been used extensively in vegetation classifications and studies [11], [12].

Scatterometers can be very useful in determining the type of vegetation over the entire world. Their advantage over other types of data such as aerial photographs includes rapid, global coverage and the ability to work through clouds.

Different methods have been employed to derive vegetation maps from different types of satellite data. Both trained and untrained algorithms have been developed. Carpenter et al. discuss using the fuzzy ARTMAP neural network in a methodology to classify vegetation over the Cleveland National Forest [12]. Green and Clark, et al. discusses different methodologies in classifying mangroves [11]. Pierce et al. uses a knowledge based classification for SAR data [13] and Yamagata et al. uses texture analysis for SAR data over wetlands [14]. Benedetti, R et al. uses an unsupervised classification using principal component analysis [15]. Many studies combine data from different instruments to exploit the advantages of each [10].

## 2.4 Overview of the ERS-1/2 AMI and NSCAT scatterometers

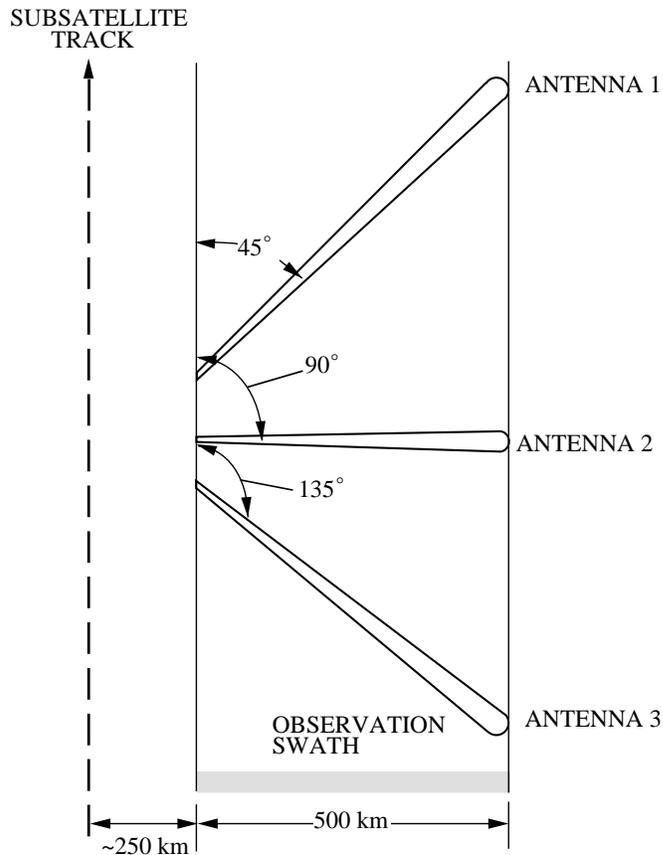


Figure 2.1: ERS-1 fan beam pattern

Data from the European Remote Sensing (ERS-1/2) AMI instruments and NASA's Scatterometer (NSCAT) are used in this study. ERS-1 is a fan beam scatterometer with a single polarization (vertical) that operates in the C-band at 5.3 GHz [16]. It has three antenna beams, mid, aft and fore. Figure 2.1 shows the antenna beam configuration of the instrument. The swath width is 500km with a spatial resolution of 50 km. Each 50 km resolution cell is measured at two angles of incidence. The range of incidence angles goes from 25-59 degrees for the fore and aft beams and 18-47 degrees for the mid-beam. ERS-1 covered a time period from July 17, 1991 thru

June of 1996. It has a nominal resolution of 50 km. ERS-2, the second generation of this satellite, is essentially the same instrument and covers a time period from May of 1996 up to the present time.

NSCAT is a dual-polarization (vertical and horizontal) scatterometer that operates in the Ku-band [17]. It has a fan-beam design with three antennas on each side, one of which is dual polarization. Figure 2.2 shows the antenna orientation. NSCAT uses Doppler processing to achieve cross-track resolution that allows the 600 km swath to be partitioned into 25 km sections. To achieve the along-track resolution of 25 km NSCAT makes measurements at intervals corresponding to 25 km. It has a nominal resolution from 25-50 km. It had a short mission from October 1996 to July of 1997. Both these instruments are used in this study in order to examine the seasonal response of vegetation at two different bands.

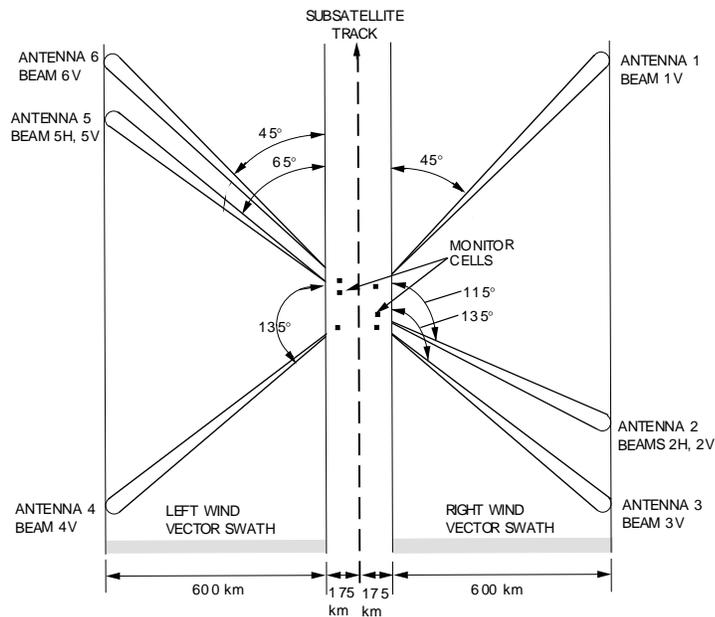


Figure 2.2: NSCAT fan beam pattern.

## 2.5 The SIRF Algorithm

Unfortunately the lower resolution of the scatterometer measurements is a limiting factor in land and vegetation studies. Various techniques have been developed to improve the resolution of these scatterometers. One such technique has been developed by Long, Hardin and Whiting [18] and is called the Scatterometer Image Reconstruction with Filtering (SIRF) Algorithm. SIRF is a nonlinear Multiplicative Algebraic Reconstruction Technique (MART). It takes the  $\sigma^o$  measurements of the same geographic location at different incidence angles and, using the linear scattering model, estimates  $\mathcal{A}$  and  $\mathcal{B}$  values for these areas.

A linear scattering model is used to relate  $\sigma^o$  with incidence angle. It has been shown that  $\sigma^o$  has a near linear relationship with incidence angle from  $22^\circ - 55^\circ$ .

$$10 \log_{10} \sigma^o = \mathcal{A} + \mathcal{B}(\theta - 40^\circ) \quad (2.4)$$

where  $\mathcal{A}$  is  $\sigma^o$  at  $40^\circ$  incidence angle and  $\mathcal{B}$  is the incidence angle dependence of  $\sigma^o$  (i.e. the slope of the line). The  $\mathcal{A}$  and  $\mathcal{B}$  values are estimated using the SIRF algorithm over a small region and then these values are used to form a greyscale image of large land regions. A detailed description and analysis of SIRF is found in [18].

$\mathcal{A}$  and  $\mathcal{B}$  images of North America and South America are used in this analysis. The images are average 14 day images. Figures 2.3 through 2.6 illustrate sample  $\mathcal{A}$  images for both the NSCAT-V and ERS-2  $\sigma^o$  data over North and South America. Figures 2.7 through 2.10 show the corresponding  $\mathcal{B}$  images for NSCAT-V and ERS-2 over North and South America.

## 2.6 Singular Value Decomposition

SVD has many uses in multivariate data analysis. It can be useful in principal component analysis and in this paper it is used to form an orthonormal basis set.

In Singular Value Decomposition (SVD), any  $M \times N$  matrix  $A$  can be written as the product

$$A = U\Sigma V^T \quad (2.5)$$

SIRF A image of North America for NSCAT-V

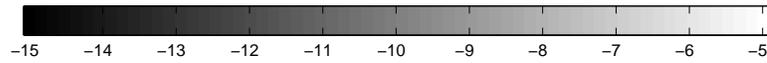
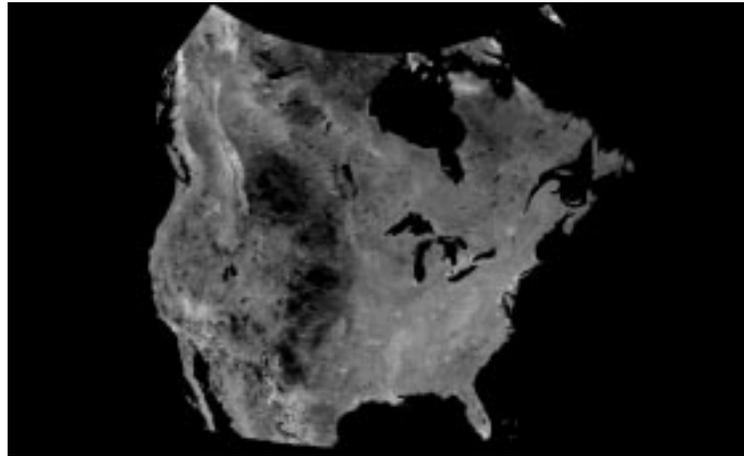


Figure 2.3: SIRF A image of North America for NSCAT-V.

SIRF A image of North America for ERS-2

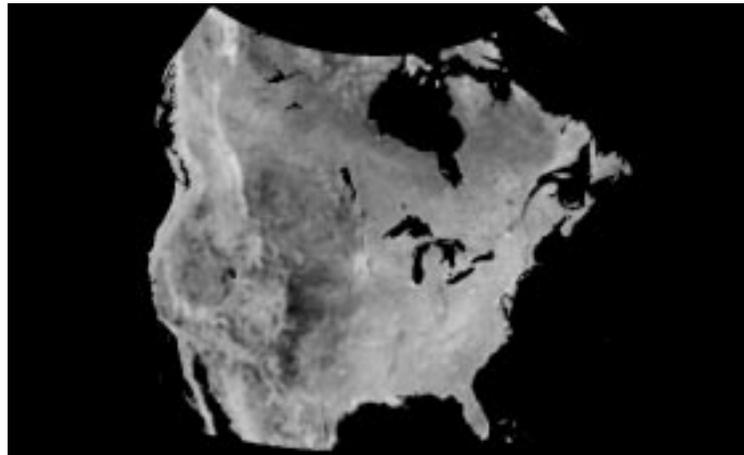


Figure 2.4: SIRF A image of North America for ERS-2.

SIRF A image of South America for NSCAT-V

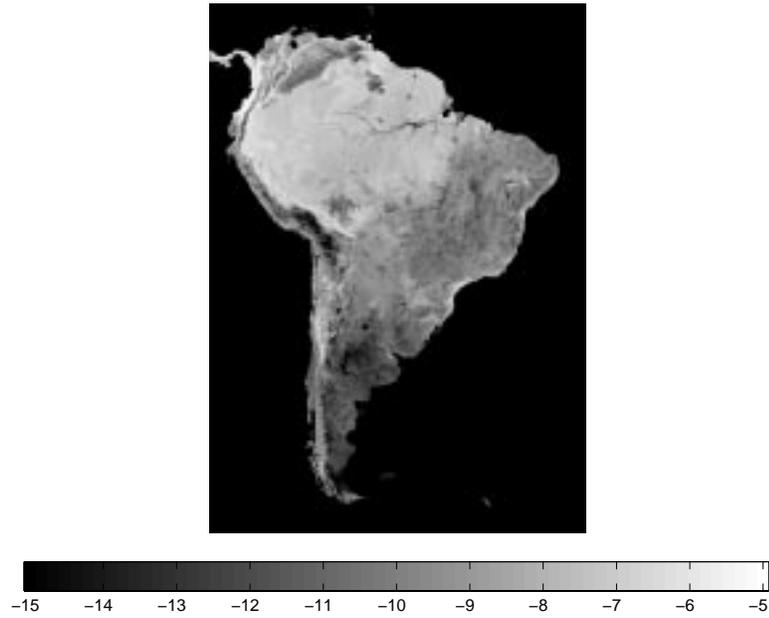


Figure 2.5: SIRF A image of South America for NSCAT-V.

SIRF A image of South America for ERS-2

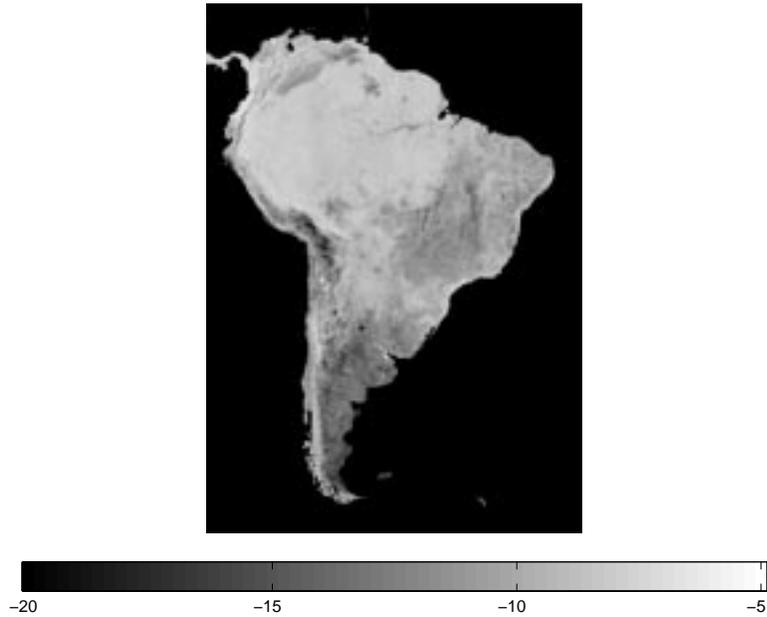


Figure 2.6: SIRF A image of South America for ERS-2.

SIRF B image of North America for NSCAT-V

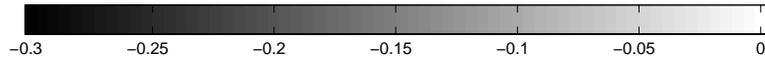
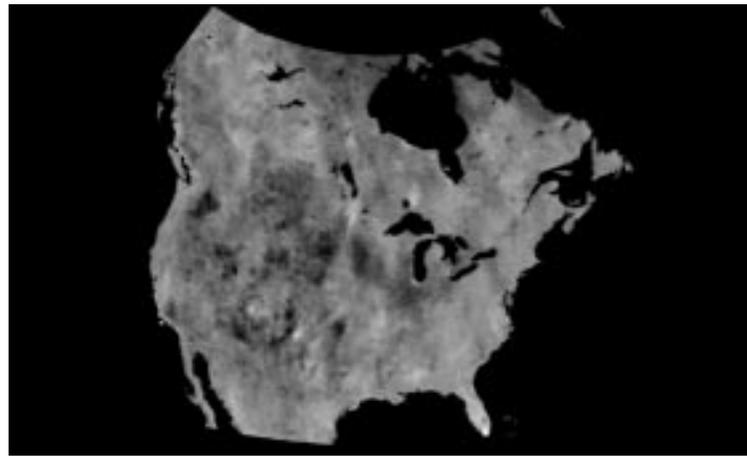


Figure 2.7: SIRF B image of North America for NSCAT-V.

SIRF B image of North America for ERS-2

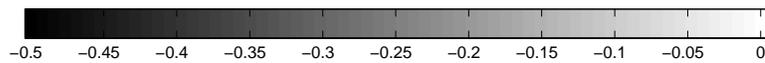
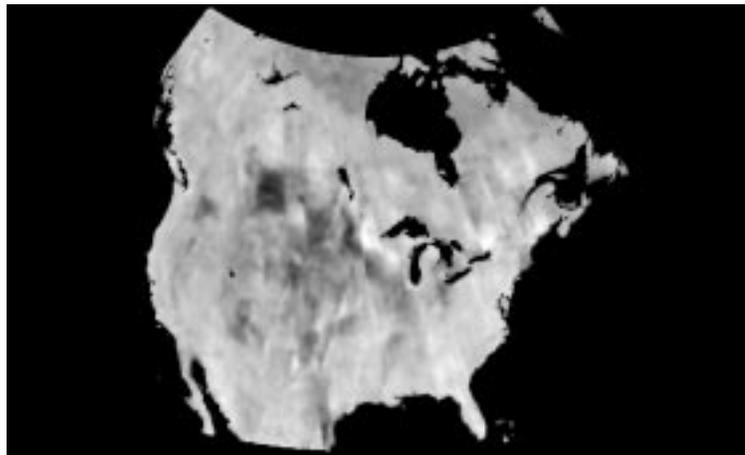


Figure 2.8: SIRF B image of North America for ERS-2.

SIRF B image of South America for NSCAT-V

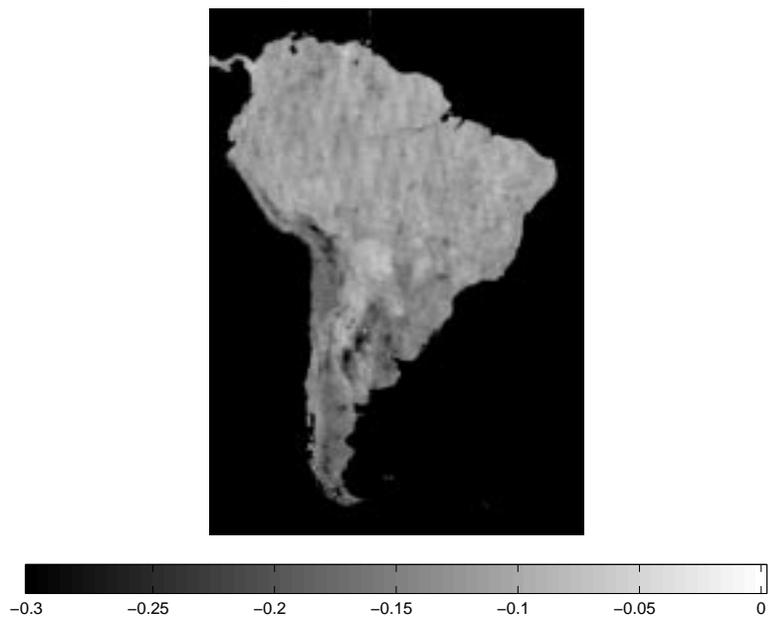


Figure 2.9: SIRF B image of South America for NSCAT-V.

SIRF B image of South America for ERS-2

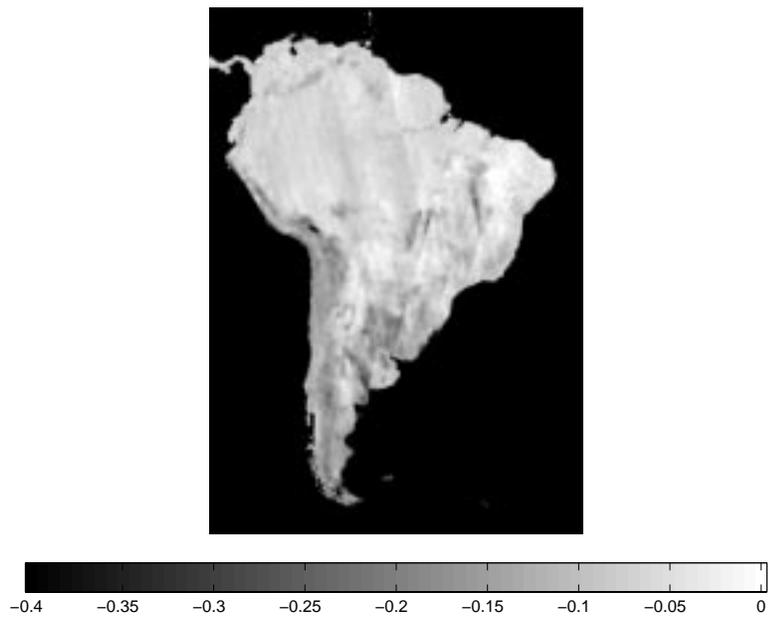


Figure 2.10: SIRF B image of South America for ERS-2.

where  $U$  is an  $M \times M$  orthogonal matrix,  $V$  is an  $N \times N$  orthogonal matrix and  $\Sigma = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_p]$  is  $M \times N$  with  $p = \min(M, N)$ . If  $M > N$  then  $\Sigma$  has the following form:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \cdots & \sigma_p & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}. \quad (2.6)$$

Here,  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$  and the values of  $\sigma_i$  are called the singular values of  $A$  and are the square root of the eigenvalues of  $A^T A$  and  $AA^T$  (i.e.  $\sigma_i = \sqrt{\lambda_i}$ ,  $\lambda_i$  being the eigenvalues of  $A^T A$  and  $AA^T$ ). Here the columns of  $U$  are the eigenvectors of  $AA^T$  and the rows of  $V^T$  are the eigenvectors of  $A^T A$ . Using the SVD we can form a new basis set from the eigenvectors of  $U$ . The columns of  $U$  form an orthonormal basis of the column space of  $A$ .

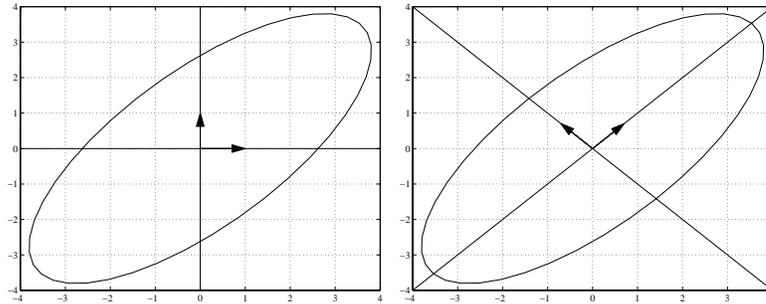


Figure 2.11: Example of SVD and how it is used to form an orthogonal basis set along lines of highest variance. The basis set  $(x,y)$  before SVD (left) and the basis set  $(u,v)$  after rotating the basis.

When this technique is applied to a covariance matrix,  $R$ , the analysis becomes a eigenvalue decomposition problem, where  $\Sigma$  contains the eigenvalues of  $R$  along

the diagonal and  $U = V$ , with the columns of  $U$  being the eigenvectors. In this case, the eigenvalues are seen as the variance of the data that lie along each of the eigenvectors. In other words, the new basis formed from the eigenvectors lie along the highest variance of the data. This technique is called principal value decomposition (PCA).

In the case where the data vector,  $A$ , does not represent a covariance matrix but is the actual data itself, the singular values of  $A$  still represent the highest variance in the data that lie along each of the eigenvectors. This is because  $AA^T$  can be seen as a sample correlation matrix and therefore the singular values of  $A$ , although not the actual variance, is an estimate of the variance.

To illustrate how we use SVD in forming a new orthogonal basis set, the following example is put forth. The ellipse on the left in Figure 2.11 represents a 2-dimensional data distribution. The highest variance of the data lies along one of the axis of the ellipse. If we take the SVD and use the eigenvectors as the new basis set, in essence the basis set is rotated to align with the axes of the ellipse as seen on the right of Figure 2.11. From this figure we can see that the primary axis lines up along the line of highest variance on the data with the second orthogonal vector lined up along the second highest variance of the data. This allows easier analysis of the data because by looking at the eigenvalues, which can be interpreted as the variance of the data in that dimension, one can determine how much information is in each dimension. Also one can reduce the dimensionality of the space by eliminating the dimensions with the lowest variance and keeping those basis that account for the most variance in the data. This can often make computations easier and more efficient.

## Chapter 3

### The Seasonal Response of $\sigma^o$

It has been shown that  $\sigma^o$  measurements over land are sensitive to freeze/thaw cycles and seasonal changes in vegetation regions [1], [3]. This chapter examines the seasonal response of  $\sigma^o$  over an annual cycle for both NSCAT and ERS-1/2  $\mathcal{A}$  and  $\mathcal{B}$  values. The sensitivities of  $\mathcal{A}$  and  $\mathcal{B}$  to temperature and precipitation are evaluated and the seasonal response of specific vegetation regions are considered.

#### 3.1 Temporal Response of $\mathcal{A}$ and $\mathcal{B}$ values over Land

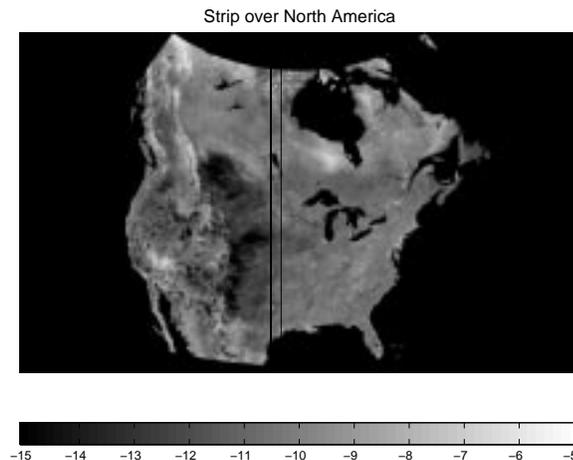


Figure 3.1: SIRF image over North America showing the strip taken to create the Hovmoller diagrams.

The temporal response of  $\mathcal{A}$  and  $\mathcal{B}$  values are first considered. To do this a strip through the Midwest region of North America was extracted from each 14-day image over a season (usually 10-12 months) and concatenated together to form a Hovmoller diagram. A Hovmoller diagram looks at the temporal changes over a particular area by concatenating the  $\sigma^o$  response of the area over time. A North American SIRF  $\mathcal{A}$  image showing the extracted strip is shown in Figure 3.1. All Hovmoller diagrams for both  $\mathcal{A}$  and  $\mathcal{B}$  images and ERS-1/2 and NSCAT images that follow use this same strip. Figures 3.2 and 3.3 show the Hovmoller diagrams as well as the average  $\sigma^o$  value over different latitude slices for  $\mathcal{A}$ -values while Figures 3.4 and 3.5 show them for  $\mathcal{B}$ -values. For NSCAT, the first image begins in October of 1996 and the last ends in July 1998. For ERS-1/2, the first image starts in January and goes to the end of December. Notice that for different latitudes there are different temporal variations in  $\mathcal{A}$  and  $\mathcal{B}$ . For example, slice one tends to be higher in the summer than the winter while slice four tends to be more flat. This suggests that over different regions we get different seasonal  $\sigma^o$  responses. Additionally, the  $\mathcal{B}$  values tend to be more noisy, especially for the ERS-1/2 case, suggesting that the  $\mathcal{B}$  images may be limited in their use. The following analysis is to determine what is causing the change. We will look at the histograms of  $\sigma^o$  over the season as well as the correlation between  $\sigma^o$  and precipitation and temperature to determine if any changes are due to these factors.

From each vertically sectioned strip, a normalized histogram was computed. A few of these plots are seen in Figures 3.6 and 3.7. These plots show that there are two distinct distributions of values, more so for NSCAT than ERS-1/2. These distributions change and fluctuate throughout the year. The fluctuations appear to correspond to vegetation growth. During the summer months when vegetation and crops are growing we see a distribution of  $\mathcal{A}$  values that are higher. During the winter months when crops and vegetation are only found in the southern regions, the distribution tends to be lower.

If a threshold is chosen between the two peaks in the histogram and two different values are chosen for values of pixels on each side of that threshold we get the images in Figure 3.8. Here we can see more clearly that there seems to be a

North America strips of NSCAT-V (Oct 1996–July 1997)

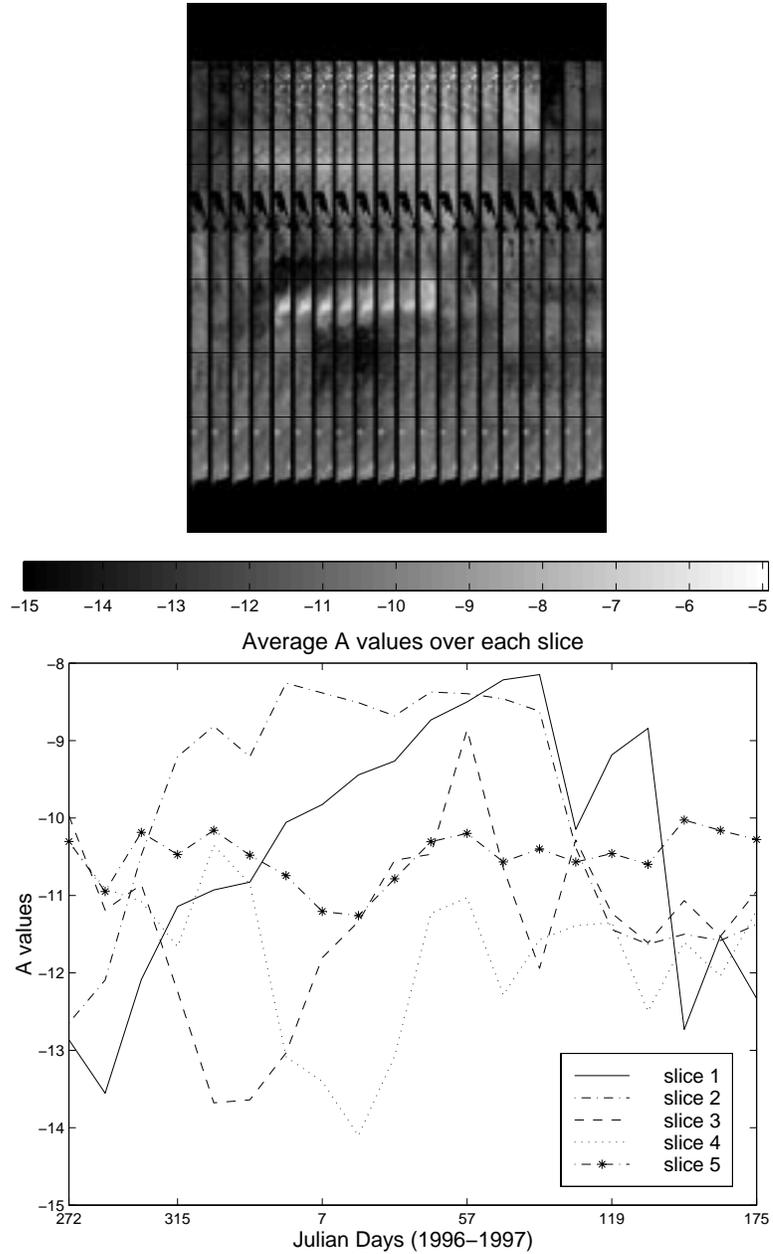


Figure 3.2: Time series strips over a part of the Midwest region of North America for NSCAT-V and the corresponding average values over each slice shown with slice one at the top.

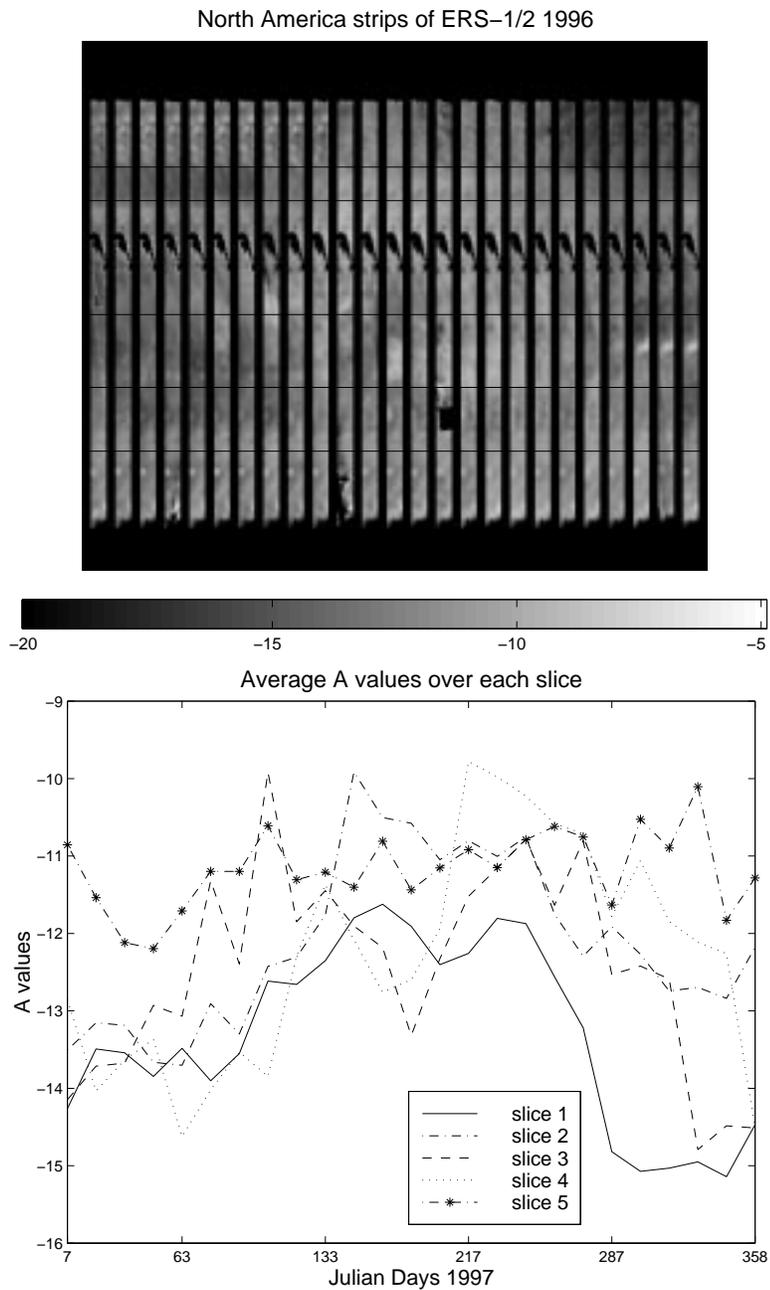


Figure 3.3: Time series strips over a part of the Midwest region of North America for ERS-1/2 in 1996 and the corresponding average values over each slice shown with slice one at the top.

North America strips of B-values for NSCAT-V (Oct 1996–July 1997)

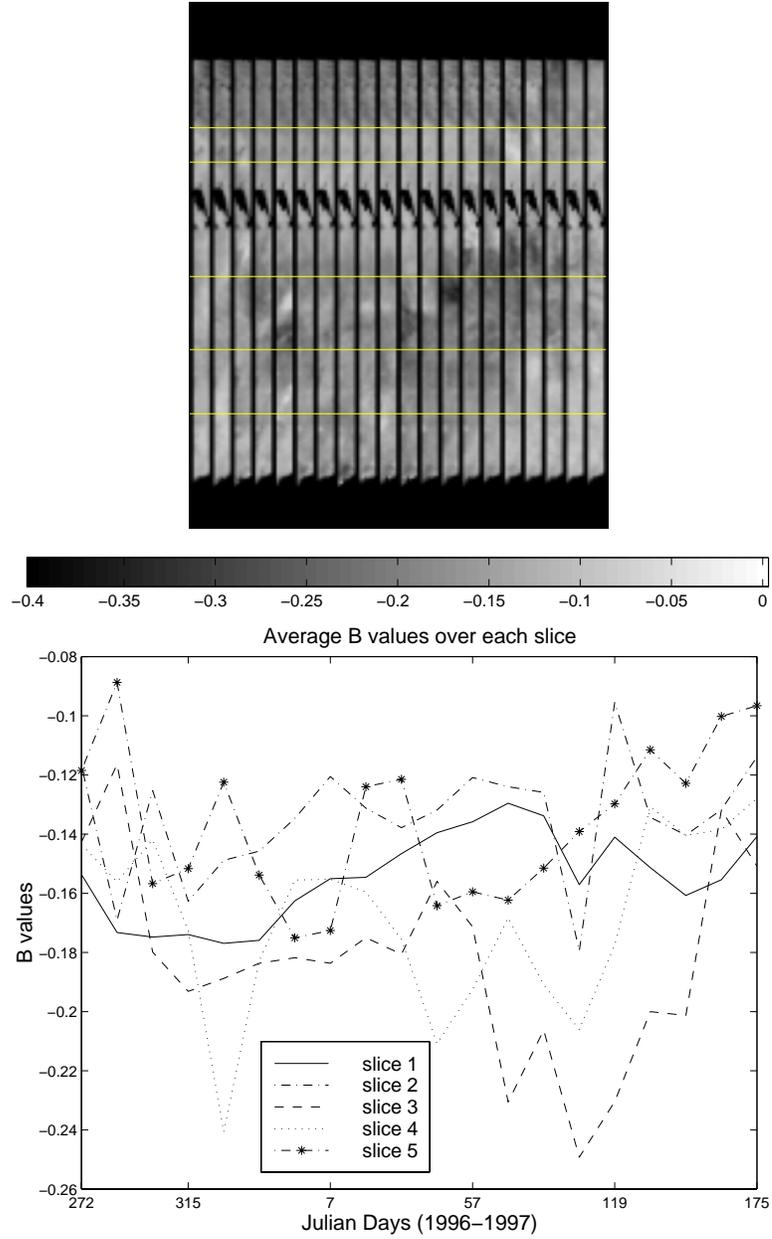


Figure 3.4: Time series strips of B images over a part of the Midwest region of North America for NSCAT-V and the corresponding average values over each slice shown with slice one at the top.

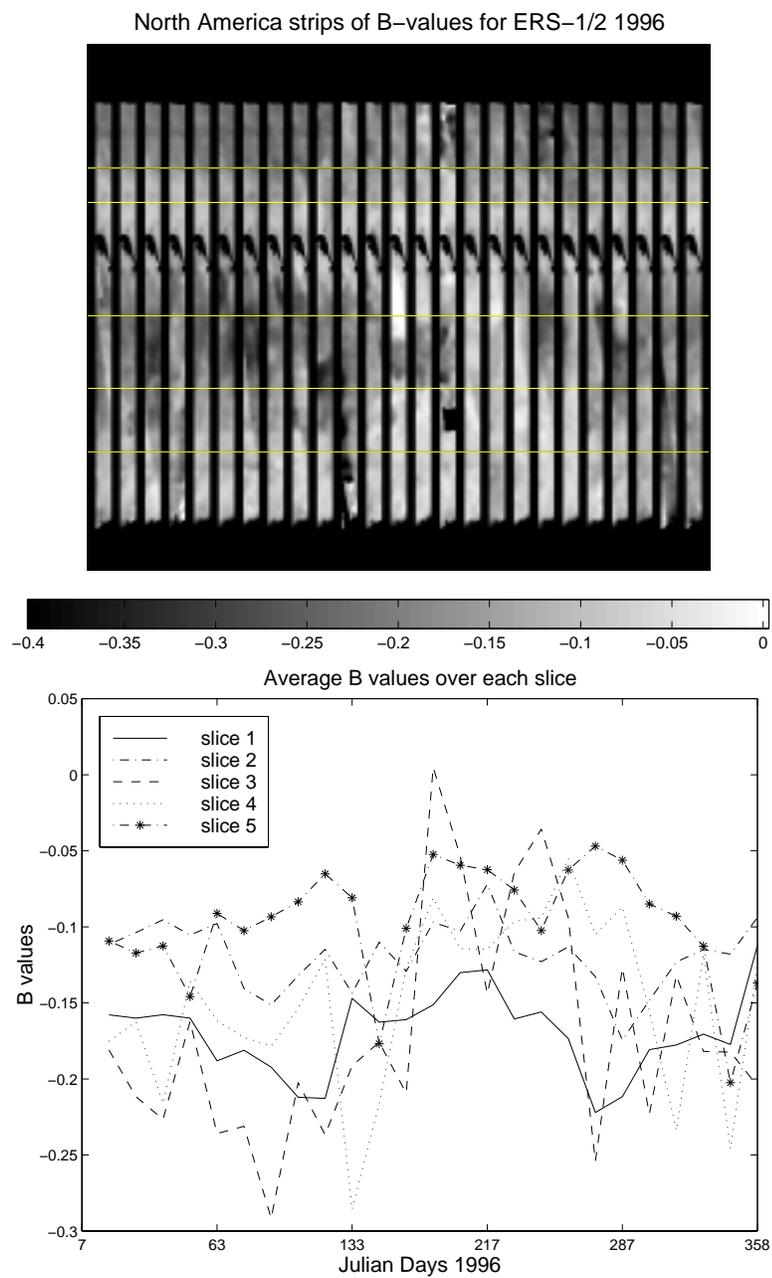


Figure 3.5: Time series strips of B images over a part of the Midwest region of North America for ERS-1/2 in 1996 and the corresponding average values over each slice shown with slice one at the top.

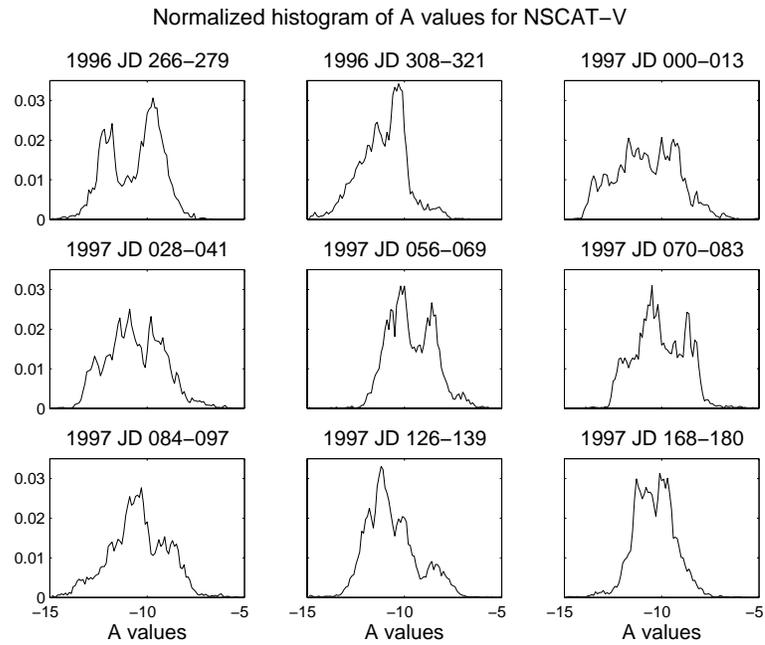


Figure 3.6: Normalized histogram plots of the strips at various times throughout the year for NSCAT-V  $\mathcal{A}$  values.

larger distribution of higher  $\mathcal{A}$  values over the summer months than over the winter at C-band and a higher distribution of  $\mathcal{A}$  values over the winter months for Ku-band. This suggests that both C-band and Ku-band are sensitive to  $\sigma^o$  temporal changes in vegetation but are seeing different seasonal changes over the same area.

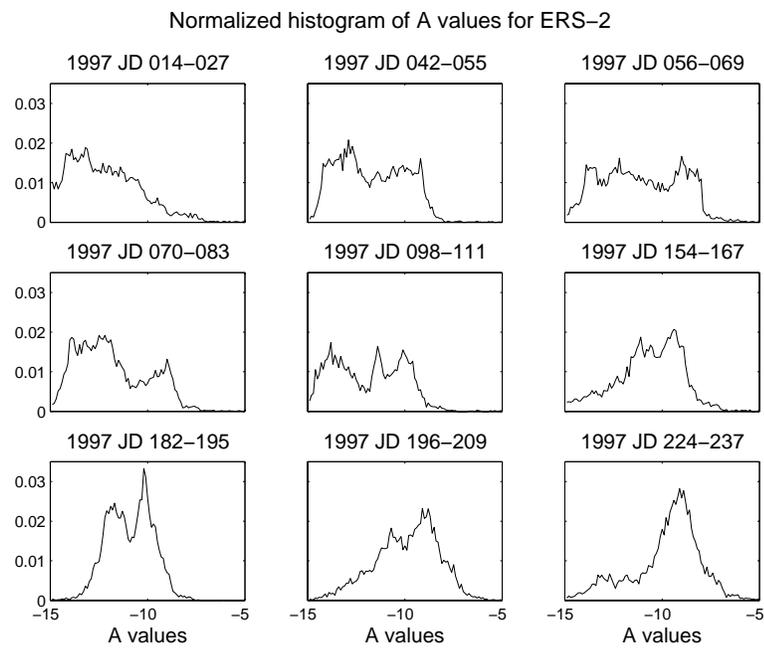
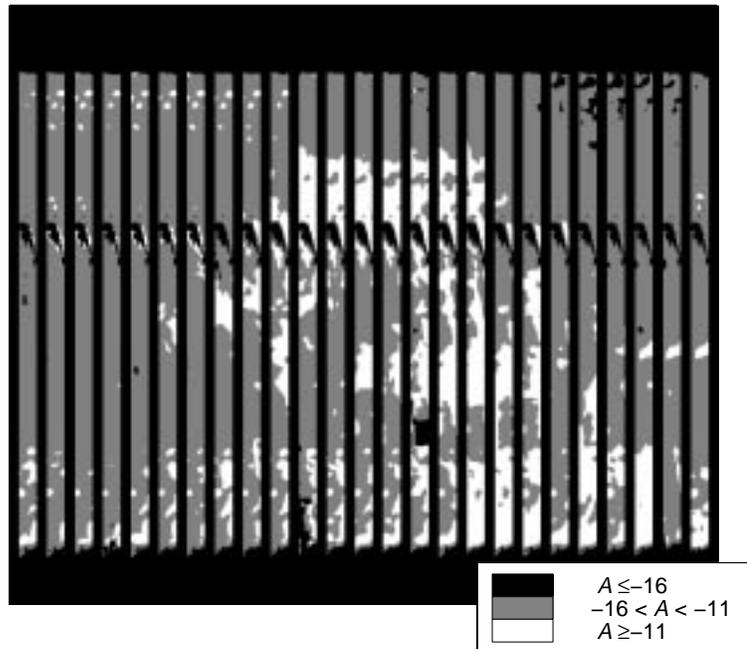


Figure 3.7: Normalized histogram plots of the strips at various times throughout the year for ERS-1/2  $\mathcal{A}$  values.

ERS-2 1996 thresholded  $A$  temporal image



NSCAT-V thresholded  $A$  temporal image

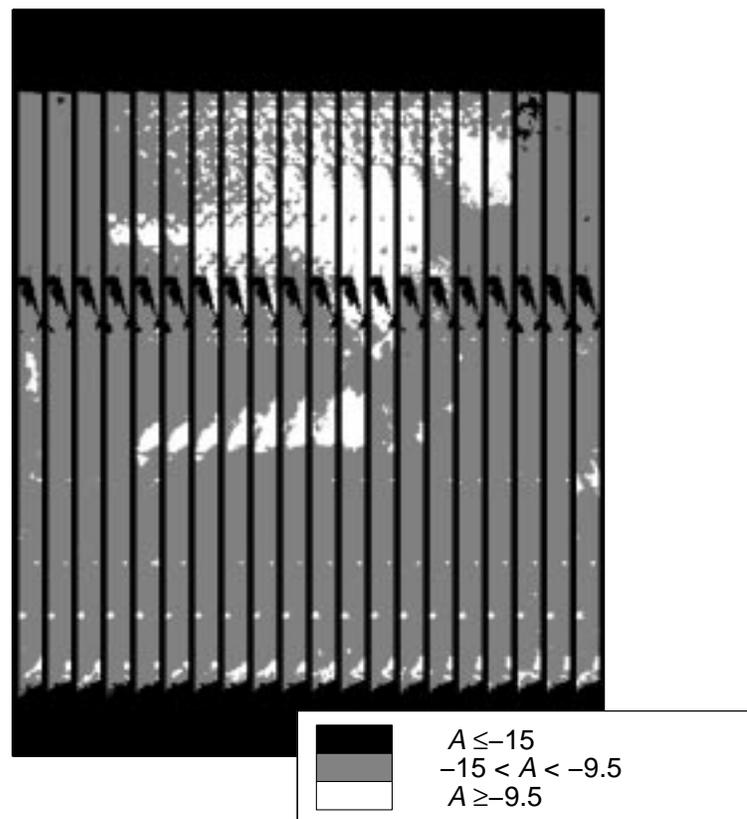


Figure 3.8: Thresholded temporal images of  $A$ .

### 3.2 $\sigma^o$ versus Precipitation and Temperature

Precipitation and temperature are obviously big factors in determining the distribution of vegetation. The following analysis is to determine the correlation between precipitation and temperature and the seasonal response over a region [19].

Annual time series of 14-day  $\mathcal{A}$  and  $\mathcal{B}$  images were created for North America.  $\mathcal{A}$  and  $\mathcal{B}$  values for a small region of the Midwest were extracted from each 14-day image and averaged. Plots of these values over several years are presented in Figs. 3.9 and 3.10. From these plots, we can see the seasonal variation of  $\sigma^o$  over a given region. The  $\mathcal{A}$  values tend to be higher during the summer months than during the winter months. The 4 dB summer/winter change is attributed to freezing and thawing of surface moisture and to vegetation growth in spring/summer.  $\mathcal{B}$  values also show a similar annual cycle. NSCAT  $\mathcal{A}$  values show less annual variation though the  $\mathcal{B}$  value range is similar.

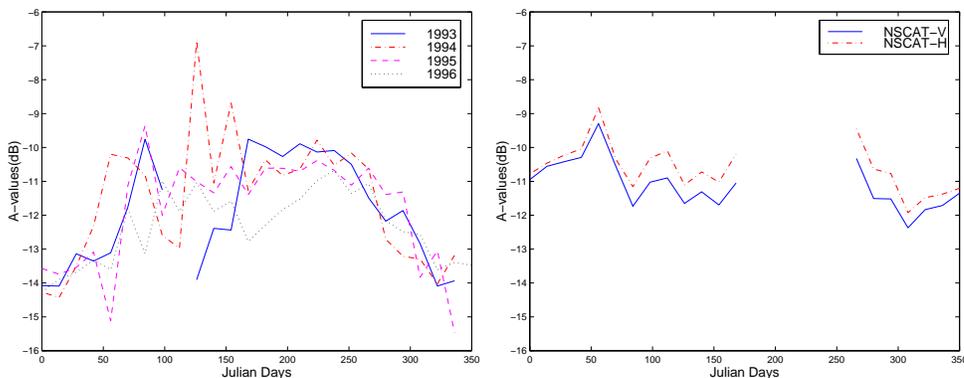


Figure 3.9:  $\mathcal{A}$  values versus time for ERS-1/2 (left) and NSCAT (right) over the small test region.

Seasonal temperature and precipitation data was then obtained from the National Climate Data Center Web Site [20] over the small study region and the  $\sigma^o$  values were plotted against them. Figure 3.11 shows this correlation between temperature and precipitation and the  $\mathcal{A}$  and  $\mathcal{B}$  values for ERS-1/2 over the study region.

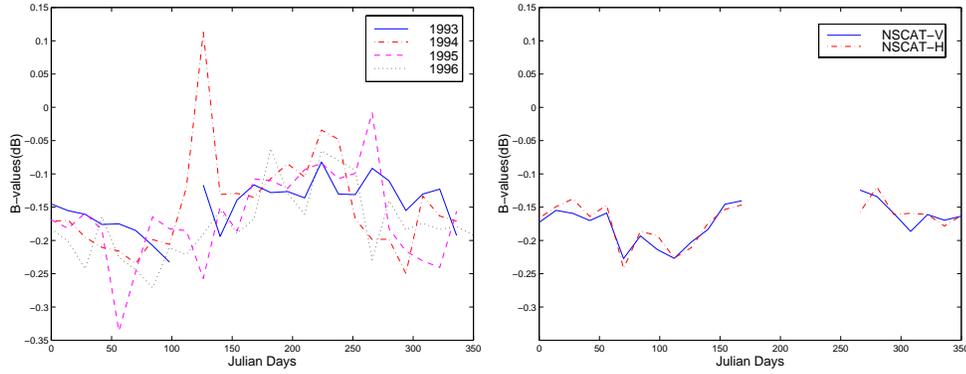


Figure 3.10:  $\mathcal{B}$  values versus time for ERS-1/2 (left) and NSCAT (right) over the small test region.

The correlation coefficient,  $\rho$ , was calculated for both cases. For temperature versus  $\mathcal{A}$  values  $\rho = 0.73$ , and for  $\mathcal{B}$  values  $\rho = 0.55$ , suggesting that there is a slight correlation between temperature and the  $\mathcal{A}$  value. For precipitation  $\rho = 0.316$  vs.  $\mathcal{A}$  values and  $\rho = 0.255$  versus  $\mathcal{B}$  values. This suggests that no real correlation exists between precipitation and  $\sigma^o$ . This may be due to the fact that precipitation is often very localized and non-uniform over a given region. Because of this, the  $\sigma^o$  values tend to be less correlated with precipitation than temperature which is more uniform over large areas.

The correlation also between precipitation and temperature and  $\sigma^o$  seasonal values from NSCAT are also shown. Figure 3.12 shows the correlation between temperature and  $\mathcal{A}$  values for NSCAT. For vertical polarization  $\rho = -0.43$  and for horizontal  $\rho = -0.10$ . This suggests that there is little relation between temperature and Ku-band  $\sigma^o$  values. For precipitation, vertical polarization  $\rho = -0.352$  while for horizontal polarization  $\rho = -0.126$ . The correlation coefficient for  $\mathcal{B}$  values in the Ku-band were for vertical polarization  $\rho = -0.054$  and for horizontal  $\rho = -0.149$ .

The correlation between C-band ERS-1/2 data and Ku-band NSCAT data is plotted in Fig. 3.13 and the correlation coefficients are  $\rho = 0.104$  for vertical polarization and  $\rho = 0.336$  for horizontal polarization. In looking at the plots in

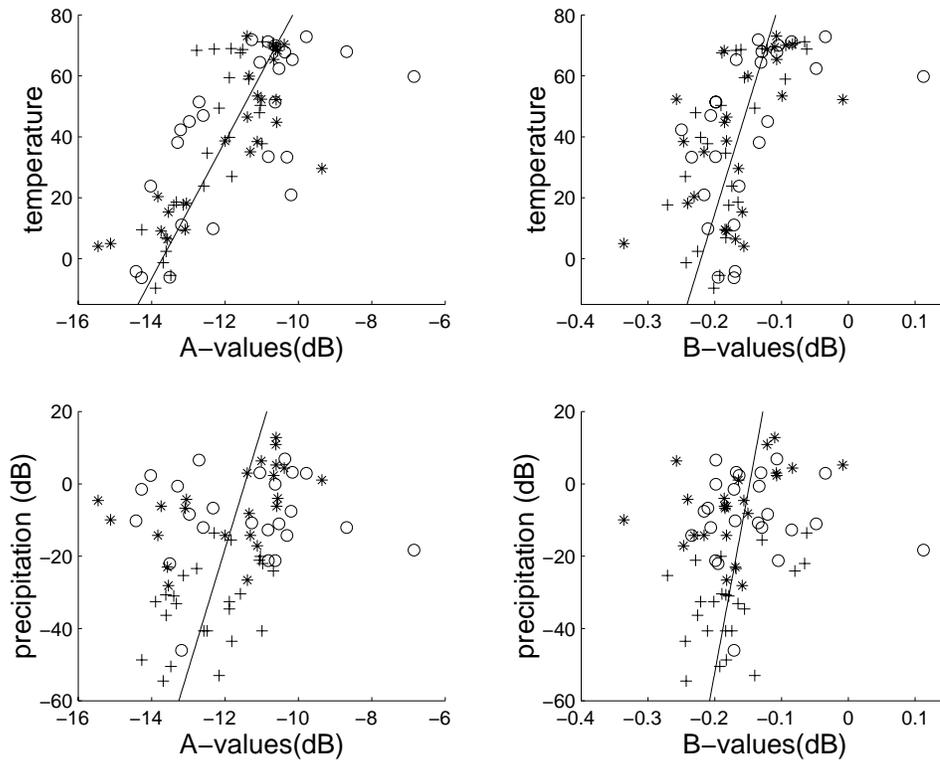


Figure 3.11: C-band  $\sigma^o$   $\mathcal{A}$  values and  $\mathcal{B}$  values vs. temperature. o-1994 \*-1995 +-1996.

correlation between C-band and Ku-band, there appears that there is an offset between the v-pol Ku-band and the h-pol Ku-band. This is seen in Fig. 3.14.

The proceeding data suggests that there is a stronger correlation between temperature and precipitation for C-band  $\sigma^o$  than for Ku-band in this particular region. There also appears to be little correlation between the two bands. The Ku-band result could possibly be improved with more data points, but due to the short time that NSCAT was operational only a portion of the annual cycle could be observed.

Note that this analysis was done over a particular region only and therefore the results could vary over different areas and vegetation classes. It should be noted that temperature and precipitation in many vegetation areas may affect the  $\sigma^o$  response in those areas and were not addressed in the above analysis.

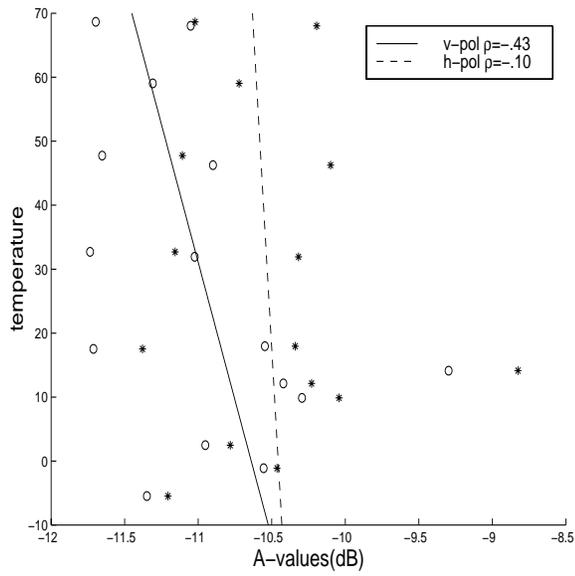


Figure 3.12: Ku-band  $\sigma^o$   $\mathcal{A}$  values for both v-pol(o) and h-pol(\*) vs. temperature.

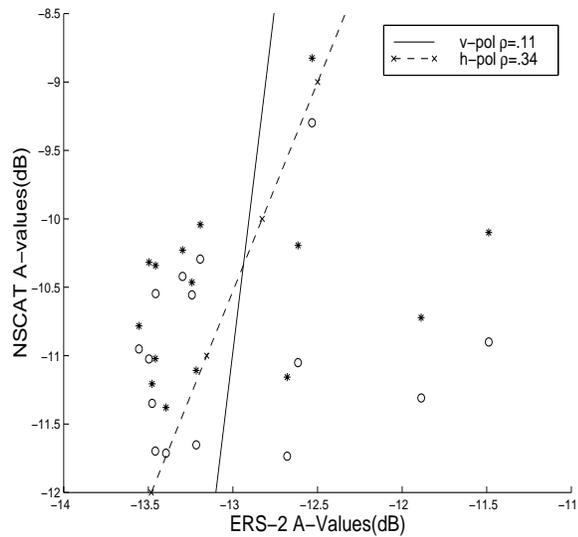


Figure 3.13: Ku-band v-pol(o) and h-pol(\*) vs. C-band  $\mathcal{A}$  during the NSCAT mission.

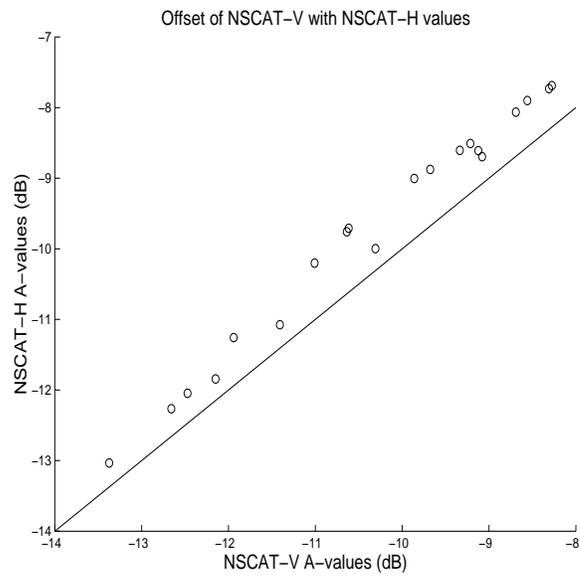


Figure 3.14: NSCAT-H  $\mathcal{A}$  values vs. NSCAT-V  $\mathcal{A}$  values over the study region showing the offset of  $\mathcal{A}$  values for horizontal polarization with vertical polarization.

### 3.3 $\sigma^o$ Variation over Vegetation Regions

To observe the backscatter coefficient,  $\sigma^o$ , over different vegetation regions, six vegetation types from South America and eight vegetation types from North America were chosen. These vegetation types, or classes, are chosen from two main vegetation maps. One is the Matthews Global Vegetation, Land Use, and Seasonal Albedo data set from the National Oceanic and Atmospheric Administration (NOAA) [21] and the other is a vegetation map derived from AVHRR data created by the Geography department at the University of Maryland [8]. One of the reasons that these two maps were chosen was due to the different characteristics of each one. The Matthews map, although much lower in resolution, is derived from around 100 published sources complimented by satellite data, while the AVHRR map, although high in resolution, is derived only from a single satellite data set. Figure 3.15 shows the Matthews world vegetation map and Fig. 3.16 shows the AVHRR-derived world vegetation map.

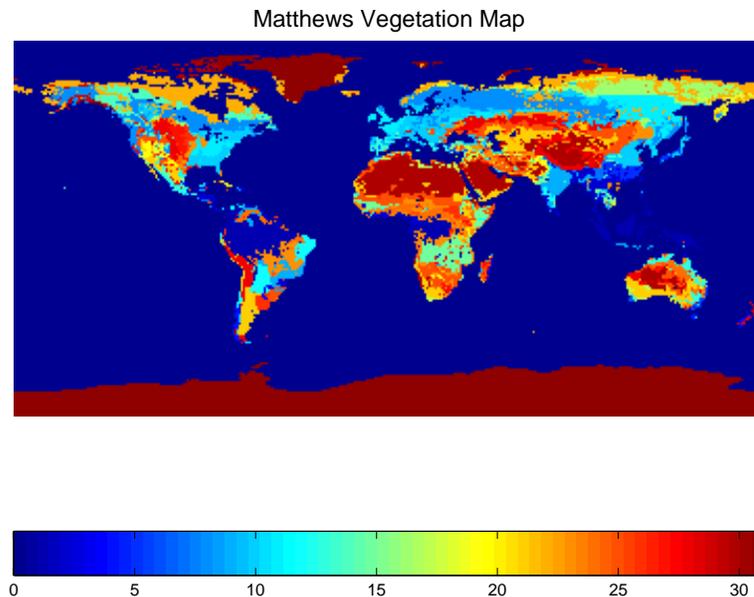


Figure 3.15: The Matthews Vegetation map of the world showing all 32 vegetation classes.

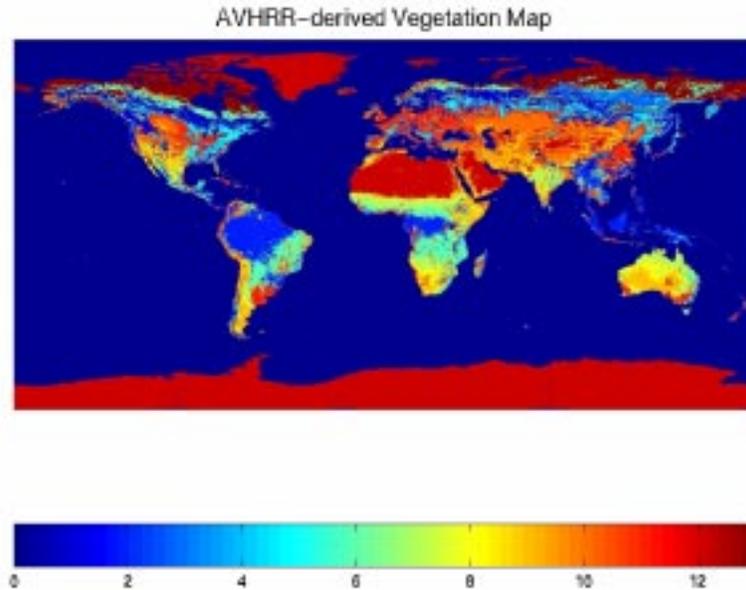


Figure 3.16: The AVHRR-derived Vegetation map of the world showing its thirteen vegetation classes.

In comparing these two maps two observations should be noted. One is the large difference in definition and number of vegetation classes and two, the difference in pixel resolution between them and between the  $\mathcal{A}$  and  $\mathcal{B}$  images. Also, the projections differ between the SIR images and the vegetation maps. Both vegetation maps are in a lat-lon projection while the SIR images are in a Lambertian projection. The next section discusses how these problems are addressed.

The first problem considered is creating maps with the same projection and pixel resolution. This allows easy comparisons between the vegetation maps and the SIR images. For each pixel in the SIR image, the corresponding latitude and longitude in the vegetation image is found and the corresponding vegetation class is determined for that pixel. This produces a vegetation map with the same resolution and projection as the SIR images. Figures 3.17 and 3.18 show the new results for both North and South America for the Matthews and the AVHRR-based maps.

To overcome the differences in the classes of each vegetation map new vegetation classes are defined by unifying the classes in the vegetation maps. The Matthews vegetation map has 32 separate categories for vegetation throughout the world as seen in Table 3.1, while the AVHRR-derived map has only thirteen categories for vegetation as listed in Table 3.2. In combining the two vegetation maps care was taken to preserve similar vegetation descriptions together. No attempt to combine the vegetation classes over the entire world was made, only the vegetation classes over the study regions of North and South America. Each region is separately considered. Table 3.3 shows the resulting classes for North America and South America and how they were combined. Only the most prominent vegetation classes of each region are used.

Table 3.4 and 3.5 shows the combined vegetation classes over both continents as well as the percentage of coverage for each vegetation over the study area. Figs. 3.19 and 3.20 show the corresponding combined vegetation maps for both North and South America.

Table 3.1: The 32 different vegetation classes found in the Matthews vegetation map.

MATTHEWS VEGETATION CLASSES			
Veg Class	Description	N. Amer	S. Amer
1	Tropical evergreen rainforest, mangrove forest		✓
2	Tropical/subtropical evergreen seasonal broadleaved forest		✓
3	Subtropical evergreen rainforest		✓
4	Temperate/subpolar evergreen rainforest		✓
5	Temperate evergreen seasonal broadleaved forest, summer rain		
6	Evergreen broadleaved sclerophyllous forest, winter rain	✓	
7	Tropical/subtropical evergreen needleleaved forests		
8	Temperate/subpolar evergreen needleleaved forest	✓	
9	Tropical/subtropical drought-deciduous forest		✓
10	Cold-deciduous forest, with evergreens	✓	✓
11	Cold-deciduous forest, without evergreens	✓	
12	Xeromorphic forest/woodlands	✓	✓
13	Evergreen broadleaved sclerophyllous woodland	✓	
14	Evergreen needleleaved woodland	✓	
15	Tropical/subtropical drought-deciduous woodland	✓	✓
16	Cold-deciduous woodland	✓	
17	Evergreen broadleaved shrubland/thicket, evergreen dwarf-shrubland	✓	
18	Evergreen needleleaved or microphyllous shrubland/thicket	✓	
19	Drought-deciduous shrubland thicket	✓	
20	Cold-deciduous subalpine/subpolar shrubland, cold-deciduous dwarf shrubland		
21	Xeromorphic shrubland/dwarf shrubland	✓	✓
22	Arctic/alpine tundra, mossy bog	✓	
23	Tall/medium/short grassland with 10-40% woody tree cover	✓	✓
24	Tall/medium/short grassland with < 10% woody tree cover or tuft-plant cover	✓	✓
25	Tall/medium/short/grassland with shrub cover	✓	✓
26	Tall,grassland, no woody cover	✓	✓
27	Medium grassland, no woody cover	✓	
28	Meadow, short grassland, no woody cover	✓	✓
29	Forb formations		
30	Desert	✓	✓
31	Ice		
32	Cultivation		

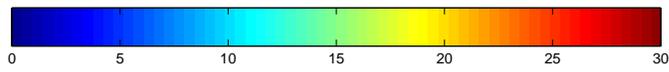
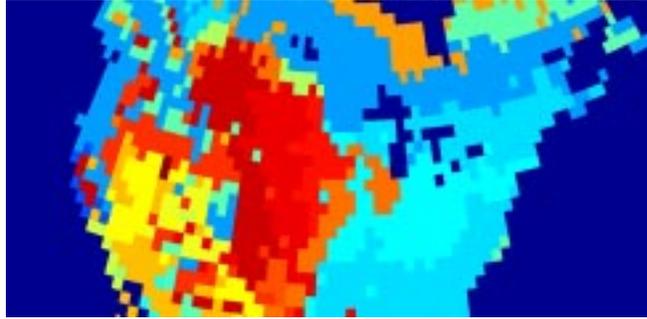
Table 3.2: The 16 different vegetation classes from the AVHRR-derived vegetation map.

AVHRR-DERIVED VEGETATION CLASSES			
Veg Class	Description	N. Amer	S. Amer
1	Evergreen Needleleaf Forests	✓	✓
2	Evergreen Broadleaf forests	✓	✓
3	Deciduous Needleleaf forests		
4	Deciduous Broadleaf forests	✓	✓
5	Mixed forests	✓	✓
6	Woodlands	✓	✓
7	Wooded grassland/shrub	✓	✓
8	Closed Bushlands or Shrublands	✓	✓
9	Open Shrublands	✓	✓
10	Grasses	✓	✓
11	Croplands	✓	✓
12	Bare	✓	✓
13	Mosses and Lichens	✓	✓

Table 3.3: New vegetation classes and how they were combined from the Matthews and AVHRR-based maps.

Combined Classes over North America		
Combined Class	Matthews Class	AVHRR-based Class
1	14,8	1
2	11	4
3	10	5
4	19,21	8
5	25	9
6	26,27,28	10,11
7	23	6,7
8	22	13
Combined Classes over South America		
Combined Class	Matthews Class	AVHRR-based Class
1	1,2,3,4	2
2	25	9
3	9	4,5
4	26,28	10,11
5	21	8
6	23,12	6

Matthews Vegetation Map over North America



AVHRR-based Vegetation Map over North America

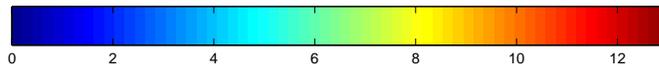
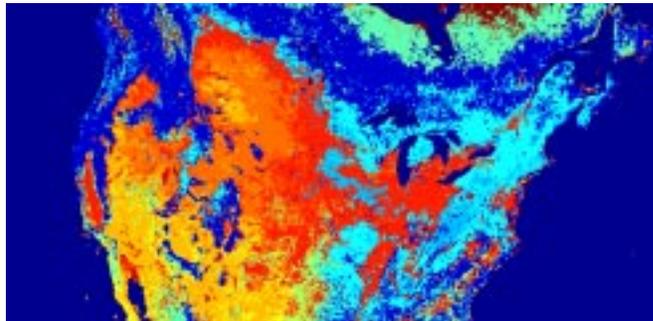
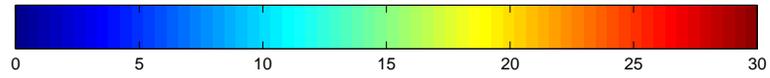
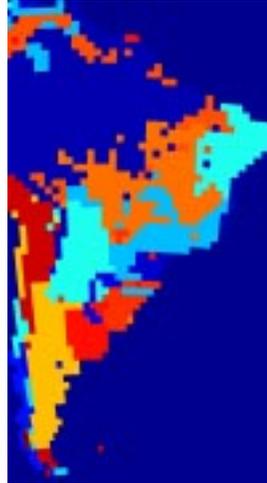


Figure 3.17: The vegetation maps over North America showing the new projection and pixel resolution.

Matthews Vegetation Map over South America



AVHRR-based Vegetation Map over South America

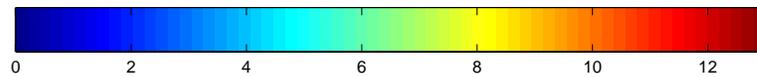
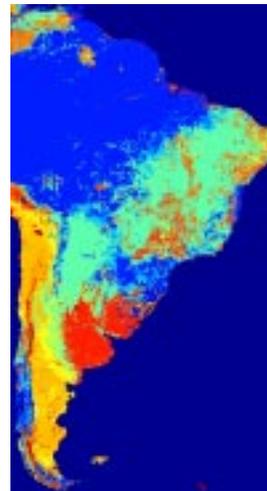


Figure 3.18: The vegetation maps over South America showing the new projection and pixel resolution.

Table 3.4: Vegetation classes over North America

Veg Class	Description	%Pixels Matthews	%Pixels AVHRR
1	Evergreen needleleaved forests	29.8	28.4
2	Deciduous forests with evergreens/mixed forests	17.7	14.8
3	Deciduous forests	10.3	4.9
4	Xeromorphic shrublands/Shrubland/Closed Bushlands	9	2
5	Grassland with shrub cover/Open shrublands	5.6	8.1
6	Woodlands/grasslands with woody cover	5.3	9.1
7	Tall, medium, short grasslands, meadows/Croplands	19	31.3
8	Arctic alpine tundra, mossy bog/Mosses and Lichens	3.2	1.4

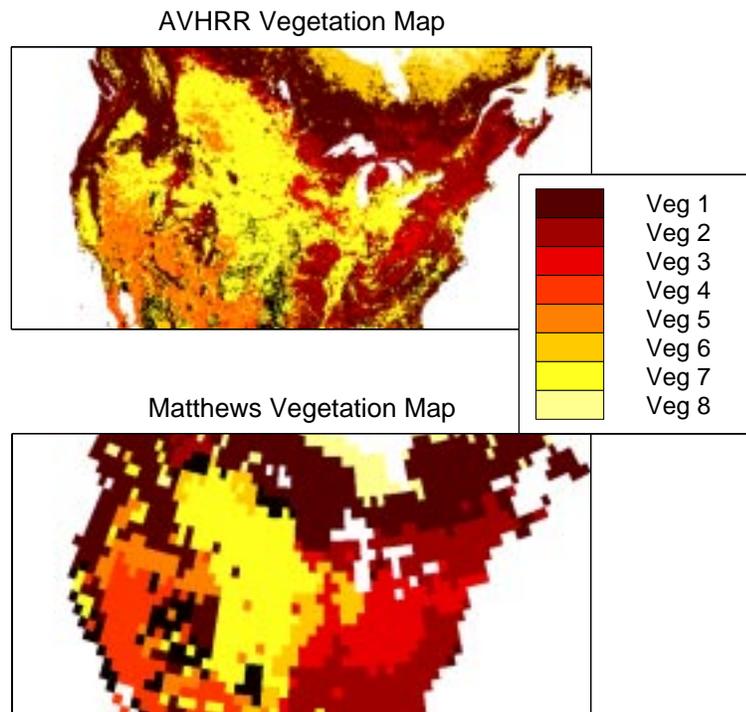


Figure 3.19: The unified vegetation classes for Matthews and AVHRR maps over North America.

Table 3.5: Vegetation classes over South America

Veg Class	Description	% Pixels Matthews	% Pixels AVHRR
1	Tropical Rainforests/Evergreen Broadleaf forests Evergreen Rainforests	40.3	42.7
2	Deciduous broadleaf forests/mixed forests Tropical/subtropical drought deciduous forests	10.2	5.8
3	Xeromorphic shrubland/shrublands/closed bushlands	9.6	3.2
4	Open shrublands/grassland with shrub cover	2.6	7
5	Xeromorphic forest/ woodlands/wooded grasslands grassland with 10-40% woody tree cover	27.7	25.4
6	Meadows/short, medium, tall grasslands	9.6	16

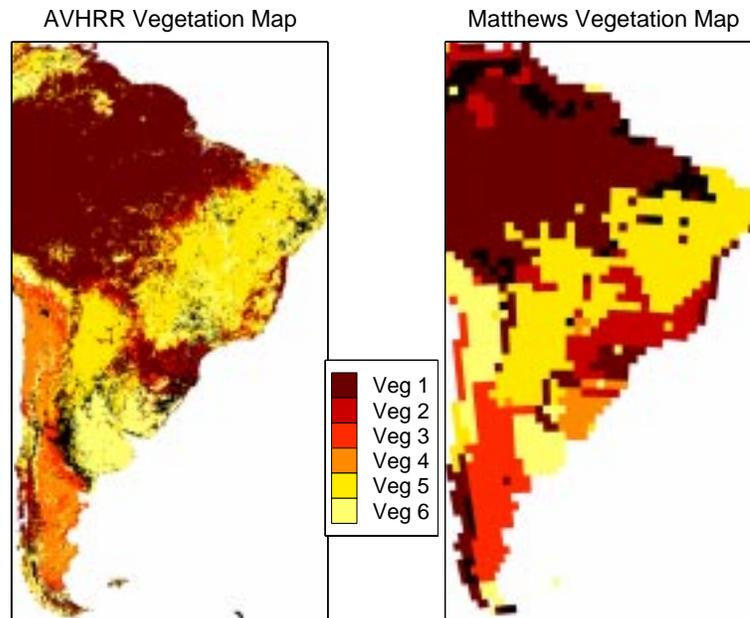


Figure 3.20: The unified vegetation classes for Matthews and AVHRR maps over South America.

### 3.4 Seasonal Response of $\sigma^o$

For each vegetation region, an average value of the  $\sigma^o$  response for each vegetation class is calculated and plotted over time for  $\mathcal{A}$  values at both C and Ku bands and  $\mathcal{B}$  values at Ku-band. Due to noise in the C-band  $\mathcal{B}$  values, they are not shown. These plots can be seen in Figs.3.21 through 3.23 for North America and Figs. 3.24 through 3.26 for South America. The response of ERS-1/2 is plotted over the same time frame as NSCAT-V so that they can easily be compared. Figures 3.27 and 3.28 show year-to-year seasonal changes over the same vegetation classes for ERS-1/2 over North and South America. These show that the yearly  $\sigma^o$  response over each vegetation region exhibits little interannual variability.

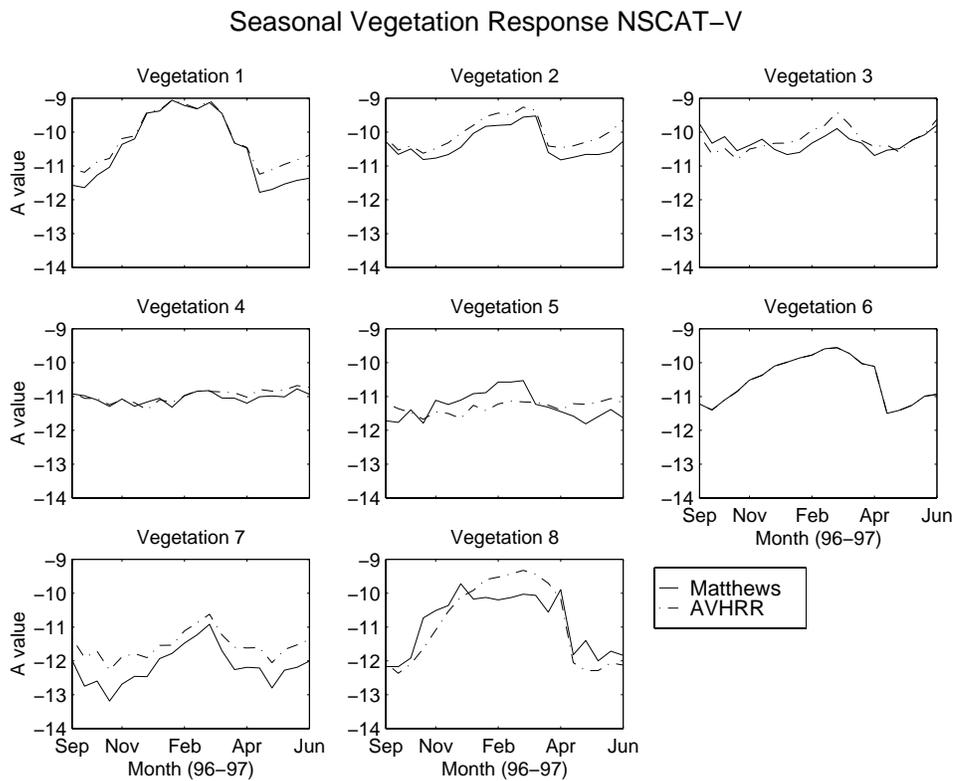


Figure 3.21: Seasonal response of  $\mathcal{A}$ -values over each vegetation region of North America for NSCAT-V.

### Seasonal Vegetation Response ERS-2

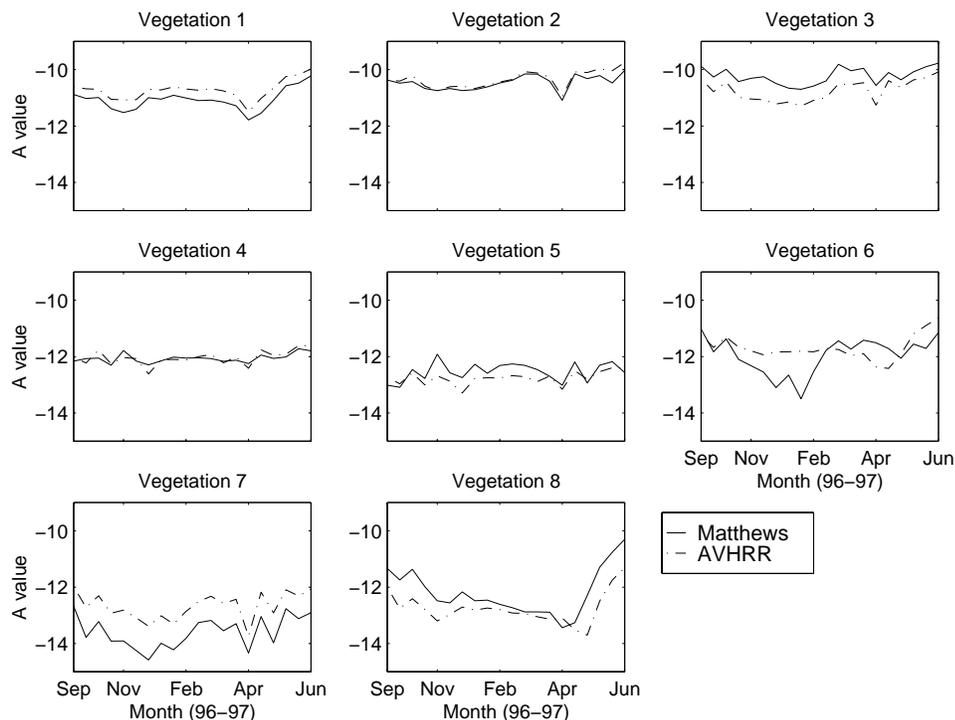


Figure 3.22: Seasonal response of  $\mathcal{A}$ -values over each vegetation region of North America for ERS-2.

Notice that for each vegetation region a different response is seen. This indicates that different vegetation classes have different seasonal  $\sigma^o$  responses. The seasonal responses also differs from C-band to Ku-band. The seasonal differences seems are much larger for the  $\mathcal{A}$ -values than for the  $\mathcal{B}$ -values. This difference in seasonal response from vegetation class to vegetation class gives motivation in creating an algorithm that takes advantage of these differences to classify vegetation areas. It also gives motivation in combining the responses from both C-band and Ku-band to try and better classify the vegetation regions. The next chapter discusses the steps of an algorithm that will take advantage of the differences seen.

In examining the seasonal response of  $\sigma^o$  over vegetation, this analysis ignores the variance of the seasonal response and assumes that a similar seasonal response is seen over the entire area covered by a given vegetation class. This analysis neglects

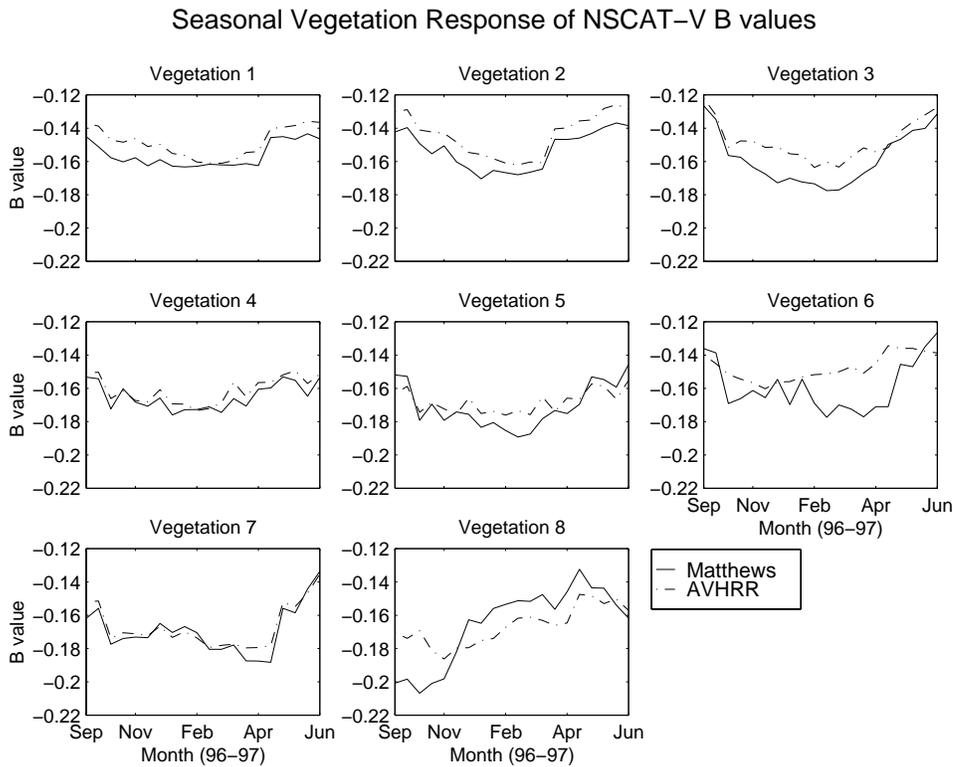


Figure 3.23: Seasonal response of  $B$ -values over each vegetation region of North America for NSCAT-V.

spatial variations in the  $\sigma^o$  response of a given vegetation class. This is seen as a limitation of this initial study.

In particular, snow and freeze/thaw cycles exhibit spatial variance with latitude and topography. No attempt is made in this study to address this issue. There is clearly correlation between vegetation types and climate conditions (i.e., winter snow and freezing) and topography. Snow cover, especially in the northern latitudes, may have strong correlation with  $\sigma^o$ . In particular, snow may dominate the seasonal  $\sigma^o$  response for all classes and increase the similarity between responses of different classes, thus increasing confusion between the classes later in the vegetation classification. In effect, this temporal variation can reduce the classification of vegetation due to inhomogeneities in climate and topography of a given class.

### Seasonal Vegetation Response NSCAT-V for South America

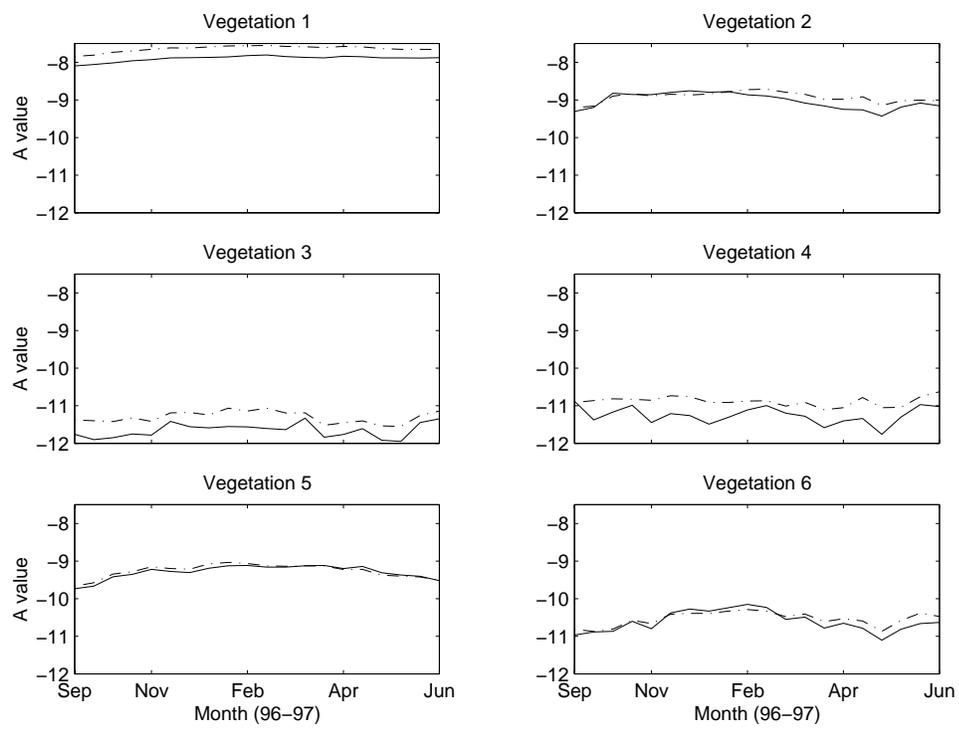


Figure 3.24: Seasonal response of  $\mathcal{A}$ -values over each vegetation region of South America for NSCAT-V.

### Seasonal Vegetation Response ERS-2 for South America

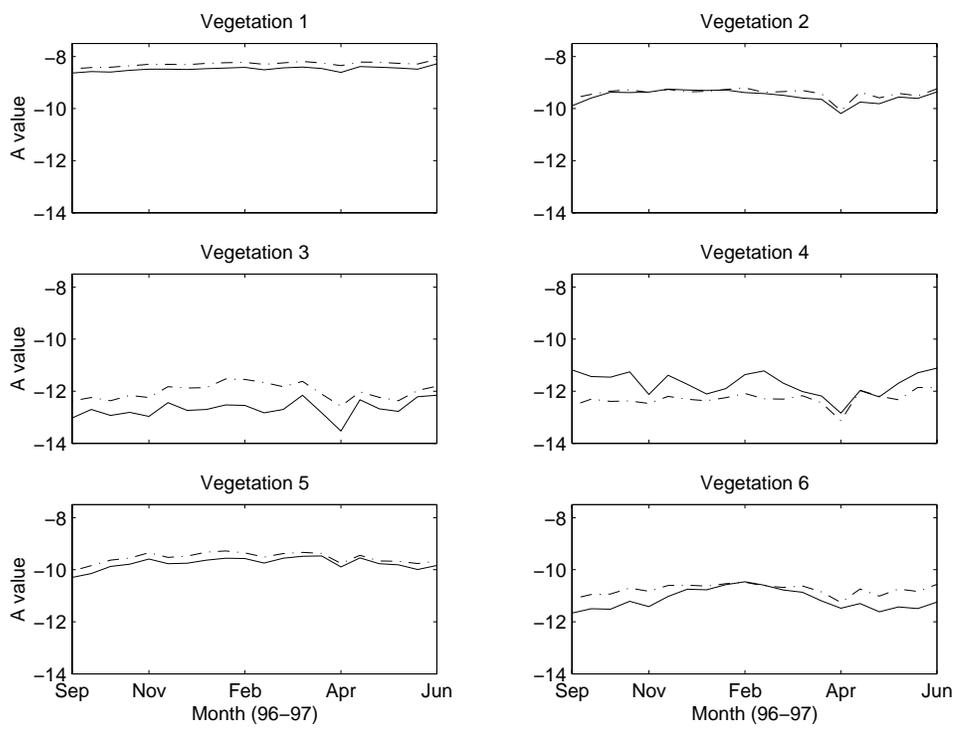


Figure 3.25: Seasonal response of  $\mathcal{A}$ -values over each vegetation region of South America for ERS-2.

### Seasonal Vegetation Response NSCAT-V B values for South America

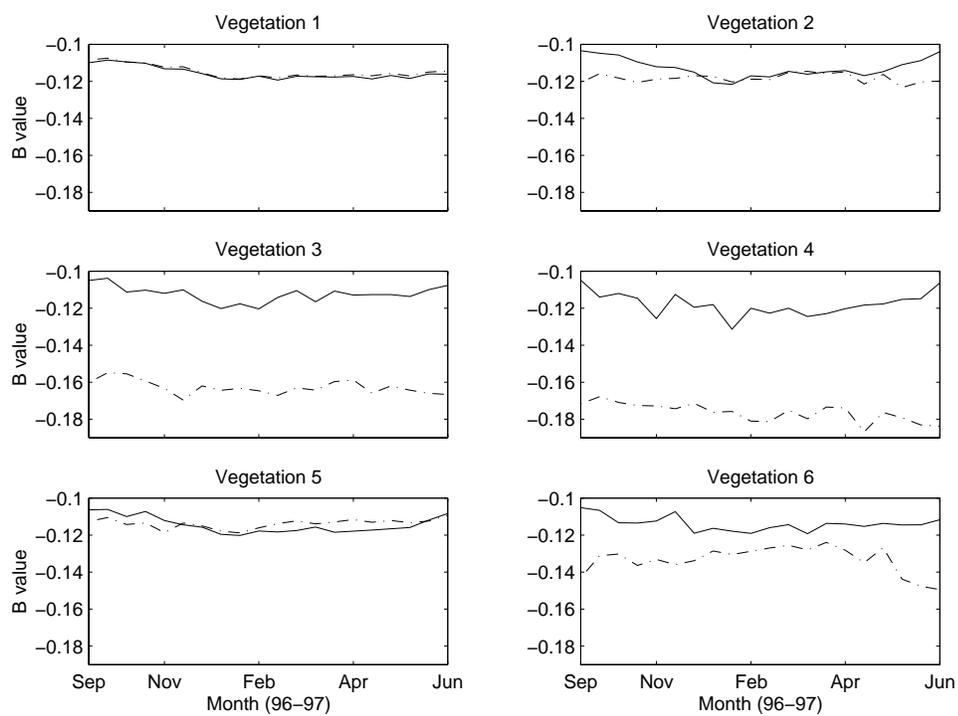
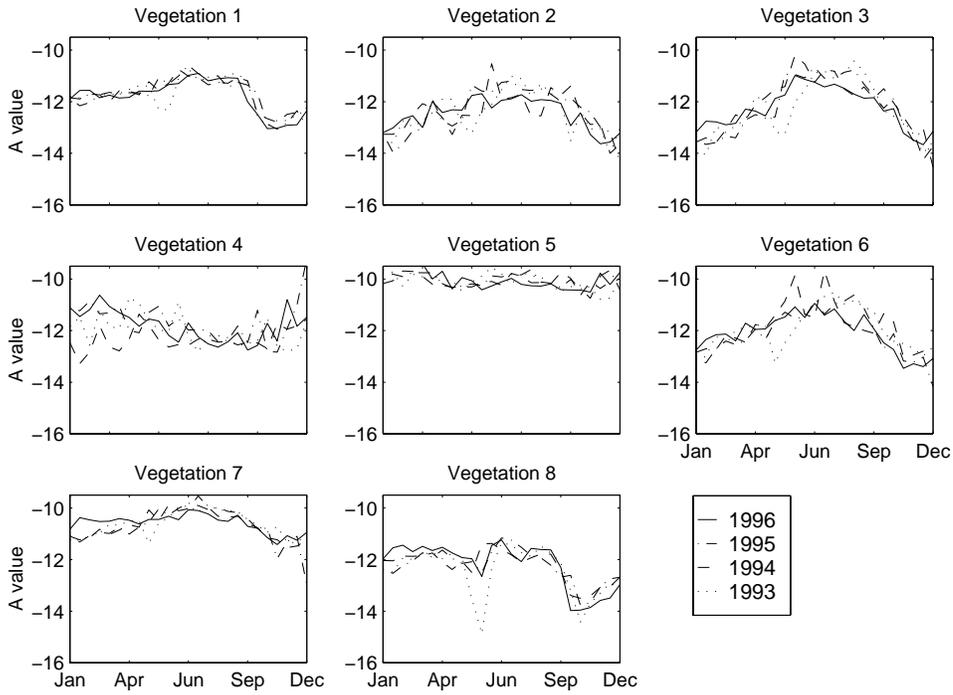


Figure 3.26: Seasonal response of  $\mathcal{B}$ -values over each vegetation region of South America for NSCAT-V.

### Vegetation Response of ERS-2 using Matthews



### Vegetation Response of ERS-2 using AVHRR

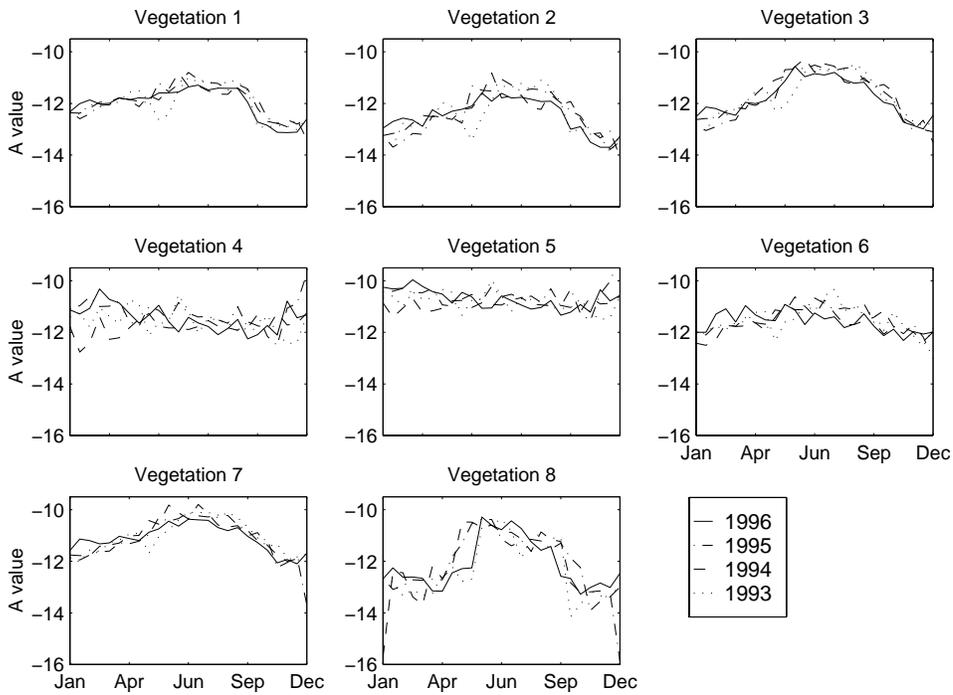
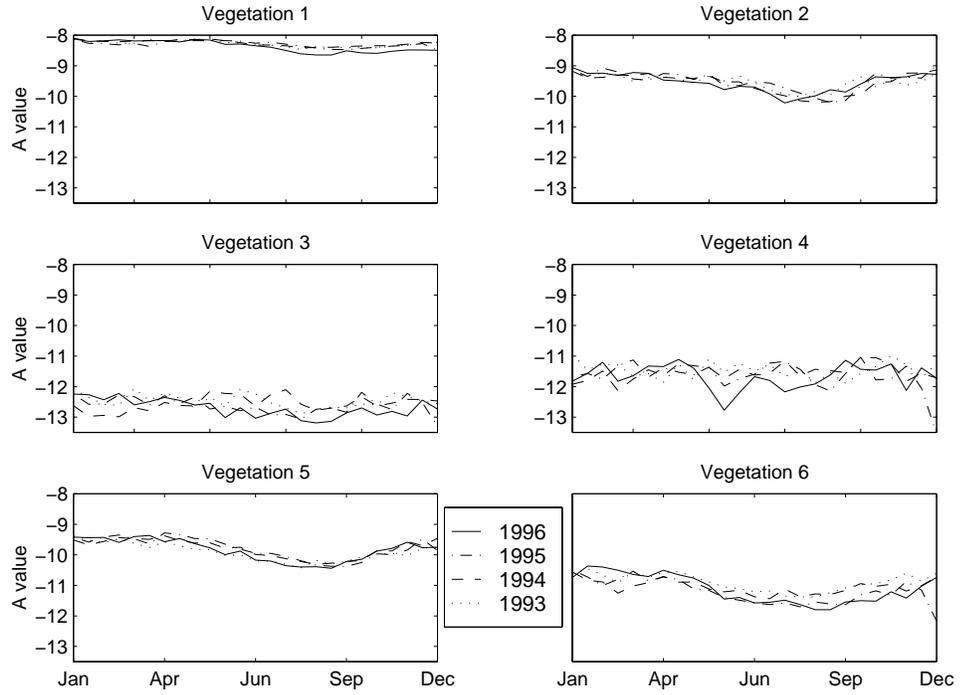


Figure 3.27: Yearly seasonal response of  $\mathcal{A}$ -values over each vegetation region of North America for ERS-1/2.

### Vegetation Response of ERS-2 using Matthews



### Vegetation Response of ERS-2 using AVHRR

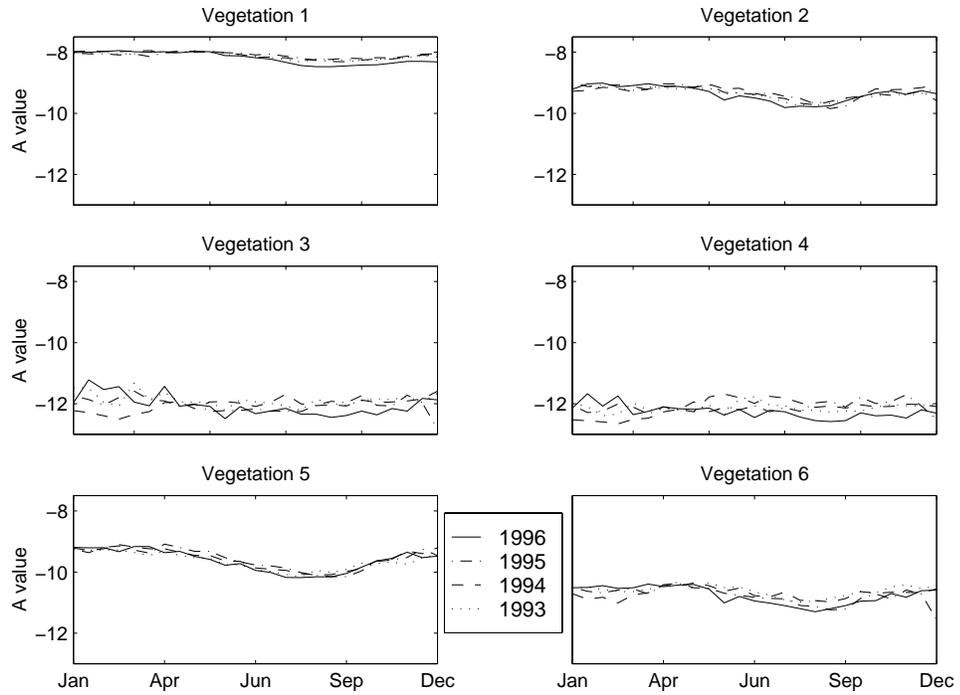


Figure 3.28: Yearly seasonal response of  $\mathcal{A}$ -values over each vegetation region of South America for ERS-1/2.



## Chapter 4

### Vegetation Classification Algorithm

In this chapter the steps of the vegetation classification algorithm are described. Each step and any assumptions and methods used are explained. The algorithm uses the images created by the SIRF algorithm in the creation of the vegetation classification maps. Singular Value Decomposition (SVD) as explained in the background chapter is used in the algorithm.

#### 4.1 The Steps

The steps of the algorithm follow. There are two divisions to the algorithm. The first part is training the algorithm; the second is the actual classification. The steps to the algorithm are:

1. Training the Algorithm

- Select a training set.
  - Select a data set.
  - Choose vegetation classes according to a selected vegetation map.
  - Choose the length of the desired seasonal response.
  - Using the vegetation map, find the average seasonal response of the selected data set over each chosen vegetation class. This measured response is the training set.
- Find a basis set using SVD on the training set.

- Create a matrix, where the columns of the matrix are the vectors of the training set and correspond to the seasonal response of the vegetation regions.
- Compute the SVD of the training set matrix.
- Select the singular values containing most of the variance.
- Take the corresponding eigenvectors as the new basis set.
- Project the training vectors onto the basis to find the coefficients. These coefficients become the centroids of each vegetation class.

## 2. Classification

- Project the seasonal response of each pixel onto the new basis.
- Using a distance classifier, compute the distance from each pixel to each centroid.
- Assign to each pixel the vegetation class closest to it.
- Repeat for each pixel.

### 4.2 Selecting a Training Set

Two data sets are used in the algorithm. The one is NSCAT-V scatterometer data, the other is ERS-1/2 scatterometer data. Using these data sets,  $\mathcal{A}$  and  $\mathcal{B}$  images are separately created using the SIRF algorithm described in Chapter two. Two different vegetation maps are also used for the training set. They come from the vegetation set described in Chapter three and have six vegetation regions over South America and eight vegetation regions over North America, based on the Matthews and AVHRR-derived vegetation maps. Figures 3.19 and 3.20 in Chapter three show the vegetation classes. Tables 3.4 and 3.5 show the names of the vegetation classes. These vegetation classes are chosen to account for the majority of the vegetation over these two areas and to be consistent between the two maps. The black areas in the vegetation maps are areas of vegetation that are not described by the defined vegetation classes.

For each vegetation class described, the corresponding pixels in the  $\mathcal{A}$  and  $\mathcal{B}$  values of the 14-day images are extracted and a randomly selected tenth of these pixels are used to obtain an average value of the  $\sigma^o$  response for each class. The average value of each vegetation class is plotted over a ten month period for both  $\mathcal{A}$  and  $\mathcal{B}$  values and is shown in Figures 3.21 through 3.26 of Chapter three. The  $\mathcal{B}$ -values of ERS-1/2 are not used for vegetation classification due to the noisy nature of the data.

### 4.3 Finding a Basis Set

In finding a basis set we let

$$X = [X_1|X_2|\dots|X_N] \quad (4.1)$$

where  $N$  is the number of vegetation regions.  $X_1, X_2, \dots, X_N$  correspond to the seasonal response of each vegetation class discussed in Section 4.2. For each data set and vegetation map, there is different training matrix,  $X$ . For example, the vegetation response over each class for the Matthews vegetation map using NSCAT-V data forms the columns of a training matrix, while the NSCAT-V vegetation response for the AVHRR-derived map forms the columns of another training matrix. For each training matrix, the singular value decomposition (SVD) is then taken on  $X$  and a set of eigenvectors,  $Q_1$ , are chosen where  $U = [Q_1|Q_2]$ , and where the columns of  $U$  are the eigenvectors of  $XX'$ .

To determine the minimum number of eigenvectors for  $Q_1$  that produces acceptable results, the log of the eigenvalues from highest to lowest are plotted and are shown in Figure 4.1 for NSCAT-V using the Matthews vegetation map. As can be seen, the top two eigenvalues account for the most variation in the vegetation responses. Choosing only the top eigenvalues reduces the space while maintaining the most information.

Figure 4.2 shows the corresponding eigenvectors for the eigenvalues. Notice that the first eigenvector, although it has some variations, is mainly a DC response. The other vectors represent different seasonal variations in the vegetation responses.

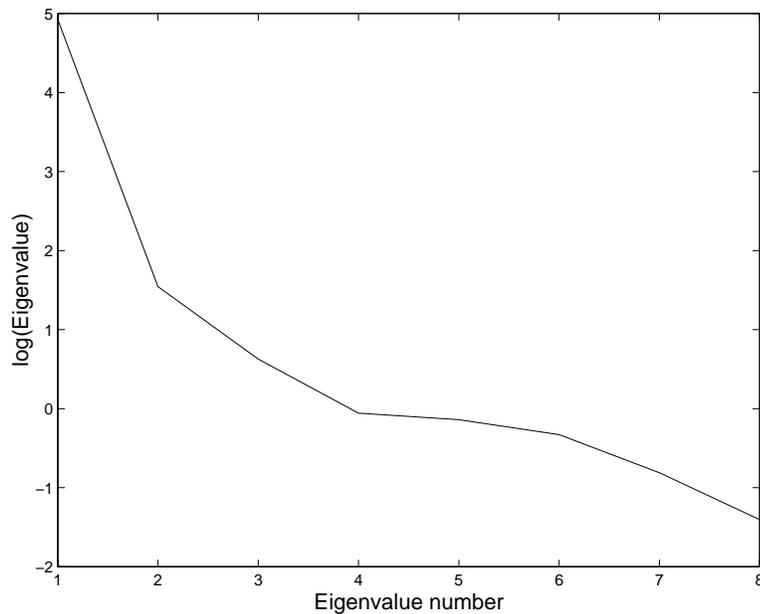


Figure 4.1: The log of the principal component values of  $X$ , showing that the highest two basis vectors account for most of the variance in the vegetation responses.

Although most of the variation is in the top two principal components, the other components contain information that may help improve the classification. Due to this and the fact that computational power was not a problem, all vectors are kept in the analysis.

The columns of  $Q_1$  are seen as a basis set that contains the most prominent seasonal responses (i.e., the principal component responses) of each vegetation region. The eigenvalues are seen as the variance along each basis set. In other words, the axes in the new vector space lie along the most prominent vegetation features of the various vegetation responses. Our motivation in doing this is to orthogonalize the most prominent vegetation characteristics between vegetation classes so as to make it easier to distinguish between one vegetation class and another. From here  $X$  can be expressed as a linear combination of  $Q_1$  or,  $X = \sum_i a_i Q_1$ . In matrix notation the equation becomes

$$X = A Q_1 \tag{4.2}$$

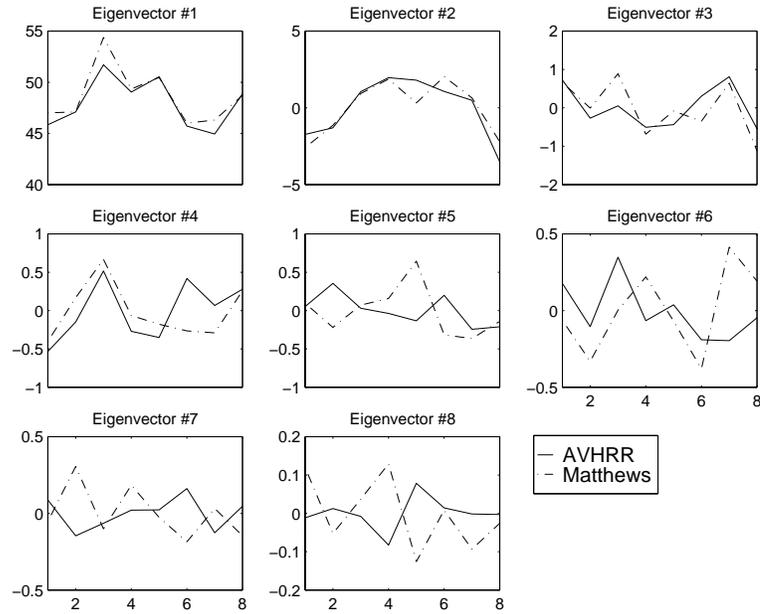


Figure 4.2: The eigenvectors of  $X$ , showing that the highest basis vectors is mainly a DC component while the other vectors account for actual variations in the vegetation responses.

where the rows of  $A$  are  $a_i$ . The  $a_i$ 's were calculated for each of the chosen vegetation regions above. These  $a_i$ 's are used as the centroids in the next step.

#### 4.4 Classification

As part of the classification, a metric or distance measure is required. Two different distance criteria are used for comparison. One is the usual  $L^2$ -norm distance and the other is a maximum likelihood metric. In both cases the seasonal response of each pixel in the image is projected onto the given basis set,  $r_i = Q^\dagger x_i$ , where  $x_i$  is the seasonal response of the  $i_{th}$  pixel.  $r_i$  is now the response of the pixel in the new basis set.

The first distance criteria, the  $L^2$ -Norm distance, is defined in the usual way, i.e.  $d = \|x - y\|$ , where  $x$  and  $y$  are two vectors in the space. In our case, for each pixel in the algorithm calculates the distance from  $r_i$ , the vegetation response of the pixel, to each vegetation centroid in the new orthogonal space. The centroid that is

the closest to the pixel response is chosen as the vegetation class of that pixel. This is shown in the following equation:

$$Im(i) = \min_j \|r_i - a_j\| \quad (4.3)$$

where  $j$  is the number of vegetation regions and  $i$  is the number of pixels in the image. Here  $Im$  is the final vegetation map.

Although the  $L^2$  norm classifier is easy to implement, it does not take into account the shape or statistics of the data. An alternate distance classifier which considers the data's statistics is also used. With this approach, we make the assumption that the seasonal response of  $\sigma^o$  has a Gaussian distribution over each vegetation class. We start with the multivariate Gaussian conditional probability distribution,

$$p(x|y) = \frac{1}{(2\pi)^{n/2} \sqrt{|R_y|}} e^{-\frac{1}{2}(x-\mu_y)^T R_y^{-1}(x-\mu_y)} \quad (4.4)$$

where  $R_y$  is the covariance matrix of  $x$  given  $y$  and  $\mu_y$  is the mean of  $x$  given  $y$ . This equation gives the probability that  $x$  occurred given that  $y$  is true. In our case, this compares the probability that the vegetation response,  $x$ , is the response over a certain vegetation class,  $y$ . Or, in other words, given a vegetation class,  $y$ , what is the probability that the response,  $x$ , is over that vegetation class. Maximizing the probability over the possible vegetation classes gives the most likely class. To make the maximization easier, the natural logarithm of the conditional probability is maximized. Taking the natural logarithm of Eq. 4.4,

$$\ln p(x|y) = -\frac{1}{2} \left[ \ln((2\pi)^n |R_y|) + (x - \mu_y)^T R_y^{-1}(x - \mu_y) \right] \quad (4.5)$$

$$= -\frac{1}{2} \left[ \ln(2\pi)^n + \ln |R_y| + (x - \mu_y)^T R_y^{-1}(x - \mu_y) \right]. \quad (4.6)$$

The  $\ln(2\pi)^n$  and  $-\frac{1}{2}$  terms are just constants that can be dropped in the maximization. Thus,

$$\max(\ln p(x|y)) = \min(\ln |R_y| + (x - \mu_y)^T R_y^{-1}(x - \mu_y)). \quad (4.7)$$

The maximization is changed to a minimization due to dropping the negative sign.

The second term of Eq. 4.7 is called the Mahalanobis distance. This term is often used as a distance classifier. In our analysis, instead of using the Mahalanobis distance, the entire term, called a maximum likelihood classifier, is used. The advantage that this approach has over the  $L^2$  norm distance is that the maximum likelihood classifier includes information about the statistics of the data; therefore, it can better create a delimiting line between different clusters of data. The classification improves using this classifier. A problem arises if we want to extend this to a non-trained algorithm where  $R$  is not known. In this case the  $L^2$  norm, or another criteria that does not use a priori information, would be used. From the maximum likelihood classifier the pixel value is chosen using the following:

$$Im(i) = \min_j [(r_i - a_j)' R_j^{-1} (r_i - a_j) + \ln |R_j|] \quad (4.8)$$

where  $R_j$  is the sample covariance of the projection of the seasonal response of each vegetation region,  $r_i$  is the pixel response and  $a_j$  is the centroid of each vegetation region. Here  $Im$  is the final vegetation map.



## Chapter 5

### Results using the Algorithm

Using the algorithm described in the previous chapter vegetation images were created over North and South America using both the AVHRR vegetation map and the Matthews vegetation map at both C- and Ku-bands. This chapter gives the results of the algorithm for each case and quantitatively compares them using several accuracy and inaccuracy assessments.

#### 5.1 Results over South America

Figures 5.1, 5.2, 5.3, and 5.4 show the results of this algorithm for Ku-band NSCAT data and for C-band ERS-2 data over South America. The values in the image correspond to the region numbers and vegetation types of the above study regions. Tables 5.1-5.8 show the percentage of pixels properly classified over each class.

Table 5.1: Percentage of NSCAT vegetation classification over South America for Matthew's using  $L^2$  Norm.

$L^2$ Norm	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	79.2	6.2	1.7	1.7	5.7	3.4
2	24.8	33.3	0.8	1.6	26	13.2
3	3.9	4.3	59.1	13.5	4.2	12.3
4	1.0	2.5	35.7	27.2	8.9	22.7
5	16.1	25.6	2.3	2.6	41.5	11.2
6	13.9	11.6	28.2	7.9	11.7	24

Table 5.2: Percentage of NSCAT vegetation classification over South America for Matthews using maximum likelihood (ML).

ML	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	78.2	6.3	2	2.2	6.3	3
2	18.6	63.9	0.8	5.4	5.4	5.4
3	4	8.3	66.4	5.9	3.5	9.4
4	1.8	11.7	3.9	76.5	1	3.1
5	17.5	30.2	3.2	4.6	40.5	3.3
6	10.2	17.2	9.6	14.0	7.3	38.9

Table 5.3: Percentage of NSCAT vegetation classification over South America for AVHRR using  $L^2$  Norm.

$L^2$ Norm	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	82	9	1	0.5	4.8	1.7
2	25.6	37	3.1	3.1	21.7	9.1
3	2.3	8.1	59.7	6.7	10.6	11.7
4	8	11.1	50.8	10.1	9.8	9.9
5	8.6	33.4	3.5	2.1	40.7	11.1
6	3.1	9.2	36.3	7.1	23.7	19.1

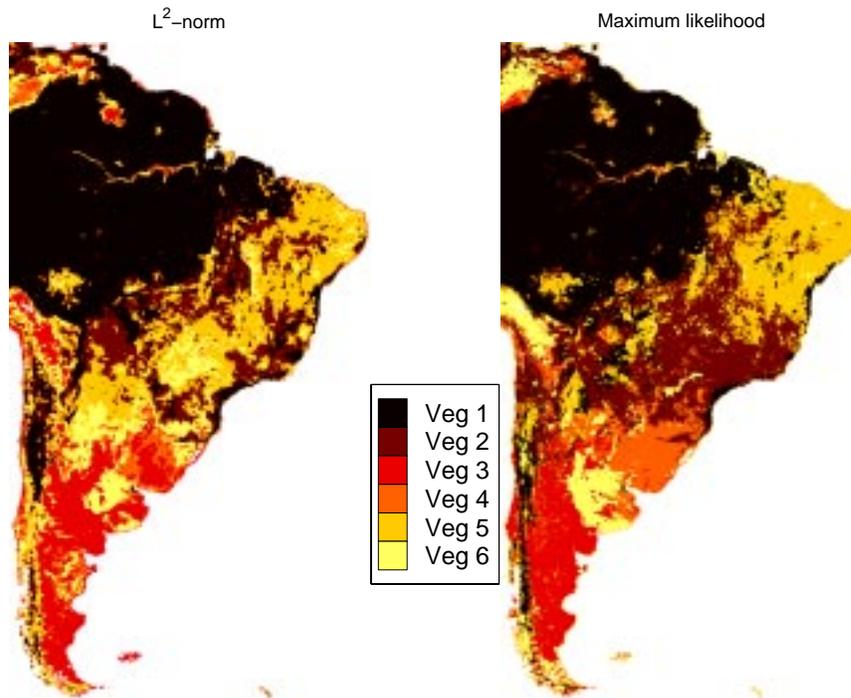


Figure 5.1: NSCAT Vegetation Classification over South America using the Matthews vegetation map. The left one is the classification using the  $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier.

Table 5.4: Percentage of NSCAT vegetation classification over South America for AVHRR using maximum likelihood (ML).

ML	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	82.9	7.8	1.1	1.3	4.7	1.4
2	27.8	40.6	2.5	4.4	18.9	5.4
3	3.0	10.9	28.8	44.7	4.8	6.9
4	3.7	10.9	11.8	61.7	6.6	4.9
5	14.1	16.3	2.7	5.7	52.9	7.8
6	2.3	10.3	10.7	16.6	20.9	37.6

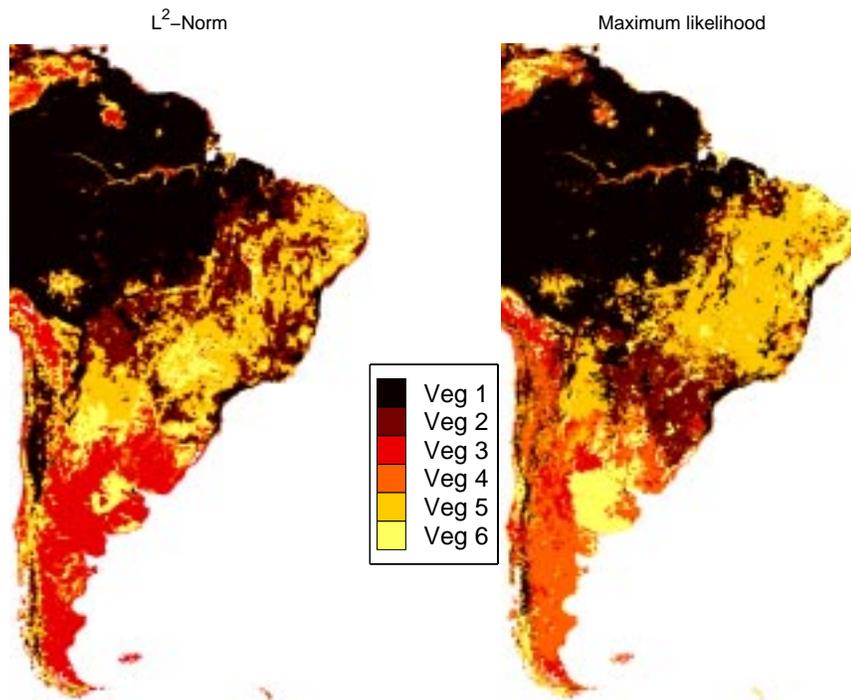


Figure 5.2: NSCAT Vegetation Classification over South America using the AVHRR vegetation map. The left one is the classification using the  $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier.

Table 5.5: Percentage of ERS-2 vegetation classification over South America for Matthews using  $L^2$  Norm

$L^2$ Norm	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	82.2	6.4	0.7	1.4	4.9	2.1
2	29.6	30.9	0.5	2.4	24.6	11.1
3	4	3.4	65.9	8.0	4.8	10.6
4	1	3.6	32.8	31.2	11.4	18.1
5	22	20.8	1.9	2.7	41.6	9.7
6	13.7	10.5	25.0	9.5	13.7	25.0

Table 5.6: Percentage of ERS-2 vegetation classification over South America for Matthews using maximum likelihood.

ML	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	82.1	4.6	1.5	1.8	5.8	2.0
2	24.6	58.4	0.8	4.8	7.2	3.3
3	9.7	4.5	64.4	2.2	3.8	12.0
4	1.5	11.4	0.8	81.1	0.2	3.2
5	23.8	18.2	2.8	2.6	47.0	4.3
6	16.7	6.5	12	10.6	8.5	43.1

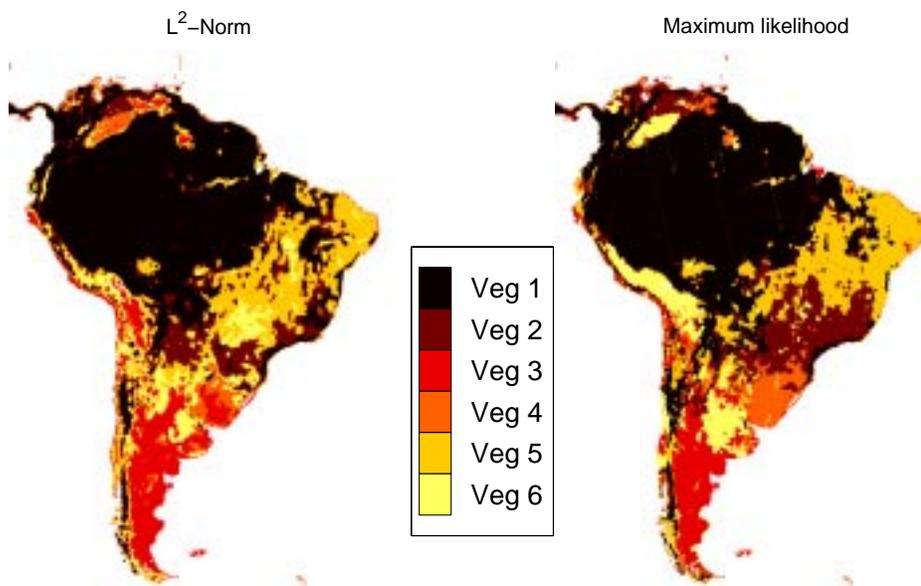


Figure 5.3: ERS-2 Vegetation Classification over South America using the Matthews vegetation map. The left one is the classification using the  $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier.

Table 5.7: Percentage of ERS-2 vegetation classification over South America for AVHRR using  $L^2$  Norm

$L^2$ Norm	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	83.9	7.3	0.4	0.4	4.9	2.3
2	31.4	28.5	2.2	1.4	21.9	14.2
3	3.0	5.9	15.2	50.7	9.9	14.3
4	2.8	4.5	12.3	56.2	6.3	17.4
5	17.8	21.9	1.9	1.5	40.3	16
6	9.6	9.0	15.4	19.3	20.3	24.9

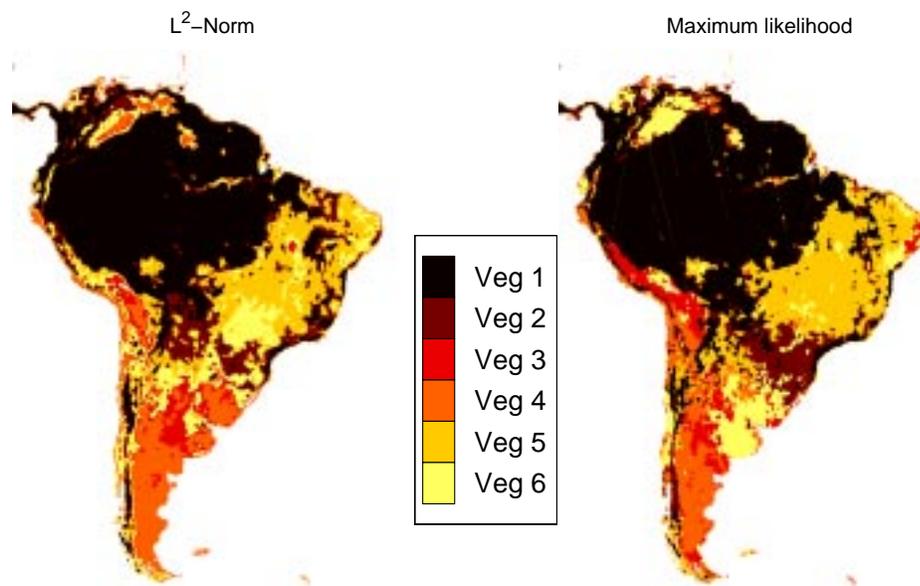


Figure 5.4: ERS-2 Vegetation Classification over South America using the AVHRR vegetation map. The left one is the classification using the  $L^2$ -norm and the right is the classification using the maximum likelihood (ML) classifier.

Table 5.8: Percentage of ERS-2 vegetation classification over South America for AVHRR using maximum likelihood (ML).

ML	Confusion Matrix					
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6
1	84.9	5.8	0.6	1.2	5.3	1.5
2	36.9	29.9	1.4	2.6	21.4	7.3
3	7.9	5.1	38.9	37.3	3.4	6.7
4	7.6	6.2	10.0	66.9	6.1	2.6
5	21.8	12.8	3.2	3.1	50.8	7.7
6	10	9.4	11.8	8.7	21.9	36.7

## 5.2 Results over North America

Figures 5.5, 5.6, 5.7, and 5.8 show the results of this algorithm for Ku-band NSCAT data and for C-band ERS-2 data over North America. The values in the image correspond to the region numbers and vegetation types of the above study regions. Tables 5.9-5.16 show the percentage of pixels properly classified over that particular region.

Table 5.9: Percentage of NSCAT vegetation classification over North America for Matthews using  $L^2$  Norm

$L^2$ Norm	Confusion Matrix							
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	61.9	7.9	6.0	3.4	4.3	6	3.8	3.9
2	7.9	20.2	56.6	5.5	1.0	2.5	2.1	0.1
3	2.7	2.0	84.1	8.0	0.2	0.3	0.8	0
4	2.6	2.8	45.5	20.7	7.4	1.3	18.5	0.1
5	11.2	1.7	9.1	13	25.9	2.1	34.6	1.9
6	1.3	4.2	23.8	42.7	4.2	0.4	22.8	0.4
7	5.2	2.6	3.3	9.3	6.8	1.2	70.7	0.8
8	52.5	3.6	0	1	0.5	26.0	1.6	11.1

Notice that using the Maximum Likelihood classifier increases the percentage of pixels correctly classified. This can be expected due to the nature of the data. From these charts it is difficult to say which classification is most correct, as some have a higher percentage of correct classification over some vegetation classes while some have a higher percentage over others. To help quantitatively determine the best classification scheme, accuracy and inaccuracy assessments are done on the classifications. In [22] a different accuracy and inaccuracy assessment is proposed that better assess the data than the usual kappa statistic. The new accuracy assessment is derived from the Kullback-Leibler information. The usual class-averaged and overall accuracies are given.

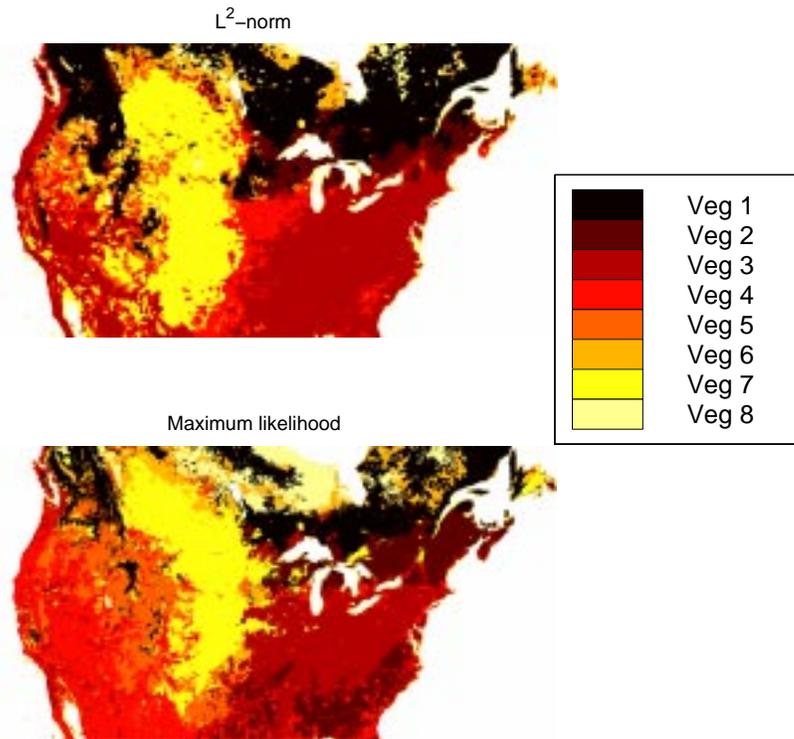


Figure 5.5: NSCAT Vegetation Classification over North America using the Matthews vegetation map. The top one is the classification using the  $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier.

Table 5.10: Percentage of NSCAT vegetation classification over North America for Matthews using maximum likelihood (ML).

ML	Confusion Matrix							
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	43.6	5.6	4.1	8.1	10.4	10.5	4.3	10.6
2	2.2	52.6	22	10.4	0.8	2.7	4.9	0.4
3	0.6	7.8	80.0	4.7	0.3	0.4	4.2	0.0
4	0.7	3.1	4.4	77.5	9.8	0.6	2.4	0.3
5	3.2	2.9	1.9	41.0	40.4	1.2	8.8	0.3
6	1.4	7.6	44.5	7.3	6.9	7.9	24.1	0.1
7	2.4	0.8	4.3	12	13.9	3	63.3	0.4
8	26.5	0.5	0.7	0.4	2	14.7	1.0	50.6

Table 5.11: Percentage of NSCAT vegetation classification over North America for AVHRR using  $L^2$  Norm

$L^2$ Norm	Confusion Matrix							
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	49.1	9.3	11.7	5.1	1.9	8.2	4.6	8.6
2	23.5	32.2	30.8	5.6	0.8	2.9	1.9	1.2
3	15.8	32.1	18.8	12.8	2.6	6.7	8.3	2.2
4	0	24.3	15.2	14.5	10.5	0.3	33.1	0.2
5	0.1	17	12.6	14.3	12.2	0.3	41.7	1.2
6	36.1	7.1	8.8	7.3	2.5	8.9	2.6	25.6
7	2.9	7.7	13.3	10.0	6.9	2.7	49.9	6.1
8	9.5	3.7	0	0.3	1.5	3.3	5.8	65.3

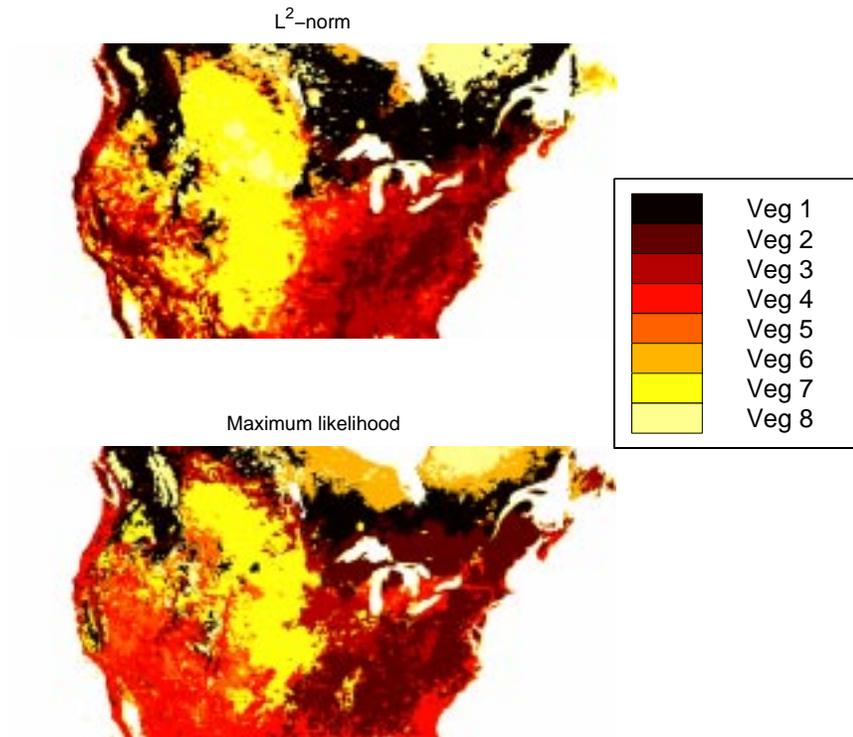


Figure 5.6: NSCAT Vegetation Classification over North America using the AVHRR vegetation map. The top one is the classification using the  $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier.

Table 5.12: Percentage of NSCAT vegetation classification over North America for AVHRR using maximum likelihood (ML).

ML	Confusion Matrix							
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	39.9	20	4.5	12.8	1.2	10.6	3.2	6.3
2	7.2	69.2	7.3	10.0	0.6	1.3	2.8	0.5
3	6.2	39.2	31.5	8.4	0.8	2.2	9	2.1
4	0.6	2.1	0.3	71.4	15.3	1.3	6.8	0.4
5	0.5	1.6	0.5	55.5	27.6	1.4	11.7	0.7
6	10.8	6	1.1	15.2	2.4	42.8	2.1	18.5
7	2.9	10.0	13.1	14.5	8.5	2.0	46.9	1.5
8	0.70	0.1	0.4	0.3	0.4	9.7	2.8	74.9

Table 5.13: Percentage of ERS-2 vegetation classification over North America for Matthews using  $L^2$  Norm

$L^2$ Norm	Confusion Matrix							
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	25	12.9	25.9	5.6	2.7	2.8	3.8	18.5
2	8.7	19.7	60.6	1.4	0.9	3.5	0.8	0.4
3	9.0	10.1	69.7	1	0.8	7.2	0.2	0
4	12.3	7.4	16.0	22.7	22.1	0.4	17.8	0
5	11.6	4.9	4.0	19.8	30.3	1.5	27.4	0.1
6	21.4	7.9	8.1	2.2	1.9	33.6	24.4	0.4
7	4.2	1.8	2.7	3.4	7.2	10.4	69.6	0.7
8	12.0	2.6	6.1	1.2	1.8	3.2	13.7	55.6

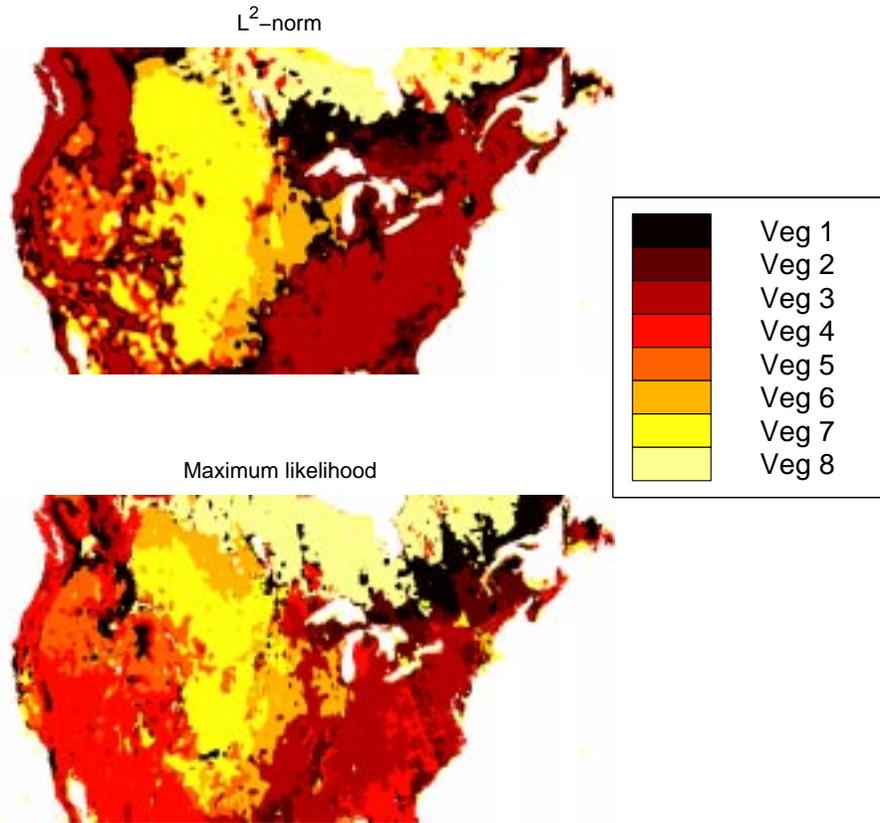


Figure 5.7: ERS-2 Vegetation Classification over North America using the Matthews vegetation map. The top one is the classification using the  $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier.

Table 5.14: Percentage of ERS-2 vegetation classification over North America for Matthews using maximum likelihood (ML).

ML	Confusion Matrix							
Vegetaion Regions	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	23.3	6.1	7.7	18.3	4.8	2.0	1.8	33.2
2	5.0	29.3	34.0	18.1	1.7	2.9	2.7	2.4
3	0.5	5.3	74.8	9.3	1	6	1.0	0.1
4	1.7	0.5	1.7	78.5	13.3	0.5	2.4	0.3
5	3.9	0.3	1.6	31.8	54.7	0.4	6.6	0.3
6	0.2	1.7	19.7	1.9	4.8	64.8	6.2	0.6
7	1.9	0.4	2.4	7.5	14.1	12.8	60.2	0.7
8	7.6	0.1	3.4	3	0.7	1.9	1.1	78.6

Table 5.15: Percentage of ERS-2 vegetation classification over North America for AVHRR using  $L^2$  Norm.

$L^2$ Norm	Confusion Matrix							
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	18.5	43.3	6.2	6.9	1.9	17.9	1.5	2.2
2	8.3	66.5	12	3.8	1.1	5.2	1.8	0.2
3	5	47.6	13.8	8.4	4.6	9.7	9.4	0.8
4	6.9	18.8	5.6	18.6	29.3	10.1	6.0	2.8
5	2.8	9.3	5.4	17.4	47.3	7.5	4.8	4.9
6	8.8	19.3	5.8	7.2	6.4	23.5	2.8	25.2
7	3.2	14.1	8.3	9.6	29	6.7	18.3	10.1
8	0.6	11.3	0	1.6	4.6	2.6	0.8	67.5

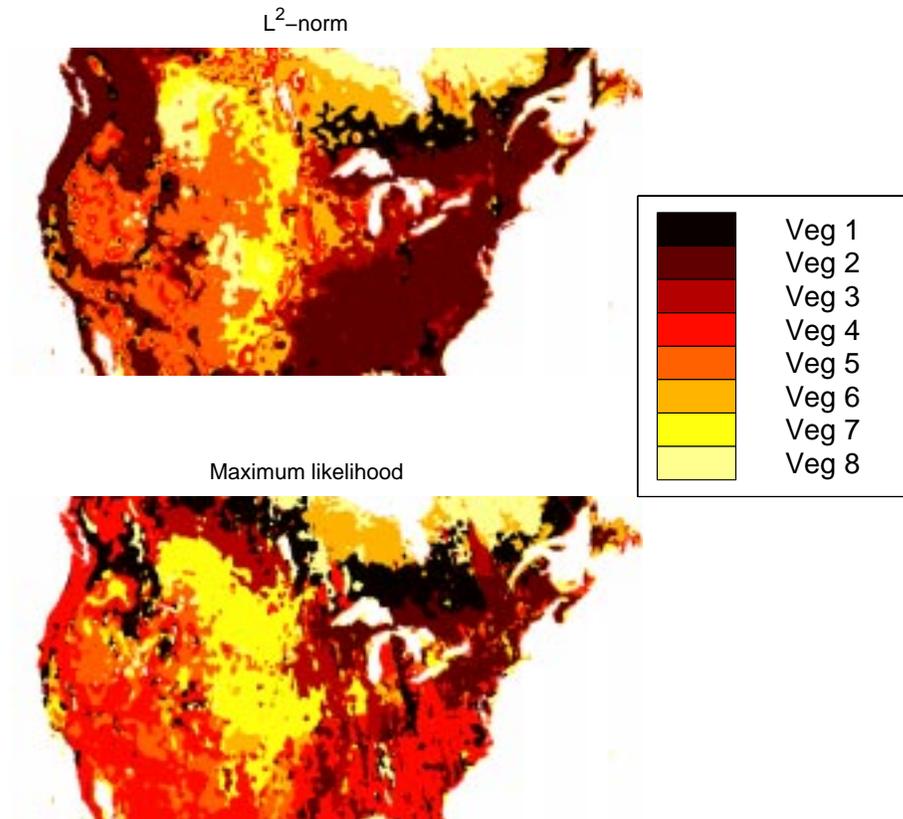


Figure 5.8: ERS-2 Vegetation Classification over North America using the AVHRR vegetation map. The top one is the classification using the  $L^2$ -norm and the bottom is the classification using the maximum likelihood (ML) classifier.

Table 5.16: Percentage of ERS-2 vegetation classification over North America for AVHRR using maximum likelihood (ML).

ML	Confusion Matrix							
Vegetation Class	Veg 1	Veg 2	Veg 3	Veg 4	Veg 5	Veg 6	Veg 7	Veg 8
1	35.6	8.7	3.9	32.1	3.3	7.2	1.9	5.9
2	12.7	40.7	7.4	29.8	3.2	1	3.1	1.0
3	10.3	21.8	25.6	28.5	3.3	1.6	7.4	0.8
4	0.7	0.6	0.8	65.7	20.3	2.2	7	0.8
5	0.2	0.9	0.6	48.7	34.9	2.2	10.6	1.5
6	11.7	2.9	2.2	26.5	8.4	23.6	4.9	18.7
7	3.3	3.8	14.1	15.3	15.3	5.4	39.7	2.4
8	3.6	1	1.7	2.5	3.6	7.5	2.5	66.6

Table 5.17: Class-Average and overall accuracy coefficients over South America for the given classification parameters used in the classification algorithm

		$L^2$ Norm			ML		
		C(X)	A(X)	$J_p(X)$	C(X)	A(X)	$J_p(X)$
NSCAT	Matthews	.4488	.5645	.5193	.6183	.6235	.5962
	AVHRR	.4175	.5356	.4477	.5115	.6294	.5979
ERS-2	Matthews	.4718	.6050	.5524	.6399	.6721	.6495
	AVHRR	.4180	.5745	.5078	.5174	.6389	.6005

Table 5.18: Class-Average and overall Accuracy coefficients over North America for the given classification parameters used in the classification algorithm

		$L^2$ Norm			ML		
		C(X)	A(X)	$J_p(X)$	C(X)	A(X)	$J_p(X)$
NSCAT	Matthews	.3751	.4929	.3504	.5301	.5455	.5036
	AVHRR	.3257	.3867	.3347	.5212	.4694	.4535
ERS-2	Matthews	.4147	.3939	.3465	.5897	.4768	.4245
	AVHRR	.3556	.2897	.2488	.4296	.3706	.3623

### 5.3 Accuracy and Inaccuracy of the Classification

Given an error matrix  $X$ , the class-averaged coefficient,  $C$ , and overall accuracy coefficient,  $A$ , are computed using:

$$C(X) = \frac{1}{r} \sum_{i=1}^r \frac{x_{ii}}{n_i} \quad (5.1)$$

$$A(X) = \frac{1}{N} \sum_{i=1}^r x_{ii}. \quad (5.2)$$

Here,  $r$  is the number of classification categories,  $n_i$  is the number of samples for each category, and  $N$  is the total number of samples. These two accuracy assessments are shown for both North and South America in tables 5.17 and 5.18.

Notice that the class-average and overall accuracies tend to be higher over South America than they are over North America. This is most likely due to two causes: 1) South America has fewer vegetation classes and 2) each vegetation class over South America covers a larger area unlike North America where many vegetation

classes are smaller in overall area. The Amazon forest in South America, for example, is a large area correctly classified. From looking at the table it is apparent that using C-band data gives a better accuracy than Ku-band data. The accuracy is also slightly higher using the Matthews vegetation map than the derived vegetation map from AVHRR. Classification using the maximum likelihood classifier is clearly more accurate than the regular  $L^2$  Norm distance.

The third accuracy assessment is computed using:

$$J_p(X) = \prod_{i=1}^r \left( \frac{x_{ii} + 1/2}{n_i + 1/2} \right)^{n_i/N} . \quad (5.3)$$

Here  $r$  is the number of classification categories,  $n_i$  is the number of samples from each category, and  $N$  is the total number of samples. The reason that this accuracy assessment is used is due to the favorable properties this assessment exhibits over the kappa statistic. For an in-depth discussion of the comparisons between the two see [22]. This assessment gives a coefficient between 0 to 1 only, where the coefficient is 0 if and only if all samples are misclassified and 1 indicates total classification, while the kappa statistic will give various values according to the types of complete misclassification. In Table 5.17 the  $J_p$  accuracy assessment coefficients for all the classifications over North America are given and in Table 5.18 they are given over South America.

Using this accuracy assessment it is apparent that over South America using C-band data in this vegetation classification algorithm tends to be more accurate than using Ku-band. The most accurate classification scheme is using the maximum likelihood discriminator with the Matthews vegetation map for C-band data. Over North America we see that the trend has switched. Using Ku-band data tends to be more accurate than using C-band data with the most accurate classification scheme using the maximum likelihood classifier with the Matthews vegetation map for Ku-band data.

Deciduous forests and mixed forest as well as evergreen forests often have similar seasonal responses and therefore often get misclassified into one another. By combining similar regions and repeating the accuracy assessment, the accuracy of the

Table 5.19: New vegetation classes over North America

	New Vegetation Class	Old Vegetation Class
1	Forests	Evergreen needleleaved forests Deciduous forests with evergreens/mixed forests Deciduous forests
2	Shrub/Woodlands	Xeromorphic shrublands/Shrubland/Closed Bushlands Grassland with shrub cover/Open shrublands Woodlands/grasslands with woody cover
3	Grasslands	Tall, medium, short grasslands, meadows/Croplands
4	Alpine Tundra	Arctic alpine tundra, mossy bog/Mosses and Lichens

Table 5.20: New vegetation classes over South America

	New Vegetation Class	Old Vegetation Class
1	Tropical Forests	Tropical Forests Tropical Rainforests/Evergreen Broadleaf forests/Evergreen Rainforests
2	Mixed Forests	Deciduous broadleaf forests/mixed forests, Tropi- cal/subtropical drought deciduous forests Xeromorphic shrubland/shrublands/closed bushlands
3	Shrub/Woodlands	Open shrublands/grassland with shrub cover Xeromorphic forest/ woodlands/wooded grass- lands/grassland with 10-40 % woody tree cover
4	Grasslands	Meadows/short, medium, tall grasslands

grouped classes can be evaluated. Taking the vegetation regions and combining the most similar vegetation regions together to reduce the number of vegetation regions over each area gives the vegetation classes in Tables 5.19 and 5.20. The accuracy assessments for the reduced vegetation region area of both North and South America are given in tables 5.21 and 5.22.

As would be expected, reducing the number of vegetation classes increases accuracy. Seeing this observation leads us to examine another useful assessment that measures the inaccuracies there are in the classification. For example, whether mixed forests are misclassified into a similar category like deciduous forests or into a category

Table 5.21: The overall accuracy,  $A(X)$ , class-average,  $C(X)$ , and  $J_p(X)$  accuracy coefficient over South America for the reduced vegetation set.

		$L^2$ Norm			ML		
		$A(X)$	$C(X)$	$J_p(X)$	$A(X)$	$C(X)$	$J_p(X)$
NSCAT	Matthews	.6075	.4878	.5699	.6561	.6046	.6422
	AVHRR	.5987	.4836	.5396	.6799	.5740	.6545
ERS-2	Matthews	.6381	.4956	.5941	.6911	.6174	.6757
	AVHRR	.6129	.4786	.5590	.6765	.5437	.6424

Table 5.22: The overall accuracy,  $A(X)$ , class-average,  $C(X)$ , and  $J_p(X)$  accuracy coefficient over North America for the reduced vegetation set.

		$L^2$ Norm			ML		
		$A(X)$	$C(X)$	$J_p(X)$	$A(X)$	$C(X)$	$J_p(X)$
NSCAT	Matthews	.6948	.5083	.6435	.6673	.6323	.6663
	AVHRR	.5755	.5551	.5291	.6504	.6943	.6377
ERS-2	Matthews	.6930	.6259	.6799	.6284	.7062	.6182
	AVHRR	.5261	.5528	.4506	.5366	.6005	.5245

that is farther away like grasslands. For an in depth discussion of this assessment scheme we refer to [22]. The inaccuracy assessment,  $R_p$  is computed in the following way:

$$R_p(X|, ) = 1/N \sum_{i=1}^r \sum_{j=1}^r \gamma_{ij} x_{ij} \quad (5.4)$$

where  $N$ ,  $r$  and  $x$  are defined as above.  $\gamma$  in this equation is the cost matrix, where for each element,  $\gamma_{ij}$ , in  $\gamma$ , denotes the cost of allocation when a sample from category  $C_i$  is classified into category  $C_j$ . This inaccuracy assessment gives values from 0 to  $M$ , where

$$M = \sum_{i=1}^r n_i/N \max(\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{ir}). \quad (5.5)$$

Zero indicates all samples are correctly classified and  $M$  indicates all samples are respectively misclassified into the largest cost categories. For the above classifications two cost matrices are chosen and the inaccuracy coefficients are computed for both

North and South America. The cost matrices are

$$, 1 = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & 1 & 2 & 3 & 4 \\ 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 4 & 3 & 2 & 1 & 0 & 1 \\ 5 & 4 & 3 & 2 & 1 & 0 \end{bmatrix} \quad (5.6)$$

$$, 2 = \begin{bmatrix} 0 & 1 & 2 & 2 & 2 & 3 \\ 1 & 0 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 & 0 & 1 \\ 3 & 2 & 1 & 1 & 1 & 0 \end{bmatrix} \quad (5.7)$$

for South America and for North America

$$, 1 = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 0 & 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ 5 & 4 & 3 & 2 & 1 & 0 & 1 & 2 \\ 6 & 5 & 4 & 3 & 2 & 1 & 0 & 1 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \end{bmatrix} \quad (5.8)$$

$$, 2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 3 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 2 \\ 2 & 2 & 2 & 1 & 1 & 1 & 0 & 1 \\ 3 & 3 & 3 & 2 & 2 & 2 & 1 & 0 \end{bmatrix}. \quad (5.9)$$

Tables 5.23 and 5.24 give the results of the inaccuracy coefficients for the two cost matrices over North and South America. The normalized coefficients are also shown for easy comparison. Examining the results we see that the maximum likelihood discriminator performs better than the  $L^2$  Norm discriminator. The algorithm tends to misclassify vegetation classes into classes with similar characteristics. This is most noticeable over North America where more initial vegetation classes are used. This indicates that this algorithm performs well over areas containing large homogeneous sections of vegetation, with large differences in vegetation type between the various classes used.

Considering the limitations of the reference vegetation maps, the classifications exhibit a high degree of accuracy and consistency. The primary confusion observed is between related vegetation classes of similar canopy density. We believe that coupling scatterometer data with AVHRR data or other types of data could improve the accuracy of global vegetation maps. This algorithm would be most useful in discriminating between a few large class types or perhaps as an initial classification to be used in conjunction with other types of schemes and data sets.

Table 5.23: The  $R_p$  inaccuracy coefficients for the given , cost functions over South America for the given classification parameters used in the classification algorithm.

		$L^2$ Norm			
		$R _{, 1}$	$(R _{, 1})/M$	$R _{, 2}$	$(R _{, 2})/M$
NSCAT	Matthews	1.1543	.2638	.5941	.2380
	AVHRR	1.0846	.2420	.5719	.2213
ERS-2	Matthews	1.0528	.2369	.5577	.2184
	AVHRR	1.0480	.2314	.5951	.2269
		ML			
		$R _{, 1}$	$(R _{, 1})/M$	$R _{, 2}$	$(R _{, 2})/M$
NSCAT	Matthews	1.0425	.2382	.5303	.2124
	AVHRR	.8513	.1899	.4729	.1830
ERS-2	Matthews	.9333	.2030	.5055	.1979
	AVHRR	.9256	.2044	.5178	.1975

Table 5.24: The  $R_p$  inaccuracy coefficients for the given , cost functions over North America for the given classification parameters used in the classification algorithm.

		$L^2$ Norm			
		$R _{, 1}$	$(R _{, 1})/M_1$	$R _{, 2}$	$(R _{, 2})/M_2$
NSCAT	Matthews	1.2824	.2160	.4284	.1632
	AVHRR	1.7360	.2916	.6048	.2426
ERS-2	Matthews	1.5817	.2665	.4797	.1827
	AVHRR	1.7849	.3013	.6171	.2490
		ML			
		$R _{, 1}$	$(R _{, 1})/M_1$	$R _{, 2}$	$(R _{, 2})/M_2$
NSCAT	Matthews	1.4128	.2380	.4750	.1809
	AVHRR	1.4958	.2513	.5091	.2042
ERS-2	Matthews	1.6954	.2856	.6302	.2400
	AVHRR	1.6738	.2825	.6060	.2445



## Chapter 6

### Conclusion

This chapter summarizes the results of this thesis, discusses the contributions made, and discusses future work to be done.

#### 6.1 Overview

This thesis has presented a new algorithm to classify vegetation based on the seasonal response of  $\sigma^o$ . The algorithm was used for  $\sigma^o$  data at two different frequencies with two different vegetation maps to train it. The seasonal signature of  $\sigma^o$  was evaluated. First the seasonal responses of  $\sigma^o$  over different regions in North America was examined. It was found that over a given region the  $\sigma^o$  response changes over time. Changes in  $\sigma^o$  exhibited some correlation with precipitation and temperature.

Second, the  $\sigma^o$  seasonal response was observed over different vegetation classes at two different frequencies over both North and South America. The results showed that there is a significant difference in seasonal response over different vegetation regions at different frequencies. Using these differences, an algorithm was developed to take advantage of this information to classify vegetation areas. The algorithm took a training set from two different vegetation maps and used singular value decomposition (SVD) to create a basis set from the training set. The results were compared with the vegetation maps. It was shown that the algorithm performed well over large areas containing homogeneous vegetation types that differ greatly from each other. This was seen over South America where large areas are covered by the same vegetation type. Also, misclassification most likely occurred with classes of similar types suggesting

that the fewer number of vegetation classes, the better the algorithm. Overall, the results were consistent with previous work done and could be combined with other methods to improve classification.

## 6.2 Contributions

This thesis has made three primary contributions to the field of scatterometry. First, an extensive assessment of the seasonal response of  $\sigma^o$  over different frequencies and over different vegetation regions. Second, the development of an operational algorithm for vegetation classification using the seasonal response of  $\sigma^o$ . Third, a method of combining two different maps at two different projections and grid resolution for easy comparison.

The assessment of the seasonal response of  $\sigma^o$  over different vegetation regions is shown to be useful in seasonal studies of similar vegetation classes over different areas. Also studies in the annual changes of areas such as the rainforest can help monitor deforestation and changes in ecosystems and environments.

This work is the first time that the seasonal response of scatterometry data has been used in classifying large vegetation areas. Previous methods mainly have used radiometers or SAR images at one particular instant in time in vegetation classification studies, and tend to be more regional in nature. This algorithm was able to use the seasonal  $\sigma^o$  response over vegetation to classify large vegetation areas. The algorithm has shown consistent results with previous work and could be combined with other methods of vegetation classification for improvement.

Finally, an important method of combining dissimilar maps for comparison arose out of the development of the algorithm and the comparison of the different vegetation maps. This method can be used in future vegetation classification work for easy pixel to pixel comparison.

## 6.3 Future Work

A number of different areas of future research were identified as part of this study. The seasonal response vegetation classification algorithm needs to be more

thoroughly compared with other types of vegetation classification techniques. A method to combine the information from other sensors such as radiometers or SARs to improve the classification. The algorithm presented in this thesis could be useful perhaps as a pre-classifier to get a general vegetation classification over a large area, then have it followed up by a more localized classification method.

A study in optimizing how best to combine the information from the seasonal response of both C-band and Ku-band data to obtain the best results should be considered. Also, new insight may be gained by adapting the algorithm to classify vegetation without first training it or by incorporating a new method such as neural networks, besides SVD to train or classify the data.

The assessment done in this thesis assumed that the seasonal response of a particular vegetation class did not vary with latitude, altitude or climate conditions. These assumptions will have to be addressed and could account for some of the misclassification, especially over North America where snow and freeze/thaw cycles are more prominent compared to South America.



## Appendix A

### Temporal Average Algorithm

#### A.1 Abstract

Spaceborne scatterometers can be used in the detection of the ice edge of the Arctic and Antarctic in order to study the effects of global warming and climate changes. Because of the non-linear properties of the SIR algorithm, errors in the imaged location of the edge can occur over areas of movement during the imaging interval. Ice growth and melting, for example, can cause errors in the detection of the actual location of an ice edge in a SIR image. A ‘time-weighted average’ algorithm of the  $\sigma^o$  values is developed to help improve the associated temporal estimate from the SIR image over areas of motion. This algorithm provides a better estimate of the temporal location of features in the SIR image for both temporally uniform and non-temporally uniform measurement distributions.

#### A.2 Introduction

To generate a SIR image from scatterometry data, several days worth of data are used. The resulting image is a weighted, non-linear average of the surface response over this period. Previously, the surface has been assumed to be constant during the days over which the SIR image was formed. Over areas of movement the location of the observed features in the SIR image has been assumed to be the location of the observed features at the mid-point of the time period over which the SIR image was created. Since the SIR algorithm is non-linear, however, this may not always be the case. For example, consider Fig. A.1. Here we see a simulated SIR image of a fast

moving edge, the linear temporal average of the edge over the six day period, and the edge at day 3 which corresponds to the location of the edge at the center of the six day period. Notice that the edges in the SIR image and in the day 3 image do not seem to correspond to each other as well as desired.

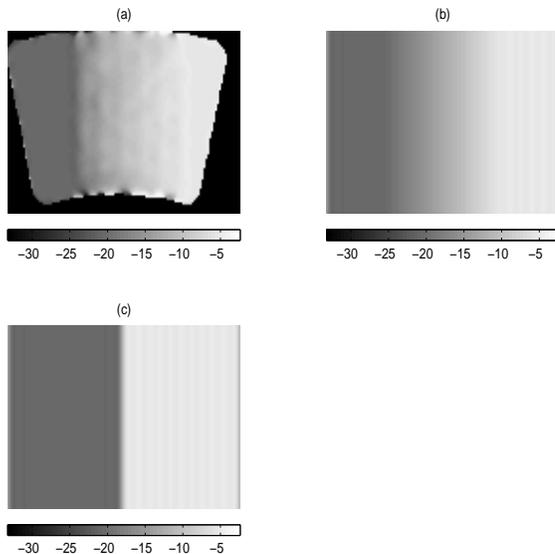


Figure A.1: (a) Simulated SIR image (b) linear average image (c) filtered image at day 3.

A simulation was created to evaluate the effects of the movement of a feature, specifically an edge. In this simulation,  $\mathcal{A}$  and  $\mathcal{B}$  truth images were created using a very fast moving edge. The speed of the edge in the simulation was set to 10 pixels ( $\sim 90km$ ) per day to more easily show the effects of the movement. A small area in the Antarctic was used to get the time and spatial distribution of real measurements. The location of the edge in the truth images was determined by the time those measurements occurred in the real data over the 6 day period. Figure A.2 shows a few of the truth images over the six day period beginning at day 0 and ending at day 6 at increments of 1/2 day. These images have been low-pass filtered to 25km

resolution to aid in the comparison with the SIR images, which have approximately 25km resolution. This filtering accounts for edge ringing in the images. The unfiltered truth images were used to create simulated measurements over a six day period. A SIR image was then generated from the simulated measurements.

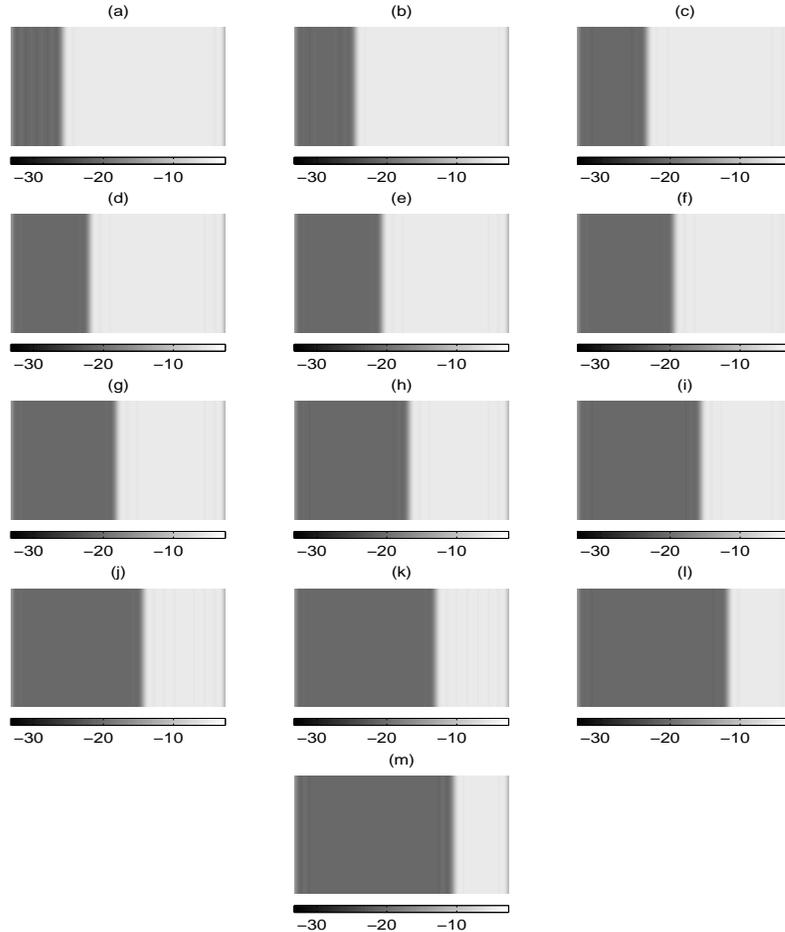


Figure A.2: Simulated ice edge images from (a) day 0 to (m) day 6, in 1/2 day increments.

Figure A.3 shows horizontal slices through the three different images shown in Fig. A.1. Notice that the magnitude of the edge of the SIR image is higher than the linear average edge over the 6 day period. This, in effect, causes the edge of the

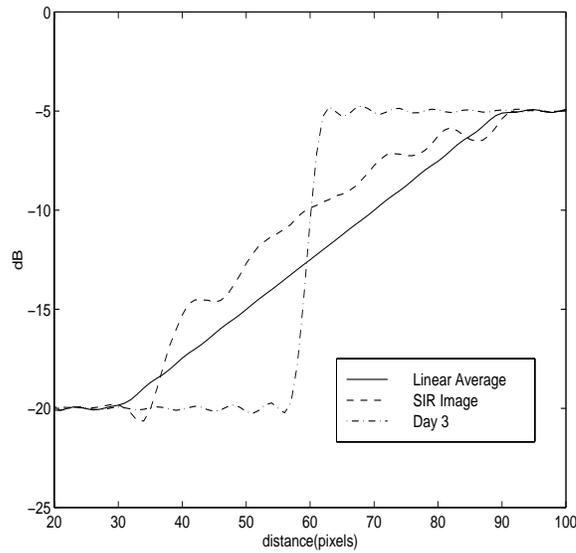


Figure A.3: Ice edge of the SIR image, linear average image and filtered image at day 3

SIR image to appear, in this case, further to the left than the average edge. (For an edge moving from right to left, i.e. growing, it would appear further to the right.) Another interpretation is shown in Fig. A.4. This shows the linear average edge shifted to the left by  $x$  days. Notice that this edge is a much closer approximation to the edge of the SIR image. In other words, We can think of the edge in the SIR image as a shifted linear average. For a receding edge the average corresponding to the SIR image would be shifted to the left. For a growing edge the average would be shifted to the right. Through this observation, a time index algorithm has been developed to help with the discrepancies seen in these images. This paper will discuss the algorithm, and explain how it works. We consider some particular cases.

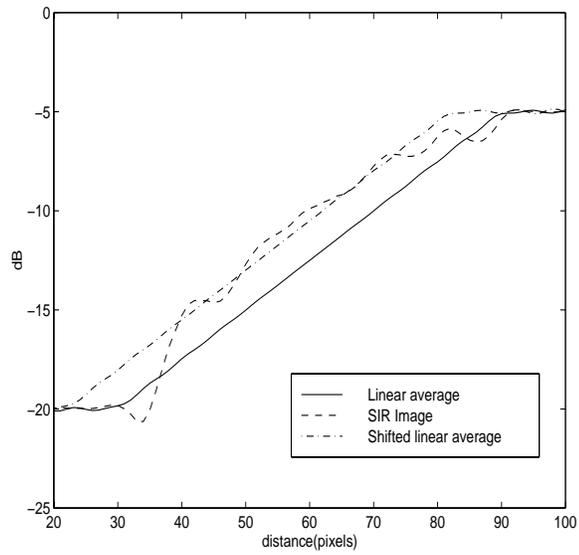


Figure A.4: SIR image and linear average image shifted from the center to an earlier period.

### A.3 The Time Index Algorithm

The time index algorithm provides a time value which closely corresponds to the observed temporal average of the image pixels. It provides a unique value for each pixel. The following is the algorithm for the weighted time index average,  $T_k$ , of the SIR image for the  $k^{th}$  pixel:

$$T_k = \frac{\sum_{i=1}^n t_i \overline{\sigma_i^o}}{\sum_{i=1}^n \overline{\sigma_i^o}} \quad (\text{A.1})$$

where  $t_i$  is the time of the  $i^{th}$  hit measurement and  $\widehat{\sigma}_i^o$  is the  $\mathcal{B}$ -corrected  $\sigma^o$  measurement of the  $i^{th}$  hit of the pixel.  $\widehat{\sigma}^o$  is computed by using the  $\mathcal{B}$  values from the SIR images and plugging them into the following equation,

$$\widehat{\sigma}_i^o = \sigma_i^o + \mathcal{B}_k(\theta - 40^\circ) \quad (\text{A.2})$$

where  $\sigma_i^o$  is the  $i^{th}$  measurement of the  $k^{th}$  pixel. The results of this algorithm is expressed as a SIR image, like that shown in Fig. A.5(b), called the time index image.

For the simulation previously considered, notice in the algorithm that as time increases from the beginning of the time range to the end (in this case 6 days) the power decreases (or increases for a growing edge) over the period where the edge motion occurs. The value of each pixels in the time index image will tend towards the time over which the highest  $\sigma^o$  values occurred over the six days, in this case, the first three days. This can be interpreted as shifting the average of the 6 day time period towards the first 3 days. In the algorithm, where no motion occurs, the time index image is the temporal average of the measurements.

To illustrate this, consider Fig. A.5 which shows a simulated region with a receding edge (i.e., the edge is moving from left to right) at a rate of 10 pixels per day. Figure A.5(a) shows the SIR image, while Fig. A.5(d) shows the filtered truth image at day 3. Notice that the SIR image in Fig. A.5(a) shows the edge to the left of the temporal average at day 3. The time index image, Fig. A.5(b), shows that

over the region where the edge movement occurred, the “effective” time is 2.2 days instead of three. This can be best seen in Fig. A.6(c), which shows a horizontal slice of the time index image. Over the region where there was movement the time index image shows that the edge corresponds to a temporal average around 2.2 days. Figure A.5(c) shows the 2.2-day filtered temporal average of the edge. This appears to give a closer estimate of where the edge of the SIR image lies. In Fig. A.6 we see this clearer. Fig. A.6(a) shows the horizontal slice of the simulated SIR image, the filtered temporal average image at day three and the linear average image. Figure A.6(b) shows the linear average image shifted to a center point that corresponds to an average day of 2.2.

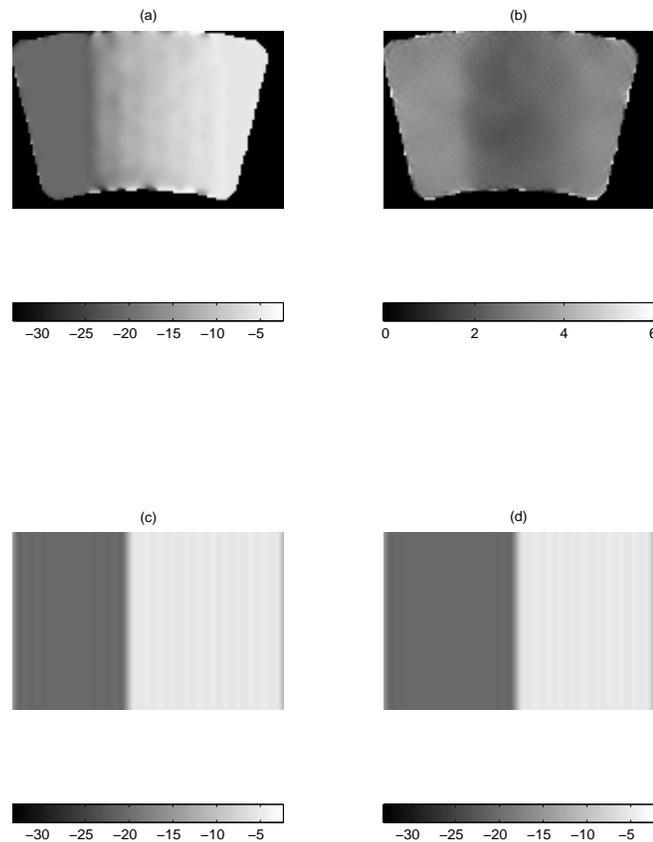


Figure A.5: (a) The simulated SIR  $\mathcal{A}$  image (b) Time index image in days (c) Filtered image at day 2.2 (d) Filtered image at day 3.

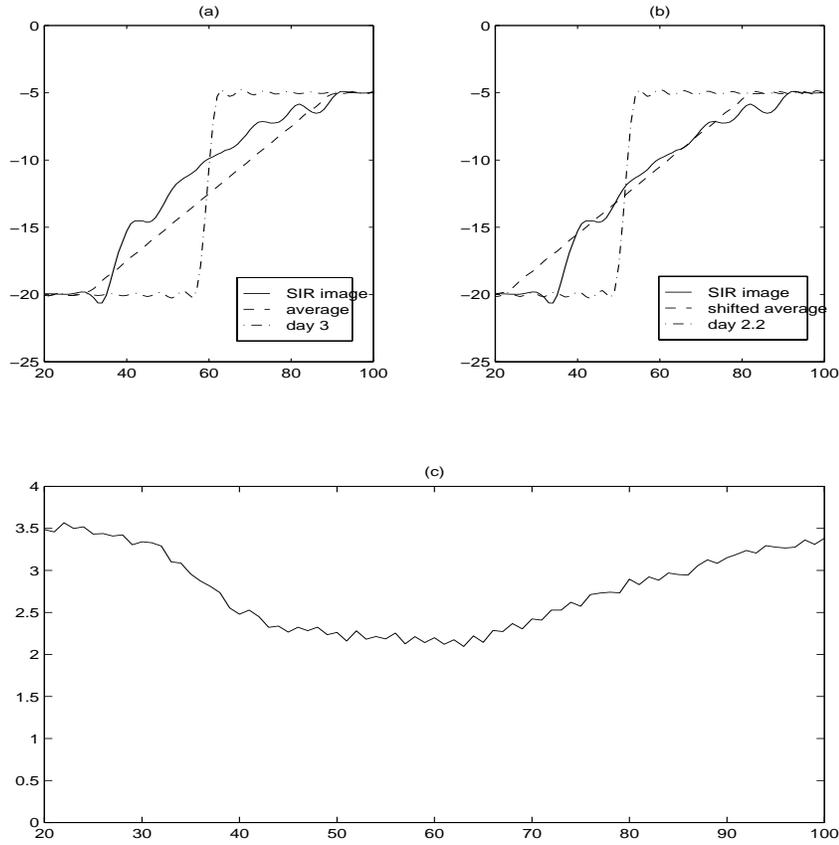


Figure A.6: (a) Horizontal slice of simulated SIR  $\mathcal{A}$  image, filter image at day 3 and shifted average image with center corresponding to day 2.2 (b) Horizontal slice of simulated SIR  $\mathcal{A}$  image, filtered image at day 2.2 and shifted linear average image with center at day 2.2 (c) Time index image in days.

#### A.4 Non-temporally Uniform Measurements

The simple case given above is where the time of the measurements used to create the image is evenly or uniformly distributed throughout the 6 day period. We now consider how the algorithm behaves when the measurements are distributed differently. Figures A.7, A.8, A.9 and A.10 show the results of the algorithm with 4 different types of temporal distributions. For all of these cases the edge moves over the same distance at the rate of 10 pixels per day. In Fig. A.7, the times of

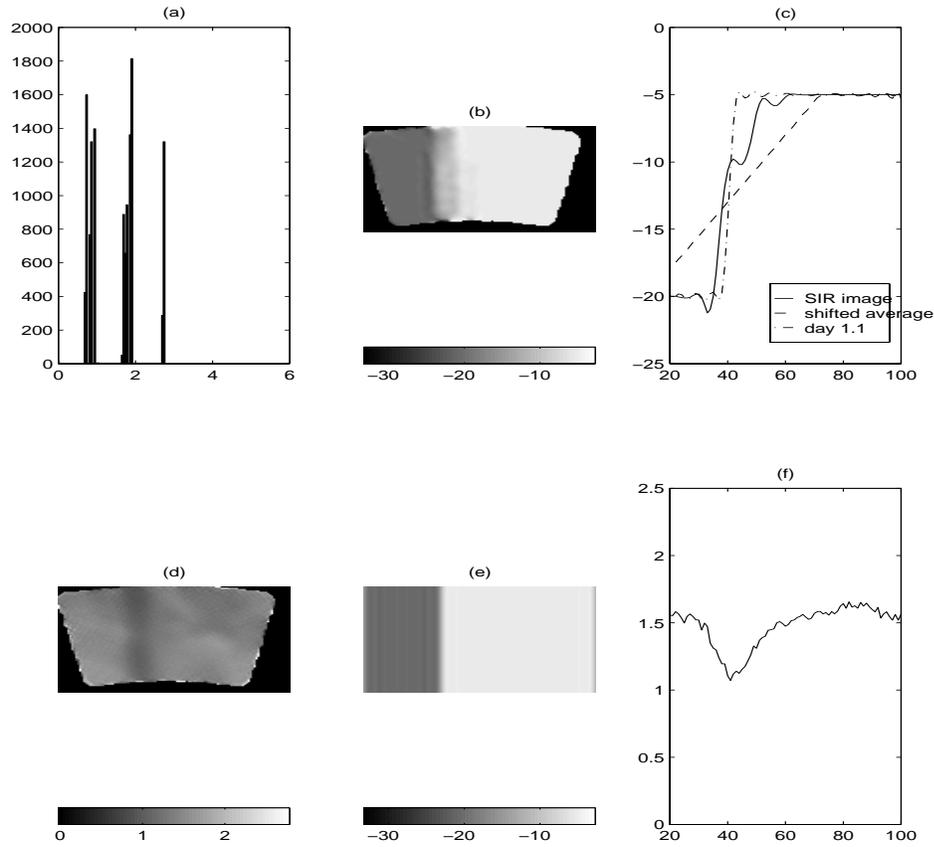


Figure A.7: (a) Histogram of the time of the measurements used to create image (b) SIR  $\mathcal{A}$  image (c) Horizontal slice of the SIR image, the filtered image at day 1.1, and the shifted linear average image with center at day 1.1 (d) Time index image in days (e) Filtered image at day 1.1 (f) Horizontal slice of the time index image

the measurements used to create the image occur in the first three days of the 6 day period. The time index image, Fig. A.7(d), shows that over the area of edge movement, the value of the image is around 1.1 days, while over the area where there was no movement, the value of the image is around 1.5 days, which corresponds to the temporal average of the measurements A.7(a). This is different than the temporal average of the edge for a uniform distribution of measurements, which occurs around three. Notice that in Fig. A.7(c), the edge of in the SIR image corresponds very closely with the temporal average image of day 1.1. The linear average image shifted to correspond to day 1.1 is also shown. Figure A.7(e) shows the temporal average

image at day 1.1, and Fig. A.7(f) shows the horizontal slice of the time index image. Figure A.8 shows the results when the measurements occur between days two and four, and Fig. A.9 shows the results for measurements between days 3 to 6. Finally, Fig. A.10 shows a ‘random’ distribution of measurements between the 6 days. It is interesting to note that when the measurements are distributed through out the 6 day period, as in Fig. A.10, the edge of the SIR image corresponds more closely with the shifted linear average image, while if the distribution of measurements are clustered around a few days, the edge of the SIR image will more closely resemble that of the temporal average image. Notice that for all four cases, while the temporal average of the edge for a uniform distribution of measurements occurs around day three, the time average algorithm depicts more accurately the location of the edge in the SIR image.

## A.5 Conclusion

Because of the non-linear aspects of the SIR algorithm, using the center of the imaging interval is often not sufficiently accurate in determining the temporal location of an image feature in the SIR image. The time-weighted average algorithm has been developed to better determine the effective temporal reference points. The value of the time-weighted average image over the area of movement tends to the direction of brightest measurements. This holds true for a variety of measurement distributions.

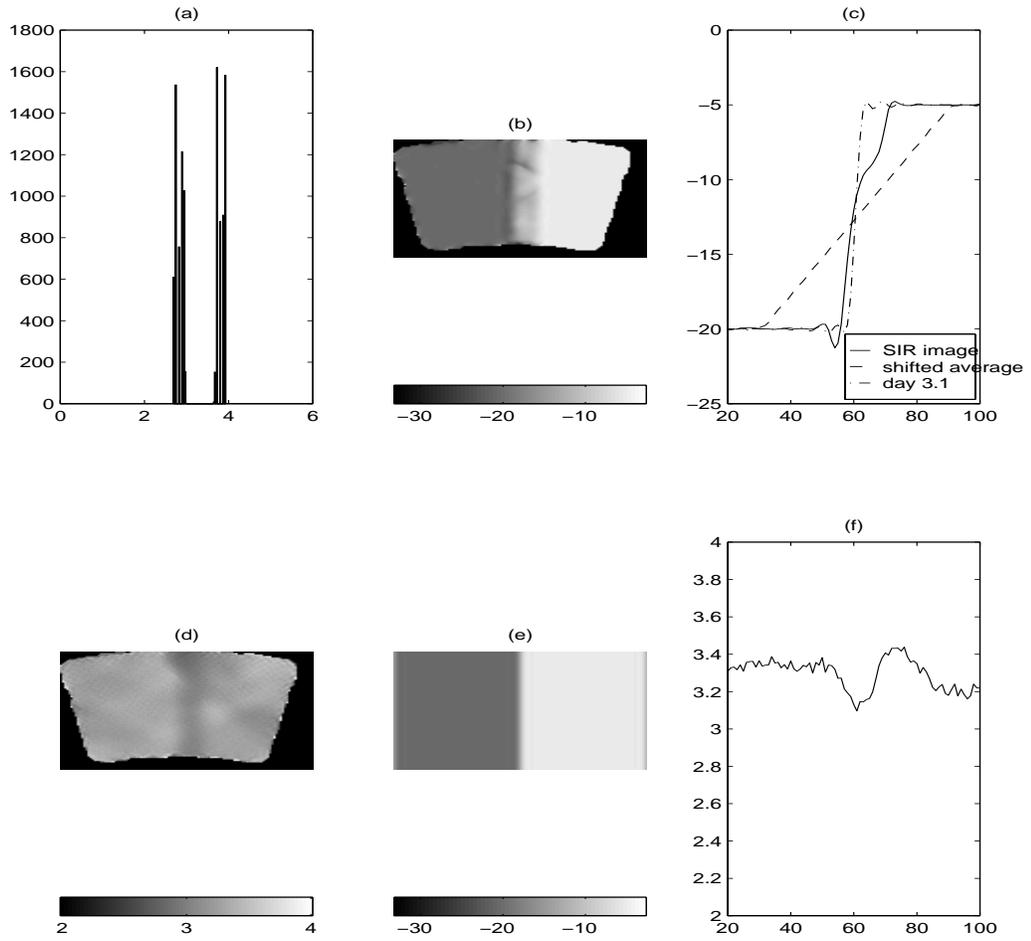


Figure A.8: (a) Histogram of the time of the measurements used to create image (b) SIR  $\mathcal{A}$  image (c) Horizontal slice of the SIR image, the filtered image at day 3.1, and the shifted linear average image with center at day 3.1 (d) Time index image in days (e) Filtered image at day 3.1 (f) Horizontal slice of the time index image

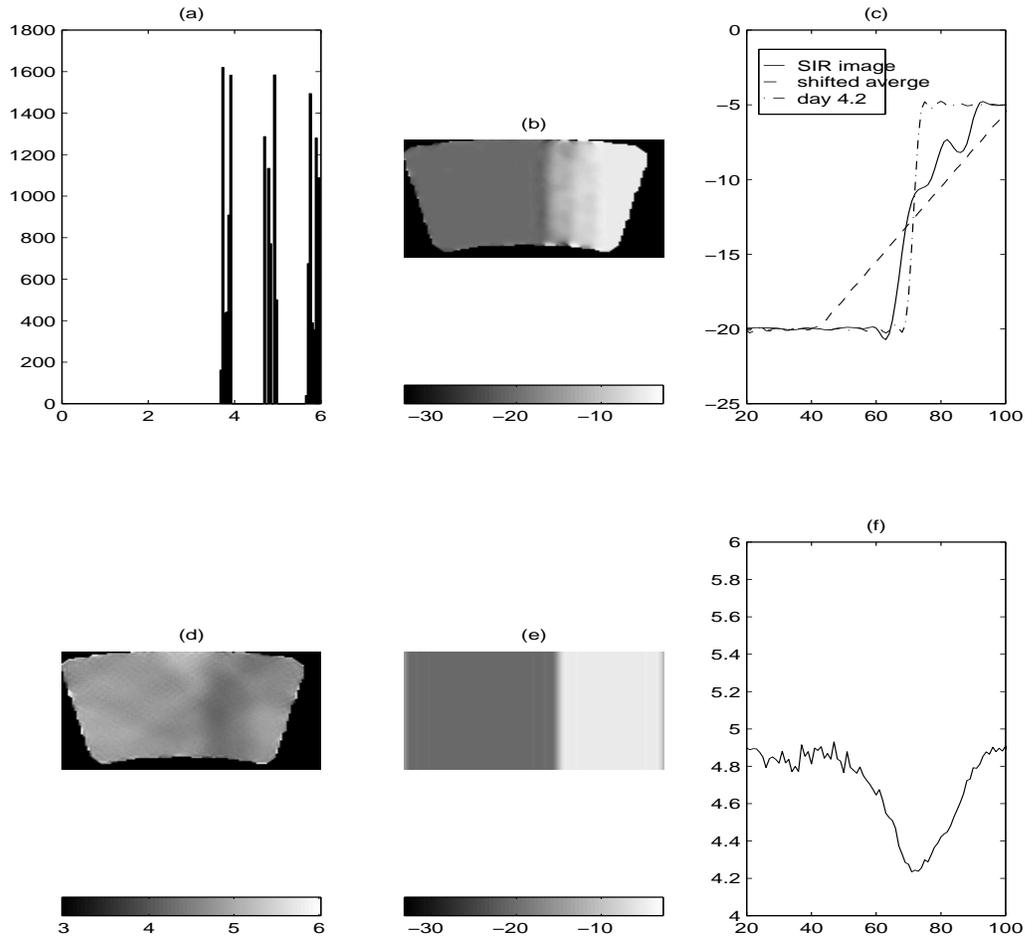


Figure A.9: (a) Histogram of the time of the measurements used to create image (b) SIR  $\mathcal{A}$  image (c) Horizontal slice of the SIR image, the filtered image at day 4.2, and the shifted linear average image with center at day 4.2 (d) Time index image in days (e) filtered image at day 4.2 (f) Horizontal slice of the time index image

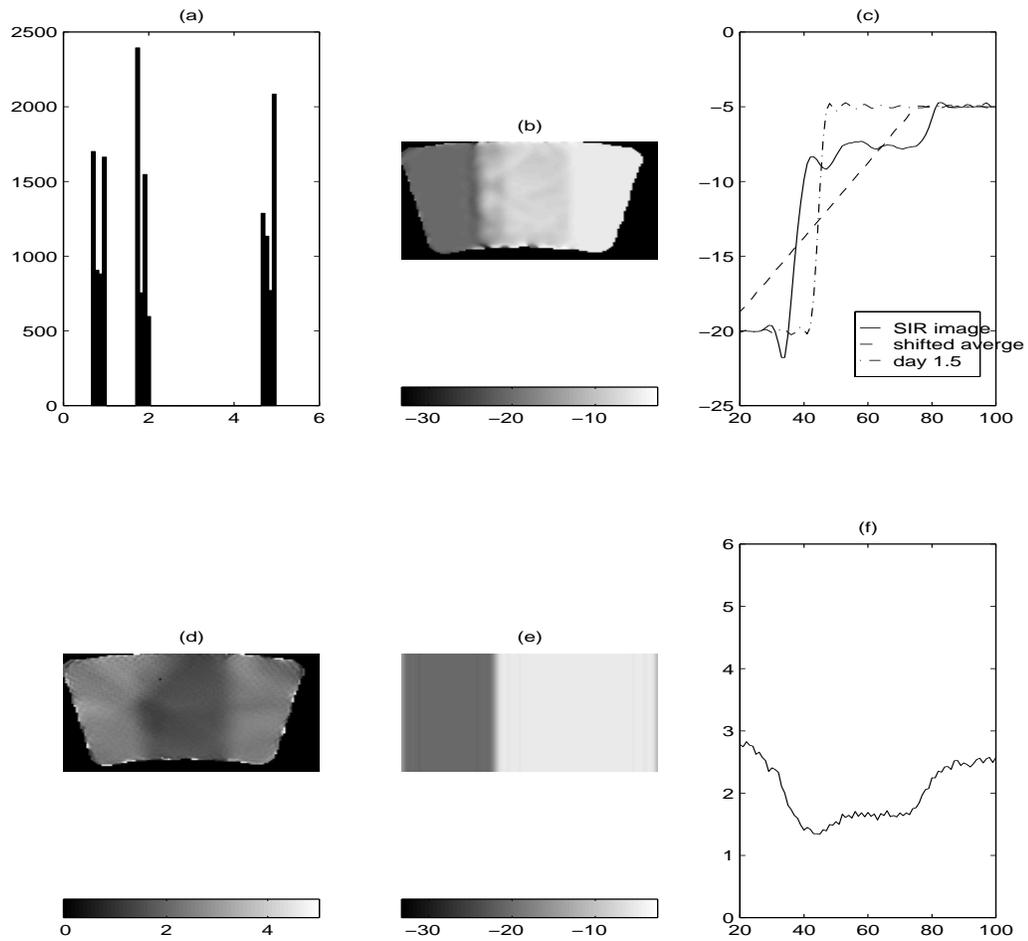


Figure A.10: (a) Histogram of the time of the measurements used to create image (b) SIR  $\mathcal{A}$  image (c) Horizontal slice of the SIR image, the filtered image at day 1.5, and the shifted linear average image with center at day 1.5 (d) Time index image in days (e) filtered image at day 1.5 (f) Horizontal slice of the time index image



## Bibliography

- [1] *Thawing Processes During Siberian Spring Observed by ERS Scatterometer and SAR*, 4-8 August 1997.
- [2] *Land Applications of ERS-1 Wind Scatterometer in Boreal Forest Zone*, 4-8 August 1997.
- [3] *Analysis of the Canadian Boreal Forest using Enhanced Resolution ERS-1 Scatterometer Imagery*, 27-31 May 1996.
- [4] J. T. Pulliainen, T. Manninen, and M. T. Hallikainen, “Application of ERS-1 wind scatterometer data to soil frost and soil moisture monitoring in the boreal forest zone”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 3, pp. 849–863, May 1998.
- [5] D. G. Long and M. R. Drinkwater, “Greenland observed at high resolution by the Seasat-A scatterometer”, *Journal of Glaciology*, vol. 32, no. 2, pp. 213–230, 1994.
- [6] D. G. Long and P. Hardin, “Vegetation studies of the amazon basin using enhanced resolution seasat scatterometer data”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 2, pp. 449–460, May 1994.
- [7] *The Assessment of Green Vegetation Coverage of Grasslands Using AVHRR Data*, 6-10 July 1998.
- [8] “Global land cover mapping using AVHRR data”, <http://www.geog.umd.edu/landcover/lc/projects.html>.

- [9] *Combined multi-temporal ERS-1/Jers-1 imagery for land-cover classification*, 6-10 July 1998.
- [10] E. Nezry, E. Mougin, A. Lopez, J. P. Gastellu-Etchegorry, and Y. Laumonier, “Tropical vegetation mapping with combined visible and SAR spacebourne data”, *International Journal of Remote Sensing*, vol. 14, no. 11, pp. 2165–2184, July 1993.
- [11] E. P. Green, C. D. Clark, P. J. Mumby, A. J. Edwards, and A. C. Ellis, “Remote sensing techniques for mangrove mapping”, *International Journal of Remote Sensing*, vol. 19, no. 5, pp. 935–956, March 1998.
- [12] G. A. Carpenter, M. N. Gjaja, S. Gopal, and C. E. Woodcock, “ART neural networks for remote sensing: vegetation classification from landsat TM and terrain data”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 2, pp. 308–325, March 1997.
- [13] *Knowledge-based classification of SAR images*, Aug 18-21 1993.
- [14] *Classification of wetland vegetation by texture analysis methods using ERS-1 and JERS-1 images*, Aug 18-21 1993.
- [15] R. Benedetti, P. Rossini, and R. Taddei, “Vegetation classification in the middle mediterranean area by satellite data”, *International Journal of Remote Sensing*, vol. 15, no. 3, pp. 583–596, February 1994.
- [16] E. Attema, “The active microwave instrument on-board the ERS-1 satellite”, *Proceedings of the IEEE*, vol. 79, no. 6, pp. 791–799, June 1991.
- [17] F. M. Naderi, M. H. Freilich, and D. G. Long, “Spaceborne radar measurement of wind velocity over the ocean– an overview of the NSCAT scatterometer system”, *Proceedings of the IEEE*, vol. 79, no. 6, pp. 850–866, June 1991.
- [18] D. G. Long, P. Hardin, and P. Whiting, “Resolution enhancement of spaceborne scatterometer data”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 3, pp. 700–714, May 1993.

- [19] *Seasonal Comparison of C-band and Ku-band Backscatter Coefficients over Vegetation Regions*, July 1998.
- [20] “National climate data center”, <http://www.ncdc.noaa.gov>.
- [21] E. Matthews, “Global vegetation and land use: New high resolution data bases for climate studies”, *Journal of Climatology and Applied Meteorology*, vol. 22, pp. 474–487, 1983.
- [22] R. Nishii and S. Tanaka, “Accuracy and inaccuracy assessments in land-cover classification”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 491–497, January 1999.