A STUDY OF THE SCATTEROMETER IMAGE RECONSTRUCTION ALGORITHM AND ITS APPLICATIONS TO POLAR ICE STUDIES

by

David S. Early

A dissertation submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

Brigham Young University

May 1998

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a dissertation submitted by

David S. Early

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

Date	David G. Long, Chair
Date	David V. Arnold
Date	Michael A. Jensen
Date	Gayle F. Miner
Date	Michael D. Rice

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the dissertation of David S. Early 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

David G. Long
Chair, Graduate Committee

Accepted for the Department

Michael D. Rice
Graduate Coordinator

Accepted for the College

Douglas M. Chabries
Dean, College of Engineering and Technology

ABSTRACT

A STUDY OF THE SCATTEROMETER IMAGE RECONSTRUCTION ALGORITHM AND ITS APPLICATIONS TO POLAR ICE STUDIES

David S. Early

Electrical and Computer Engineering

Doctor of Philosophy

A spaceborne microwave scatterometer is designed for the remote sensing of near-surface ocean winds, but rapid repeat coverage and immunity to meteorological effects makes it ideal for studying polar ice as well. Unfortunately, the intrinsic resolution of the scatterometer is insufficient for many geophysical studies.

Using multiple, overlapping passes of the scatterometer, the scatterometer data can be reconstructed into enhanced resolution images with resolutions better than the intrinsic 25-50km instrument resolution. This research establishes a theoretical backdrop for resolution enhancement of scatterometer data using the irregular sampling characteristics of multiple passes of the scatterometer. An optimization of the previously developed SIR algorithm for the ERS-1 scatterometer is presented. Issues of algorithm initialization, cubic estimation and acceleration of the iterative updates are addressed. The work is concluded with an application of the SIR resolution enhancement algorithm to Southern Ocean sea ice, including an examination of azimuthal modulation of backscatter and a simple sea ice classifier.

ACKNOWLEDGMENTS

I express my thanks to Dr. David Long for his guidance and support from before I began my graduate studies through to the bitter end. Also, Dr. Mark Drinkwater of the Jet Propulsion Laboratories provided much needed direction and experience in the later stages of my research. His encouragement was welcome. Dr. Perry Hardin was also very supportive and gave me some much needed encouragement in the middle of my studies. His faith in me and sound advice were gratefully received.

Most importantly, I thank my wife, Karen, for the moral and physical support during these hectic years in graduate school. This work is dedicated to her and to my parents, who kept telling me I had better finish my Ph.D. or else.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Remote Sensing and the Environment

Of growing concern today is the possibility and implication of rapid climate changes induced by industrialization, pollution, and the dumping of waste materials. Contributing to the controversy that surrounds these issues is a fundamental lack of knowledge of how the Earth acts as a system, and thus how man's actions will perturb that system. We know relatively little about how the Earth's different systems interact to produce weather on a global scale. Models used for predicting local and national weather patterns are not always correct, and are of little use for long term climate forecasts. The farther in the future our present models look, the less accurate they become.

More complex models are being developed to not only increase the accuracy of weather predictions, but to incorporate what we do understand of the Earth as a system and the influences of things like pollution, CFCs and deforestation. But even using models developed in recent years, results of different models vary and can predict either a rise in average global temperature, no change at all, or even a drop in average temperature, depending on assumptions made in the model about how the Earth's various systems react to human activities. Accurate modeling requires understanding the Earth as a system, and the range of predictions from current theories on global climate change reflect the fact that our understanding is still imperfect.

Experience has shown that certain parts of the Earth's system are more sensitive to human activities and to global climate change. Deforestation of significant

portions of the Amazon rain forest has raised speculation about whether this will affect carbon dioxide levels in the atmosphere. If levels of this greenhouse gas increase, some climate models predict significant temperature changes all over the globe. Changes in atmospheric layers, like the ozone hole that has developed over Antarctica, have been linked to man-made chemicals released into the atmosphere. Another, less well known indicator of climate change is Arctic and Antarctic ice. The vast, time-varying ice fields at the North and South poles are very sensitive to temperature changes and influence ocean circulation and global weather patterns.

Over the past three decades, the scientific community has had an increased interest in polar ice and the role it plays in atmospheric and oceanic processes. Examples of the influence of polar ice include:

- Polar ice radically changes the albedo of the Earth's surface, modifying the planetary radiation budget.
- Polar sea ice acts as an insulating layer between the warmer ocean and the cooler atmosphere (heat loss or absorption through an ice layer is 100 times slower than from open water).
- Ocean circulation is affected by the formation and melting of sea ice because of heat transfer and thermohaline effects [1].
- Both glacial and sea ice masses act as heat sinks and sources for atmospheric heat, affecting atmospheric circulation.

Because of the effects of polar ice on atmospheric and ocean circulation, the interannual variations and intraseasonal dynamics of sea ice in the Northern and Southern hemispheres contribute to the climate of the entire globe. Ocean currents, wind patterns and storm systems are all affected by the presence and condition of sea ice. Careful and detailed monitoring of polar ice is a key element in understanding and monitoring global climate change. As effects of human actions are propagated through the Earth's weather system, subtle changes in the sea ice may allow us to

monitor the severity of the affects of pollution, deforestation and industrialization on the global climate and in which direction – warmer or cooler – we are heading.

One of the most cost efficient ways of monitoring polar ice characteristics is remote sensing. Satellite remote sensing avoids the cost and hazards of *in situ* measurements, and has the potential to provide a much larger data base than could ever be achieved by *in situ* measurements alone. Remote sensing platforms allow us to synoptically view polar ice in space and time in order to study polar ice in detail.

There are three space-based microwave instruments which have been widely used to study polar ice: the synthetic aperture radar (SAR), the radiometer, and more recently, the scatterometer. The SAR is an active radar that has high resolution but relatively poor spatial coverage. Radiometers are passive radar systems and are capable of scanning large areas in a short period of time, but have intrinsicly low resolution and are significantly influenced by atmospheric effects (e.g., clouds and precipitation). The third spaceborne microwave instrument, the scatterometer, is designed to study near-surface ocean winds [2]. The scatterometer has low resolution, but has wide and frequent spatial coverage and is not significantly affected by atmospheric processes. The recent development of a resolution enhancement algorithm for scatterometer data, the Scatterometer Image Reconstruction (SIR) algorithm [3], has made possible the generation of medium-scale resolution images of polar sea and glacial ice and promises to fill the gap between SAR and radiometer imagery. The purpose of this research is to expand the utility of the spaceborne microwave scatterometer in polar ice studies.

1.2 Review of Polar Remote Sensing

Polar ice studies using remote sensing systems are currently being carried out using a wide variety of instruments and platforms. Radiometers, scatterometers and SAR systems have all been used in various frequency bands on platforms ranging from sleds and ships to airplanes and satellites. Platforms mounted near the earth's surface, such as ship-mounted radars, play an important role in providing detailed ground truth studies, coupling radar measurements with data about surface

conditions. Space-based platforms offer the most comprehensive spatial and temporal coverage, although the resolution is much less than similar systems flown on airplanes or other airborne systems. SAR systems overcome some of the resolution issues with images with scales on the order of meters. However, high data rates and spacecraft power capabilities associated with the SAR can limit SAR spatial coverage in remote areas like Antarctica. This research is directed towards developing a compromise between high resolution and frequent, wide-area coverage.

1.2.1 What Is the Goal of Polar Remote Sensing?

Polar ice is a very complex material. A photograph of sea ice near the Antarctic ice pack edge is shown in Figure 1.1. Remote sensing of polar ice would be useless unless geophysical parameters (e.g. ice type, ice concentration, open water concentration, ice age and thickness) can be extracted from the satellite data. According to [4], there are ten desirable geophysical parameters associated with polar sea ice. In the following list, a desired error limit has been included for some properties:

- Sea Ice extent (to within ± 5 km)
- Sea Ice concentration (open water percentage: $\pm 0.5\%$ over 100km)
- Sea Ice growth/melt rates
- Sea Ice-thickness distributions
- Sea Ice motion (± 1 km/day)
- Ice-type distribution
- Ice properties (e.g. polynyas, ridges, mesoscale oceanic features, etc.)
- Snow cover
- Ice roughness
- Internal Stress

Figure 1.1: Photograph of pancake ice taken near the Antarctic ice pack edge. Note the edges on the pancakes which are formed by wind and wave action forcing pancakes together and piling up the edge. Courtesy of Dr. Mark Drinkwater, JPL.

Each parameter has a use for determining sea ice characteristics that may affect the climate and is valuable for scientific study of polar ice processes. However, there is little consensus on the accuracies needed to adequately model there effects on climate processes. Also, basin wide characteristics are needed so that all the ice parameters are included in any model. Note that the parameters that are specified on scales of kilometers, not meters. SAR imagery provides scales on the order of meters, which is much higher resolution than currently required for most studies. Coarse resolution instruments have potential for studying these parameters. The main spaceborne sensors (SAR, Radiometer, Scatterometer) are compared in the following sections.

1.2.2 SAR

Synthetic aperture radar (SAR) uses ground processing to obtain resolutions on the order of several meters, much finer than the nominal resolution of the antenna used. The best resolutions available from satellite SAR systems range from 10m per pixel to 100m per pixel, making SAR imagery the most detailed of any space-based platform. SAR image products have proven a valuable tool for studying polar regions. Figure 1.2 shows a SAR image that has been classified into four categories using the Alaska SAR Facility (ASF) Geophysical Processor System (GPS). The GPS was developed using data gathered from *in situ* studies to create statistical characterizations of radar backscatter from various polar ice types. Coupled with meteorological data, the GPS uses a minimum distance or maximum likelihood classifier to group statistically similar pixels into the various ice types [5].

Sea ice motion is another product available from SAR imagery. Figure 1.3 illustrates how ice motion and deformation is derived from SAR imagery. First, a grid is overlayed on an 'original' image. This acts as a starting point for the motion/deformation. A second image taken over the same region two or three days later is processed using feature extraction algorithms so that the grid corners from the first image are placed over the same parts of the sea ice pack. The result is a deformation grid as shown in Figure 1.3(b). Using the scale of the image, motion vectors and velocities can be derived.

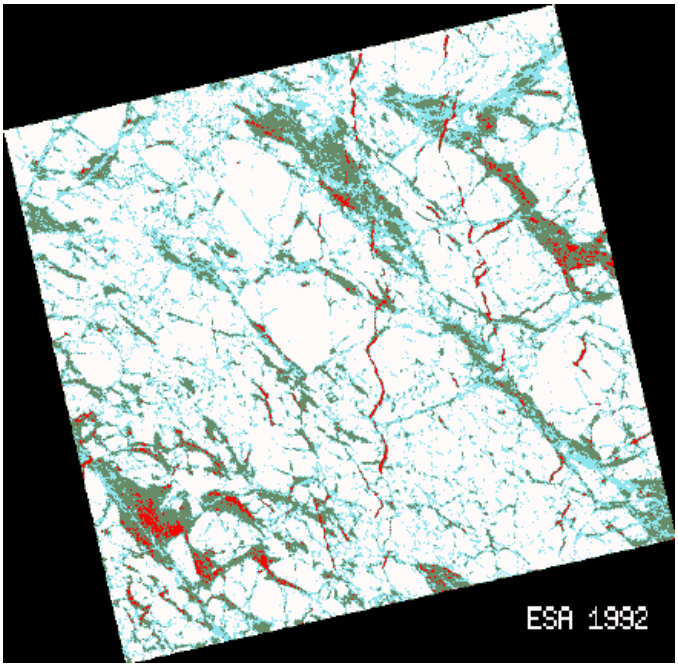
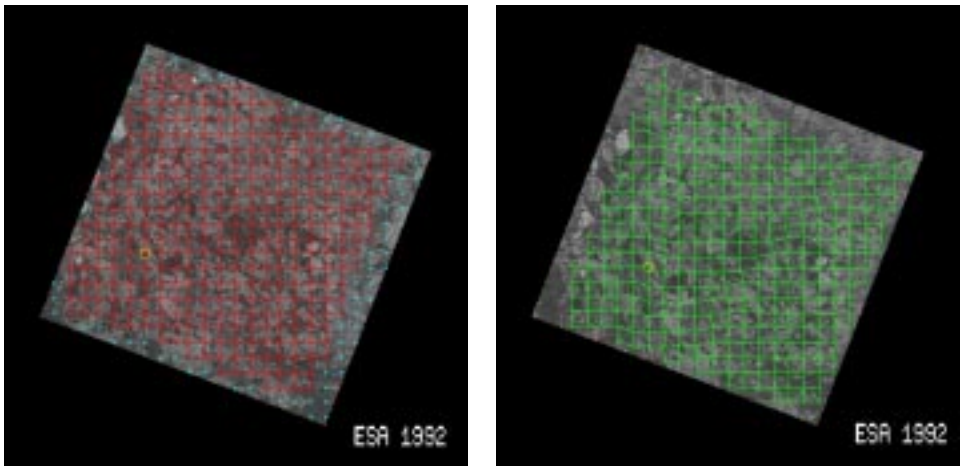


Figure 1.2: Ice classification example using an Arctic SAR image from March 26, 1992. In this image, the following color legend applies: (black) = background; (red) = new ice/open water (1.2%); (green) = smooth first year ice (12.8%); (light blue) = rough first year ice (16.1%); (white) = multiyear (69.8%). Percentages are the total percentage of each type in the image. This image was obtained from the Alaska SAR Facility WWW page. SAR data is copyright ESA 1992.

SAR systems do have limitations. SAR systems generate huge quantities of data for each image and require a receiving station available for concurrent downloading of the collected data. Near populated areas, where ground stations are inexpensive and easier to maintain, this is not a problem, but in Antarctica there are not enough receiving stations to cover the entire continent and existing stations are not online all year long, leaving spatial and temporal gaps in the Antarctic data set. Power capabilities limit the number of images the SAR is capable of generating per orbit. Also, because images are made over a swath of approximately 100km, the area covered in each orbit is limited.



(a) Original Image (b) Deformed Ice Pack

Figure 1.3: These figures illustrate ice motion and deformation using SAR imagery. In figure (a), the grid shows the original position of the ice. Figure (b) shows the same grid deformed so the corners of the grid are over the same features as in figure (a) for an image taken three days after the first. These images were obtained from the Alaska SAR Facility WWW page. Copyright ESA 1992.

1.2.3 Radiometers

Radiometers are passive microwave radar systems [6]. They measure microwave emissions from objects at various frequencies and polarizations. Microwave emissions from ice can be associated with the composition and structure of the sea ice; saline content, age, snow cover and wetness of the ice, all contributing to changes in the emissivity.

Spaceborne radiometers typically have many data channels (e.g. the Scanning Multichannel Microwave Radiometer (SMMR) has 10 data channels: five frequencies each with H-pol and V-pol data streams). However, the high correlation between data channels does not allow ten different parameters to be extracted from the data set. In fact, principle component analysis of SMMR data reveals that there are at most three independent channels of information [7]. The three channels of independent information are used to create a Polarization Ratio (PR) (the normalized difference of two measurements at different polarizations but the same frequency) and

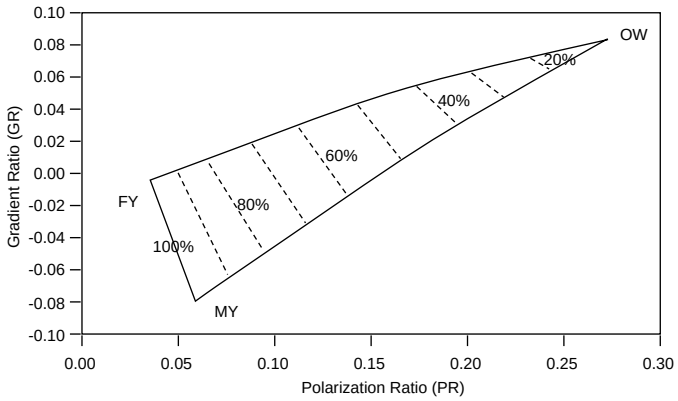


Figure 1.4: Plot showing a geometrical representation of the NASA SMMR algorithm. The polarization and gradient ratios are used to determine the percentage of ice cover by plotting the ratios on the graph in the figure. The percentages listed on the plot are total ice percentages, with open water (OW) as the 0 ice percentage point. The percentage of first year (FY) and multiyear (MY) ice can be calculated based on where the GR/PR point falls within the graph This figure was taken from *Gloersen et al.* [1992], p. 32.

a Gradient ratio (GR) (the normalized difference of two measurements at the same polarization but two different frequencies), which are used in the main classification scheme for radiometric data [8]. Figure 1.4 shows a graphical representation of the algorithm developed in [8]. The concentration of sea ice is determined using the PR and GR. Figure 1.5 shows an example of radiometric data classified by first year ice concentration. Note how well defined the ice edge is in this figure.

Radiometric data has two significant drawbacks. First, the resolution is very low, about 25-50km depending on the frequency used. The intrinsicly low resolution of the instrument limits the utility of the instrument for polar ice studies. Second, the radiometer is very sensitive to atmospheric effects. Clouds and precipitation significantly modify the microwave signature which introduces error into any classification scheme. Examples of this type of data corruption can be seen as dark patches over the open ocean in Figure 1.5. Even though radiometers can cover the polar regions in two or three days, the data is usually averaged over many days or on a monthly basis to remove the effects of clouds and other atmospheric disturbances. Nevertheless, radiometer data is widely used in polar ice studies, even though many of

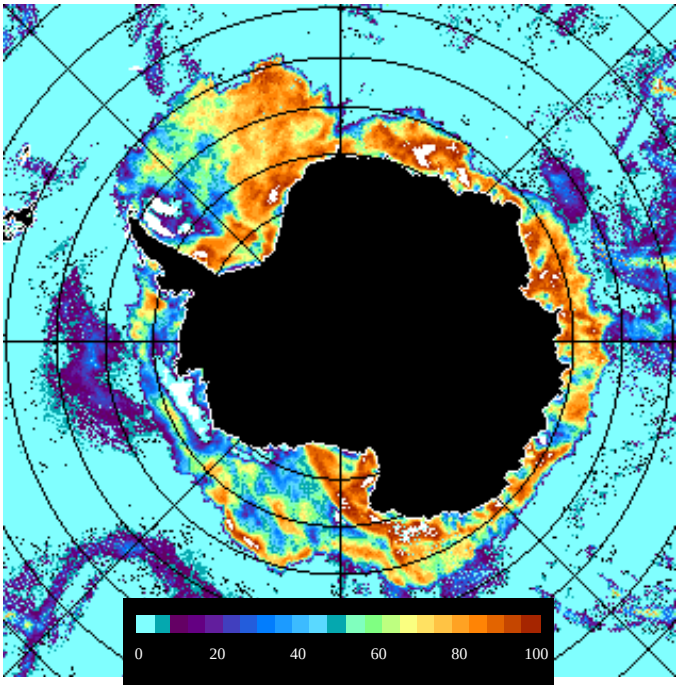


Figure 1.5: A Special Sensor Microwave Imager (SSMI) First year Ice concentration map, JD 120 1994. This image was created using radiometric temperature data and the algorithm developed in *Gloersen et al.*, 1992.

the desired parameters in the list in section 1.2.1 cannot be determined with adequate resolution using radiometer data.

1.2.4 Scatterometer

Scatterometers are active radar systems originally designed to measure near-surface winds over the ocean. Typically, the instruments have nominal resolutions of 25-50 km which is sufficient for the development of ocean wind fields but, like the radiometer, is too coarse for determining most of the desired ice parameters in sections 1.2.1. The parameters in section 1.2.1 require accuracies much finer than the nominal scatterometer resolution. However, because scatterometers can cover large

areas in a matter of two or three days and is relatively immune to clouds and precipitation, there is interest in using the scatterometer for polar ice study. For instance, in [9], ERS-1 scatterometer data is used to create a sea ice mask, the equivalent to finding the ice extent, to eliminate ice infested areas from near surface ocean wind maps. In other studies, the scatterometer is used to study Greenland glacial ice and map the ice faces [10], and to determine rudimentary sea ice type information at the nominal scale of the scatterometer [11, 12].

Because the scatterometer is relatively immune to atmospheric effects, it will be more useful than the radiometer for providing good temporal resolution. Also, the scatterometer provides data from multiple incidence angles which can be used for identifying ice information. For example, in [10] the incidence angle information was used to extract ice information. But, the scatterometer still has a low nominal spatial resolution (25-50km) and spaceborne scatterometers are usually one (single polarization) or two (dual polarization) channel systems, providing much less independent information compared to the radiometer. To overcome the resolution limitation, the SIR resolution enhancement algorithm has been developed for scatterometer data. This algorithm is described in Chapter 2. The ERS-1 scatterometer measurement geometry is described in Appendix A.

1.3 Problem Definition and Contributions

This dissertation presents the theoretical background of scatterometer resolution enhancement and addresses some of the practical issues of applying the previously developed SIR resolution enhancement algorithm to study of polar ice using ERS-1 data, with some data from the short-lived NSCAT scatterometer. The goal is the establishment of a basin-wide monitoring platform capable of rapid and complete coverage of the polar regions at a scale of 10 to 20 km per pixel. The work herein describes a theoretical backdrop for SIR and other reconstruction algorithms, the tuning of SIR specifically for the ERS-1 scatterometer, issues of azimuth modulation which affect the utility of SIR over polar ice, and the application of SIR imagery to create sea ice classified images.

The contributions of this work to the remote sensing community are:

1. The establishment of a theoretical foundation for resolution enhancement for the scatterometer with the possibility of extending this theory to other remote sensing instruments.

2. The qualification of the relationship between aperture function, noise and sample spacing to signal recoverability.

3. The tuning of SIR for the ERS-1 scatterometer.

4. A quantification of the azimuthal modulation over Southern Ocean sea ice.

5. The application of the SIR algorithm to Southern Ocean sea ice classification.

These contributions, in conjunction with past efforts in land scatterometer imaging, establish the validity of the SIR resolution enhancement technique and help justify future research efforts in this field.

1.4 Dissertation Overview

We begin in Chapter 2 with a brief review of the SIR resolution enhancement algorithm through the summary of the material published in [13] and [3]. This provides the backdrop for the theoretical presentation in the next chapter.

In Chapter 3 the theoretical fundamentals of signal reconstruction from irregular samples is presented. Here, the scatterometer measurements are demonstrated to be equivalent to irregular sampling of the surface and a relationship between the measurement aperture function, sample spacing and noise to the frequency content of the recoverable signal is made. Several examples and simulations are presented to support the theory.

In Chapter 4 tuning procedures are presented to aid in refining SIR for the ERS-1 scatterometer. This tuning reduces the error introduced in the signal estimate because of limited iterations in the reconstruction.

In Chapter 5 the azimuthal modulation characteristics of Southern Ocean sea ice are studied. Using actual ERS-1 data, selected regions are studied for variation of the backscatter with azimuth. This is an important consideration for the application of SIR to polar ice because SIR combines multiple measurements with various azimuth angles to create an image.

In Chapter 6, a classification scheme for the SIR imagery is presented. Using *in situ* and satellite measurements, signatures for various ice types are developed and applied to SIR imagery to create a classified sea ice image product.

Finally, in Chapter 7 the conclusions from this work are presented along with an outline of recommended future work to extend the utility of SIR for general surface studies.

Several appendices provide derivations, expanded background material, and some auxiliary results associated with this research. In Appendix A, a description of the ERS-1 satellite geometry is presented. Appendix B presents a previously published analysis of the effects of motion on SIR images. Appendix C presents a comparison of two map projections (one equal area, one conformal) and a derivation of a correction factor that can be used with the conformal projection to allow area comparisons such as those presented in Chapter 6.

Chapter 2

SIR Primer

Although the SAR has high resolution, its narrow swath and high data processing overhead give an edge to scatterometers for basin wide monitoring tasks if the scatterometer data can be reconstructed into enhanced resolution images. Many of the desired parameters in Section 1.2.1 do not require resolution on the order of tens of meters. Resolutions on the order of kilometers are sufficient to characterize many of the parameters listed in Section 1.2.1. Further, these parameters are needed on a basin wide basis. With resolution enhancement, the scatterometer is capable of providing adequate resolution and spatial coverage for most polar studies.

Using a modified algebraic reconstruction technique, an algorithm for increasing the resolution of scatterometer images has been developed [3]. The algorithm is called the Scatterometer Image Reconstruction (SIR) algorithm. If an additional modified median filter is used at each iteration of the algorithm to reduce noise effects, the algorithm becomes SIR with Filter, or SIRF. Originally developed for use with the SEASAT Ku-band scatterometer (1978), data from multiple overlapping cells are algebraically recombined and back projected onto a higher resolution grid. The effective resolution enhancement is a function of how closely the cells are spaced, the noise level and the aperture function. In order to have sufficient cell spacing, several days worth of data must be extracted from the satellite data record for each image. In this presentation, the notation and general structure used in [3] and [14] are adopted to maintain continuity.

2.1 Reconstruction Problem Statement

The scatterometer basically sends out a pulse and measures the backscattered power. The measured backscatter power is represented by σ^o, if in decibels, σ^o_{dB} (The symbol σ^o is also used to refer to the radar cross section. The relationship of radar cross section to backscatter power has led to the adoption within the remote sensing community of σ^o to represent backscatter power). The radar return may have both incidence and azimuth angle dependence, but as shown later in Chapter 5, we may ignore the azimuthal modulation over Antarctic sea ice. However, there will be incidence angle dependence, and data collected over several days will have many different incidence angles and direct comparison of all measurements in a data set necessitates the removal of incidence angle dependence.

Backscatter power is a nonlinear spline which can be modeled by the following non-linear equation [15, 16]:

$$\sigma^o_{dB} = \mathcal{A} + \mathcal{B}(\theta - 40^\circ) + \mathcal{C}(\theta - 40^\circ)^2 + \mathcal{D}(\theta - 40^\circ)^3 + \ldots \tag{2.1}$$

where σ^o is the measure of the backscatter power, $\mathcal{A}$ is the value of σ^o at 40^o and where the $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ and other factors model the higher dimensional dependence of the surface scattering on incidence angle (The spline shape is illustrated in Figure 4.28). Over a limited incidence angle range, between 25° and 55° for the ERS-1 C-band scatterometer, the incidence angle dependence of backscatter is approximately linear in dB. Although the actual incidence angle dependence of σ^o in dB is nonlinear, the linear model works very well for the mid-range incidence angles as observed in the actual and theoretical backscatter results illustrated in [17], [18] and [19] (see also [15, 16]). The linear model is given by the equation

$$\sigma^o_{dB} = \mathcal{A} + \mathcal{B}(\theta - 40^\circ) \tag{2.2}$$

where σ^o_{dB} is the received backscatter power in dB and θ is the incidence angle of the measurement. Note that the measurements are 'normalized' to 40°, i.e. the actual measurement is adjusted to represent a measurement taken at 40°. Although 40° is used here, the data can be normalized to any incidence angle value. Thus $\mathcal{A}$ is the

backscatter power from the surface at 40°, and $\mathcal{B}$ represents the slope, or dependence, of σ^o_{dB} with respect to the incidence angle θ of the measurement. In the reconstruction, we will estimate from the data both the $\mathcal{A}$ and $\mathcal{B}$ values.

Note that we want to determine $\mathcal{A}$ and $\mathcal{B}$ from the measurements, and that we can create an $\mathcal{A}$ image and a $\mathcal{B}$ image for any surface we care to study. If both images are continuous, we have a relatively accurate model of the backscatter characteristics of a given surface within the limitations of the linear model in Eq. (2.2). However, since we are dealing with discrete image reconstruction, we will assume that the $\mathcal{A}$ and $\mathcal{B}$ images are not continuous, but at a higher resolution than the measurements. The goal of the reconstruction algorithm for scatterometer data is therefore the solution of the equation

$$\mathcal{S} \;=\; H(\mathcal{A} + \mathcal{B}(\theta_i - \phi)) + \mathcal{V} \tag{2.3}$$

where

$\mathcal{S}$	is an Nx1 measurement vector (indexed by i)
H	is an NxM point spread matrix (representing the effects of aperture blurring)
$\mathcal{A}$	is an Mx1 pixel vector representing the incidence angle normalized $\mathcal{A}$ image (indexed by j)
$\mathcal{B}$	is an Mx1 pixel vector representing the $\mathcal{B}$ image
$\mathcal{V}$	is an Nx1 noise vector for the noise associated with each measurement
θ_i	is the scaler incidence angle associated with measurement i in $\mathcal{S}$
ϕ	is the scaler normalizing incidence angle (we use 40°, the mean value of the incidence angle for ERS-1)

The terms, $\mathcal{S}$, H, θ and ϕ are known, leaving $\mathcal{A}$ and $\mathcal{B}$ the values to solve for. The noise term $\mathcal{V}$ is included for a complete model and is assumed to be small.

2.1.1 A Note on Noise

In general, the noise in a scatterometer measurement is dependent on the returned backscatter. A noisy measurement can be accurately modeled by the equation

$$\sigma_{\text{noisy}} = (1 + kpN(0,1))\sigma_{\text{true}} \tag{2.4}$$

where kp is an emperically determined percentage dependent on the instrument (for ERS-1, $kp = 5\%$) and $N(0,1)$ is a zero mean unit variance Gaussian random variable.

The dependence of the noise component on the actual measurement greatly complicates the removal of the noise from the system. However, in the material presented herein, we do not attempt to compensate for the noise in the system. Generally, the noise level for scatterometers is quite low, and we will simply ignore it through most of the following discussion. When it is referenced, we will make the simplifying assumption that the noise is additive and a simple Gaussian random variable:

$$\sigma_{\text{noisy}} = \sigma_{\text{true}} + \mathcal{V} \tag{2.5}$$

as in Eq. (2.3). Note that $\mathcal{V} = kpN(0,1)\sigma_{\text{true}}$. In the case where the surface is relatively homogeneous, the noise term is nearly a Gaussian random variable, a fact demonstrated in Chapter 5 (see particularly Figures 5.9 and 5.10). Also, since the original SIR algorithm development assumed an additive independent Gaussian noise component, no attempt is made here to modify that assumption. It will be noted, however, that greater overall accuracy could be obtained if future work were to integrate the dependent noise component into the algorithm.

2.2 The SIR Algorithm

The resolution enhancement algorithm SIR is based on the Algebraic Reconstruction Technique. The Algebraic Reconstruction Technique, or ART, was proposed by Gorden, Bender and Herman in 1971 [20] as a solution to a medical imaging

problem. Computer Aided Tomography (CAT) scans or similar medical imaging systems generate a series of projections through the body. A known signal is sent, and the attenuated signal is received on the other side of the object being imaged. The problem is reconstructing these projections into an image. The basic idea of ART is this: begin with an estimate of the object. In many cases, this will be a constant image, or if some *a priori* knowledge about the object is known, a better first guess can be used. For each measured projection, a simulated measurement is created using the estimated object. The values of the object where the projection passes through are modified to match the measured projection. This process is performed for each measurement to complete one iteration, and then it is repeated iteratively using the updated estimate object.

The formulation of the correction factor used to update the estimate object will determine the type of estimate developed. If an additive correction factor is used (Additive ART), the resulting estimate is a least squares estimate of the image from the data in the limit for infinite iterations [21]. If a multiplicative factor is used (Multiplicative ART), then the estimate is a maximum entropy estimate in the limit [21] [22]. As the correction factors go to zero (Additive ART) or one (Multiplicative ART) with increasing iterations, the algorithm converges for that data set. How convergence is measured and defined for iterative processes is reviewed in [23]. The convergence of ART in a noiseless system has been documented in [20], [24] and [25].

The motivation for the SIR algorithm is the generally poor performance of ART in the presence of noise, especially when applied to the relatively noisy data from the Seasat Scatterometer SASS [13] (Also see Chapter 3). SIR is a modified multiplicative block Algebraic Reconstruction Technique (ART) with square root damping and soft limiting of the update scale factor as described in [3]. To better understand the SIR algorithm, we begin with a brief review of ART, show the development of single variable ($\mathcal{A}$ only) SIR and complete this primer with the extension to the two variable ($\mathcal{A}$ and $\mathcal{B}$) SIR algorithm.

2.2.1 ART

For the presentation of multiplicative ART, the model in Eq. (2.3) will be simplified to a one variable estimate. This is equivalent to assuming no dependence on incidence angle, so the model reduces to

$$\mathcal{S} \;=\; H(\mathcal{A}) + \mathcal{V}. \tag{2.6}$$

Let s_i represent the ith element of the measurement vector $\mathcal{S}$, a_j the jth pixel of the truth image $\mathcal{A}$, and a_j^n the jth pixel of the nth iterative update of the estimate. The individual elements of the matrix H are given by h_{ij}. Multiplicative ART can now be represented as follows: the updated value of the jth pixel due to the ith measurement is given by [13, 21]:

$$a_j^{k+1} \;=\; a_j^k \left(s_i \frac{\sum_{n=1}^{M} h_{in}}{\sum_{n=1}^{M} h_{in} a_n^k} \right) h_{ij}. \tag{2.7}$$

One iteration is defined as one pass through all N measurements s_i, $1 < i < N$. After all measurements have been used to update the a_j^{k+1} pixels, k is incremented and the process is repeated. The term in parentheses is called the *update*. Because this multiplicative factor is used to modify the present pixel value to obtain the new pixel value. For multiplicative ART, the initial image a_i^0 must be non-zero, otherwise the product of the update and previous pixel values will always be zero.

Note that each pixel may be updated several times in each iteration if there are multiple measurements that cover the pixel. It has been shown that ART does not necessarily converge if there is significant noise in the system [20, 21, 25]. By using a simultaneous iterative reconstruction technique (SIRT), the performance of ART in the presence of small amounts of noise can be improved. The idea is relatively simple: All measurements covering a pixel are used to create the update value, effectively averaging out the noise in the system. The modified algorithm is

called block ART and is given by [13, 21]:

$$a_j^{k+1} = a_j^k \left(\sum_{i=1}^{N} s_i \left[\frac{\sum_{n=1}^{M} h_{in}}{\sum_{n=1}^{M} h_{in} a_n^k} \right] h_{ij} \right) \Bigg/ \sum_{i=1}^{N} h_{ij}. \tag{2.8}$$

Let us consider what is happening in this equation. First, note that if the ith measurement covers the jth pixel, h_{ij} is non-zero, so only those measurements affected by the jth pixel are used in the sums. Next, the value in the square bracket is the inverse of the back projection, p_i^k, of the ith measurement:

$$p_i^k = \frac{\sum_{n=1}^{M} h_{in} a_n^k}{\sum_{n=1}^{M} h_{in}}. \tag{2.9}$$

Now define a variable $d_i^k = s_i / p_i^k$, called the *scale factor*. The scale factor is a measure of how close the estimated value of the measurement (p_i^k) is to the actual measurement (s_i). The closer this value is to 1, the less correction is needed. Equation (2.8) thus represents the weighted average of the updates contributed by each measurement covering the jth pixel. Equation (2.8) can be rewritten as:

$$a_j^{k+1} = \frac{a_j^k \left(\sum_{i=1}^{N} d_i^k h_{ij} \right)}{\sum_{i=1}^{N} h_{ij}} \tag{2.10}$$

which represents the algorithmic formulation of block multiplicative ART.

2.2.2 Single Variable SIR

For the development of single variable SIR, we again assume that $\mathcal{B}$ is known and that all the measurements have been normalized to 40° using the linear relationship in Section 2.1. To motivate the development of SIR, consider a situation where the actual measurement s_i is very different than the corresponding back projection p_i^k. The scale factor d_i^k will be either quite large or quite small, causing

a significant change in the estimate image. Noise will be a major contributor to the difference between the back projection and actual measurements.

In order to moderate the size of the scale factor and thus moderate the amplification of noise in the system, a damping factor is introduced and applied to the scale factor. It is desirable to have the damping factor be symmetric about 1 in log space (see Figure 2.1) so that a scale factor less than one is not emphasized more or less than a scale factor greater than one. This in effect spreads the error created by the noise uniformally through the constraint space [13]. It is easily seen that raising the scale factor d_i^k to any power w meets the symmetry requirement. Through a series of subjective evaluations, square root damping was chosen as an 'optimum' damping factor for a system with significant noise, and we continue to use a square root damping in the current study [3, 13].

Experimentation has shown that scale factor damping alone is not sufficient to prevent the system from going unstable in the presence of large amounts of noise [13]. To further stabilize the system, a non-linear smoothing is performed by emphasizing either the back projection or the current estimate in the update. The algorithm is best illustrated by presenting it mathematically. The jth pixel value is calculated by the set of equations

$$a_j^{k+1} = \frac{\sum_{i=1}^{M} update_{ij}^k h_{ij}}{\sum_{i=1}^{M} h_{ij}} \tag{2.11}$$

$$update_{ij}^k = \begin{cases} \left[\frac{1}{2p_i^k}(1-\frac{1}{d_i^k})+\frac{1}{a_j^k d_i^k}\right]^{-1} & d_i^k \geq 1 \\ \frac{1}{2}p_i^k(1-d_i^k)+a_j^k d_i^k & d_i^k < 1 \end{cases} \tag{2.12}$$

$$d_i^k = \left(\frac{s_i}{p_i^k}\right)^{\frac{1}{2}} \tag{2.13}$$

where d_i^k is the square root damped scale factor. Note that when d_i^k is very large or very small (meaning a large change in the current pixel value), the back projection p_i^k influences the update more than the scale factor. Correspondingly, if the scale factor is near 1,then the back projection is deemphasized.

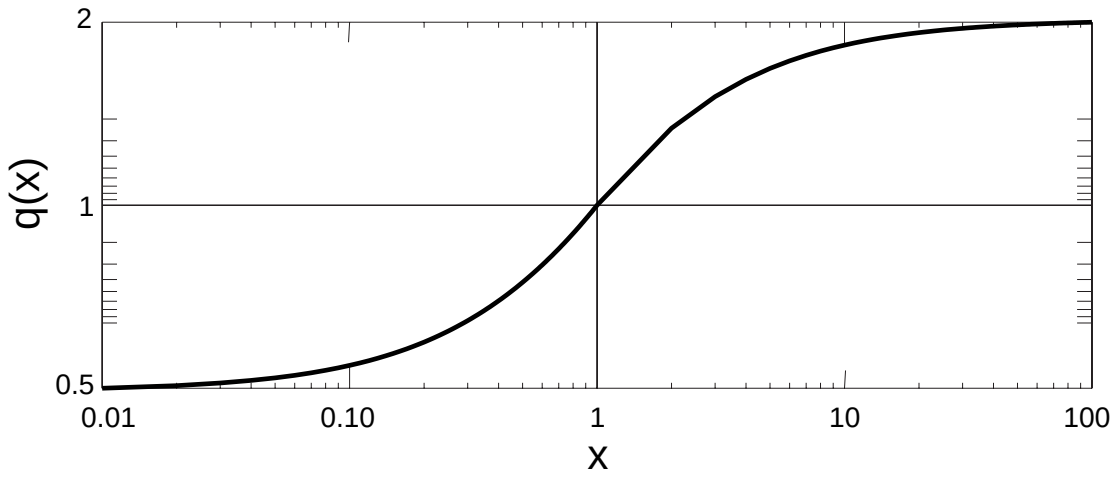


Figure 2.1: A log-log plot of the non-linear scale factor used in SIR. Note the non-linear damping affect added to limit the scale factor to between $\frac{1}{2}$ and 2.

Consider a situation where the back projection is relatively close to the current pixel value, so $a_j^k \approx p_i^k$. Under this assumption, the update term in Eq. (2.12) can be approximated by:

$$update_{ij}^k \approx a_j^k q(d_i^k) \tag{2.14}$$

$$q(x) = \begin{cases} \frac{2}{(1+\frac{1}{x})} & x \geq 1 \\ \frac{1}{2}(1+x) & x < 1 \end{cases} \tag{2.15}$$

where $q(x)$ is the approximation of the non-linear update in Eq. (2.12) (See Figure 2.1). What is important to note here is the behavior of the function $q(x)$, which has limits of $\frac{1}{2}$ to 2. This non-linear limiting of the multiplicative scale factor deemphasizes noise in the resulting imagery, but causes a slowed development of the $\mathcal{A}$ value. The soft-limiting behavior of the function is illustrated in the log-log plot in Figure 2.1.

2.2.3 The Full SIR Algorithm

The main modification made to single variate SIR algorithm for scatterometer image reconstruction is the estimation of the $\mathcal{B}$ term in Eq. (2.3) between

$\mathcal{A}$ estimate iterations. This is essential to the scatterometer image reconstruction since multiple measurements at various incidence angles must be combined to create the image, and if there is any incidence angle dependence it must be accounted for in the $\mathcal{B}$ value. SIR adopts a two stage process for estimating the $\mathcal{A}$ and $\mathcal{B}$ images. First, the $\mathcal{A}$ value is estimated under the assumption that the current $\mathcal{B}$ estimate is correct. Then, the update values from the $\mathcal{A}$ image are used to create an updated $\mathcal{B}$ image.

The $update_{ij}^{k}$ term in Eq. (2.12) is an estimate of the $\mathcal{A}$ value for a given measurement, with the weighted average of all the $update_{ij}^{k}$ terms for a given pixel used as the final iterative update. To obtain the $\mathcal{B}$ estimate, the $update_{ij}^{k}$ values for each pixel are *unnormalized* using the current $\mathcal{B}$ estimate:

$$(\hat{\sigma^o})_{ij}^{k} = update_{ij}^{k} + b_j^k(\theta_i - 40^\circ). \tag{2.16}$$

As the scale factor approaches 1, the $\mathcal{B}$ estimate is simply a linear regression estimate of the $\mathcal{B}$ value from the unmodified measurements (for details on this process, see [13]). The linear regression is performed for each pixel and is given by:

$$\hat{b}_j^k = \frac{\sum_{i=1}^{N} h_{ij} \sum_{i=1}^{N} \theta_i (\hat{\sigma^o})_{ij}^{k} h_{ij} - \sum_{i=1}^{N} \theta_i h_{ij} \sum_{i=1}^{N} (\hat{\sigma^o})_{ij}^{k} h_{ij}}{\sum_{i=1}^{N} h_{ij} \sum_{i=1}^{N} \theta_i^2 h_{ij} - \left(\sum_{i=1}^{N} \theta_i h_{ij} \right)^2}. \tag{2.17}$$

This estimate of $\mathcal{B}$ is approximate, and is influenced heavily by the number and distribution of incidence angles in the measurements covering a pixel. If there are few measurements, or the spread between the incidence angles in the measurements is low, the result may be a spurious estimate of $\mathcal{B}$. To avoid these problems, a weighting is performed for each pixel using the variance of the incidence angles of the measurements covering the pixel. When the variance of the incidence angles is low, the estimate is less likely to be accurate so the current value of $\mathcal{B}$ is emphasized and

the update value deemphasized. The equations are

$$weight = \frac{\sum_{i=1}^{N} h_{ij} \sum_{i=1}^{N} \theta_i^2 h_{ij}}{\left(\sum_{i=1}^{N} \theta_i h_{ij} \right)^2} \tag{2.18}$$

$$b_j^{k+1} = \frac{weight \times \hat{b}_j^k + b_j^k}{1 + weight}. \tag{2.19}$$

The combination of Eqs. (2.11) and (2.19), performed alternately, constitute the SIR algorithm. For a complete explanation of the SIR algorithm, the reader is referred to [3] and [13].

2.3 Summary

The SIR algorithm is a multivariate variation of multiplicative ART with square root damping and a non-linear scale factor. SIR was developed as a two stage, iterative reconstruction algorithm for use with scatterometer data. The single iteration of the algorithm is summarized as follows:

1. **The $\mathcal{A}$ Update:** The $\mathcal{A}$ values are updated first using the equations:

$$a_j^{k+1} = \frac{\sum_{i=1}^{M} update_{ij}^k}{\sum_{i=1}^{M} h_{ij}} \tag{2.20}$$

$$update_{ij}^k = \begin{cases} \left[\frac{1}{2p_i^k}(1 - \frac{1}{d_j^k}) + \frac{1}{a_j^k d_i^k} \right]^{-1} & d_j^k \geq 1 \\ \frac{1}{2} p_i^k (1 - d_j^k) + a_{ij}^k d_j^k & d_j^k < 1 \end{cases} \tag{2.21}$$

$$d_i^k = \left(\frac{s_i}{p_i^k} \right)^{\frac{1}{2}}. \tag{2.22}$$

2. **The $\mathcal{B}$ Update:** The $\mathcal{B}$ update uses the $update_{ij}^{k}$ terms calculated during the $\mathcal{A}$ update phase to calculate the new $\mathcal{B}$ values using the following equations:

$$(\hat{\sigma^o})_{ij}^{k} = update_{ij}^{k} + b_j^k(\theta_i - 40^\circ) \tag{2.23}$$

$$\hat{b}_j^k = \frac{\left(\sum_{i=1}^{N} h_{ij}\right)\left(\sum_{i=1}^{N} \theta_i(\hat{\sigma^o})_{ij}^{k} h_{ij}\right) - \left(\sum_{i=1}^{N} \theta_i h_{ij}\right)\left(\sum_{i=1}^{N} (\hat{\sigma^o})_{ij}^{k} h_{ij}\right)}{\left(\sum_{i=1}^{N} h_{ij}\right)\left(\sum_{i=1}^{N} \theta_i^2 h_{ij}\right) - \left(\sum_{i=1}^{N} \theta_i h_{ij}\right)^2} \tag{2.24}$$

$$weight = \frac{\left(\sum_{i=1}^{N} h_{ij}\right)\left(\sum_{i=1}^{N} \theta_i^2 h_{ij}\right)}{\left(\sum_{i=1}^{N} \theta_i h_{ij}\right)^2} \tag{2.25}$$

$$b_j^{k+1} = \frac{weight \times \hat{b}_j^k + b_j^k}{1 + weight}. \tag{2.26}$$

The influence of noise on the result is improved (compared to multiplicative block ART) through reducing the size of the multiplicative update term in the $\mathcal{A}$ value development based on the relative difference between the back projection and actual measurements and by coupling the incidence angle variance to the $\mathcal{B}$ value update.

Chapter 3

Enhanced Resolution Imaging From Irregular Samples

3.1 Introduction

In this Chapter, we discuss the theory behind resolution enhancement algorithms with specific interest in how SIR achieves an enhanced resolution scatterometer image. Because we wish to apply this theory to scatterometers, we will develop this theory around a model of the surface response that describes the microwave backscatter from a point. We model backscatter from the surface as a function of location and incidence angle. Let $f(x, y, \theta)$ be the function that gives the backscatter from a point (x, y) on the surface at an incidence angle of θ. For brevity, the θ term is dropped and unless otherwise stated, is implicit in the following sections. The measurement system can be modeled by

$$\sigma^o = Hf + \mathcal{V} \tag{3.1}$$

where H is an operator that models the measurement system (sample spacing and aperture filtering), f is the original surface functions defined above, $\mathcal{V}$ is the noise term and σ^o represents the measurements made by the instrument. Note that the set of measurements σ^o are a discrete sampling of the function f. For the moment, noise is ignored. The function $f = \mathcal{A} + \mathcal{B}(\theta - \phi)$ in Eq. (2.3).

The development will be presented as follows: First, a model of the desired solution is presented, followed by a discussion of sampling and reconstruction from sampling. Next, the effects of the measurement aperture function on the sampled data are discussed, and then a theory of reconstruction from irregular samples is presented, including a development of an arbitrarily band limited Banach space. Next,

the discussion of irregular sampling theory is related to the ART (both additive and multiplicative) reconstruction algorithms and the SIR algorithm. Finally, the effects of noise are considered with an emperical review of performance of the ART algorithms and SIR for the noisy case.

3.2 Solution Model

Equation (3.1) offers a very simple but powerful model for the measurement system. Many things can be included in the point spread function H, such as sample spacing and the aperture function. For enhanced resolution image reconstruction, we are interested in the inverse problem:

$$\hat{f} = \hat{H}^{-1}\sigma^o \tag{3.2}$$

where $\hat{f}$ is an estimate of f from the σ^o samples. The inverse of the point spread function, $\hat{H}^{-1}$, is exact only if H is invertible. If H is invertible, then $\hat{f} = f$. In general, the direct inversion of H is not practical. For large areas such as the Antarctic basin, the number of samples will be very high (on the order of 1 million). As such, H, whose size is (number of samples) x (number of pixels in reconstruction) becomes quite large, increasing the computational complexity of a direct inversion. Also, in many cases H is simply not invertible, and we must make an estimate of the inverse. Introduce noise and this further complicates the estimate of the inversion. In the following discussion, we propose the use of an iterative reconstruction, allowing us to control to some degree the computational difficulties, and to reduce the effects of noise on the resulting image.

3.2.1 The Nyquist Sampling Theorem

A traditional approach to sampling and reconstruction is based on the Nyquist sampling theorem. The theorem simply states that a band limited function can be completely reconstructed from regularly spaced samples if the sample rate exceeds the Nyquist rate of twice the maximum frequency in the signal. The reconstruction is a simple low pass filter that matches the band limits of the original

signal, which is equivalent to using a sinc function as an interpolating function (for a detailed explanation, see [26]). A block diagram of this system is given in Figure 3.1 for a case where the measurement system has an aperture function. Sinc interpolation reconstructs the aperture filtered version of the signal, not the unfiltered original. Since all real measurement systems have an aperture function, it is usually designed to act as a prefilter to eliminate high frequency components of the signal that might otherwise cause aliasing in the reconstructed signal based on the sample spacing. This approach is very similar to the ERS-1 scatterometer design: A desired sample spacing of 25 km dictated an aperture function that, to avoid aliasing, eliminates wavelengths smaller than 50 km. Note in this approach that the sample spacing is the single largest limiting factor in the reconstruction.

3.2.2 A Modified Approach

Figure 3.2 illustrates a different approach. Suppose that a system is designed with a sample spacing of 25 km and an aperture function created to minimize aliasing (much like ERS-1). Later, a method of halving the sample spacing to 12 km is developed, but the aperture function remains unchanged. If the aperture function is an ideal low pass filter, then the high frequency information in the original function is unrecoverable and decreasing the sample spacing does not produce any new information. However, most aperture functions have side lobes and, depending on the noise in the system, the side lobes still contain enough information to recover at least some of the higher frequency content of the original signal because of the new sample rate. As illustrated in Figure 3.2, this is equivalent to inverting the effects of the aperture function in the reconstruction in Eq.(3.2). Notice that the main difference between Figures 3.1 and 3.2 is the operator being inverted (H, denoted by the dotted box). In Figure 3.2, the operator includes the aperture function as well as the sampler, and the inversion algorithm for this illustration is more involved than simply applying a low pass filter.

Figure 3.3 presents a chart comparing the two methods outlined above. The rows of the chart show the Fourier transforms of (top to bottom) the original function,

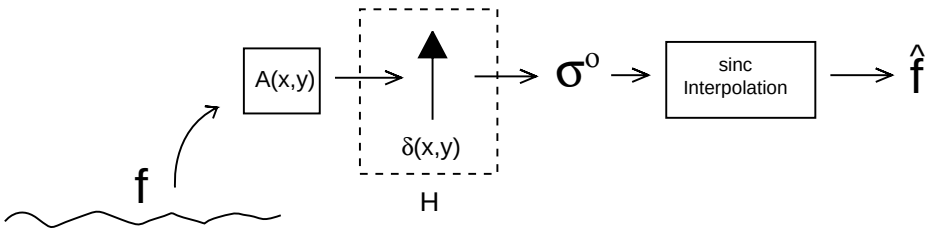


Figure 3.1: A block diagram illustrating a simple Nyquist sampling problem. The original surface, f, is filtered by the system aperture function, $A(x, y)$. This is sampled on a regular grid using the delta function $\delta(x, y)$, giving the measured values σ^o. To recover the surface function, sinc interpolation is used. According to the sampling theorem, this is equivalent to low pass filtering the impulse train created in sampling step. Note that the operator being inverted by the sinc interpolation in this scheme is the sampling step: the signal recovery algorithm does not attempt to invert the effects of the aperture function. In fact, the aperture function is usually designed to band limit the original surface function to avoid aliasing in the recovered signal.

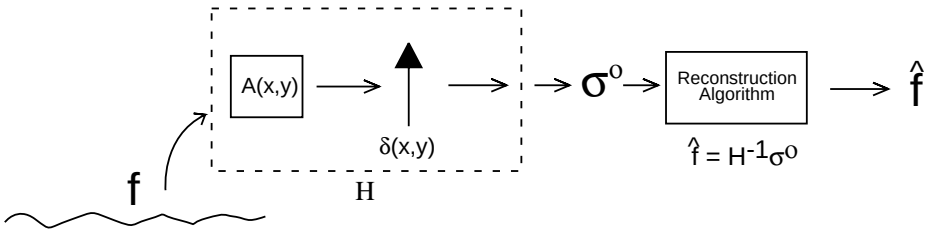


Figure 3.2: A block diagram illustrating a reconstruction algorithm that attempts to invert the effects of the aperture function. Compare this figure with the block diagram in Figure 3.1. The primary difference between the two figures is the operator that is being inverted, denoted in each figure by H and the dotted box. In this scheme, the aperture function filters the surface, but the inversion given by $\hat{f} = H^{-1}\sigma^o$ inverts the effects of the aperture function according to a regularization scheme (e.g., least squares estimate, maximum entropy estimate, etc.).

the aperture function and the reconstructed estimate of the function. Traditional sampling has a desired frequency range dictated by the sample spacing of the system. For the example in Figure 3.3, the aperture function is designed to limit the frequency content of the signal to the recoverable band defined by a given sample spacing. For a reconstruction, the desired frequency range is increased by better sampling, although the aperture function limits the reconstruction due to nulls in the frequency domain and attenuation of some frequencies. The effects of the aperture function are discussed in the next section.

The latter process of reconstruction is used by SIR. SIR utilizes multiple passes of the satellite to build up a dense sampling of the surface. As is illustrated in the following sections, this decreases the sample spacing, allowing the effects of the aperture filter to be inverted using an iterative reconstruction method and thus achieving higher resolution than the nominal instrument resolution.

3.2.3 The Aperture Function

Any real-life sampling usually involves a non-ideal sampler with a finite aperture which affects the frequency content of the sampled data. Since most real aperture functions are window-type functions, the result of applying the aperture is to low pass filter the data. Window-type aperture functions generally have frequency nulls that result in information loss, and this affects the information that can be recovered from samples gathered with the aperture function.

To understand the effect the aperture function has on the original function, we can show that the aperture sampled data is equivalent to ideal delta sampling of a function convolved with the aperture function. While the aperture function may not be circularly symmetric and vary from sample to sample, here we assume it is symmetric and the same for every sample for simplicity. Let the aperture function be represented by $g(x, y)$, and the convolution of two functions be denoted symbolically by $*$. For any sample point, say at a point on the surface (x, y), the sample value is

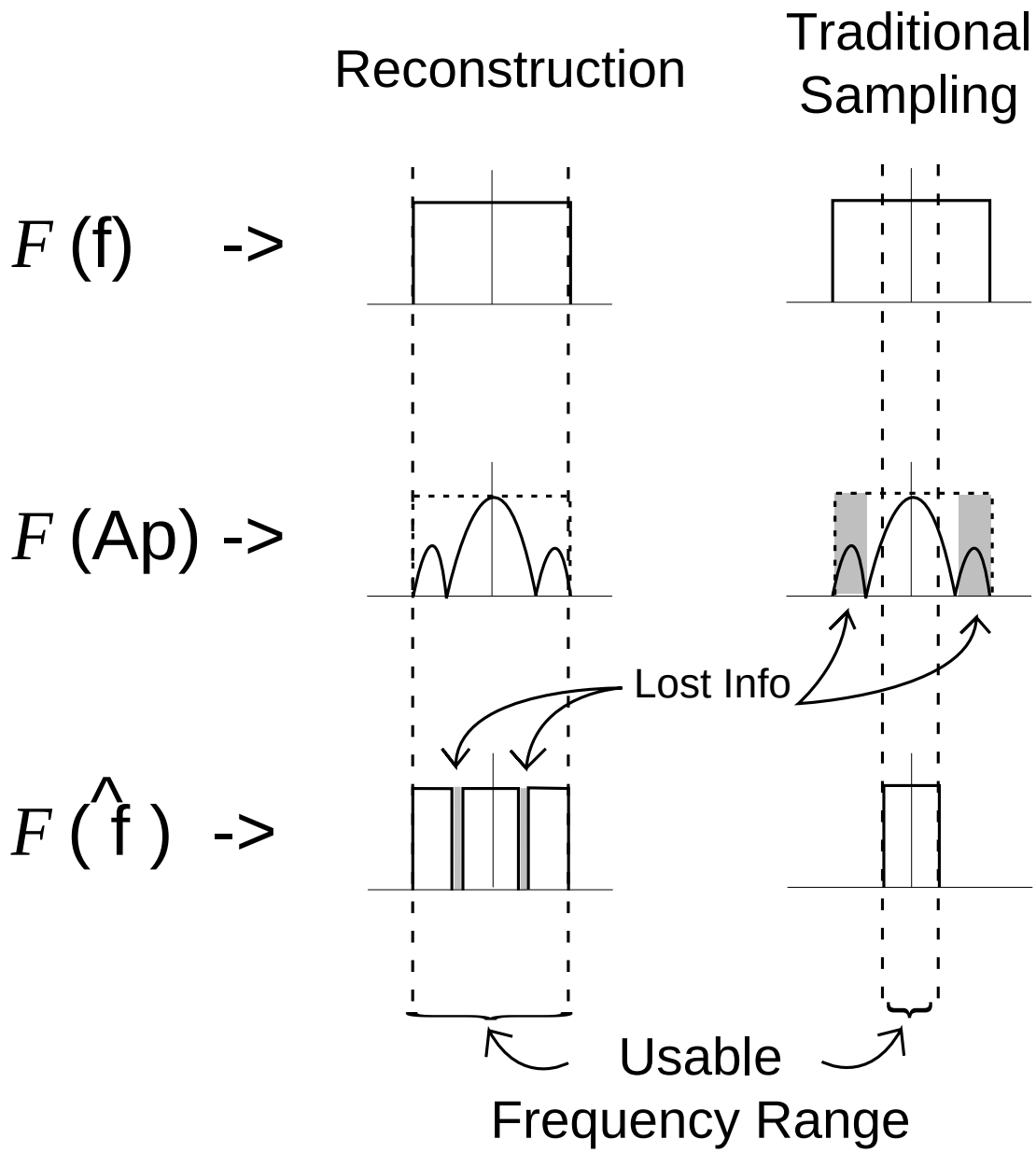


Figure 3.3: This chart shows a comparison of the frequency content at various stages of the reconstruction illustrated in Figure 3.1 and 3.2. The rows of the chart represent the Fourier transforms of (top to bottom) the original function, the aperture function and the reconstructed estimate of the function. In "Reconstruction", the estimate is limited by the nulls in the aperture function frequency response. In "Traditional Sampling", the usable range is limited by the sample spacing. The aperture function is used to filter the signal so that aliasing is minimized.

given by the convolution

$$f_s(x,y) = g(x,y) * f(x,y) \tag{3.3}$$

$$= \iint_A g(x-x', y-y') f(x',y')\, dx'\, dy' \tag{3.4}$$

where $f(x,y)$ is the true surface function and A is the region of support for the aperture function (i.e., the spatial region over which the aperture has non-zero values). Now formulate a new continuous function $h(x,y)$ which is the convolution of the aperture function and the original surface function. Sampling this filtered function with ideal impulses leads to the same sample values as sampling the original surface using a sampler with an aperture function. The value of this new function at a point (x,y) is given by

$$h(x,y) = g(x,y) * f(x,y) = f_s(x,y). \tag{3.5}$$

Note that the resulting sampled set $h_s(x,y)$ is equivalent to the original data sampled with the aperture function as in Eq. (3.3). Thus, from a functional perspective, we can imagine that the measurement cells of a scatterometer are ideal samples of a surface filtered by the aperture function.

The effects of the aperture function are two-fold: First, the low side lobes levels (compared to the main lobe) associated with typical aperture functions (e.g., Hanning or Hamming windows for very low side lobes or rectangular windows with higher side lobes) substantially attenuate high frequency information that we desire to reclaim through an appropriate algorithm. If the noise in the system is high relative to these side lobe levels, the information is lost in the noise since the enhancement algorithm amplifies the noise as well as the desired signal in the reconstruction process. However, as shown later, even very low side lobe levels can be recovered if there is little or no noise in the system. Second, typical aperture have frequency nulls in their spectrum. These nulled frequencies represent lost information no matter how good the reconstruction algorithm. The recoverability of high frequency information in the side lobes is dependent on the amount of attenuation relative to the noise in the system.

3.3 Irregular Sampling Theory

Irregular sampling problems have been examined since the early 1960's (see [27] for a review of the problem and key references). However, in most studies, very limiting requirements are placed on the irregular sampling grid. There are essentially two views on the nature of the irregular sampling grid: The first view is that there is some structure to the grid so that it can be described with some function or periodicity. Included in this class of irregular sample grids are wholly random grids described by a probability distribution. The second view is one of an arbitrary irregular grid parameterized by δ which describes the maximal spacing in the grid.

There are advantages and disadvantages to each viewpoint. One implication of a structured grid is that the density of the samples is relatively uniform over the entire sampling grid, i.e. any given two areas have about the same number of samples in them and the standard deviation of the distance between sample points is relatively small and constant when computed over small sets of the data. This leads to good statistical characteristics of the reconstructed signal relative to the original bandwidth. The main drawback is that satellite data does not generally fit this description of irregular data as the spacing can vary significantly through latitude and from pass to pass.

The arbitrary grid, however, places no restrictions on the structure of the sampling grid. The biggest problem with this description is that the recoverable frequencies are usually defined by the largest sample gap on the grid. Even if samples are closely spaced in all regions but one, that one region with the large sample spacing determines the frequencies completely recoverable by an appropriate algorithm. While a regional description of resolution and frequency content can be employed to reduce the effects of a single bad sample gap, this can be difficult and tedious to apply over large regions with changing sampling characteristics. Nonetheless, the arbitrary grid is prefered in this study because it is a more accurate model of the constantly changing satellite sampling grid.

To establish the validity of reconstruction from irregular samples, some recent work using arbitrary sampling grids is cited. Specifically, Karl Gröchenig [28]

presents a methodology that places no restrictions on the sample grid, allowing for density changes in the sampling grid.

3.3.1 Preliminaries

In order to present the important lemmas and theorems of [28], certain definitions and groundwork must be established. This includes definitions of the spaces and functions used and some auxiliary information not presented in the cited text but that make the development here complete. The definitions in this section are usually for the two dimensional case, since that is the case of interest for satellite remote sensing applications.

First, the definitions of the Hilbert space and the subspace of band limited functions on the Hilbert Space are presented (as in [28]):

DEFINITION 1 ($L^2(R^2)$)
Let $L^2(R^2)$ denote the Hilbert Space of square-integrable functions on R^2 with the norm $\|f\| = (\int_{-\infty}^{\infty} |f(x)|^2\, dx)^{\frac{1}{2}}$.

DEFINITION 2 (Ω)
Let $\Omega \subseteq R^2$ be a compact set where Ω denotes the cube $\prod_{i=1}^{2}[-\omega_i, \omega_i]$. The $\omega = (\omega_1, \omega_2)$ define the **extension** *of Ω.*

DEFINITION 3 ($B^2(\Omega)$)
Let $B^2(\Omega)$ be a closed subset of $L^2(R^2)$ such that $B^2(\Omega) = \{f \in L^2(R^2) : supp\, F \subseteq \Omega\}$ where F is the Fourier Transform of f. $B^2(\Omega)$ is, by definition, a Banach space.

These three definitions provide a mathematical formalism for the theorems to follow. The first definition establishes a general set of well-behaved functions. The term Ω describes a region of support in the frequency domain, bounded by some ω_i (Def. 2), and $B^2(\Omega)$ is the set of functions in $L^2(R^2)$ whose Fourier Transform is contained within the region defined by Ω.

Next are two definitions that are used to describe an operator used in the reconstruction algorithm:

Definition 4 (Bounded Operator)

An operator A is bounded on the space B if there exists a constant C such that

$$\|Au\| \leq C\|u\| \quad \forall u \in B. \tag{3.6}$$

Definition 5 (Operator Norm)

The operator norm, denoted $\|\cdot\|'$ is the smallest constant C such that

$$\|A\|' = sup_{\|u\|=1}\|Au\|. \tag{3.7}$$

Finally, the irregular sampling grid can be described as δ-dense for the one dimensional and two dimensional cases as follows [28]:

Definition 6 (δ-Dense: One Dimension)

*A sampling sequence $X = (x_i)_{i\in Z}, \ldots < x_{i-1} < x_i < x_{i+1} < \ldots$, is said to be δ-**dense** if $sup_i(x_{i+1} - x_i) \leq \delta$.*

Definition 7 (δ-Dense: Two Dimension)

*A sampling sequence $X = (x_i)_{i\in Z}$ in R^2 ($x_i = (\xi_1, \xi_2)$) is $\delta = (\delta_1, \delta_2)$-**dense** if $\bigcup_{i\in Z} B_\delta(x_i) = R^2$ where $B_\delta(x_i)$ represents the square $\prod_{i=1}^{2}[\xi_i - \frac{\delta_i}{2}, \xi_i + \frac{\delta_i}{2}]$ centered at x_i.*

The one dimensional case is presented for use as background to the more applicable two dimensional case, but in this treatment, any reference to δ-dense is to the two dimensional case unless otherwise specified. We note that the definition for δ-dense in the two dimensional case can also be generalized to larger dimensions.

In the one dimensional case, δ is determined by the largest separation between two sample points. Thus one large gap in the sampling grid severely limits the frequencies recoverable from the sampled data. For an intuitive insight, consider a one dimensional case and recall that the reconstruction from regular samples is an interpolation using the $sinc(x)$ as an interpolation function. If the samples are not close enough together, the interpolation function cannot reconstruct the signal properly between samples. So, even if the samples are fairly dense everywhere else,

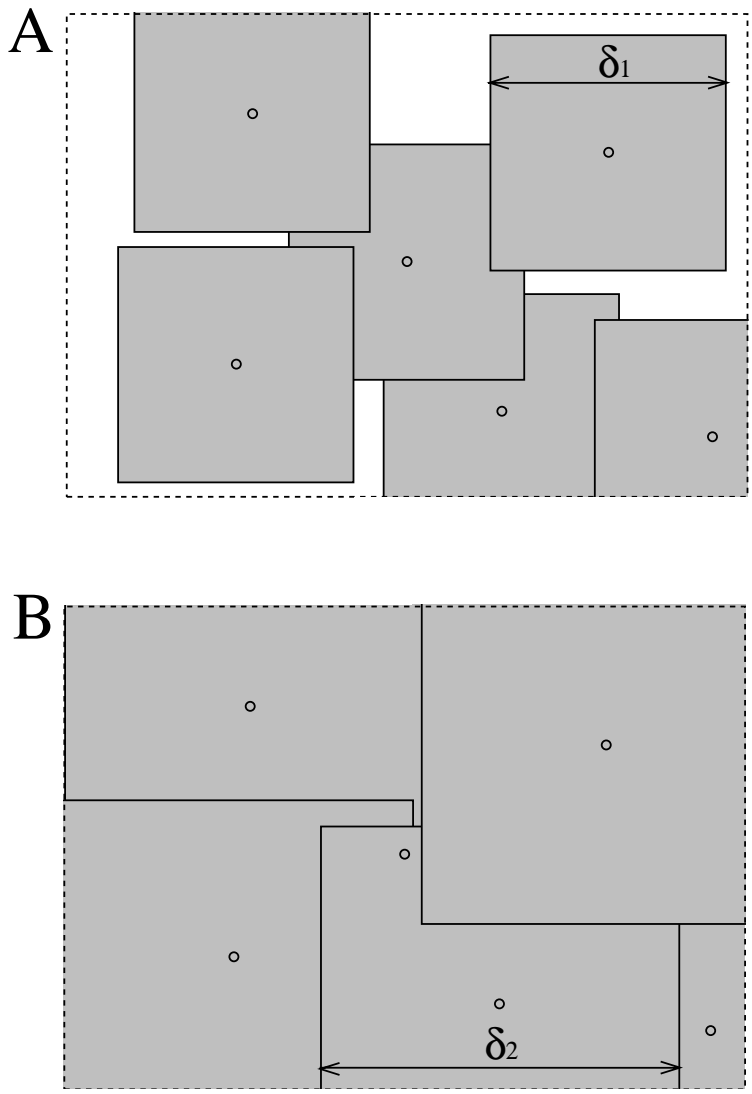


Figure 3.4: A graphical illustration of δ-dense in two dimensions. Figure A shows a situation where δ is too small for the union of the boxes around each sample point to cover the entire R^2 space. Figure B shows a larger δ where the union of the boxes does span the R^2 space. δ-dense corresponds to the smallest δ for which the R^2 space is spanned. A smaller delta yields better frequency reconstruction.

a single large gap limits whether or not the entire signal and be *completely* recovered from the sample set.

For the two dimensional case the definition is slightly different. Here, δ-dense is defined as the minimum value of δ around each sample point that completely fills the R^2 space. In Figure 3.4(A), a set of sample points is shown with a δ_1 box surrounding each sample point as described in Definition 7. Note that the union of the boxes is not sufficient to cover the entire R^2 space. In Figure 3.4(B), a larger δ value, δ_2, is used that does cover the entire space. Thus, this particular sampling set is δ-Dense with $\delta = (\delta_2, \delta_2)$. In reality, since a sampling set is limited to some finite space in R^2, we assume that the sampling set is periodic in space with a period determined by the x and y dimension of the finite sample space. By extension, the entire R^2 space is covered by the union of the δ_2 boxes for the periodic grid.

3.3.2 Iterative Reconstruction from Irregular Samples

With all the preliminaries established, the main theorems may now be presented. The bulk of the work in [28] is based on the following Lemma:

LEMMA 1 (GRÖCHENIG'S LEMMA: ITERATIVE RECONSTRUCTION)
Let A be a bounded operator on a Banach space B such that $\|I - A\|' < 1$ (I is the Identity Operator), where $\|\cdot\|'$ denotes the operator norm on B. Then A is invertible on B and $A^{-1} = \sum_{n=0}^{\infty}(I - A)^n$.

1. *Moreover, every $f \in B$ can be reconstructed by the iteration*

$$\phi_0 = Af \tag{3.8}$$

$$\phi_{n+1} = \phi_n - A\phi_n \tag{3.9}$$

$$f = \sum_{n=0}^{\infty} \phi_n \tag{3.10}$$

   with convergence in B.

2. *Setting $f_n = \sum_{k=0}^{n} \phi_k$, the error is*

$$\|f - f_n\|_B \leq \|I - A\|'^{\,n+1} \frac{1 + \|I - A\|'}{1 - \|I - A\|'} \|f\|_B \tag{3.11}$$

The proof is provided in [28]. The iterative procedure is based on the series $A^{-1} = \sum_{n=0}^{\infty}(I - A)^n$, the Neumann expansion for the inverse of an operator. The only limitations on the operator A are that it be bounded (Definition 4) and that $\|I - A\|' < 1$, required for the Neumann expansion to be valid. Note that the δ term (i.e. δ-density) defined in the previous section is not included in this Lemma. The relationship of δ to the recoverable frequency content is determined by the specific choice of A.

An Example Operator

As an example of a valid operator A and the subsequent relationship between the δ-dense sampling grid and the recoverable frequency content of the original signal, the algorithm given in [28] is presented:

THEOREM 1 (GRÖCHENIG'S THEOREM)
Given $\Omega \subseteq R^2$ with linear extension $\omega = (\omega_1, \omega_2)$ and the appropriate Banach space $B^2(\Omega)$, choose $\delta = (\delta_1, \delta_2)$ such that $\delta \cdot \omega = \sum_{i=1}^{2} \delta_i \omega_i < ln(2)$. If $\chi = (x_i)_{i \in I}$ is a δ-dense sampling set in R^n, then every $f \in B^2(\Omega)$ can be reconstructed from its sampled values $f(x_i)$ by the iteration

$$\phi_0 = Af \tag{3.12}$$

$$\phi_{n+1} = \phi_n - A\phi_n \tag{3.13}$$

$$f = \sum_{n=1}^{\infty} \phi_{\cdot} n \tag{3.14}$$

The operator A is defined by

$$Af = P(\sum_i f(x_i)\psi_i) \tag{3.15}$$

where $(\psi_i)_{i \in Z}$ is a partition of unity with the properties supp $\psi_i \subseteq B_\delta(x_i)$ ($B_\delta(x_i)$ as defined in Definition 7), $0 \leq \psi_i(x) \leq 1$, and $\sum_i \psi_i(x) = 1$ and P is the orthogonal projection from $L^2(R^2)$ onto $B^2(\Omega)$.

Note that the original image f is band limited. The operator is, in this case, very simple. It can be visualized easily for a regular grid, which is a special case of the

general irregular grid. Consider a regular Cartesian sampling grid with equal spacing in the x and y directions, the distance between two samples in the x or y direction denoted by Δ. One choice for the partition of unity function ψ_i is a simple indicator function where at a point $x_i = (x, y)$, $\psi_i = 1$ on the square $(x \pm \Delta/2, y \pm \Delta/2)$ and zero elsewhere. The result is a step-type image. The operator P is simply the low pass filter of this step image where the low pass filter corresponds to the frequency region of support Ω. The process is illustrated for a one dimensional example in Figure 3.5.

Theorem 1 establishes a relationship between the sample grid parameter δ and the frequency content of the original signal defined by Ω by the definition

$$\delta \cdot \omega \quad = \quad \sum_{i=1}^{2} \delta_i \omega_i < ln(2). \tag{3.16}$$

The $ln(2)$ term in Eq.(3.16) is determined in the course of the proof of the Theorem and is required for the operator A to be invertible on $B^2(\Omega)$ (See [28] for the proof). If the spectrum of the original signal is the cube $\Omega = [-\omega_0, \omega_0]^2$, and the δ-dense sample grid has $\delta_1 = \delta_2$, then the density must satisfy

$$\delta_0 < \frac{ln(2)}{2\omega_0}. \tag{3.17}$$

Note that this requires the sampling density be higher than the Nyquist sampling density for a given ω_0. In fact, it requires that the sampling frequency be nearly 1.5 times the Nyquist rate and is determined (essentially) by the largest distance between samples. While this is a significant performance hit relative to the Nyquist rate, this theorem establishes that a function can be reconstructed from irregular samples. The relationship in Eq. (3.17), however, is specific to Gröchenig's operator A. For instance, in the one dimensional case, this same operator A requires that $\delta_0 < \frac{\pi}{\omega_0}$, the Nyquist criteria. Other valid operators for one dimensional or two dimensional applications may have slightly different relationships between the δ-dense grid and the recoverable frequency content, but the upper limit is the Nyquist limit which occurs for uniform sampling with δ sample spacing.

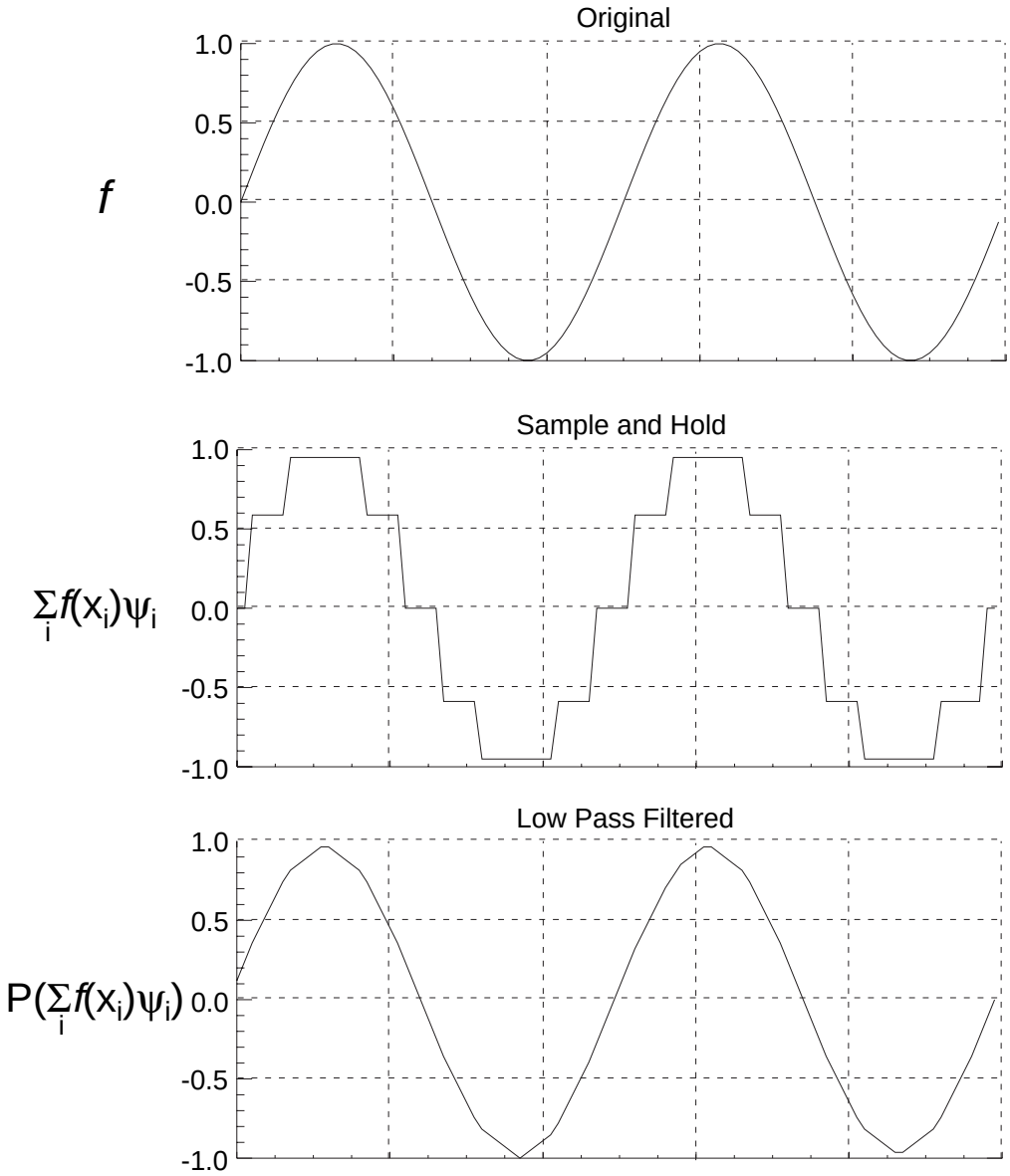


Figure 3.5: This illustrates the basic idea of the algorithm in [28]. The original image is sampled and then an indicator function is used to generate the "Step" function. This function is then low pass filtered to obtain and estimate of the original function. Note that the low pass filtered version is very similar to, but not exactly the same as, the original function.

3.4 SIR, ART and Gröchenig's Algorithm

The previous section establishes the validity of reconstruction from irregular samples (see [28] for a more detailed review). The question now is how this relates to SIR. SIR is a modified multiplicative ART algorithm. Insight into the behavior of SIR can be gained by examining the relationship of ART to irregular sampling theory. In this section, a band limited Banach space is given as a background for the discussion of SIR, ART and Gröchenig's algorithm. A comparison is then made between SIR, ART and Gröchenig's algorithm to establish the link between the methods.

3.4.1 The Band Limited Banach Space

First, because Lemma 1 requires the operator to be invertible on a Banach space, a Banach space of appropriately band limited functions must be established. A simple case is the space of low pass filtered images such that the spectrum of the functions in the space is limited to $|\Omega| < \omega$ as in Definition 2. More generally, the spectral content could be limited by other constraints such as those illustrated in Figure 3.6(A). This figure illustrates the band limits that might be imposed to circumvent the nulls that occur in a real life aperture, such as the ERS-1 aperture filter illustrated in Figure 3.6(B). In the remainder of this section, it will be shown that such a sub-band limited space is a Banach space.

Let $L^2(R)$ denote the Hilbert space of square integrable functions on R with the norm $\|f\| = (\int_{-\infty}^{\infty} |f(x)^2|dx)^{\frac{1}{2}}$. Now define a frequency range Ω' as a compact set such that for all $\omega \in \Omega'$, $|\omega| \leq \omega_0$. Further, Ω' may have arbitrary sub-band divisions as illustrated in Figure 3.6(A). Then $B^2(\Omega') = \{f \in L^2(R) : supp\, F \subseteq \Omega'\}$ is the closed subspace of all square-integrable functions with spectrum contained in Ω' (where F denotes the Fourier Transform or spectrum of f).

We will now show that $B^2(\Omega')$ is a Banach space. By definition, a *complete, normed vector space* is a Banach space. First, $B^2(\Omega')$ is a vector space since this space

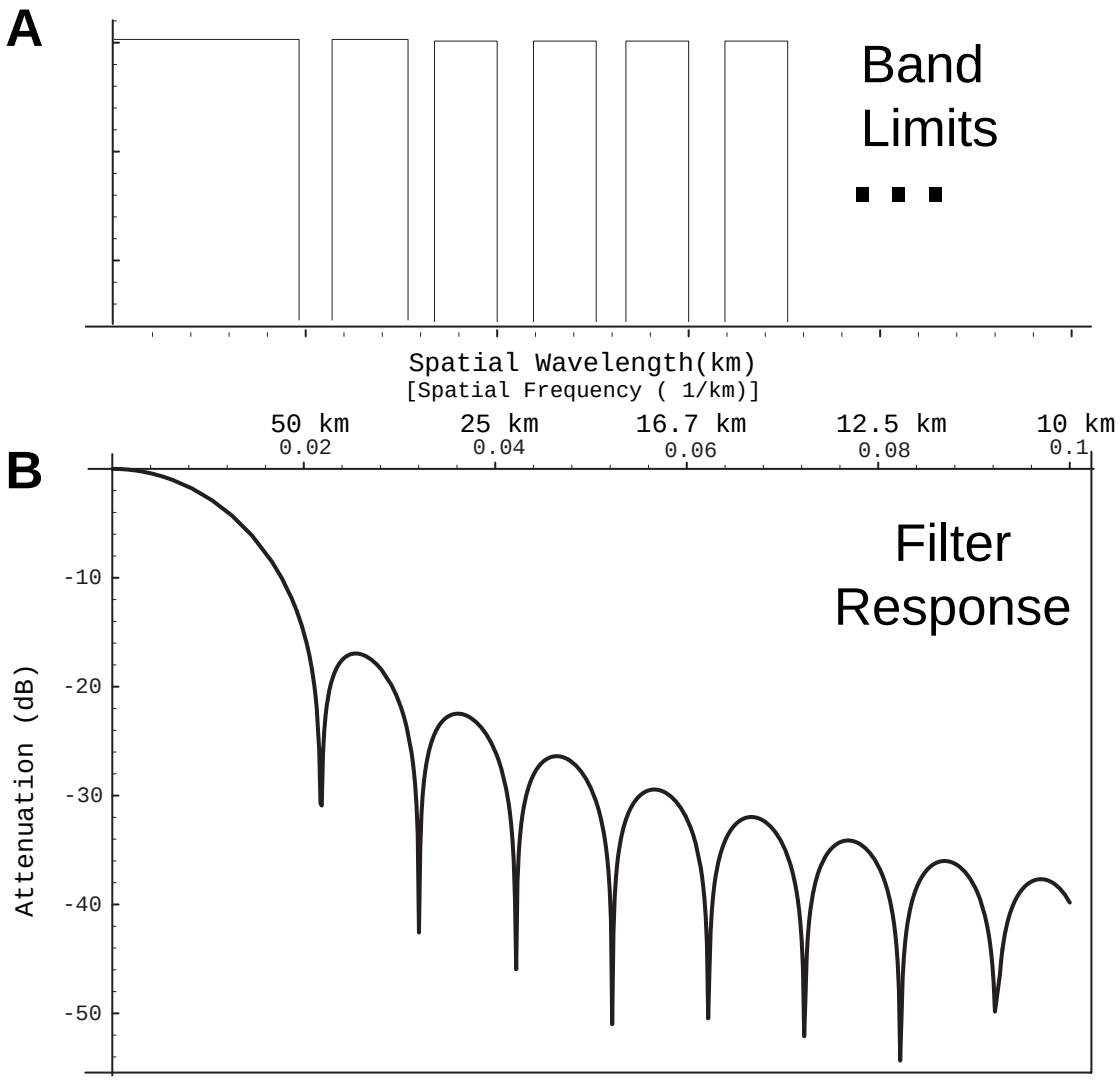


Figure 3.6: This figure shows a band limiting scheme in (A) that cuts out the nulls generated by the filter response in (B). In a noiseless case, the sub-bands could continue indefinitely due to the infinite extent of the filter response.

is closed under addition and multiplication by scalers. Recall the Fourier relations

$$af(x) \quad \overset{\lfloor}{0}pt]\mathcal{F} \Longleftrightarrow \quad aF(\omega) \tag{3.18}$$

$$af(x) + bg(x) \quad \overset{\lfloor}{0}pt]\mathcal{F} \Longleftrightarrow \quad aF(\omega) + bG(\omega) \tag{3.19}$$

and note that any scaler multiple or addition of the spectrums of two functions in the space results in a function with spectrum support in Ω'. Next, since $B^2(\Omega') \subseteq L^2(R)$ and $L^2(R)$ is a normed space, then $B^2(\Omega')$ is a normed space as well.

Next, there is the issue of completeness. For the space to be complete, every Cauchy sequence in $B^2(\Omega')$ must converge to a function in $B^2(\Omega')$. A sequence of functions $\{f_n\}_1^\infty$ in $B^2(\Omega')$ is a Cauchy sequence if $\lim_{m,n\to\infty} \|f_m - f_n\| = 0$, or in other words, for an $\epsilon > 0$ there exists an N such that $\|f_m - f_n\| < \epsilon$ for all $m, n > N$. Each Cauchy sequence $\{f_n\}_1^\infty$ converges to some function f (i.e. $\|f_n - f\| \to 0$ as $n \to \infty$). Assume that there exists an arbitrary Cauchy sequence on $B^2(\Omega')$ such that $f_n @ >> n \to \infty > f$ where $f_n \in B^2(\Omega')$, all n but f is not in $B^2(\Omega')$. This assumes that f has frequency content outside Ω'. Recall Parseval's Identity:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx \quad = \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 d\omega \tag{3.20}$$

and note the similarity of the norm $\|\cdot\|$ on $B^2(\Omega')$ to the left hand side of this equation. Also note that by our assumptions the frequency content of $f_n - f$ always has some component outside of Ω'. Thus the following is true of Fourier transform of the limit function f under our assumptions:

$$\|F(\omega)\| = \int_{-\infty}^{\infty} |F^2(\omega)| d\omega \quad = \quad \int_{\Omega'} |F^2(\omega)| d\omega + \int_{\Omega'^C} |F^2(\omega)| d\omega \tag{3.21}$$

$$\int_{\Omega'} |F^2(\omega)| d\omega \quad \geq \quad 0 \tag{3.22}$$

$$\int_{\Omega'^C} |F^2(\omega)| d\omega \quad > \quad \epsilon \tag{3.23}$$

where Ω'^C is the complement of Ω'. Equation (3.23) is always greater than some ϵ since f has frequency content outside of Ω'. By Eq (3.20), $\|f_n - f\| = \frac{1}{2\pi}\|F_n - F\|$ for all n. Finally,

$$\|F_n(\omega) - F(\omega)\| = \underbrace{\int_{\Omega'} |F_n(\omega) - F(\omega)|^2 d\omega}_{\geq 0} + \underbrace{\int_{\Omega'^C} |F_n(\omega) - F(\omega)|^2 d\omega}_{> \epsilon} \tag{3.24}$$

where $\int_{\Omega'^C} |F_n(\omega) - F(\omega)|^2 d\omega = \int_{\Omega'^C} |F(\omega)|^2 d\omega > \epsilon$ because $F_n(\omega)$ has zero value on Ω'^C. This implies that there exists $\epsilon > 0 : \|F_n(\omega) - F(\omega)\| > \epsilon$ which contradicts the assumption of the limit function f not included in $B^2(\Omega')$.

Thus, the Cauchy sequence limit function f cannot contain energy outside of the frequency band defined by Ω' since no function in the sequence $\{f_n\}_1^\infty$ has energy outside of Ω'. Therefore, $\{f_n\}_1^\infty$ must converge to a function $f \in B^2(\Omega')$, and thus $B^2(\Omega')$ is a Banach space.

3.4.2 Equivalence of Additive ART and Gröchenig's Algorithm

Because the SIR algorithm has a non-linear multiplicative update term, it is difficult to analyze directly. However, the SIR algorithm is a modified multiplicative ART algorithm. Thus, showing the equivalence of ART and Gröchenig's algorithm is helpful in understanding the behavior of SIR. We begin by examining the relationship between block additive ART and Gröchenig's algorithm.

Gröchenig's iterative algorithm given in Lemma 1 can also be written as [28]:

$$f_{n+1} = f_n + A(f - f_n) \tag{3.25}$$

with A an operator meeting the requirements of Lemma 1 and f representing the function to be recovered. Block additive ART can be written as [21]

$$\begin{aligned} a_{n+1}^j &= a_n^j + \frac{\sum_i (s_i - p_i) h_{ij}}{\sum_i h_{ij}} \\ i &\in \{1, 2, ..., M\} \\ j &\in \{1, 2, ..., N\} \end{aligned} \tag{3.26}$$

where a represents the image to be estimated, a_n is the nth iterative estimate of a, j is the pixel index and i is the measurement index with N total pixels and M total measurements. The essence of this equation is that all measurements that touch a pixel a^j are summed and normalized to create the per pixel update value. We wish to have the equation for block additive ART (Eq. (3.26)) have the same form as Gröchenig's algorithm in Eq. (3.25). The desired form for block additive ART is

given by:

$$a_{n+1} = a_n + \mathcal{H}(a - a_n) \tag{3.27}$$

where the a's are now vectors with a being the 'true' image, a_n the nth iterative estimate of a and $\mathcal{H}$ is an $N \times N$ matrix operator with the properties prescribed in Lemma 1 for A. Note that $\mathcal{H}$ can be separated into two distinct operators:

$$\mathcal{H} = H'H \tag{3.28}$$

each of which will be described in detail below. We now proceed to show that block additive ART is equivalent to Gröchenig's algorithm (i.e., Eq. (3.26) is equivalent to Eq. (3.27) and therefore also equivalent to Eq. (3.25)).

In Eq. (3.26), the normalized sum on the right hand side is a function of the vectors s, the measurement vector and p the back projection vector based on the n^{th} iterative estimate. The vector s represents the actual measurements which, as shown in Section 3.2.3, are just samples of the aperture-filtered true image. The sampled convolution of the true image a and the aperture function can be reduced to a matrix multiplication, an $M \times N$ matrix denoted by H in Eq. (3.28). This matrix is sometimes called the point spread matrix. Thus, H is the back projection operator and $s = Ha$ and $p = Ha_n$. Combining these equations with Eq. (3.28), Eq. (3.27) now becomes

$$a_{n+1} = a_n + H'(Ha - Ha_n) \tag{3.29}$$

$$a_{n+1} = a_n + H'(s - p) \tag{3.30}$$

where, as shown below, H' is a matrix operator that performs the summation and normalization in Eq. (3.26).

The form of the matrix H' can be determined by examining the form of H and $\mathcal{H}$. Figure 3.7 illustrates the matrix multiplication $s = Ha$ where the columns of H are denoted by vertical lines. The index i is the measurement index with each measurement corresponding to a row in H. Refer to Eq. (3.26) and note that for each update, only those measurements where h_{ij} is non-zero for the j^{th} pixel are used

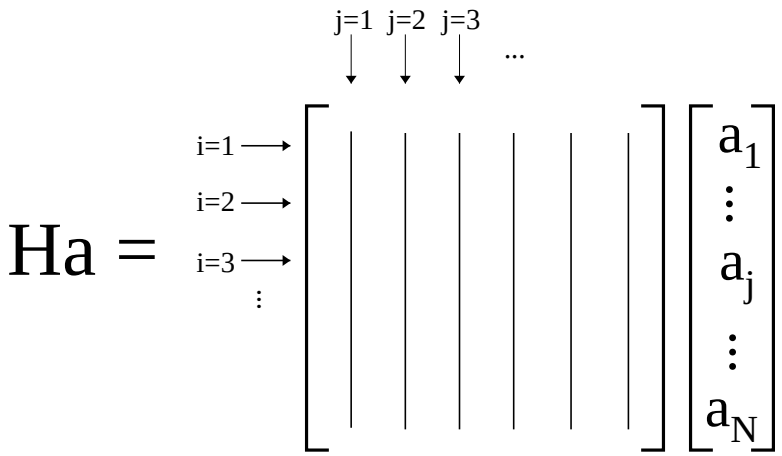


Figure 3.7: This graphic illustrates the matrix multiplication of the true image a by a general point spread matrix H. The columns of H are shown as lines. The terms i and j represent the measurement index and pixel index respectively.

to update that pixel. The measurements where h_{ij} is non-zero for the j^{th} pixel all have non-zero values in the j^{th} column of H.

The update for the j^{th} pixel is the normalized sum of the elements touching that pixel. Take the transpose of H and normalize each row by the sum of the weights in that row and call this new $N \times M$ matrix H'. Note that by multiplying H' by the vector $H(a - a_n) = (s - p)$, the individual elements of the resulting vector correspond to the update values in Eq. (3.26). Adding $H'(s - p)$ to the current estimate a_n gives the updated a_{n+1} image vector as shown in Eq. (3.30). Thus we see that the operator $\mathcal{H}$ in Eq. (3.27) can be written as $\mathcal{H} = H'H$ with H' and H as defined above. This establishes the equivalence of Eq. (3.25) and Eq. (3.26).

The Operator $\mathcal{H}$

With the successful derivation of a valid $\mathcal{H} = H'H$ for Eq. (3.27), the equivalence of block additive ART and Gröchenig's algorithm now depends on the invertibility of $\mathcal{H}$ on an appropriately defined Banach Space as per the requirement in Lemma 1. We wish to demonstrate the invertibility of $\mathcal{H}$ on the sub-band limited Banach space defined in Section 3.4.1. We must define invertibility on a sub-band

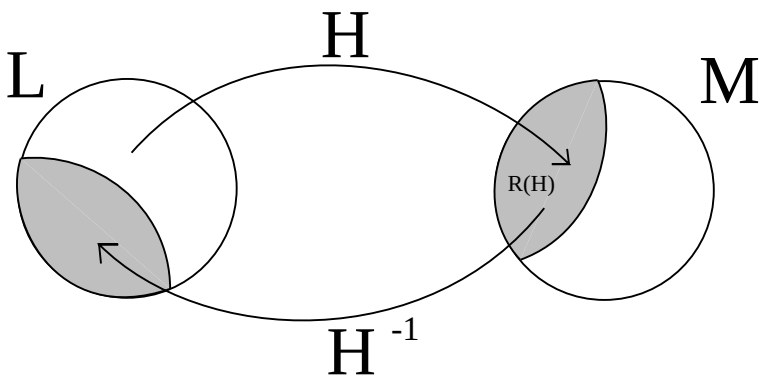


Figure 3.8: An illustration that shows the relationship of the domain and range spaces of the operator H. Note that this is H, the point spread function, not $\mathcal{H}$.

limited Banach space because we know that by definition of the point spread function H, an arbitrary input is not necessarily recoverable because of frequency nulls created by the point spread function. If the system input is selected so that it has no frequency content at the H nulls, the input function will be attenuated by the point spread function but no information will be lost. Recall that the operator $\mathcal{H}$ is defined as $\mathcal{H} = H'H$, where H' and H are defined in Section 3.4.2. H as defined in Eq. (3.28) incorporates the sample spacing and aperture function characteristics. Here we assume that the sample spacing is adequate for recovering the original signal and deal strictly with aperture function affects on invertibility. How the sample spacing affects the signal recoverability is discussed in Section 3.4.2.

Figure 3.8 shows the functional operation of H in Eq. (3.1). Let L be the space of band limited finite functions (Definition 2) represented by vectors $u \in L$. The point spread function H maps $u \in L$ into a range space $v \in R(H) \subset M$. The other half of $\mathcal{H}$, H', is the row normalized transpose of H. The row normalization consists of multiplying each row by the sum of the row elements and since elementary row operations do not affect the rank of the matrix, $\operatorname{rank} H = \operatorname{rank} H'$.

We now assume that the domain of H, $u \in B^2(\Omega')$ consists of all functions with a sub-band limited frequency response as illustrated in Figure 3.6(A). The low

pass characteristics of the aperture function built into the point spread function H indicate that certain frequencies of an arbitrary input are nulled out and therefore unrecoverable in any reconstruction. But since we define a domain $B^2(\Omega')$ that consists exclusively of functions without those frequencies, no information is lost for the new problem definition. So for $v = Hu$, v is the projection of u onto the columns of H. While v may have attenuated frequency components, all the original frequency components of u will be present in v. Now consider the left multiplication by H'. This new value, $u' = H'v = H'Hu \in B^2(\Omega')$ will also be in the original Banach space. Thus, $\mathcal{H} = H'H$ is a bounded operator on the sub-band limited Banach space, meeting the first requirement of Lemma 1.

With $\mathcal{H}$ bounded, we must next show that $\mathcal{H}$ is invertible on the Banach space for this to be a valid operator according to Lemma 1. It should be obvious that if H is full rank, than $H'H$ is full rank and the operator $\mathcal{H}$ is invertible on the Banach space. However, it is not necessarily required that H be full rank for the signal to be recoverable. Consider that by definition no information is lost in the process of applying $\mathcal{H}$ to an appropriately sub-band limited input function even if the matrices are not full rank. Simple attenuation is easily corrected and therefore $\mathcal{H}$ is invertible *on the sub-band limited Banach space.*

Thus, the block additive ART reconstruction in Eq. (3.27) is equivalent to Gröchenig's algorithm in Eq. (3.25) (both repeated here for convenience):

$$\begin{aligned} a_{n+1} &= a_n + \mathcal{H}(a - a_n) \\ f_{n+1} &= f_n + A(f - f_n) \end{aligned}$$

and both are *completely* invertible (by Lemma 1) on the sub-band limited Banach space. Therefore Eq. (3.27) represents a valid algorithm for the *complete* recovery of the original vector a for an appropriate choice of input function Banach space $B^2(\Omega')$ based on the aperture function. Note, however, that *complete* recovery is only possible if the assumption is made that the original function is contained in the space spanned by the operator inverse $\mathcal{H}^{-1}$. If this isn't the case, then we apply some conditions, or regularization, on the reconstruction in order to develop a unique

solution. Additive ART, for instance, produces the least squares minimum norm solution, while multiplicative ART gives the maximum entropy solution. In theory, we could use any of a number of regularization schemes (least squares, for example) to generate an estimate of the signal for a case where the original function is outside the space spanned by $\mathcal{H}^{-1}$. This is addressed further later in this chapter.

Signal Recoverability from δ-dense Sample Spacing

In the preceding sections, we have assumed that we have adequate sampling. We now examine the limitations of the irregular sampling on the reconstruction. Given a set of irregular samples that are δ-dense, the natural question is what frequency content can be recovered using this grid and an algorithm with the form given in Eq. (3.25). Unfortunately, $\mathcal{H}$ varies from image to image, and (for our application) is highly dependent on satellite geometry, location of the sampled surface, and instrument activity. It is thus difficult to determine an exact prediction of frequency recoverability for SIR similar to the one given in Eq. (3.17) for Gröchenig's algorithm.

However, there is a limit to the best that can be done, which is the Nyquist sampling rate. Recall the definition of δ-dense in Definition 7 and note that the δ of a δ-dense sampling set is determined by the largest gap between samples. If this δ spacing were for a regular sampling grid, it would limit the recoverable signals to a class of functions with spatial wavelengths greater than 2δ by the Nyquist criterion. This is the upper limit of any reconstruction algorithm. As shown in [28] for Gröchenig's operator, the recoverable frequency range using an irregular δ-dense sampling grid in two dimensions is smaller than for a regular grid with δ sampling. So for an arbitrary operator, it is expected that the recoverable frequency range using an irregular δ-dense sampling grid will be less than the frequency range recoverable by a regular δ spaced grid as determined by the Nyquist criterion. Experimental results for the ERS-1 scatterometer show that in the polar regions, the sampling sets are δ-dense with $\delta = 10\,\text{km}$ to $13\,\text{km}$.

Using an ART-derived reconstruction technique (SIR), 30 km to 35 km resolution is achieved [29], which is consistent with the relationship in Eq. (3.17).

For the purposes of evaluation, the upper limit of frequency recoverability is assumed to be a function of the largest sampling gap. Although actual performance likely will be slightly less (depending on the operator used), using the Nyquist rate as an upper performance limit gives a reasonable estimate of the best possible recovery using an irregular grid.

Gröchenig's algorithm was originally presented as a method for *complete* recovery of a certain class of irregularly sampled band limited functions [28]. In the previous sections, we have shown that Gröchenig's algorithm is functionally equivalent to block additive ART. In theory, block additive ART would completely recover an arbitrary band limited function for the noiseless case after infinite iterations if all conditions of Gröchenig's lemmas and theorems are met (e.g. proper initialization, sample spacing, etc). An important observation is that theoretical recoverability is irrelevant to the aperture function for the noiseless case. In other words, no matter how low the side lobes of the data are, in theory the original signal (less the appropriate nulls) can be recovered in the limit of infinite iterations. This will become important in the next chapter where practical application of this theory is discussed. We now examine the relationship of additive and multiplicative ART.

3.4.3 Relationship of Additive ART to Multiplicative ART

With the equivalence of additive ART and Gröchenig's algorithm established, we look at the relationship of additive and multiplicative ART. The difference between additive and multiplicative ART is the regularization assumed in the estimate. Additive ART is a least squares estimate in the limit of infinite iterations [21], and is based on the minimization problem:

$$\begin{aligned} &\text{Minimize} \quad \|x^2\| \\ &\text{Subject to} \quad y = Hx. \end{aligned} \tag{3.31}$$

Multiplicative ART with damping is a maximum entropy estimate in the limit of infinite iterations [21, 22], and is based on the maximization problem:

$$\begin{aligned} \text{Maximize} \quad & -\sum_{j=1}^{n} x_j \ln x_j \\ \text{Subject to} \quad & y = Hx. \end{aligned} \tag{3.32}$$

Additive ART assumes no *a priori* knowledge about the data, and fits the estimate strictly on the measurements available through minimizing the error of the back projection of the measurement onto the H space. In the context of the sub-band limited spectrum illustrated in Figure 3.4(A) this means that the reconstruction is strictly contained within the space spanned by the measurements (in this case, the sub-band limited space).

Multiplicative ART assumes a maximum entropy model. The subjective result of the maximum entropy model is sharper images. In the frequency domain, the reconstruction is not strictly restricted to the band limited frequency domain spanned by the measurement space. Additional high frequency content is added to create a 'sharper' image. However, the constraint $y = Hx$ remains, and the reconstruction is based on a projection of measurements onto the H space just like additive ART.

The choice of one method over another is a hotly debated issue. We can, in principle, select any regularization to use in the reconstruction if the regularization fits with the *a priori* knowledge we have about the surface. In [21], this decision may be based on the nature of the sampling mechanism (reflection, absorption or emission) and the nature of the approximation that the algorithm will develop for under-determined systems. The choice depends somewhat on what regularization gives the best results for the user. Least squares estimates produce a maximally smooth estimate where edges tend to be softened and blurred. A maximum entropy estimate produces a generally "sharper" image than least squares, at least for images with high contrast [30]. The assumption of sharper edges fits well with known characteristics of the Southern Ocean sea ice such fissures, high contrast areas of open water and the hard edges of ice bergs.

We now show a fundamental relationship between additive ART and multiplicative ART based on the common constraint $y = Hx$. Note that since both forms of ART have the same constraint, the resulting solutions x are of the general form

$$x = P + Q \tag{3.33}$$

where P is an element of the row space of H, or equivalently, the range space of the transpose of H, H^t, denoted $P \in R(H^t)$; and Q is an element of the null space of H, denoted $Q \in N(H)$. Any solution derived from either additive or multiplicative ART will have a component P. However, the solution derived by using additive ART results in $Q = 0$, while the solution from multiplicative ART will generally have a non-zero Q component [31, 32]. Since the constraint $y = Hx$ is the same for both algorithms, the solution for both AART and MART *will be the same in the range space of* H^t in the limit of infinite iterations. The only difference between the AART and MART solutions will be the Q component from the null space of H.

In the case of a non-band limited input function, each algorithm recovers what it can based on the band limits imposed by the filtering. From a functional point of view, the algorithm attempts to reconstruct the output of the filtering operation, *not* the original function because of the nulls introduced by the aperture function. For the scatterometer and similar remote sensing devices, this means that the aperture filtered surface is recovered, contingent on sufficient sampling and sufficiently low noise. The use of multiplicative ART rather than additive ART may result in some component from the null space being added in to better meet the maximum entropy constraint, but in the range space of H^t, the solution for either additive or multiplicative ART is the same. We conclude that both additive and multiplicative ART are viable reconstruction techniques based on the theory from [28], with the understanding that in the null spaces, additive and multiplicative ART may produce slightly different answers based on the different regularizations.

3.4.4 SIR: Blocked Damped MART

As stated previously, SIR may be formulated from either additive or multiplicative ART, but SIR as presented herein is based on multiplicative ART [13]. Recall the basic equations for block multiplicative ART (from Eq. (2.10):

$$a_j^{k+1} = \frac{a_j^k \left(\sum_{i=1}^{N} d_i^k h_{ij} \right)}{\sum_{i=1}^{N} h_{ij}} \tag{3.34}$$

and for SIR (from Eq. (2.11):

$$a_j^{k+1} = \frac{\sum_{i=1}^{M} update_{ij}^k h_{ij}}{\sum_{i=1}^{M} h_{ij}} \tag{3.35}$$

where $d_i^k = (s_i/p_i^k)^n$, where n is a damping factor ($n = \frac{1}{2}$ for SIR). Note that the scale factor for MART is simply d_i^k, whereas the scale factor for SIR is a function of not only d_i^k, but also a_j^k given by (from Eq. (2.12):

$$update_{ij}^k = \begin{cases} \left[\frac{1}{2p_i^k}(1 - \frac{1}{d_i^k}) + \frac{1}{a_j^k d_i^k} \right]^{-1} & d_i^k \geq 1 \\ \frac{1}{2} p_i^k (1 - d_i^k) + a_j^k d_i^k & d_i^k < 1 \end{cases} \tag{3.36}$$

$$d_i^k = \left(\frac{s_i}{p_i^k} \right)^{\frac{1}{2}}. \tag{3.37}$$

The scale factor for SIR is limited to between 0.5 and 2 as illustrated in Figure 2.1. Note that this graph is a function of d_i^k. This sigmoid soft-limit function helps reduce the effects of noise in a SIR reconstruction by limiting the development of regions where the difference between the measurement and the back projection are large. MART has no such limitation. Also note that if the SIR update is small, in other words, if the update is along the linear region near 1 on Figure 2.1, then the SIR and MART updates will be the same if MART uses square root damping ($n = \frac{1}{2}$). Also note that the SIR update does not change the direction of the update indicated by MART, i.e. if the MART update causes the pixel value to increase, then SIR also

causes an increase, but possibly a smaller increase. The same is true for a decrease. We can thus draw the following conclusion: In the noiseless case, the damped SIR update may slow the development of the final image in some regions compared to MART, but in the limit of infinite iterations, SIR will reach the same limit as MART because it has the same constraints and does not change the fundamental direction of the MART update. The damped update factor may cause the interim images to differ (i.e., images from different iterations of MART and SIR will not necessarily be equivalent), but the end result is the same. Because SIR may damp different part of the image in different ways, the trajectory that SIR takes to reach the limit may be different than the path taken by MART.

Thus we have the following progression showing the relationship of Gröchenig's algorithm and SIR: (i) Gröchenig's algorithm is a complete reconstruction; (ii) Additive ART is a valid (complete) operator under Gröchenig's theory; (iii) Additive ART and Multiplicative ART differ by an element in the null space based on the regularization; and (iv) SIR is a block damped version of MART, reaching the same solution in the limit of infinite iterations.

We now introduce two new considerations. The first is a finite number of iterations for any iterative algorithm; the second is noise. The former is a limitation of time and practicality: no iterative process can forever. Therefore, even if a particular algorithm is supposed to be the least squares estimate of a function, it can only approach this estimate. The result is something close to the optimal reconstruction, but not a complete reconstruction [21]. Truncation of the iterations is ultimately another form of regularization [32].

Noise further complicates this since noise is also amplified through most iterative processes, and increasing iterations will generally increase the noise. Although steps can be taken to minimize noise, it will to one degree or another limit the number of iterations that can be executed before noise overtakes the reconstruction. As will be shown in the following sections, SIR was developed to reduce the overall effects of noise on the output, and generally performs better in the presence of noise than either ART algorithm. The effects of noise are examined in the next section.

3.5 Noise

With the establishment of additive ART as a valid reconstruction algorithm from irregular samples, we now turn to the issue that motivated the development of the SIR algorithm: noise. All real systems have noise, and dealing with that noise is a significant processing consideration. Gröchenig's Lemma shows that complete reconstruction of a signal can be made, though it may take an infinite number of iterations. Experimental data shows that even highly attenuated frequency components are effectively recovered with finite iterations for a noiseless case. However, the addition of noise changes the problem because noise is amplified along with the desired signal during the reconstruction.

Unlike ART and Gröchenig's algorithm, SIR is designed to reduce the overall effects of noise in the reconstruction. In the following sections, a methodology is established for comparing the reconstruction algorithms, and then the performance results for each are compared.

3.5.1 A Note on SNR

As mentioned, the noise power will be affected by the reconstruction. If the signal power is amplified, so will any noise be amplified. Let us examine the behavior of the signal and noise power in the reconstruction. The noise model is assumed to be the additive model presented in Eq. (2.5), repeated here for reference:

$$\sigma_{\text{noisy}} = \sigma_{\text{true}} + \mathcal{V} \tag{3.38}$$

where $\mathcal{V}$ is Gaussian white noise.

In order to understand the effects of the reconstruction on the noise and thus the signal to noise ratio (SNR) of the resulting imagery, we examine the signal and noise power in the frequency domain. The noise is assumed to be Gaussian white noise, so in the frequency domain the power is spread evenly over the spectrum. The signal is attenuated by the aperture function, and will look figuratively like the lower line in Figure 3.9 if we assume the original function spectrum was a rect function, an

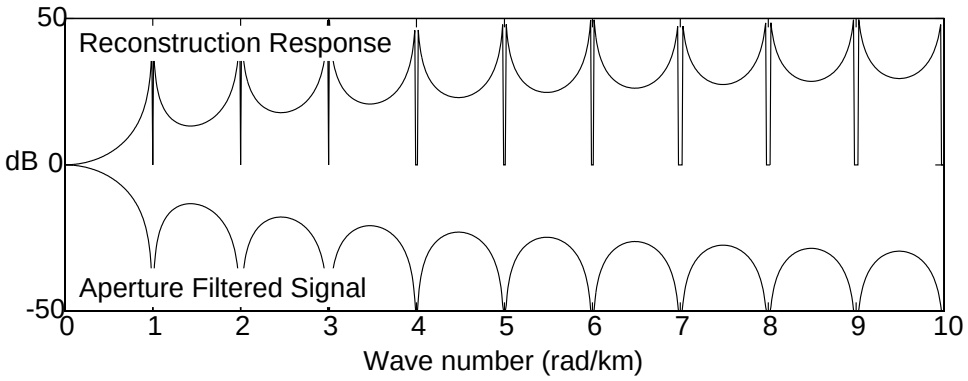


Figure 3.9: A plot showing the aperture filter response (bottom line) and the reconstruction impulse response. Note that at higher frequencies, the reconstruction amplifies the signal more. The higher the noise floor, the more dominant the noise becomes at higher frequencies.

assumption used for illustration purposes. The noise power is effectively added after sampling and so it is not affected by the aperture function.

We now assume that the reconstruction is a complete reconstruction. The reconstruction frequency response will look like the upper line in Figure 3.9 (which is the inverse of the lower line with the null regions zeroed out). In the absence of noise, the reconstruction would restore the original function except for the null regions. However, the additive noise will be amplified by the reconstruction response function just as the signal is. Because the amplification is greater at higher wave numbers, the noise power at higher frequencies will be higher relative to the signal power. Thus the signal to noise ratio at a given frequency will decrease as the wavenumber increases.

For a simple original sinc function, the reconstruction result will be the original frequency domain rect function (except the nulls) plus a noise function that in the frequency domain will look like the reconstruction response function. (Recall that the noise signal is flat in the frequency domain, so the reconstruction result will be the flat signal times the impulse function, resulting in a shifted impulse function.) Note that as the wave number increases, the noise power increases. Therefore increasing the bandwidth of the reconstruction will result in greater and greater noise power at

high frequencies lowering the SNR. If the initial SNR is good enough, this will not be a problem, but at some point the noise power will exceed the signal power, and the resulting imagery will be dominated by noise at high frequencies. In the current applications, however, the sample spacing limits the reconstruction to a sufficiently low bandwidth that noise amplification does not dominate the resulting SNR.

3.5.2 Experimental Evaluation in the Presence of Noise

While it can be shown that all three of the algorithms (ART, MART and SIR) can reconstruct a signal with no noise, we need to compare the performance of the algorithms when noise is added to the problem. In order to qualitatively evaluate the three methods, each is applied to a simple one dimensional signal. The one dimensional test signal is a sinc function which readily shows the quality of the frequency domain reconstruction from the various methods. The test signal is sampled with an irregular sampling grid. The largest sample space in the grid is equivalent to the Nyquist sample spacing for the test signal (i.e., the signal is fully recoverable if sampled on a regular grid with spacing equal to the largest spacing in the irregular grid).

The test function is illustrated in Figure 3.10. Figure 3.11 shows the frequency content of the test signal. The aperture function used for this analysis is a rectangular, constant aperture. This was chosen for convenience in algorithm implementation. The primary difference between the rectangular aperture used here and the actual cos^2 ERS-1 aperture illustrated in Figure 4.3 is width of the main lobe and the level of the first side lobes. For the cos^2 aperture, the side lobes are somewhat lower and the main lobe slightly wider.

The relationship between the frequency content of the aperture and input sequence is illustrated in Figure 3.12. The rectangular aperture for this study is chosen so that the first side lobe of the aperture is just inside the frequency content of the test signal as illustrated in Figure 3.12, allowing the reconstruction of the attenuated frequencies in the side lobe to be evaluated. We see that in Figure 3.12 the aperture function nulls out some frequencies and significantly attenuates other

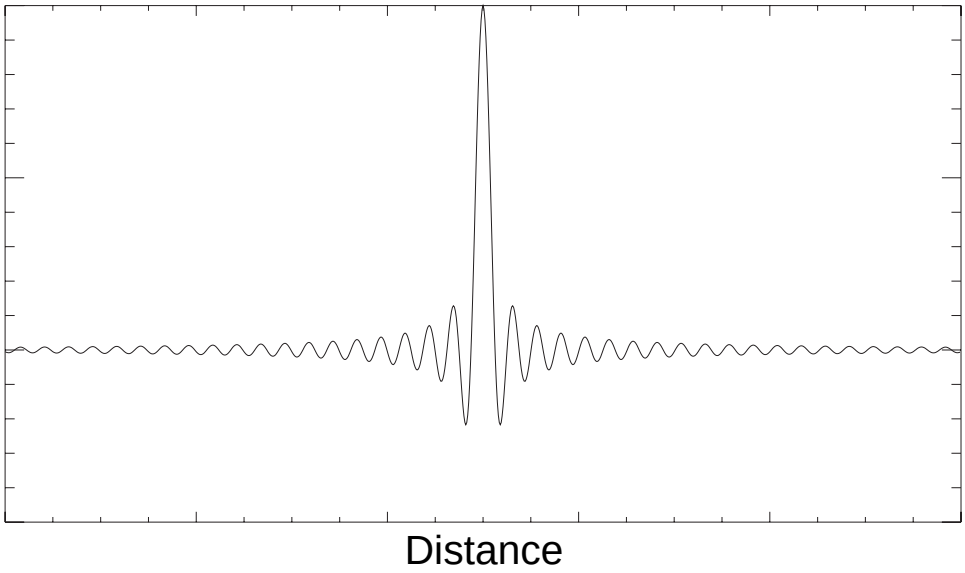


Figure 3.10: This figure shows the sinc function used as the test input for the three algorithms. The sinc function was chosen for its nicely band limited frequency content (see Figure 3.11) that readily shows the behavior of the various algorithms when certain frequencies are nulled by the aperture function.

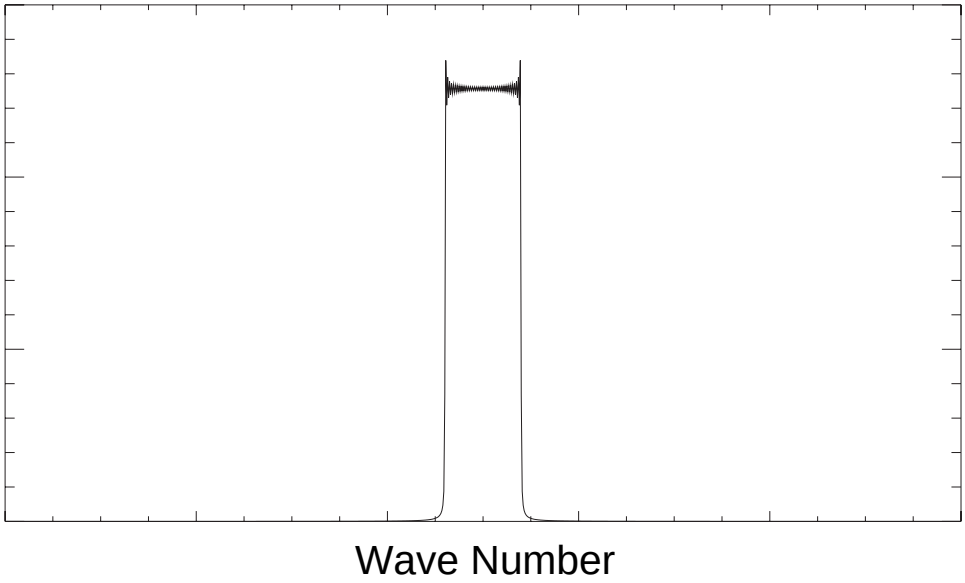


Figure 3.11: This figure shows the frequency content (spectrum) of the sinc function test signal. The sinc function was chosen for its nicely band limited frequency content to show the behavior of the various algorithms when certain frequencies are nulled by the aperture function.

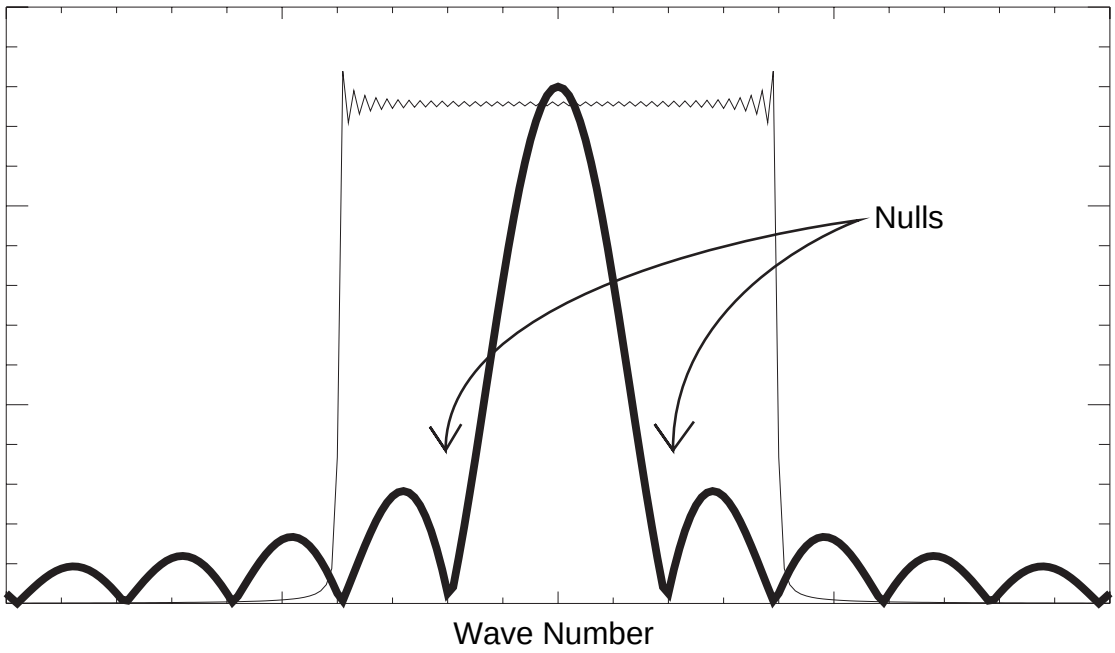


Figure 3.12: This figure illustrates the overlay of the test signal frequency content (light) with the frequency response of the aperture function (bold). The aperture function in this case is a uniform aperture. The primary difference between this uniform aperture used here and the aperture of the ERS-1 scatterometer is the side lobe levels for the ERS-1 aperture are lower due to the cos^2 roll-off and the main lobe slightly wider.

frequencies. Figure 3.13 shows the original signal filtered with the aperture function. Note how the main lobe is much larger than the main lobe of the original function in Figure 3.10. Performance of the resolution enhancement algorithms is judged by how well these problem regions (main lobe shape and side lobe levels) in the frequency domain can be corrected.

For each algorithm, a noisy case is also considered. To be as accurate as possible in creating the noisy images, the actual multiplicative noise model for ERS-1 is used to create the test images using the relationship

$$a_{\text{noisy}} = (1 + 0.05N(0,1))a_{\text{orig}} \tag{3.39}$$

where $N(0,1)$ is a zero mean unit variance Gaussian random variable. This is the

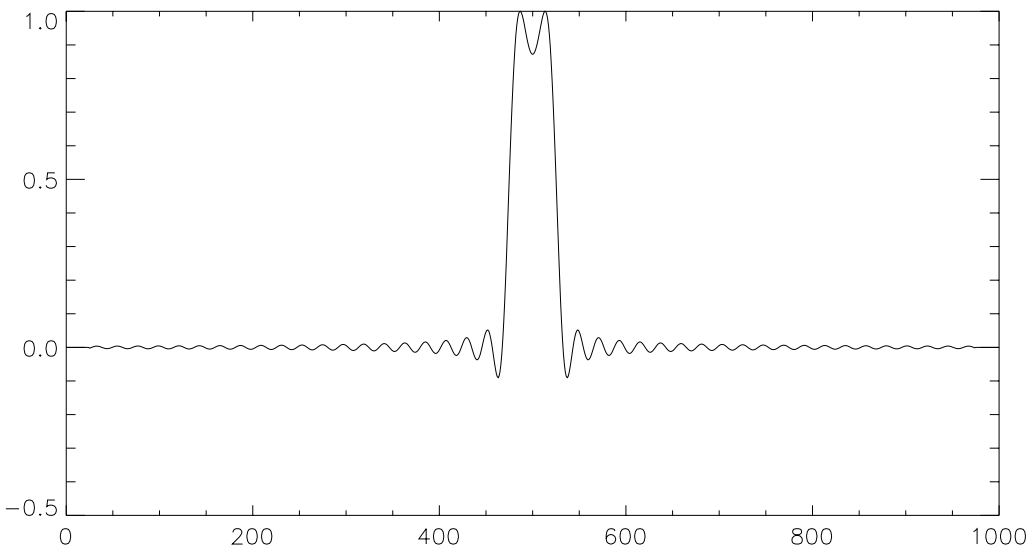


Figure 3.13: This figure illustrates the original signal filtered by the aperture function. Compare to the original signal in Figure 3.10.

noise model for the ERS-1 scatterometer system, as referenced in Section 2.1.1. The algorithm, however, assumes the additive noise model as discussed in Section 2.1.1.

In the next section, individual examples for each algorithm are presented and comparisons of the outputs are used to judge performance of the reconstruction.

3.5.3 Algorithm Performance Comparison

To demonstrate the resolution enhancement capabilities of the algorithms, we begin with a comparison of the noiseless case. One other consideration is that the algorithms can only be run a finite number of times so we examine performance as a function of iterations as well. Figure 3.14 is comparison of the output of the three algorithms at 25 iterations with no noise. Figure 3.15 is a similar output comparison but at 100 iterations while Figure 3.16 is for 1000 iterations. There are two significant observations to be made. First, there is little difference between the Additive ART

and Multiplicative ART reconstructions for an equal number of iterations at 25, 10 or 1000 iterations. There are some small numerical differences as would be expected from the use of two very different algorithms, but the differences are negligible. Second, the SIR output lags (as a function of the number of iterations) the output of the other algorithms. In fact, the SIR output at 100 iterations and the output of Multiplicative and Additive ART at 25 iterations compare very well. This lag is a result of the bounded multiplicative scale factor used to generate SIR output. As is illustrated later, the bounded scale factor results in much better performance in the presence of noise.

Next, we examine several examples for ART and SIR with and without noise. Figure 3.17 illustrates the output from Multiplicative ART both in a noiseless and in a noisy case. The original test signal is shown as a dotted line. Figure 3.18 shows the corresponding frequency content for the 25 and 100 iterations cases.

While the noiseless case shows very good frequency content recovery and acceptable time domain signals for just a few iterations, the performance of the ART algorithms in the presence of noise is not acceptable. As Figure 3.18 shows, after only 100 iterations the energy in the noise outside the desired band is increasing rapidly. Additive ART has similar poor performance with noise as multiplicative ART. The poor performance of ART in the presence of noise motivated the development of SIR [14]. For SIR, the multiplicative scale factor is damped so that a large scale factors do not overly magnify the noise at any one iteration. The main drawback to the scale factor damping is the slowing of the reconstruction.

As a comparison to the previous ART example, several outputs of the SIR algorithm are presented. Figures 3.19 and 3.20 show the time and frequency domain outputs of SIR for 25 and 100 iterations. While similar in many respects to the output of the ART algorithms, there is a noticeable lag in development of the reconstruction under SIR. However, a comparison of the frequency content of the SIR output for the noisy case in Figure 3.20 to the corresponding examples for multiplicative ART in Figure 3.18 shows that SIR performs better in the presence of noise than ART. Note also that for increasing iterations, the side lobes within the desired frequency band

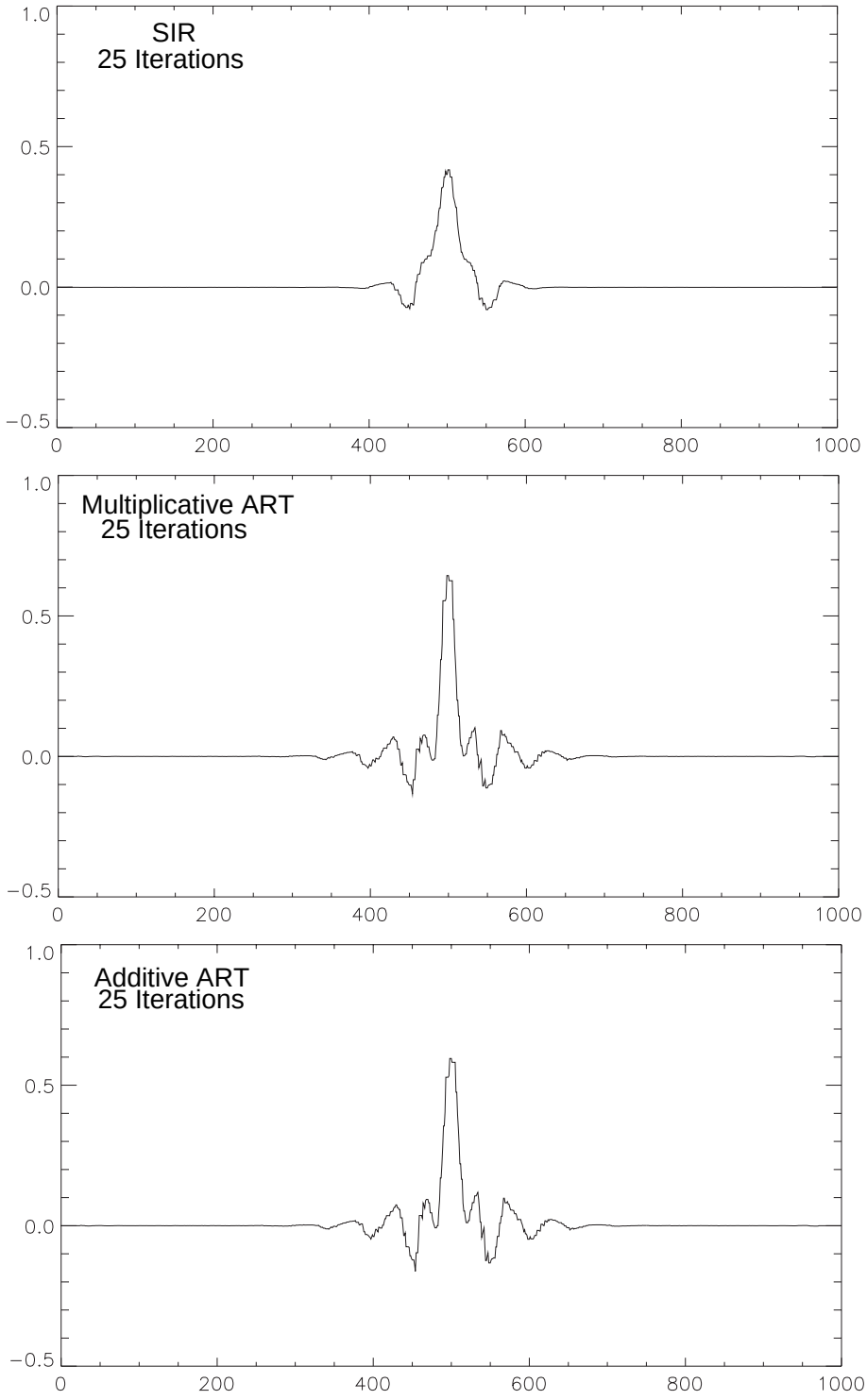


Figure 3.14: A comparison of SIR, Additive ART and Multiplicative ART outputs after 25 iterations for no noise. Note how there is little difference between Additive ART and Multiplicative ART. SIR lags behind the the other two methods, but this is part of the algorithm design.

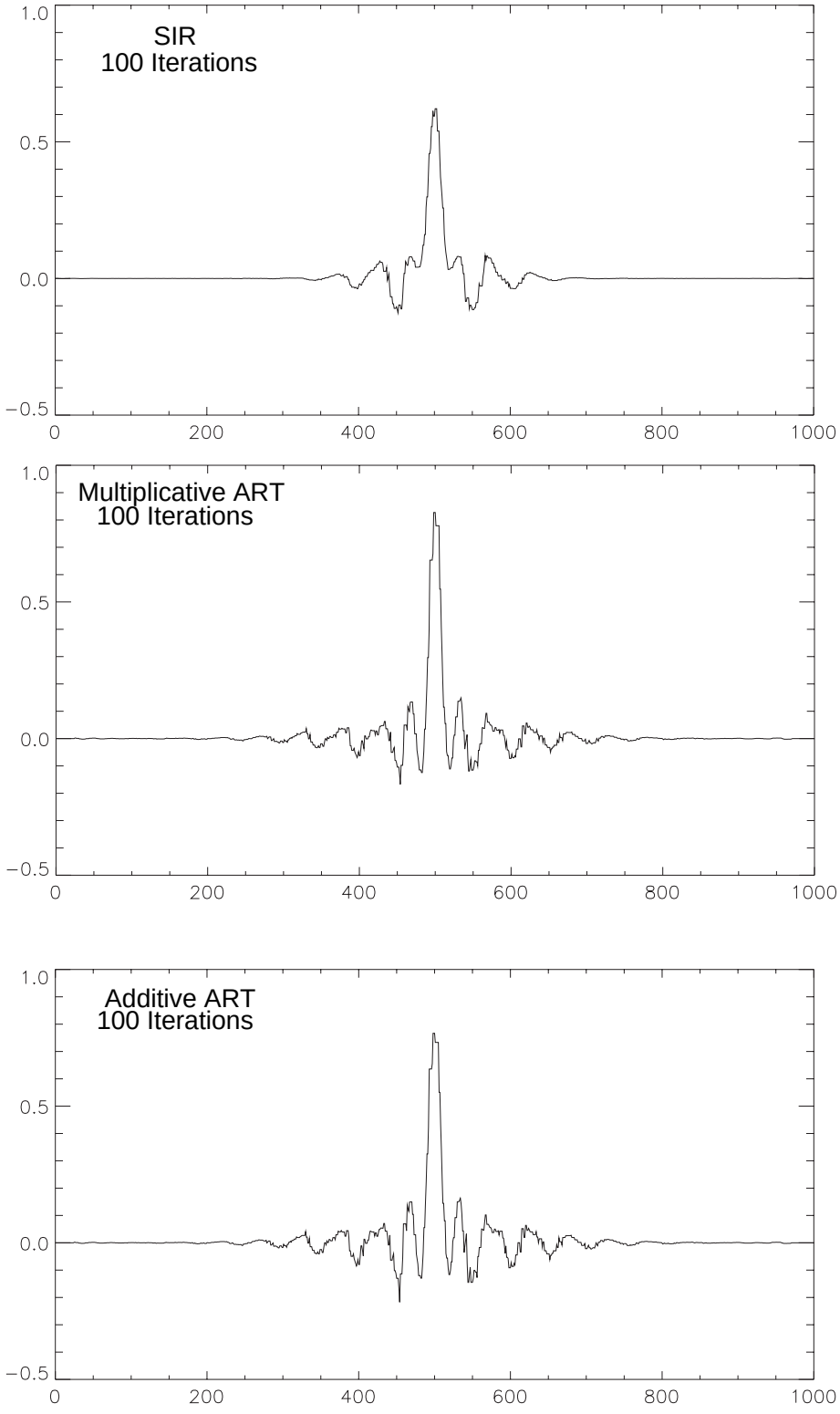


Figure 3.15: A comparison of SIR, Additive ART and Multiplicative ART outputs after 100 iterations for no noise. Note how there is little difference between Additive ART and Multiplicative ART. Compare the SIR output with the Multiplicative and Additive ART outputs in Figure 3.14

.

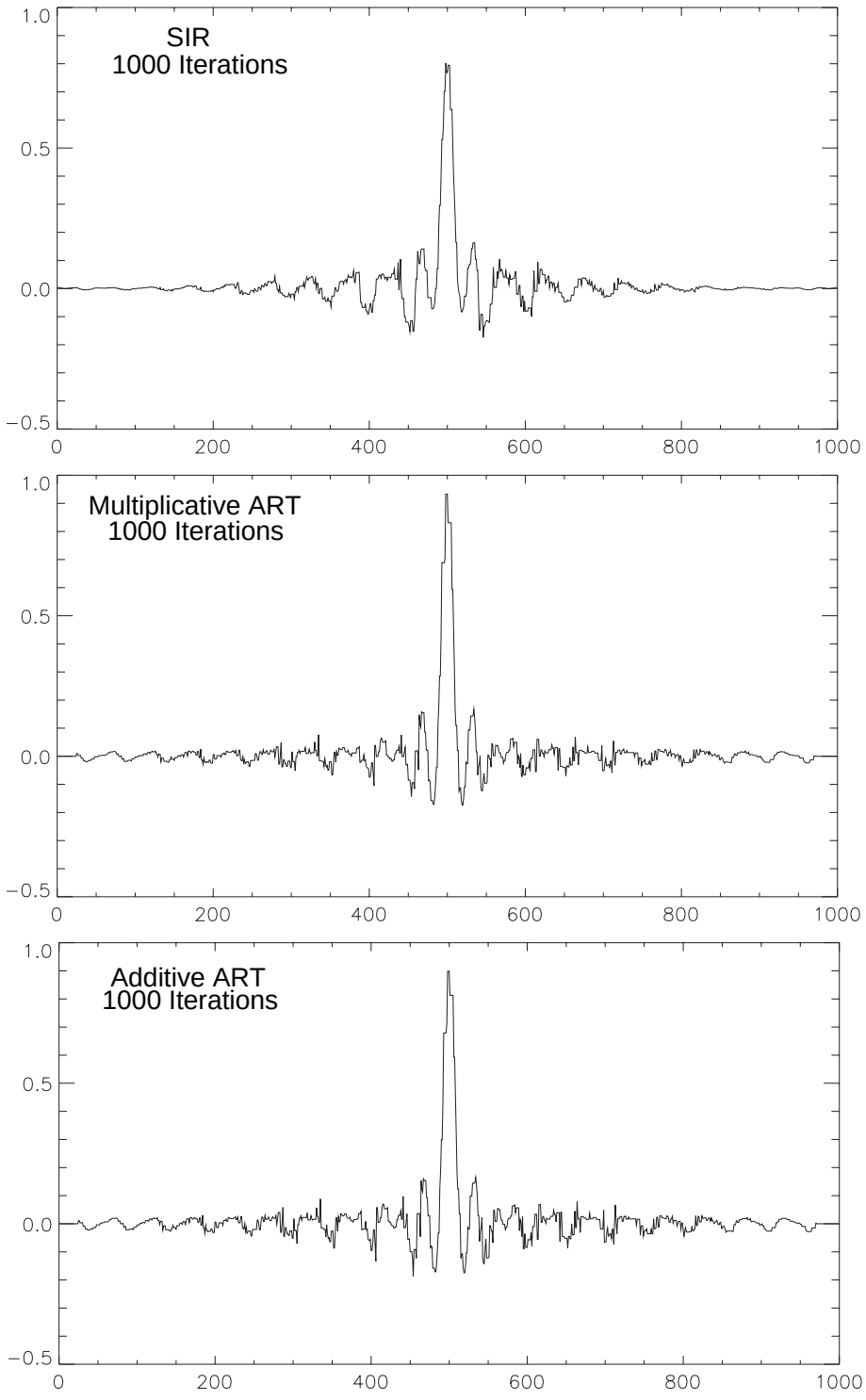


Figure 3.16: A comparison of SIR, Additive ART and Multiplicative ART outputs after 1000 iterations for no noise. Note how there is little difference between Additive ART and Multiplicative ART. Compare the SIR output with the Multiplicative and Additive ART outputs in Figure 3.15

.

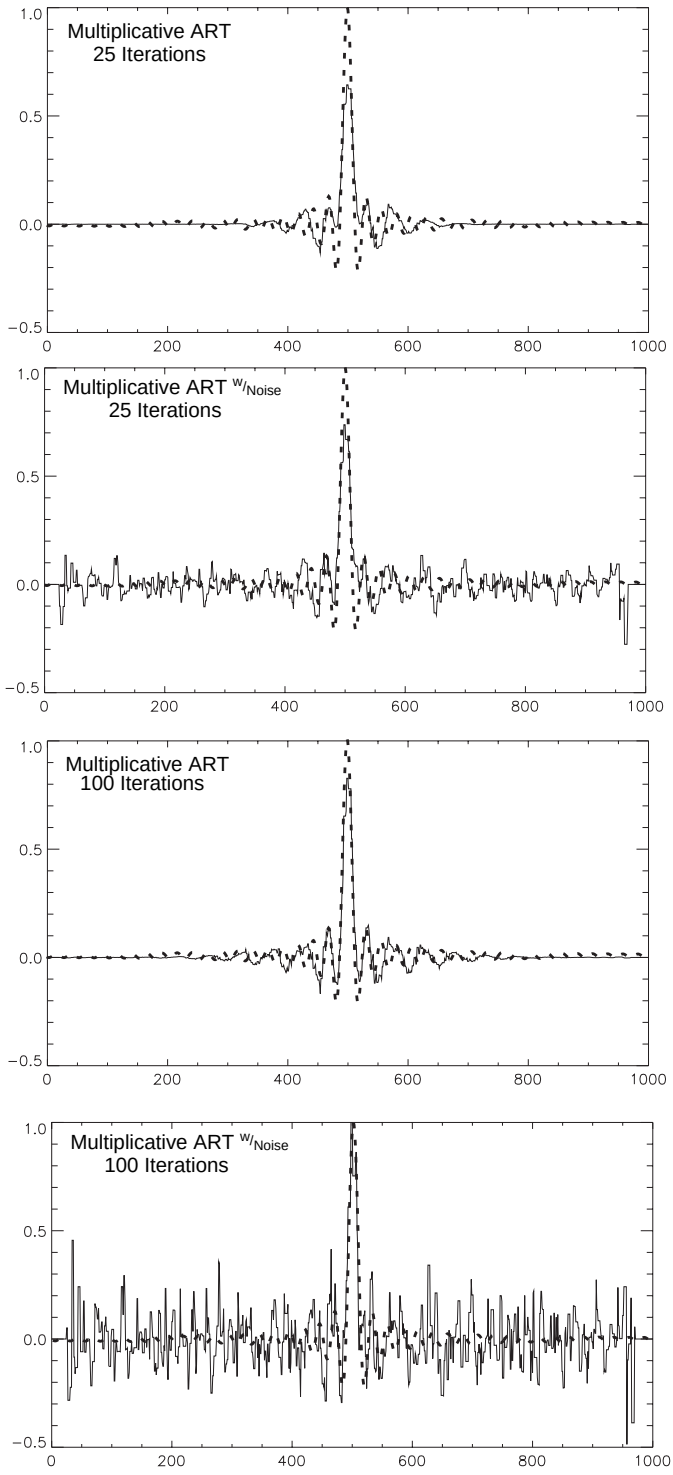


Figure 3.17: Examples of Multiplicative ART for 25 and 100 iterations with and without noise in real space.

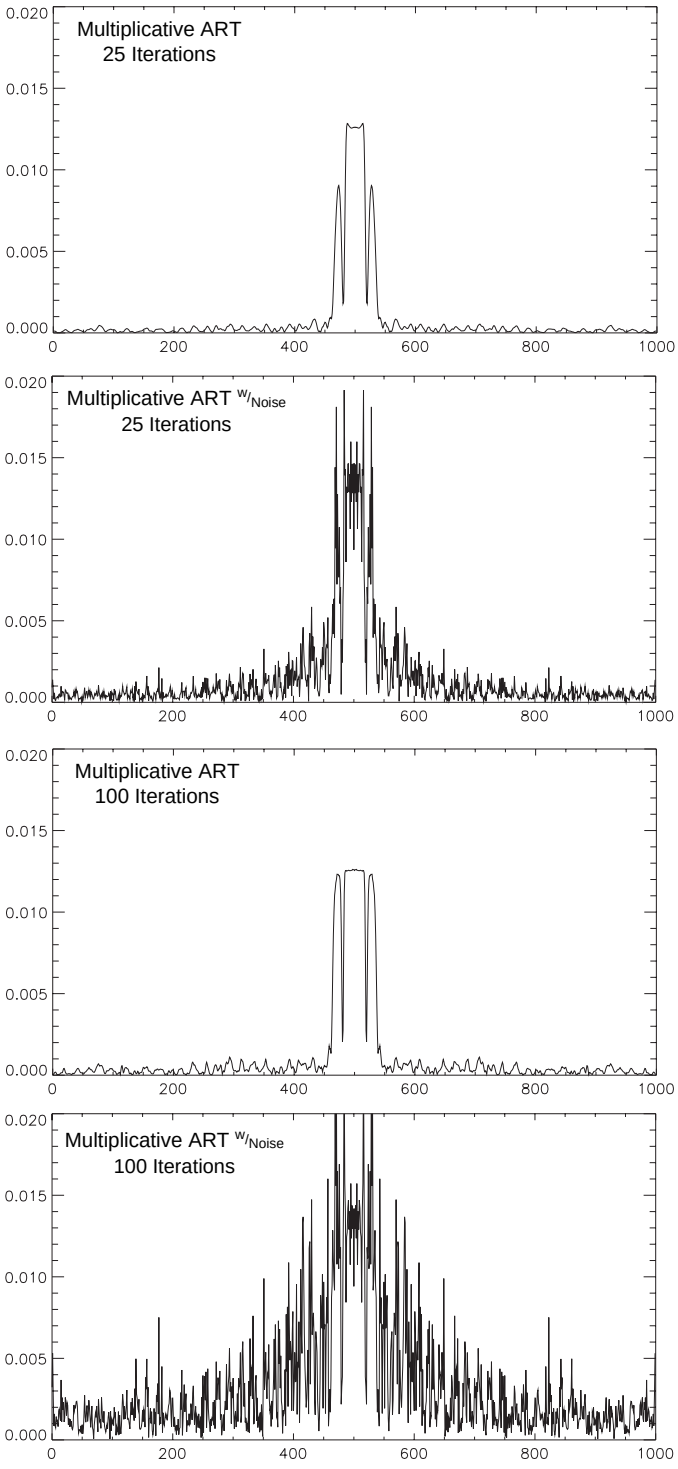


Figure 3.18: An example of Multiplicative ART with and without noise in the frequency domain. These plots correspond to the Fourier transform of the plots in Figure 3.17. Note that for a few iterations, the noise is manageable, but the 100 iterations example shows the noise becoming large relative to the signal.

continue to rise with use of the SIR algorithm just as they do for the ART examples (Refer to Figures 3.18 and 3.20).

Finally, we present some frequency domain example outputs after a large number of iterations. While it is not possible to run the algorithms for an infinite number of iterations, a comparison of algorithmic outputs for a small (25) number of iterations to a large (1000) number of iterations is instructive. Figure 3.21 shows the frequency domain representation of the output for 1000 iterations of all three algorithms. The important thing to note here is the behavior of the side lobes within the desired frequency band. Refer back to Figure 3.12 and note that the aperture function has very low side lobes relative to the test signal frequency band. But the side lobes within the test signal frequency band in Figure 3.21 show significant improvement over the very low expected side lobes of the aperture function. All three algorithms successfully reconstruct the original signal within the limits of the nulls in the aperture function as expected by the earlier theoretical development for the noiseless case. These examples show that even a few iterations are sufficient to reconstruct a significant portion of the desired signal with a residual low pass filter effect lowering the power in the higher frequencies ranges. In previous work, an inverse filter to compensate for the lower side lobes resulting from an incomplete reconstruction has been used with success [33].

3.5.4 Reconstruction Error

We now consider the performance of SIR and MART with noise added. The original motivation for SIR was desire for an iterative algorithm that performs well in the presence of noise. In order to understand the performance characteristics of SIR, the relationship between the number of iterations and the quality of the reconstructed image must be established.

At any iteration, there will be an error between the iterative reconstruction and the true image. In general, iterative reconstructions suffer from two forms of error: reconstruction error and noise amplification. The reconstruction error is the difference between the iterative image and the noiseless true image. Noise amplification, a

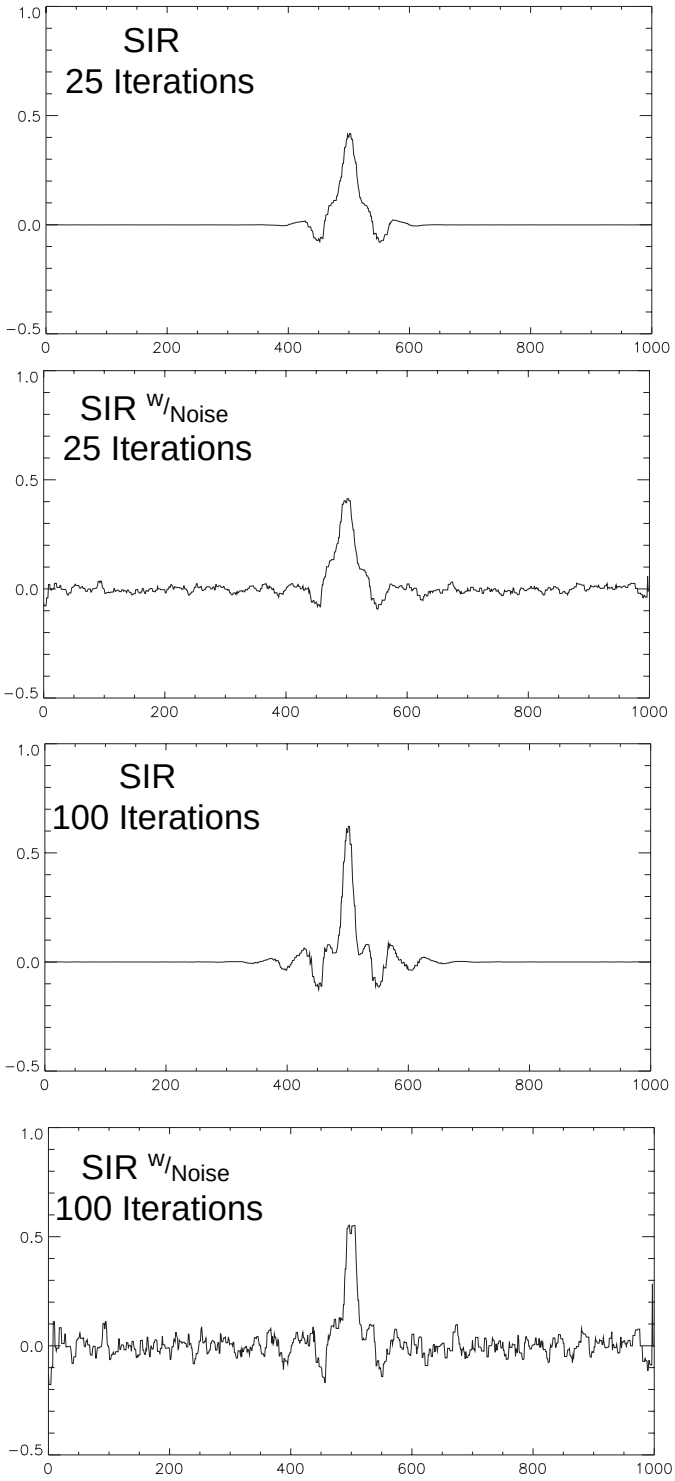


Figure 3.19: Time domain comparison of SIR output after 25 and 100 iterations. Cases with and without noise are shown. Notice how well behaved the noise is compared to a similar number of iterations used in Figure 3.17 for multiplicative ART.

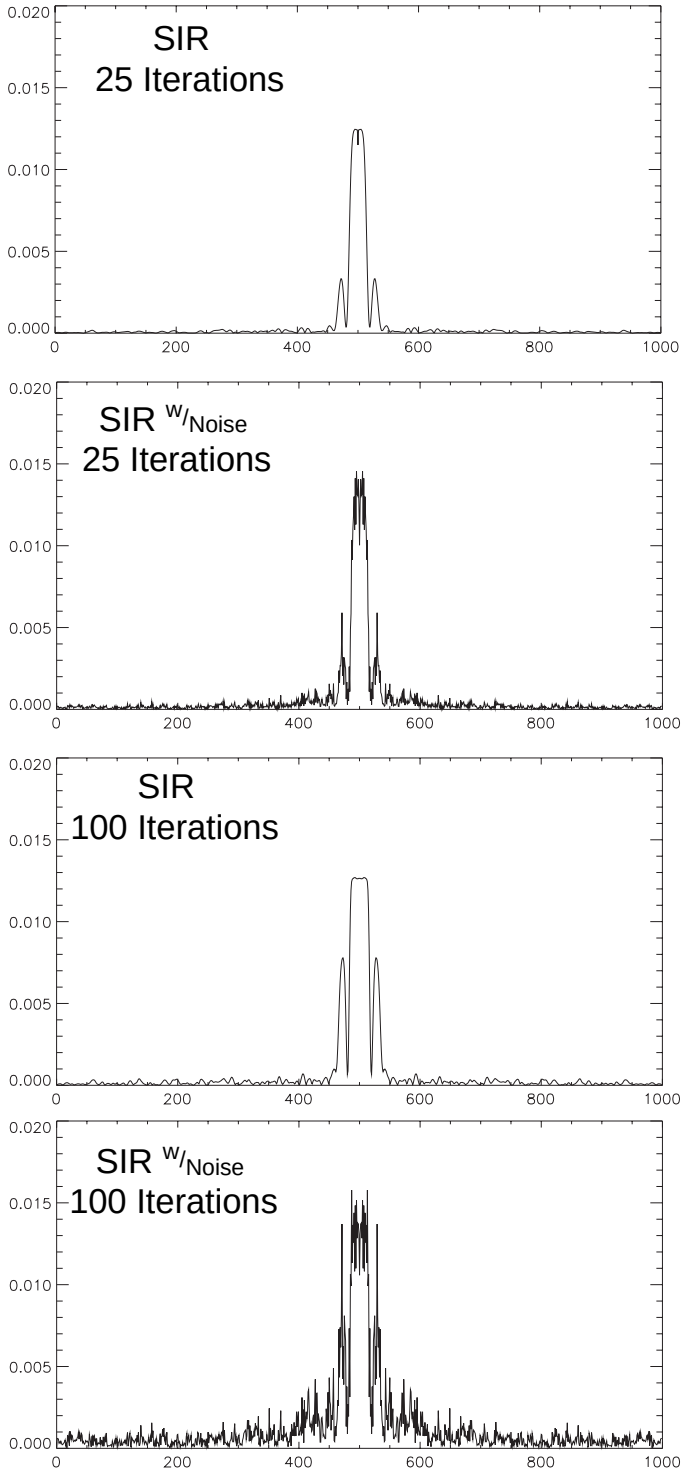


Figure 3.20: Frequency domain comparison of SIR output after 25 and 100 iterations. Compare with Figure 3.18 and note that the noise is amplified less for the SIR examples than the ART examples.

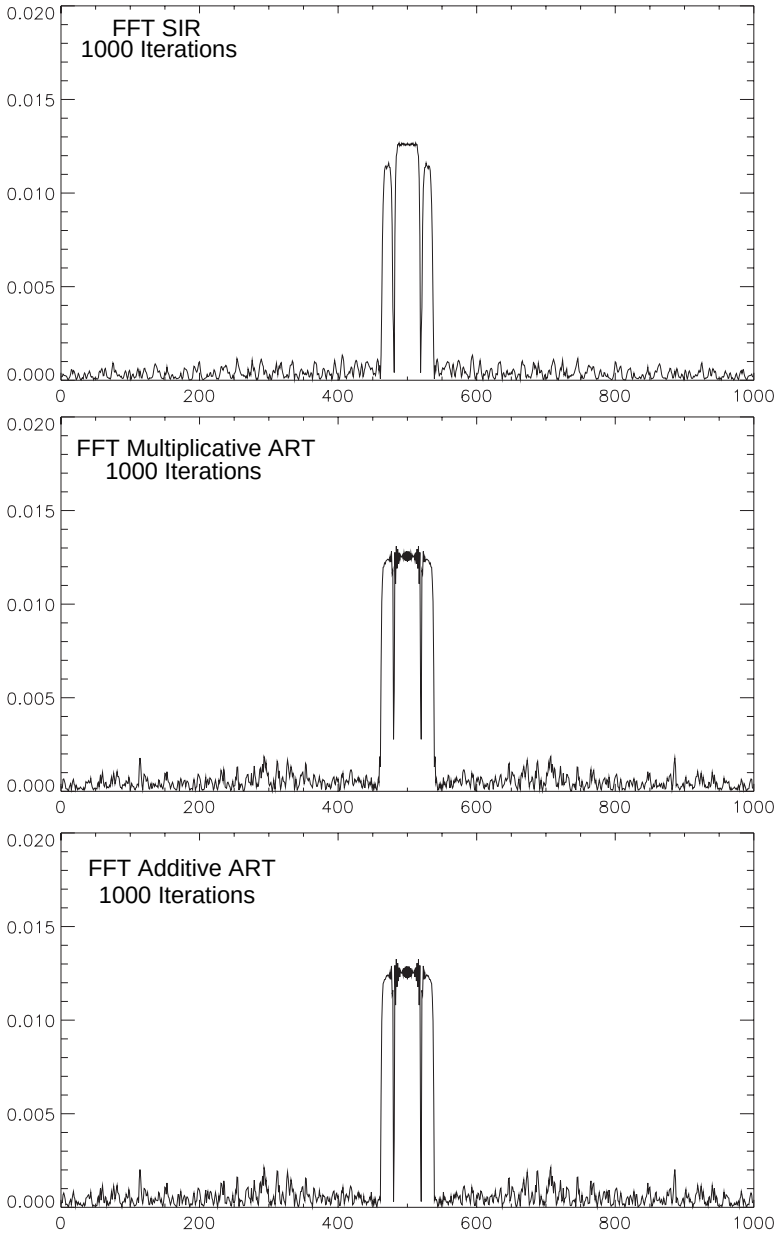


Figure 3.21: A frequency domain comparison of the outputs of all three algorithms after 1000 iterations for the noiseless case. Note how all three examples show significant improvement of the side lobe levels compared with the original levels in Figure 3.12. If SIR continues to run, it shows similar levels of side lobe reconstruction after approximately 4000 iterations.

function of the noise and the number of iterations, results from the inverse filtering of the reconstruction algorithm [21,32].

If the true image is known, the two error components (noise amplification and reconstruction error) can be calculated separately, with the sum equaling the total error for a reconstruction from noisy samples [32]. To illustrate, a graph showing the total squared error versus iterations for the individual components (noise, reconstruction and combined) is presented in Figure 3.22 for the simulation used in previous examples. The graph in Figure 3.22 shows the noise amplification error (En), the reconstruction error (Es), and the total error for the signal plus noise (E) for both MART and SIR. The total error is computed as the sum of the pixel by pixel difference between the true image and the reconstructed image at each iteration. The basic equation is

$$\text{Total Error at Iteration k} \quad = \quad \sum_{j=1}^{n} |t_j - a_j^k| \tag{3.40}$$

where n is the number of pixels, j is the pixel index, t is the true image and a^k is the reconstructed image at iteration k. For the noise amplification error, the true image is the null image (i.e. 0), not the original noise signal. The noise model used is the model in Eq. (2.4). In Figure 3.22, note that the noise amplification is much greater for MART than for SIR. The reconstruction error for SIR, however, is larger than for MART at the same iteration.

The power of SIR is its ability to go further in reducing the reconstruction error than MART without significantly increasing the noise. We note in Figure 3.22 that although MART reaches a minimum total squared error in just a few iterations, the noise amplification dominates the total error in any further iterations. The noise amplification for SIR, however, is much lower. The trade off is a slower reduction in the reconstruction error which continues to decrease with increasing iterations without a large penalty in noise amplification. It is thus possible to get to a lower point on the reconstruction curve compared to MART while adding less noise. This is illustrated in Figure 3.23. This graph shows a log-linear plot of the noise amplification versus the reconstruction error. Note that for any level of reconstruction error, SIR has a lower

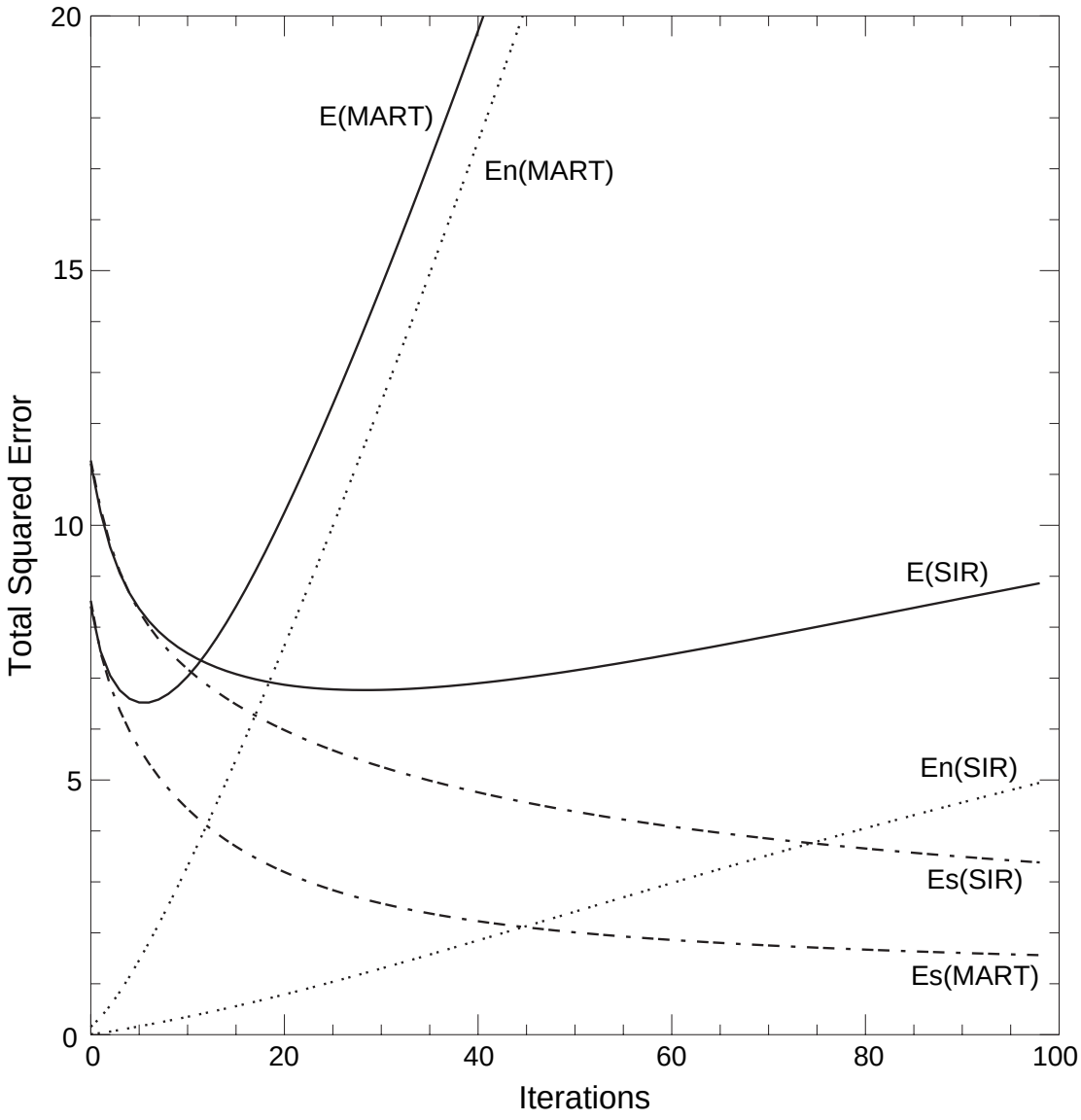


Figure 3.22: A comparison of total squared error for MART and SIR. The error component is split into the contribution from noise error (En) and reconstruction error (Es) [32]. The total error (E) is the sum of the two sub-components. Note that the noise amplification for SIR is much smaller than for MART.

noise amplification than either MART or AART. Thus, for a given signal quality, we have objective evidence that SIR has less noise than MART. This enables a simple tradeoff between reconstruction error (i.e., signal error power) and noise amplification (i.e. noise power).

It should be noted that while total squared error is a rough indicator of the accuracy of the reconstruction, the size and location of the error changes over the course of the reconstruction, migrating, for instance, to regions around edges, etc., depending on the regularization [32]. The quality of the resulting imagery is therefore not a direct function of total squared error [30]. Because SIR damps out areas where high noise level might otherwise destroy the image quality, the image quality for SIR at any given reconstruction error level is better than corresponding MART or AART products. Thus, as shown in previous studies, SIR results in a qualitatively and subjectively better reconstruction than ART as a result of reducing the reconstruction error as illustrated in Figure 3.22 [3,13,34].

3.6 Summary

In this chapter, we have presented the theory behind scatterometer image resolution enhancement. First, the aperture function was shown to be a filter on the original data and influenced the recoverable signal by nulling and attenuating some frequencies in the original data. Next, a theory of irregular sampling was presented. The concept of δ-dense was presented as a measure of the sampling density and we found that the largest sample spacing in the sample set influences the recoverability of the original signal. Gröchenig's algorithm was presented to demonstrate that a signal can be completely recovered from irregular samples, and then Additive ART was shown to be equivalent to Gröchenig's algorithm. Multiplicative ART and Additive ART solutions were shown to be identical in the Banach space defined by the aperture function.

SIR was developed to reduce the influence of noise in the reconstruction of scatterometer data. In the final section, known test data are used to demonstrate the reconstructive abilities of additive ART, multiplicative ART and SIR for noise

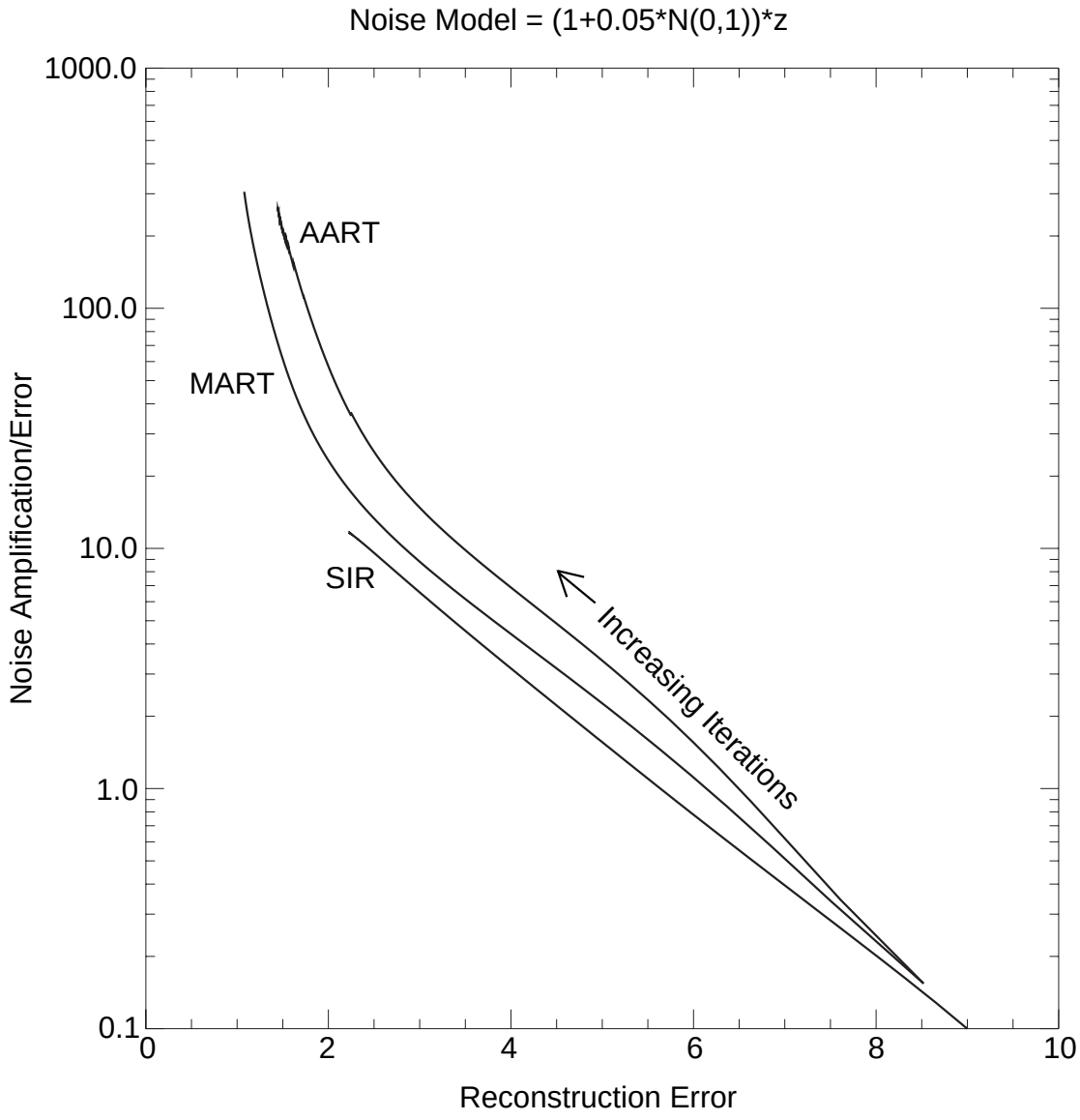


Figure 3.23: A log-linear plot of noise amplification (i.e. noise power) and reconstruction error (i.e. signal noise power). Note that for any given reconstruction error, SIR outperforms MART and AART, allowing a simple tradeoff between noise level and reconstruction accuracy.

free and noisy cases. Multiplicative ART, which SIR is based on, and additive ART typically show very poor performance in the presence of noise. Outputs of the three algorithms using a simple test signal as input demonstrated that SIR indeed has better noise handling characteristics that the ART algorithms. SIR does have a significant lag because of the damping imposed on the scale factors, but this is the tradeoff for noise control, SIR performs substantially better in the presence of noise than comparable ART algorithms.

Chapter 4

Application and Optimization of SIR for the ERS-1 Scatterometer

4.1 Introduction

Spaceborne scatterometers are designed to produce near-surface wind maps over the ocean. These maps consist of estimates of wind speed and direction for each cell across the scatterometer swath. Since the winds are rapidly changing, only a single pass of the satellite is used to construct the wind estimate for a given ocean area. An example wind map from ERS-1 data over the Southern Ocean is shown in Figure 4.1. Each arrow in the figure represents the estimate of the wind speed and direction at that point using all beams of the scatterometer. Because geostrophic winds vary slowly over the 25km spacing of the scatterometer readings, this map accurately reflects the near-surface characteristics of the local geostrophic winds at the time of this 'snapshot' of the surface.

A scatterometer does not measure the wind directly. Instead, it measure the backscattered power σ^o, the normalized radar cross section of the surface, which can then be related to the wind through a geophysical model function. In a similar manner, σ^o over land can be related to surface conditions, making maps of σ^o over land useful for geophysical studies.

In this Chapter, we examine the behavior of the Scatterometer Image Reconstruction (SIR) algorithm applied to ERS-1 scatterometer data. The SIR resolution enhancement algorithm was originally developed and optimized for the SEASAT scatterometer (SASS). Subsequent modifications of the SIR algorithm for use with

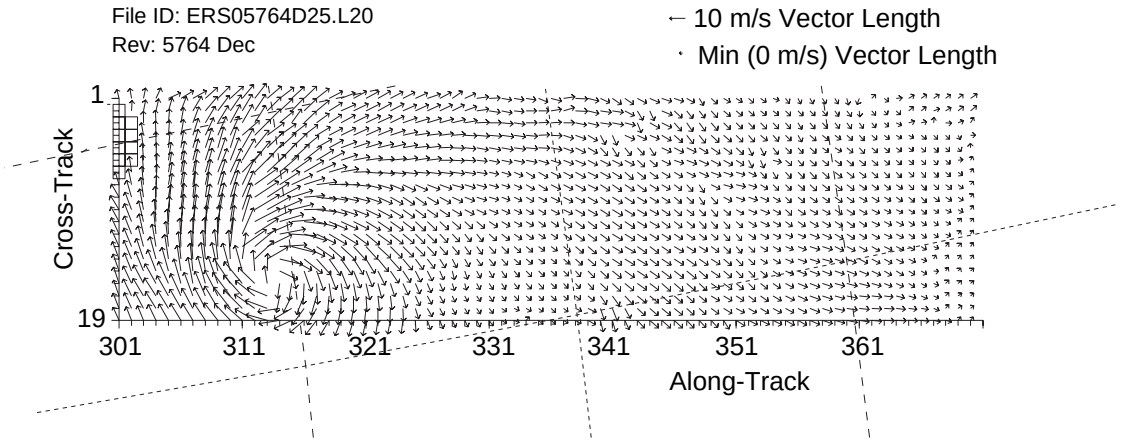


Figure 4.1: This is an example wind map generated from ERS-1 scatterometer data. Each arrow represents the wind speed and direction estimate at that point and the arrows are arranged on a 25km grid. The wind estimate is based on an established geophysical model function that relates azimuth and incidence angle and scattered power to the wind speed and direction. The cross track numbering indicates the along-beam cell number (there are 19 cells across the swath) and the dotted lines overlaying the map are longitude and latitude lines.

ERS-1 data have perpetuated, in general, iterative weights originally established for SEASAT data. In previous studies, enhanced resolution scatterometer $\mathcal{A}$ value images from ERS-1 data have been used to study earth surfaces [18, 35–41]. The $\mathcal{B}$ images, however, have had little utility because they characteristically have had little dynamic range in the $\mathcal{B}$ value. An example of an $\mathcal{A}$ and $\mathcal{B}$ image created from ERS-1 data is shown in Figure 4.2. Notice that there is very little information in the $\mathcal{B}$ image because of the narrow range of the $\mathcal{B}$ values.

In the original SASS SIR algorithm, the $\mathcal{B}$ value update is heavily damped (i.e. the $\mathcal{B}$ value was not allowed to change very much in any one iterative step) by reduction of the weighting value in Eq. (2.18) to control noise in the $\mathcal{B}$ image [14]. Using synthetic images, it has been determined that for ERS-1 data, the $\mathcal{B}$ value is so damped by the original SASS weighting that the SIR algorithm is not able to iteratively update the $\mathcal{B}$ value fast enough to reach the original value of the synthetic image. This is especially true if the true value differs greatly from the value used to

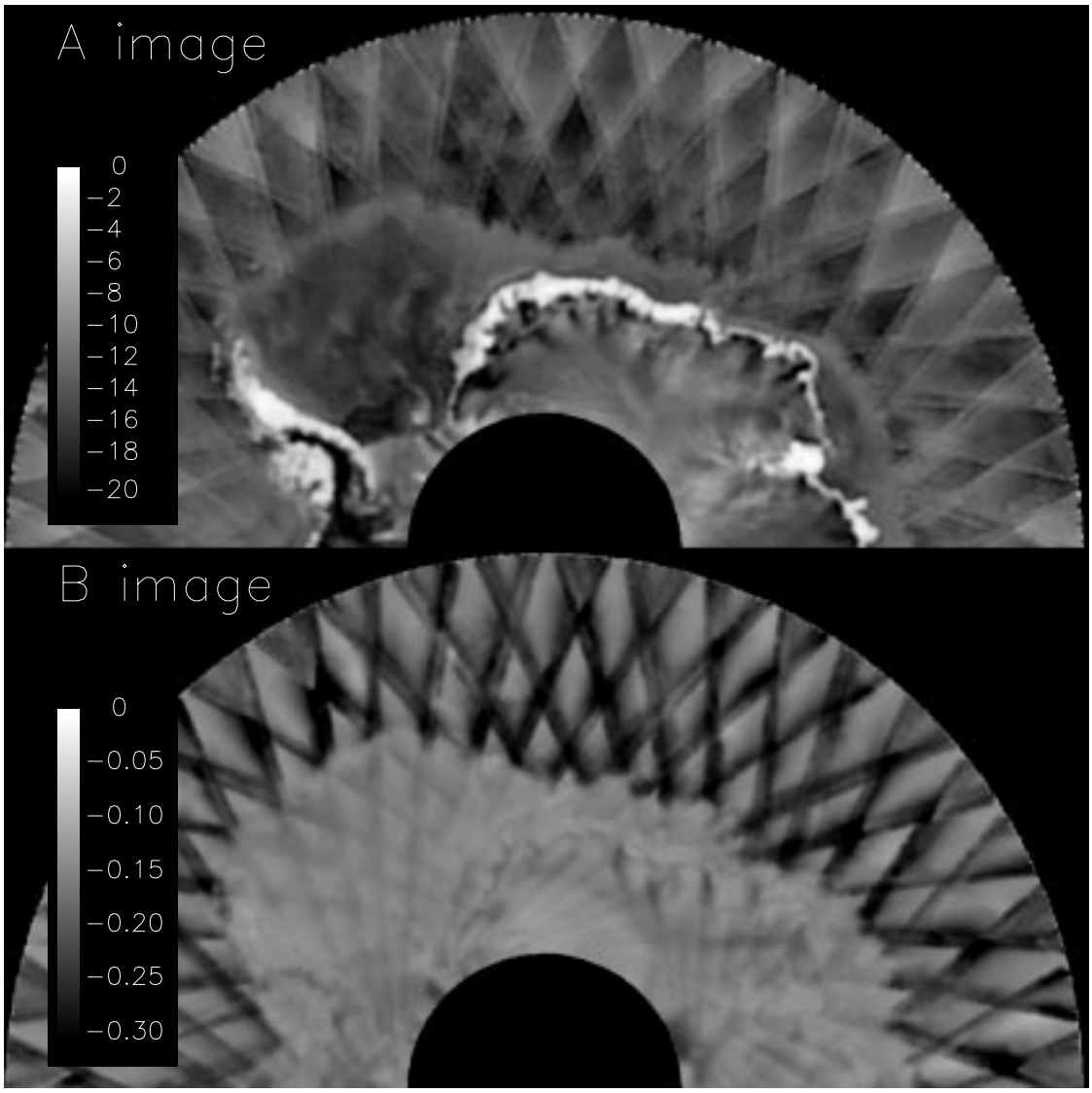


Figure 4.2: A comparison of $\mathcal{A}$ and $\mathcal{B}$ SIR images. This image is JD 120 1994. The $\mathcal{B}$ image is normalized from 0 to -0.3 and the $\mathcal{A}$ image is normalized from 0 to -20. The $\mathcal{B}$ image has very little detail, and the cross hatches over the open ocean are quite pronounced. Some cross hatches also appear in the sea ice in both the $\mathcal{A}$ and $\mathcal{B}$ images.

initialize the SIR iterative process. Other potential issues regarding the application of the SASS SIR routine to ERS-1 data include the initialization value and the limitations of the linear model discussed in Section 2.1. Also, as mentioned in the previous chapter, the theoretical reconstruction requires an infinite number of iterations, but noise considerations dictate that only a few (typically 25-50) iterations are possible. Slower than possible development of the $\mathcal{B}$ value is accentuated by the small number of iterations, and we wish to maximize the utility of the iterations we do run.

This Chapter is divided into three main sections: First, we present specifics about the ERS-1 scatterometer, including the aperture function and a discussion about how multiple overlapping measurements are combined to create a dense sample set similar to the one discussed in Chapter 3. The second section presents the results of several experiments on the effectiveness of increasing the accuracy of the SIR $\mathcal{A}$ and $\mathcal{B}$ values by modifying the algorithm. The following areas are discussed: SIR initialization values, using a cubic rather than linear model for backscatter, and modifying the $\mathcal{B}$ development weighting. The final section examines the cumulative effects of these optimization modifications.

4.2 The ERS-1 Scatterometer

4.2.1 The ERS-1 Aperture

The ERS-1 scatterometer has a circular aperture function with a $\cos^2$ roll off associated with each measurement cell. The aperture function is illustrated in Figure 4.3. The aperture function is assumed to have a region of support 100 km in diameter. To get the filter response of the aperture function, the two dimensional Fourier Transform is used on the $\cos^2$ circular aperture function and the resulting frequency response is plotted by numerically evaluating the Fourier Transform of the aperture function equation (derived in Appendix A):

$$F(\xi_1, \xi_2) \;=\; 2\pi \int_0^{50} d\rho\, \rho \cos^2(\beta\rho) J_0(k\rho). \tag{4.1}$$

The frequency response is shown in Figure 4.4. Note that the Fourier Transform is a low pass filter with the first null at approximately the 45 km wavelength. Because the

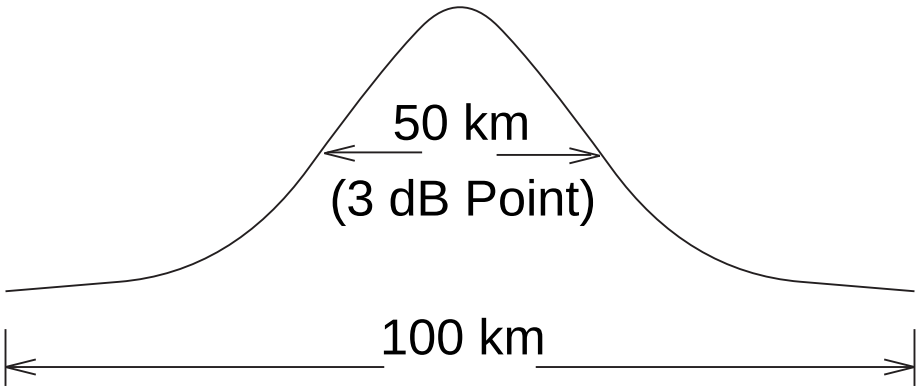


Figure 4.3: Geometry of the ERS-1 Footprint. Each measurement cell has a latitude and longitude that specify the center of the cell. This figure shows a slice through the circular footprint. The σ^o is assumed to have a cosine-squared roll off characteristic, with the 3 dB point at 50km. [42]

ERS-1 scatterometer has a very high signal to noise ratio, some of the highly attenuated higher frequencies are still far enough above the noise floor for the information to be recoverable. Once outside this small band of recoverable frequencies, the roll off of the aperture function is valuable in that it prevents aliasing.

The aperture function, in effect, defines the recoverable frequency content by determining what frequencies have enough energy above the noise floor to be useful in a reconstruction. An examination of Figure 4.4 leads to some conclusions about the band limits of recoverable signals using the ERS-1 data model. The figure indicates that a reasonable band limit would be about 30 km, or the location of the second null in the frequency plot. It should be emphasized again that it is the noise level with respect to the aperture, not the aperture function alone, that determines the recoverable frequencies. A carefully chosen aperture function with high side lobe levels has a better recoverable bandwidth for a given noise level than an aperture with very low side lobes. This is demonstrated later in this chapter.

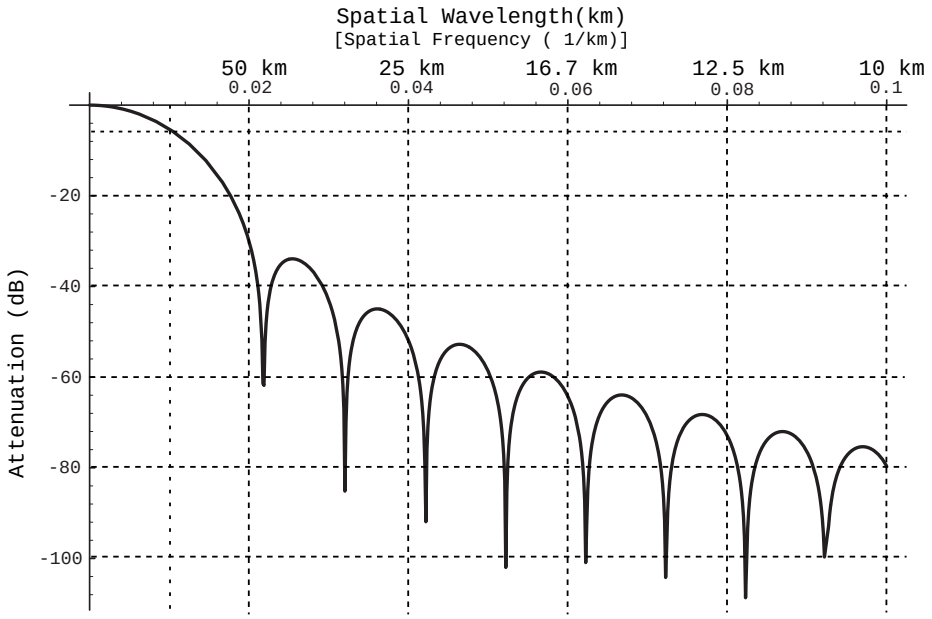


Figure 4.4: This figure illustrates the frequency response of the ERS-1 aperture function. This graph was created by a numerical computation employing Bessel functions and based on the exact Fourier Transform of the $\cos^2$ roll off function. The first null occurs at approximately the 45 km wavelength, and the 3 dB point is at the 100 km wavelength.

4.2.2 Sampling of the Surface

In this section, we consider how sampling affects the frequency content of the reconstructed images. First, consider a single pass of the satellite over a surface to be sampled. Figure 4.5 shows the ERS-1 cell locations for a single pass over a 700 by 700 km region near the south pole. The cells are spaced 25 km apart both in the along track and cross track directions.

Assume that the surface is static (i.e. unchanging) as the satellite passes overhead for a single pass. Each measurement cell in a single pass represents a "sample" of the surface on a regular 25x25 km grid. Because the satellite has a finite aperture, there is an aperture function associated with each sample of the surface. For the moment we ignore the aperture function and assume that sampling is accomplished with a delta function.

Theoretically, if the surface is band limited and sampled uniformally on a grid with Δx and Δy spacing, the original image can be recovered from its samples if the sample rate is greater than the Nyquist rate given by

$$\frac{1}{\Delta x} = \xi_{xs} > 2\xi_{x0} \tag{4.2}$$

$$\frac{1}{\Delta y} = \xi_{ys} > 2\xi_{y0} \tag{4.3}$$

where ξ_{x0} represents the bandwidth of the surface in the x direction and ξ_{y0} represents the bandwidth of the surface in the y direction. According to the Nyquist sampling criterion, sampling on a grid with 25 km spacing results in 50 km wavelengths being the smallest recoverable wavelength in the satellite data. If the surface has wavelengths smaller than 50 km, sub-sampling at a rate less than the Nyquist rate causes aliasing of the higher frequency components of the image. So, the very best resolution we can obtain from a *single* pass of the satellite is 50 km per pixel. Consideration of the aperture filtering and high frequency aliasing could further reduce this resolution.

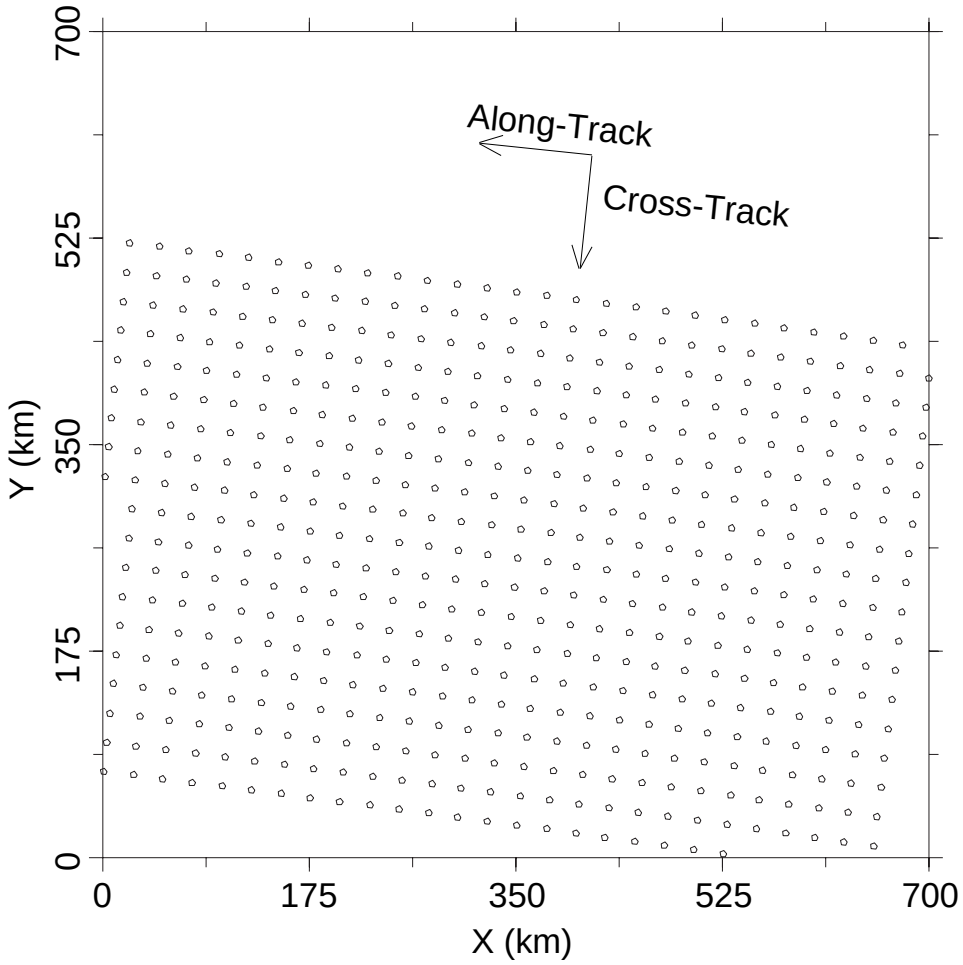


Figure 4.5: An illustration of one pass of the ERS-1 scatterometer over a 700 km square region in the Antarctic. The circles are centered on the cell center locations. Note how the measurements are on a regular 25 km grid. The actual footprint size is approximately 70 km.

4.2.3 Multiple Passes

While there is no way to reduce the cell spacing and thus increase the resolution on a single pass, multiple passes can be used to create sample sets with more closely spaced samples. We now use the fact that the ERS-1 scatterometer observes the same region on another pass and combines the results of multiple passes and assume that the surface remains motionless between passes, even though there is some motion.

The effects of motion on the SIR algorithm are presented in Appendix B. This previously published study demonstrated that motion had a low pass filtering effect on the resulting images. Large motion (greater than 15 km per day) tended to create more low pass filtering than would be evident from the effects of the aperture filter alone. The general conclusion is that edges are blurred in region with high motion. The only real problem would be for small features such as ice bergs which, if moving at sufficient velocities, may not appear in the resulting imagery. However for large areas of one ice type, the main result of high speed motion would be to blur the edges between ice classes.

For the purposes of this study, we will make the assumption of a motionless surface, with the understanding that there is some motion that may result in some blurring within the image. Each pass of the satellite now samples the same surface on a 25x25 km grid that is oriented to the satellite track. Subsequent passes along different tracks are also on a regular 25x25 km grid, but the grids from the various passes are not aligned (see Figures 4.6 and 4.7 for examples of multiple tracks). Because the satellite does not take cell readings at exactly the same points from pass to pass, combining multiple passes results in a more dense but irregular sampling grid. Figure 4.6 shows how two passes combine to give a more dense sampling than only one pass is capable of. Note that the sampling is irregular and that it is more dense in some places than others.

Figure 4.7 illustrates samples of the surface accumulated over a 5 day period (representing data from approximately 32 different passes) where the circles in Figures 4.5 and 4.6 have been replaced by points for clarity. Note how the sampling

density is much better than the nominal 25 km grid of a single pass although the new sampling is very irregular. The irregularity of the sampling is a function of orbit geometry and which instruments are being used on the satellite (e.g. use of the SAR precludes the use of the scatterometer).

4.3 Achievable Resolution using SIR

In the previous chapter, SIR is established as a valid operator for reconstructing irregularly sampled data. In this section, we discuss the achievable resolution of SIR. The theoretical limit of SIR resolution enhancement will be no greater than the theoretical limit of resolution enhancement achieved by the underlying ART algorithm since SIR is a non-linearly damped version of multiplicative ART. Therefore, as a first cut at a summary of SIR resolution enhancement, we present the limits of the ART algorithms.

As discussed in Chapter 3, additive ART is an operator under Gröchenig's theorem and therefore results in a complete reconstruction in infinite iterations. We acknowledge that multiplicative ART, on which SIR is based, does not strictly result in a complete reconstruction due to the regularization imposed in the algorithm. However, since the reconstruction results in two components, one from the measurement space and one from the null space, we assume that the reconstruction is complete if the null space component is ignored. Recall the two basic conditions for the reconstruction to be complete according to Theorem 1 in Chapter 3: First, that the operator be bounded on an appropriate band limited Banach space, and second that the sample spacing be sufficient for the frequency content of the original surface (e.g. a minimum of the Nyquist rate).

The first requirement, that the operation be limited to functions within a Banach space, is straightforward and implies that the original surface function is band limited. In practice, the earth's surface is not a band limited surface, but we apply a band limiting filter to create a band limited surface function. When SIR is used, a median filter is applied to the data during processing, effectively band limiting the data.

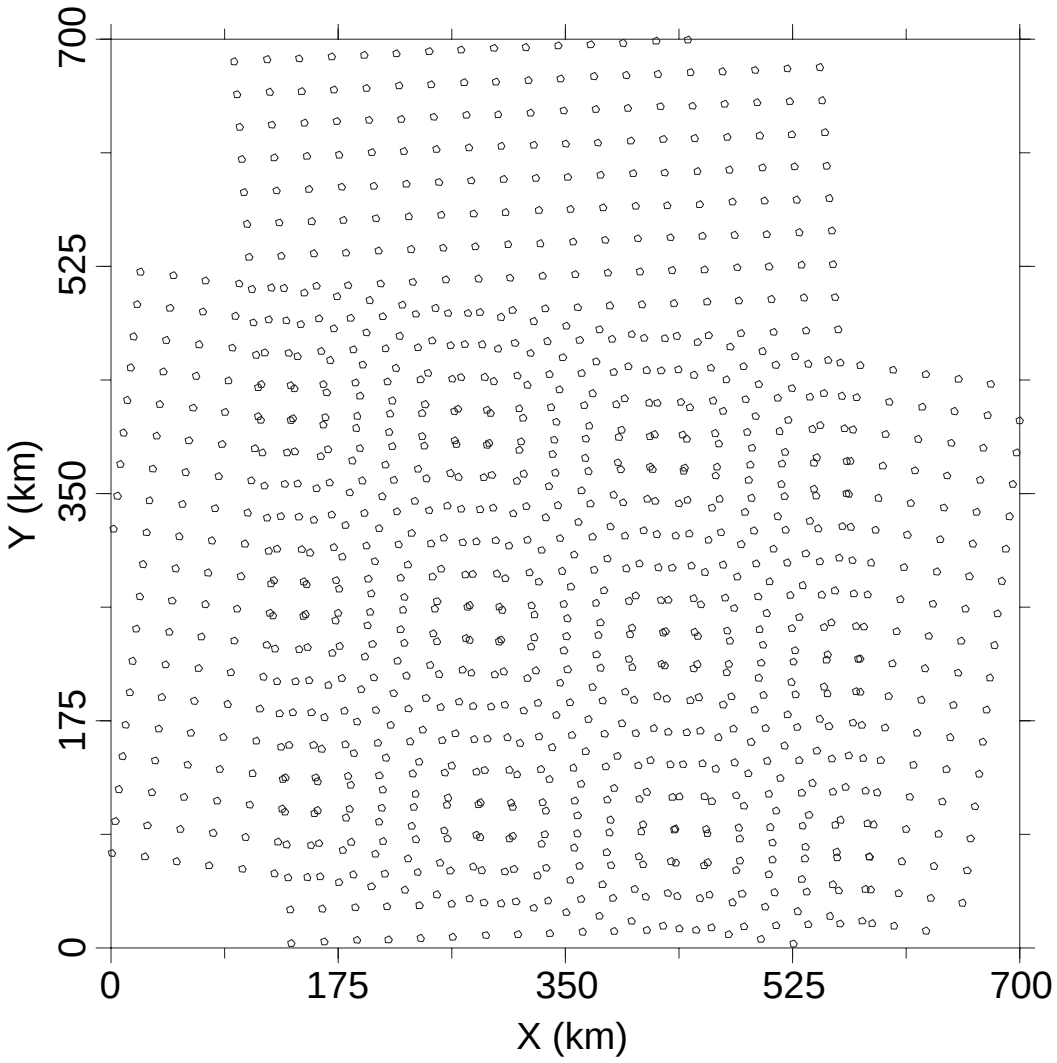


Figure 4.6: An illustration of two passes of the ERS-1 scatterometer over the same 700 km square region as in Figure 4.5. Compare to Figure 4.5 and note the areas of denser "sampling" where the two passes intersect.

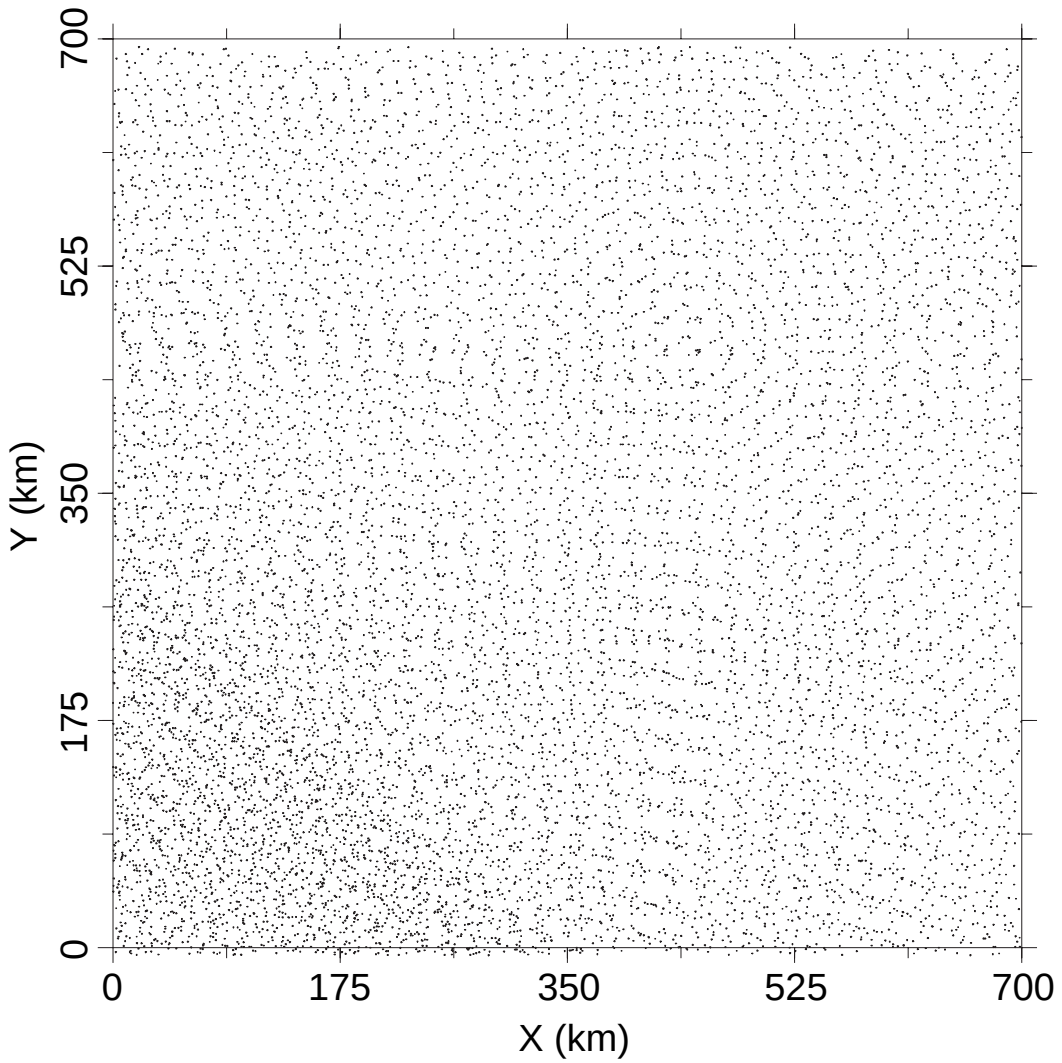


Figure 4.7: This figure illustrates all passes of the ERS-1 scatterometer over a 700 km square region over a 5 day period. The circles have been replaced by points for clarity. The density of the surface sampling is significantly better than for the single, 25km spaced sampling of a single pass.

The second requirement turns out to be the most limiting factor in this reconstruction scheme. Note that even if we severely attenuate portions of the spectrum of the original surface function, Gröchenig's theorem states that a complete reconstruction occurs in the limit. Therefore, no matter what the aperture filter function does to the original surface function, with the exception of the nulls created by the aperture function, as long as the sample spacing is sufficiently close, a complete reconstruction is possible.

However, in actual use we do not perform infinite iterations and therefore fall somewhat short of a complete reconstruction. Note in Figure 3.20 how the highly attenuated side lobes of the sample data fall short of reaching the correct peak in limited iterations. If an aperture function severely attenuates a particular frequency, limited iterations may not be able to fully recover the original image. For ERS-1, this represents a significant limitation since the side lobes of the aperture function are quite low.

Other scatterometers, such as NSCAT (1996-1997) and SASS (1978), have very different aperture functions [2] [3]. In both cases, the aperture functions are long and thin and vary in shape and orientation depending on the cell location on the antenna beam and orbit location of the instrument. This provides more high frequency information through higher side lobes along the short axis of the cell, and therefore offers better reconstruction possibilities than ERS-1. We now examine the theoretical limits on resolution enhancement for both ERS-1 and NSCAT.

4.3.1 ERS-1 Resolution

We begin by considering the aperture function. The aperture function for ERS-1 is circularly symmetric, as is the frequency domain response (see Figure 4.4), and the side lobe levels are quite low. As noted in Chapter 3, theoretically the side lobe levels are irrelevant to the information recovery process for a noiseless case. Since infinite iterations are not generally possible in a practical application, low side lobe levels may not be fully recovered in finite iterations. Nevertheless, simulations show that resolution enhancement is ultimately limited by sample spacing.

We now analyze the sample spacing and note that by using multiple passes over a 6 day period, the average sample spacing in ERS-1 data taken in the Southern Ocean is 12 km in both the x and y directions. Note that sample spacing varies by location and number of days used in the reconstruction. Using Eq.(3.17), the limit for Gröchenig's algorithm, the recoverable wavelengths λ_o are

$$\lambda_0 = \quad \tfrac{1}{\omega_0} > \quad \frac{2\delta_0}{ln(2)} = 34.6\,\text{km} \tag{4.4}$$

$$\lambda_0 > \quad \tfrac{2(12\,\text{km})}{ln(2)} \tag{4.5}$$

$$\lambda_0 > \quad 34.6\,\text{km} \tag{4.6}$$

which is consistent with experimental results for the ERS-1 scatterometer published in [29] and in Appendix B where the resolution is approximately 35 km from experimental images. This is an improvement over the nominal 50 km wavelength limit imposed by the 25 km grid and the Nyquist Criteria, even considering the nulls introduced by the aperture function.

4.3.2 NSCAT Resolution

While the expected increase in resolution of the ERS-1 data is an improvement, the resulting images are still fairly large scale. The two limitations of ERS-1 data in resolution enhancement are the aperture function and the sample spacing. An improvement in side lobe levels would result in more information recovered in limited iterations. Denser sampling is needed to further increase the resolution enhancement. The NSCAT scatterometer, which recently completed collecting nine months worth of data, has potential for much greater resolution enhancement with SIR than the ERS-1 scatterometer because of much higher side lobe levels and more densely spaced samples [2].

We begin by examining the sample spacing and note that multiple passes over 6 days result in an average spacing in NSCAT data taken in the Southern Ocean of about 4 km in both the x and y directions, again varying with the exact location and the number of days used. Using Eq.(3.17), the limit for Gröchenig's algorithm,

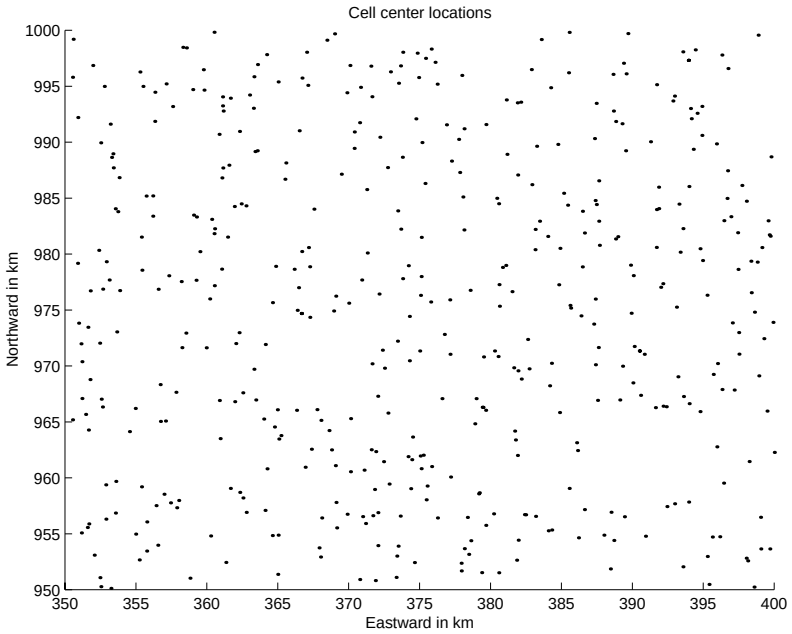


Figure 4.8: An example scatter plot of NSCAT cell centers over the Weddell sea using 6 days of data. Note variation of density of the cell centers over the area, ranging from 1 to 8 km and sometimes more in some regions.

the recoverable wavelengths λ_0 are

$$\lambda_0 = \frac{1}{\omega_0} > \frac{2\delta_0}{ln(2)} \tag{4.7}$$

$$\lambda_0 > \frac{2(4\,\mathrm{km})}{ln(2)} \tag{4.8}$$

$$\lambda_0 > 11.5\,\mathrm{km}. \tag{4.9}$$

However, the density of measurements varies somewhat more than for ERS-1, and ranges from about 1 km to 8 km, depending on the region. A sample scatter plot of NSCAT data over the Weddell sea is presented in Figure 4.8. The variation in sample spacing will cause some spatial variation in the actual wavelengths recovered in the reconstruction, but for the purposes of this study, we will concentrate on an average response based on the average spacing of the cell centers. A graph of the variation of smallest recoverable wavelength versus nominal sample spacing according to Eq.(3.17) is given in Figure 4.9 for reference. Note that the limit is about 3 times the sample spacing, where Nyquist is 2 times the spacing.

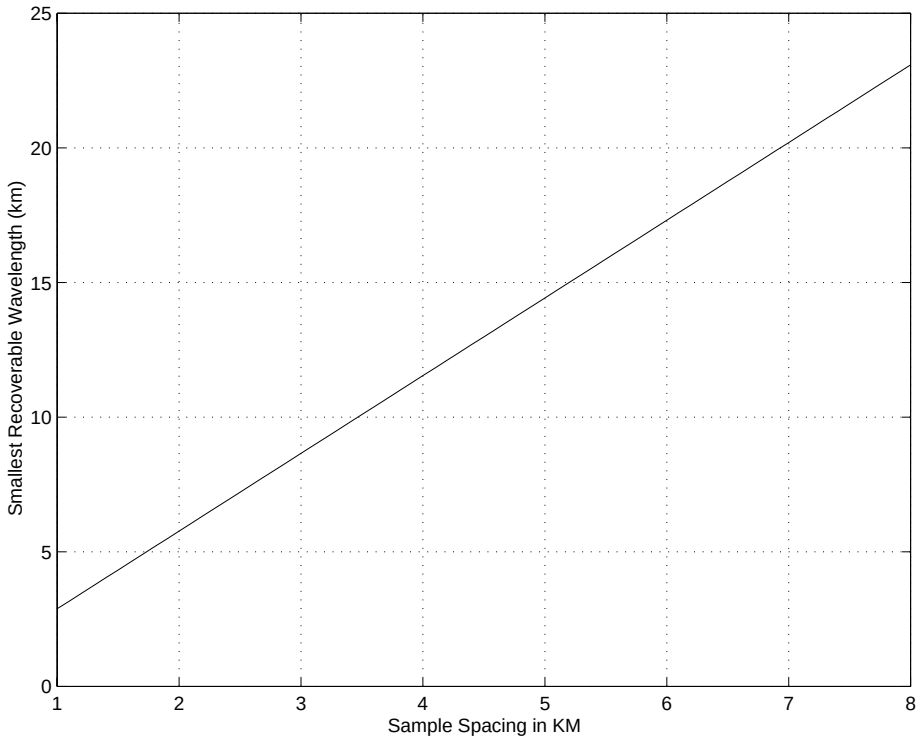


Figure 4.9: This plot shows the variation of the smallest recoverable wavelength as a function of the sample spacing based on Eq.(3.17). Note that the variation is linear and is about 3 times the nominal sample spacing where Nyquist is 2 times nominal spacing.

Unlike the ERS-1 aperture function, the NSCAT aperture is not symmetric. The NSCAT aperture is an elongated six sided cell similar to the ones illustrated in Figure 4.10. In this figure, six different cells from various antennas and locations along the antenna are presented for illustration. It should be noted that while the cell geometry changes slightly along the measurement beam, and varies with the orbit location, the general cell shape is fairly consistent: long on one axis, short on another with the width and height determined by the exact cell location. The cells are very narrow (8 - 15 km) on one axis, so the frequency domain response is quite wide, allowing more high frequency content to be retained in the measured data compared to ERS-1.

Recall that as the satellite orbits, measurements are made at several different azimuth angles. For ERS-1, this results in additional surface information for

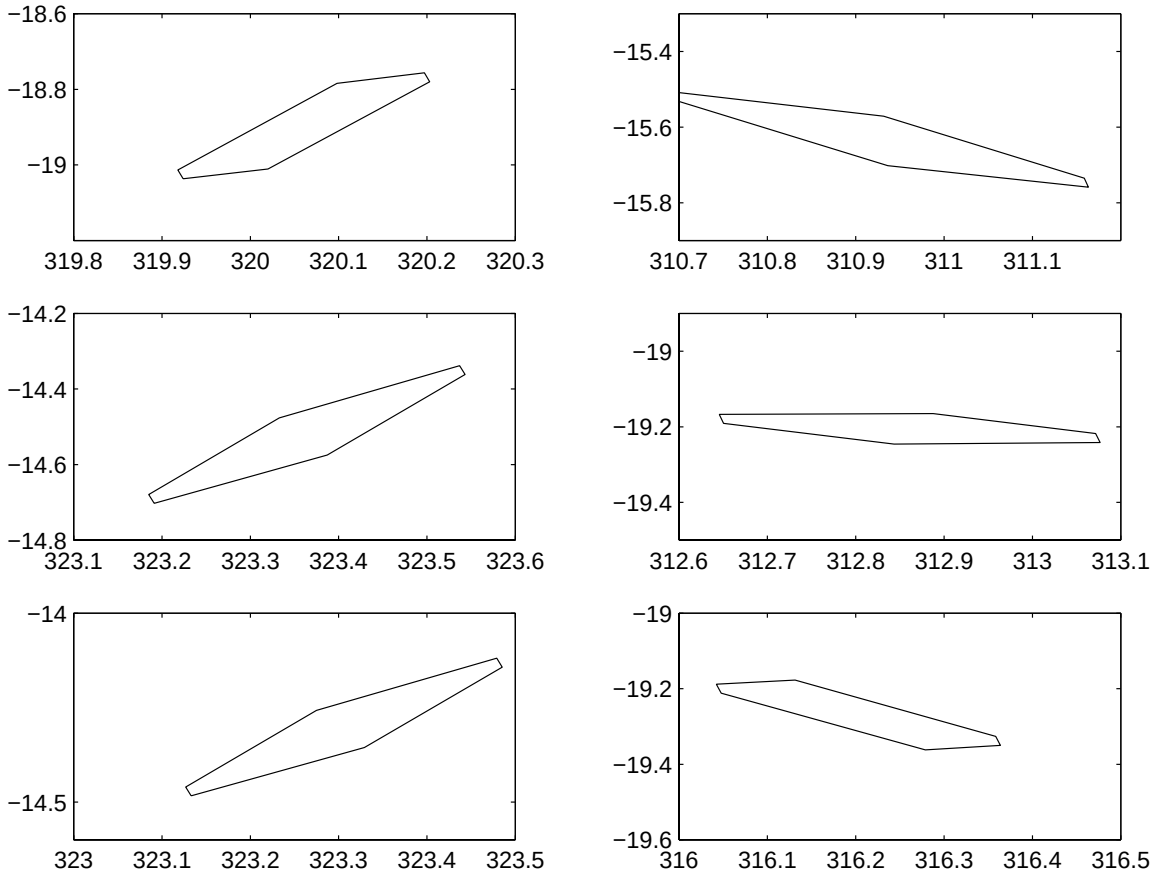


Figure 4.10: This figure illustrates several examples of NSCAT measurement cells. In the spatial domain, the cell is quite narrow on one axis (8 - 15 km). In the frequency domain, this translates to much wider passband, resulting in more high frequency content in the measured data.

azimuthally modulated surfaces, but does not affect the filtering of the surface image by the aperture function since the aperture is circularly symmetric. The oblong cells of NSCAT, however, combine in various orientations to vary the filtering of the data depending on the orientation of the cells.

SIR uses these multiple, overlapping, varied orientation measurement cells to create images. While one measurement cell retains some higher frequency content, it is directionalized, oriented on the narrow axis of the measurement cell. Multiple measurements, however, retain more high frequency content by way of higher side lobe levels over a wider range of the two dimensional space than a single measurement cell orientation because of the varying azimuth angle. The higher side lobe levels, combined with denser sampling, contribute to the better resolution enhancement of NSCAT data compared to ERS-1 data in limited iterations.

Another side benefit of the rotation of the elongated measurement cell over the surface is the potential elimination of the nulls caused by the aperture function. Consider that a rectangular aperture will have frequency nulls that are spaced differently along the short axis compared to the long axis. If multiple readings are made at various azimuth angles, one reading cell provides frequency information in the null of another reading at a different orientation. In the theoretical limit, this allows for a complete reconstruction even though the individual aperture functions have nulls since some frequency information is obtained through varying the orientation of the measurement cells. This further improves NSCAT resolution enhancement using SIR.

4.3.3 Comparison of NSCAT and ERS Imagery

The preceding discussion leads to one conclusion: NSCAT data resolution enhancement should be significantly better than for ERS-1 data. As a basic comparison of the difference in detail between NSCAT and ERS-1, Figure 4.11 presents three images over the Amazon basin. The upper left 'NSCAT' image is of NSCAT data processed with SIR. The upper right 'ERS-1' image is of ERS-1 data processed with SIR, and the lower right 'Non-enhanced ERS-1' image is of raw ERS-1 data (nominal 25 km spacing, 50 km resolution). A comparison of the unprocessed ERS-1 data and

the ERS-1 data processed with SIR shows an improvement in resolution, albeit a limited improvement. The NSCAT data, however, gives marked improvement of the corresponding ERS-1 data, resulting in more detail in the SIR image.

Figures 4.12 and 4.13 show an example ERS-2 image and NSCAT image of the Weddell sea region of the same day range in 1996 (JD 337-342). ERS-2 is a satellite identical to ERS-1 that replaced ERS-1 in 1996. Similar to the results shown in Figure 4.11, the detail level in the NSCAT figure is remarkably better than for ERS-1, reflecting the resolution improvements that result from closer sample spacing and higher side lobe levels.

4.4 SIR Accuracy

We now turn to the accuracy of the SIR images. Throughout this development, synthetic images are used to illustrate and establish SIR algorithm behavior. The synthetic images are created using a true image that has been processed to create synthetic ERS-1 data. This data is then processed to create SIR images.

4.4.1 Simulation Images

To account for incidence angle dependence, both $\mathcal{A}$ and $\mathcal{B}$ true images are created. The true images used throughout this study are mostly one level images, meaning both the $\mathcal{A}$ and $\mathcal{B}$ true images are constant images. Several different $\mathcal{A}$ and $\mathcal{B}$ values are selected to analyze SIR behavior over a wide range of $\mathcal{A}$ and $\mathcal{B}$ values. All test images are created at a much higher nominal resolution (1 km per pixel) than the nominal satellite resolution of 50 km. The size of the true images in this study is 700 km square. The images are combined with cell location and incidence angle information to create a data stream similar to the actual ERS-1 data stream that can be processed by SIR.

For the synthetic images to be as similar to actual ERS-1 SIR imagery as possible, actual cell locations and incidence angle information from the ERS-1 data are used. An area is selected that matches the size of the true images and a 7 day period of ERS-1 data is processed to extract the cell location, incidence angle, and

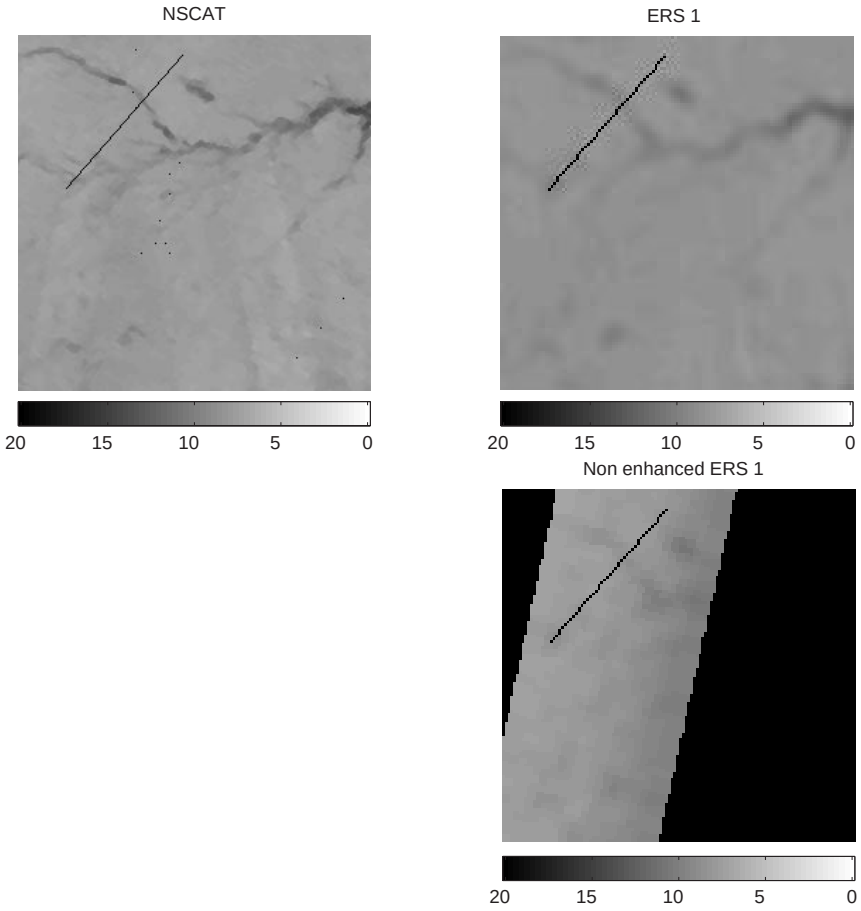


Figure 4.11: Example set of images showing comparison of an NSCAT SIR image ('NSCAT'), and ERS-1 SIR image ('ERS-1') and unprocessed ERS-1 data '('Non-enhanced ERS-1') over the Amazon basin. Note how the river feature is basically visible in the unenhanced image and becomes more detailed in the ERS-1 SIR image and very detailed in the NSCAT image. These images illustrate that resolution enhancement of ERS-1 data is slightly better than raw, and much better resolution enhancement is achieved using NSCAT data.

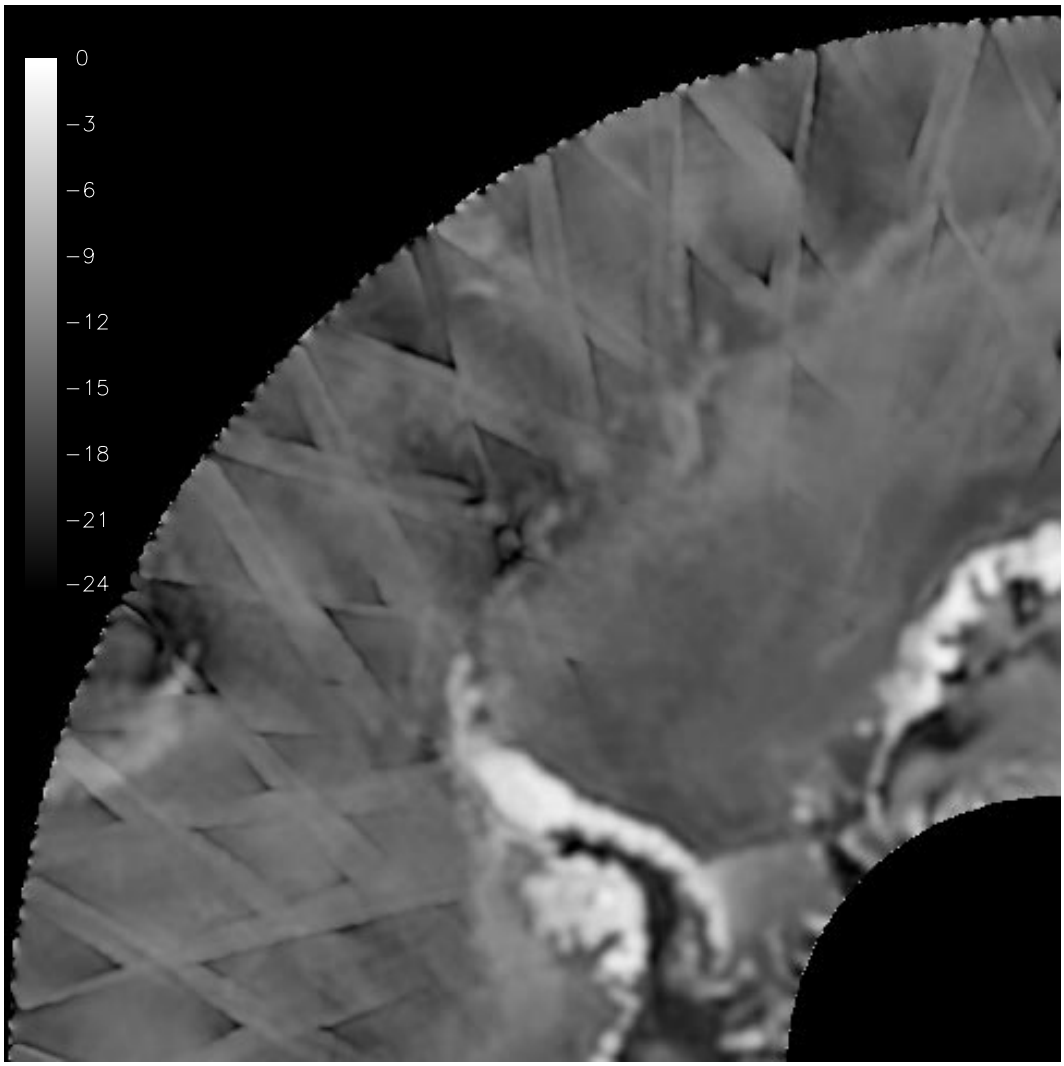


Figure 4.12: An ERS-2 image (ERS-2 is functionally identical to ERS-1) of the Weddell Sea from JD 337-342 1996. Note the lower resolution of this image compared to the NSCAT image of the same time period shown in Figure 4.13.

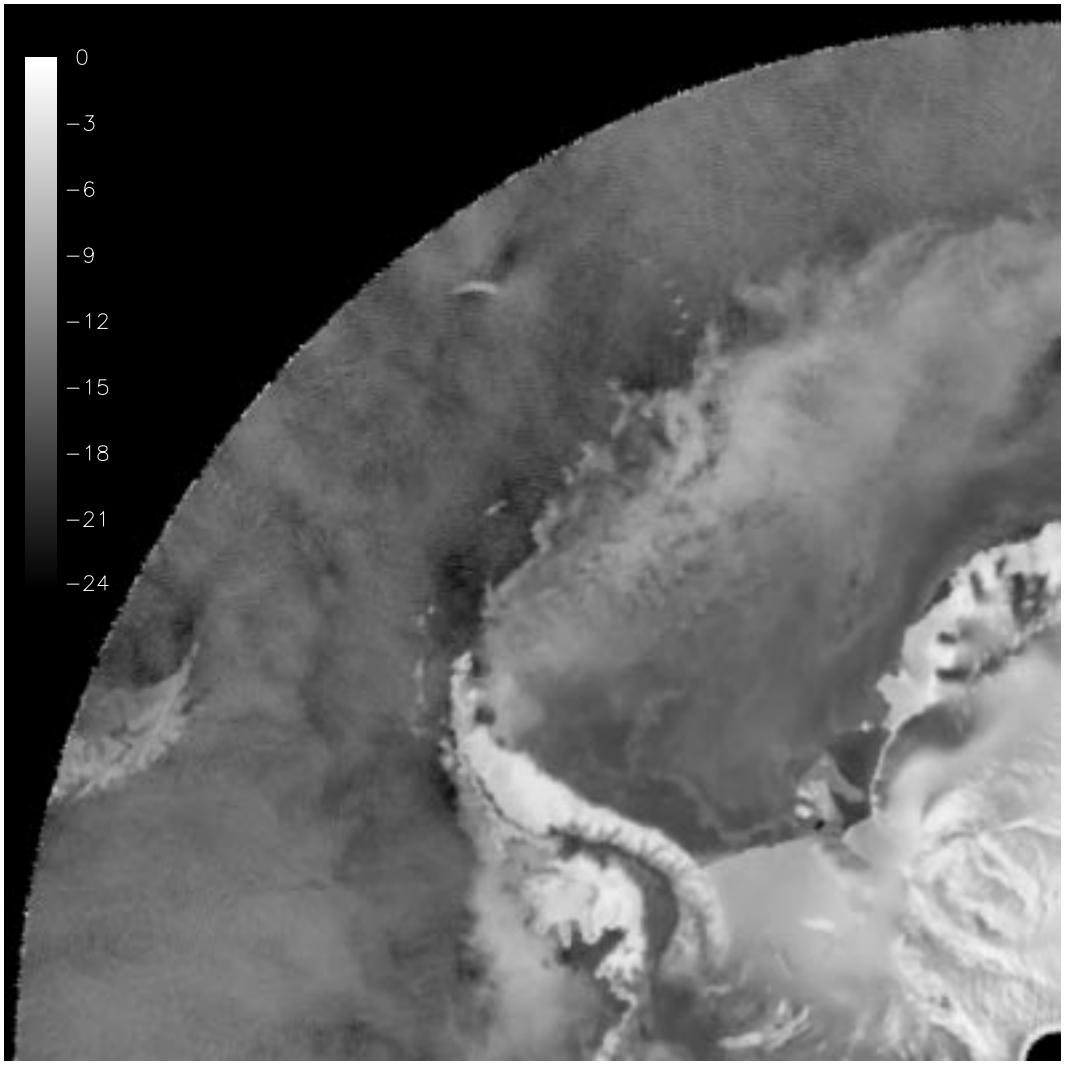


Figure 4.13: An NSCAT image of the Weddell Sea from JD 337-342 1996. Note that this image carries much more detailed information compared to the ERS-2 image from the same time period in Figure 4.12. Note that some differences between this and the example ERS-2 image may result from NSCAT being a KU-band instrument while ERS-2 is C-band. This will result in some brightness differences, but the enhancement of detail is a result of cell geometry rather than a change of frequency band.

other information. Eventually, only the σ^o value in the real ERS-1 data is replaced by a synthetic value based on the true images and the cell response information.

The cell response was shown previously in Figure 4.3. Since the variation of the incidence angle is small within a single footprint, it is assumed that the incidence angle reported in the ERS-1 data is constant over that footprint. To calculate the synthetic σ^o value, the contribution from pixel in the test image covered by the footprint is calculated by

$$\sigma_i^o = \mathcal{A}_{\mathrm{T}true} + \mathcal{B}_{\mathrm{T}true}(\theta_{\mathrm{ERS}-1} - 40.0) \tag{4.10}$$

which represents the contribution to the synthetic σ^o value of the ith pixel covered by the footprint. Recall that this calculation is in dB. The σ^o values are summed in real space using a weighted sum. The weight corresponding to each pixel is determined by the cosine squared roll off of the footprint pattern. The synthetic σ^o value is given by

$$\sigma^o_{\mathrm{S}ynthetic} = \frac{\sum_i w_i 10^{\frac{\sigma_i^o}{10}}}{\sum_i w_i} \tag{4.11}$$

and is illustrated in Figure 4.14. The grid on the left of Figure 4.14 illustrates how the different weight values are assigned to pixels based on the circular footprint shown on the right of the figure. The σ^o value replaces the σ^o value in the ERS-1 data stream. The result is a data set with synthetic σ^o values but actual cell information that can be processed with the SIR algorithm.

4.4.2 SIR Initialization

SIR generates images from some initial image using iterative updates. The initial image must be set to a pre-selected non-zero value before the first iteration. The general flow of the SIR algorithm is shown in Figure 4.15, where a_j^t is the current pixel value and Δ represents the non-linear combination of the raw σ^o values and the current pixel values to form the new pixel value, a_j^{t+1}. This processes is repeated for N iterations.

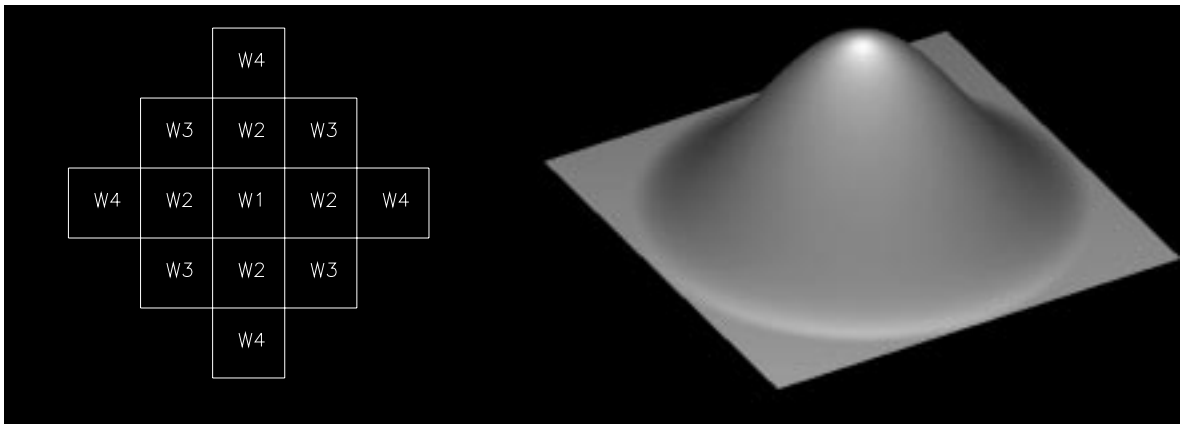


Figure 4.14: Illustration of the ERS-1 footprint as applied to a set of pixels. The grid on the left illustrates an example of the weight assignments to a particular grid of pixels. The actual pixel grid used by SIR to create the imagery presented here uses more pixels to model the footprint response.

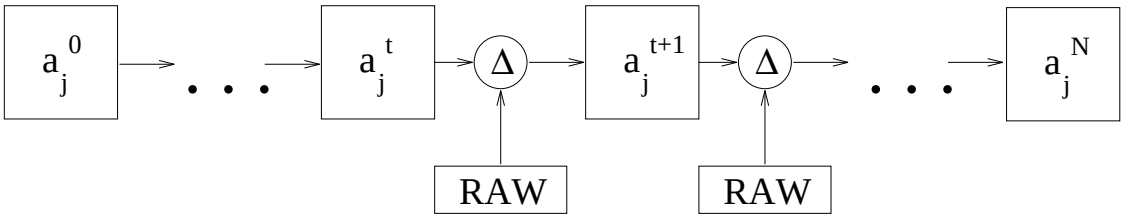


Figure 4.15: This figure illustrates the general flow of the SIR algorithm. From an initial pixel value, a_j^0, each iterative step updates the current pixel value based on information from the raw data. Δ represents the non-linear combination of the current pixel value and the raw data to form the new pixel value

The importance of the selection of the initial image value is coupled to the damping of scale factors used to generate the updates. Repeating here the equations from Chapter 2, the update for the jth pixel of the $\mathcal{A}$ image estimate is given by:

$$a_j^{t+1} = \frac{\sum_{i=1}^{L} update_{ij}^k}{\sum_{i=1}^{L} m_{ij}} \tag{4.12}$$

$$update_{ij}^t \approx \begin{cases} a_j^k \frac{1}{q(d_i^k)} m_{ij} & d_j^t \geq 1 \\ a_j^k q(d_i^k) m_{ij} & d_j^t < 1 \end{cases} \tag{4.13}$$

where d_i^k is the square root damped scale factor calculated from the raw measurements and back projections of the raw measurements. What is important to note here is the behavior of the function $q(s)$ as illustrated in Figure 2.1. This non-linear limiting of the multiplicative scale factor deemphasizes noise in the resulting imagery, but causes slowed development of the $\mathcal{A}$ value. Because too many iterations may cause the algorithm to go unstable due to noise and cause ringing in the resulting images, implementing iterations beyond a minimal level (e.g. between 20 and 30 for ERS-1) to further update the $\mathcal{A}$ value is not practical. If the initialization value of the $\mathcal{A}$ or the $\mathcal{B}$ image is too far from the true value, a finite number of iterations may not be sufficient to reach the true value. Also, as previously stated the $\mathcal{B}$ value development is heavily damped in SEASAT scatterometer processing. Processing ERS-1 data with the unmodified SEASAT weightings results in extremely slow $\mathcal{B}$ value development, which will be addressed in the next section. We now pose the questions: Will the unmodified SIR algorithm produce accurate and viable imagery from ERS-1 data? How sensitive is SIR to various initialization values for $\mathcal{A}$ and $\mathcal{B}$? These questions will be evaluated using synthetic test imagery as described below.

The effects of various initialization levels are investigated to determine the sensitivity of SIR to different initialization levels. Two true images are used in this experiment to illustrate the effects of initialization on SIR imagery. The test data sets represent relative extremes that are expected in actual image reconstructions, and are the image pairs ($\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1$) and ($\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3$). Various initialization ranges for $\mathcal{A}$ and $\mathcal{B}$ are tested with $\mathcal{A}$ ranging $-30\text{dB} < \mathcal{A} < -1\text{dB}$ and $\mathcal{B}$ ranging $-0.3 < \mathcal{B} < 0$, both representing expected extremes in the $\mathcal{A}$ and $\mathcal{B}$ values.

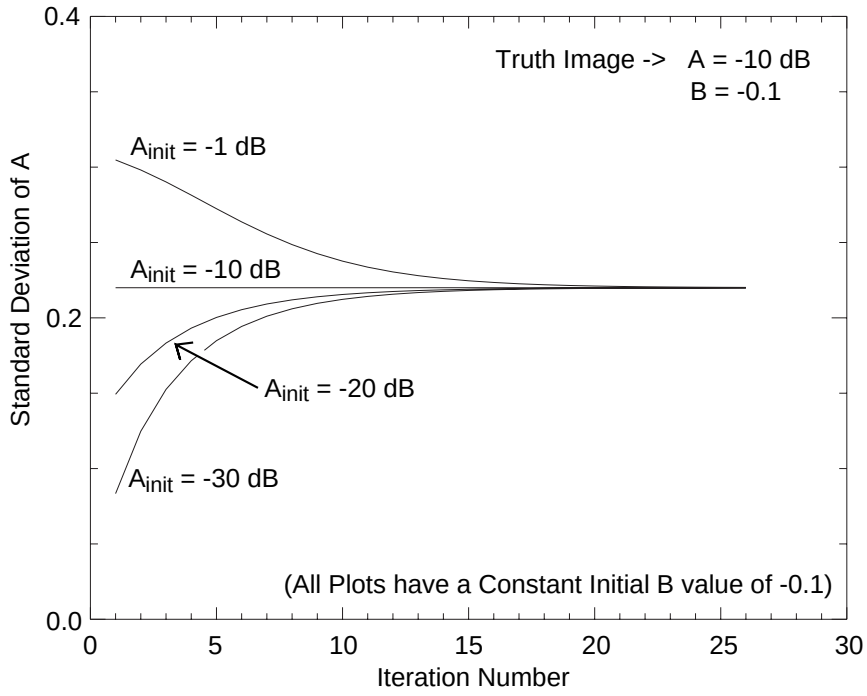


Figure 4.16: A comparison of the standard deviation of the $\mathcal{A}$ value for a fixed initial $\mathcal{B}$ value and various $\mathcal{A}$ initial values. The true image is $(\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1)$, and the initial $\mathcal{B}$ value is $\mathcal{B} = -0.1$. Note that the standard deviation is constant if the initial value of $\mathcal{A}$ matches the true value and that all other initial values used converge to this level.

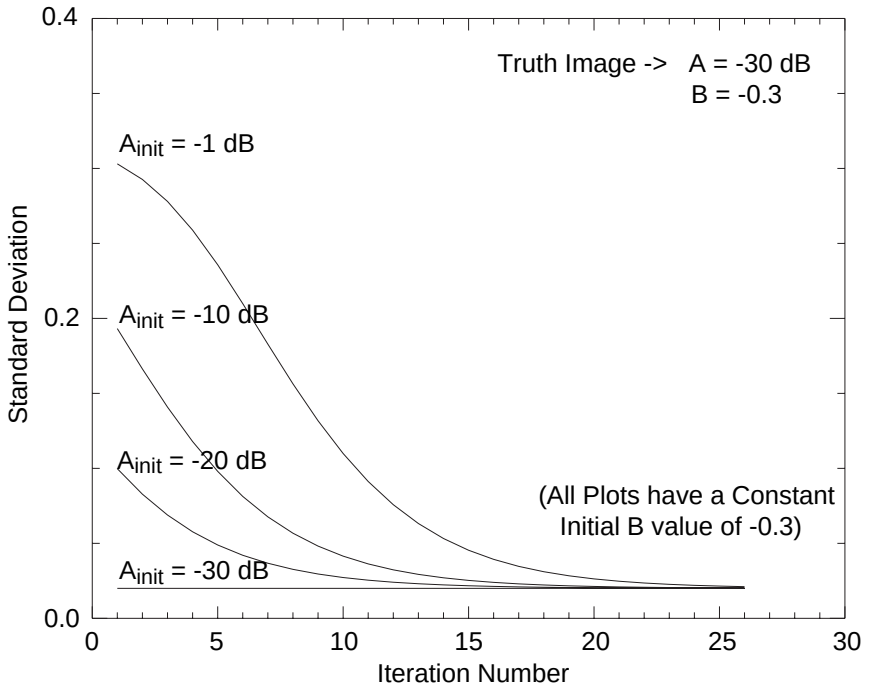


Figure 4.17: A comparison of the standard deviation of the $\mathcal{A}$ value for a fixed initial $\mathcal{B}$ value and various $\mathcal{A}$ initial values. The true image is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$, and the initial $\mathcal{B}$ value is $\mathcal{B} = -0.3$. Note that the standard deviation is constant if the initial value of $\mathcal{A}$ matches the true value and that all other initial values used converge to this level.

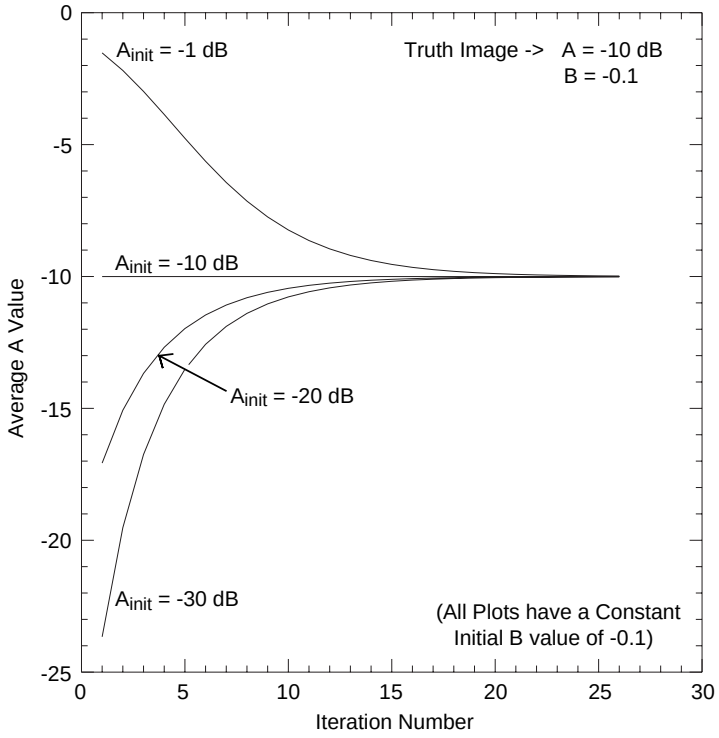


Figure 4.18: Convergence of $\mathcal{A}$ value for correct initial $\mathcal{B}$ value estimate. The true image is $(\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1)$, and the initial $\mathcal{B}$ value is $\mathcal{B} = -0.1$

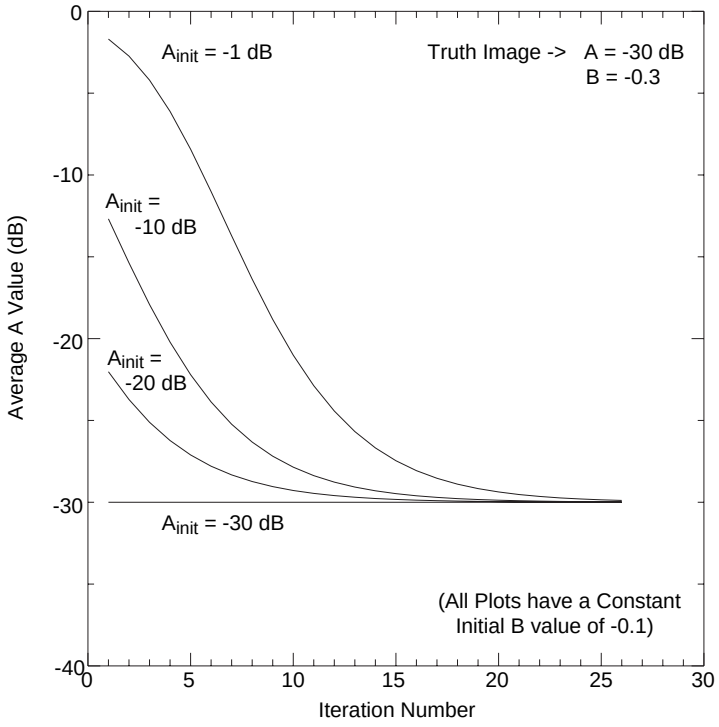


Figure 4.19: Convergence of $\mathcal{A}$ value for correct initial $\mathcal{B}$ value estimate. The true image $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$, and the initial $\mathcal{B}$ value is $\mathcal{B} = -0.3$

In the first experiment, the $\mathcal{B}$ value is initialized to the true value and the $\mathcal{A}$ value is varied between -30 dB to -1 dB. No matter what the true $\mathcal{A}$ value or $\mathcal{B}$ value was, SIR correctly predicts the $\mathcal{A}$ value with little distortion. Figures 4.16 and 4.17 illustrate the standard deviation of the $\mathcal{A}$ value for two different true image sets, while Figures 4.18 and 4.19 show the convergence of the SIR $\mathcal{A}$ value to the true value. Note that the standard deviation decreases with each iteration, converging to a very small level (about 0.2 dB). Note also in Figure 4.17 that initializing the $\mathcal{A}$ value to -1 dB when the true value was -30 dB does not prevent convergence of the standard deviation of the $\mathcal{A}$ value. As the standard deviation of the $\mathcal{A}$ value suggests, all SIR $\mathcal{A}$ images looked identical, although negligible differences do exist between images with different initialization levels. It should be noted that since the $\mathcal{B}$ initialization value was set to the true $\mathcal{B}$ value, all SIR $\mathcal{B}$ images are virtually identical as well.

Another important observation is that even when the SIR algorithm is initialized to the true $\mathcal{A}$ value and true $\mathcal{B}$ value there is a non-zero standard deviation of the $\mathcal{A}$ and $\mathcal{B}$ values. In Figure 4.16, the $\mathcal{A}_{\text{init}} = -10$ dB line represents the true value and the standard deviation is constant throughout all iterations. In Figure 4.17 the same is true for the $\mathcal{A}_{\text{init}} = -30$ dB line. In both cases, the standard deviation was less then 0.2 dB. The standard deviation does not go to 0 because of some non-linear effects of the $\mathcal{B}$ value development on the $\mathcal{A}$ value. Because Figures 4.18 and 4.19 show that the average value of the SIR $\mathcal{A}$ images converges to the true value, for a correct initial $\mathcal{B}$ value. This demonstrates that for a noiseless case, 27 iterations are sufficient for fair reconstruction of the image. For the noisy case, the number of iterations was determined by empirical review of images for a subjectively best result, with 27 iterations being judged sufficient without unduly amplifying noise.

Next, the $\mathcal{A}$ initialization value is fixed to the true value and the $\mathcal{B}$ initial value is varied from 0 to -0.3. The resulting SIR $\mathcal{A}$ images are shown in Figures 4.20 and 4.21 and show the behavior of the $\mathcal{A}$ value for improper $\mathcal{B}$ value initialization. Ideally, all these images should be monotone. These figures give a comparison of several $\mathcal{B}$ initialization value levels for the two different true image sets. When the

$\mathcal{B}$ initial value matches the true value, the resulting $\mathcal{A}$ value image has a small standard deviation. When the $\mathcal{B}$ initial value varies from the true value, however, the resulting imagery begins to show hash marks that correspond to the satellite track and reflect the varying dependence on incidence angle along the beam. This error increases the further the $\mathcal{B}$ initial value is from the true value.

Figure 4.22 shows the corresponding $\mathcal{B}$ images for Figure 4.21 for the true image set ($\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3$) with the initial $\mathcal{A}$ value of -30 dB and various $\mathcal{B}$ values. Note how the $\mathcal{B}$ value does not progress fast enough toward the true value. This is further illustrated in Figures 4.23 and 4.24. The $\mathcal{B}$ value is unable to converge, although the Figures indicate that the $\mathcal{B}$ value is moving in the right direction. Figures 4.25 and 4.26, however, show that the standard deviation of the $\mathcal{B}$ values is diverging rather than converging. This can also be seen in the $\mathcal{B}$ images in Figure 4.22. The impact on the $\mathcal{A}$ value of slow $\mathcal{B}$ development is shown in Figure 4.27 for the true image set ($\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3$). Recall that the initial $\mathcal{A}$ value is equal to the true $\mathcal{A}$ value of -30 dB and then note how the average $\mathcal{A}$ value does begin to converge toward the true value.

The results of this experiment indicate that when the unmodified SIR algorithm is used on ERS-1 data, it is very sensitive to the initial $\mathcal{B}$ value estimate, but not as sensitive to the $\mathcal{A}$ estimate. Ideally, the initial $\mathcal{B}$ value should be very near to the true value; however, in practice it will be nearly impossible to make an estimate that is close enough to the true value over a large surface area and avoid the problems of convergence illustrated above. However, as mentioned before, the $\mathcal{B}$ value development has been damped in previous applications, so increasing the speed of $\mathcal{B}$ development is addressed in the following sections.

4.4.3 Cubic Parameterization

In Chapter 2, a model for backscatter was presented that included higher order terms. While the linear model is good, there may be some utility in determining if the use of higher order terms would result in a more accurate reconstruction. In this section, we consider the influence of the linear model versus a cubic model of incidence

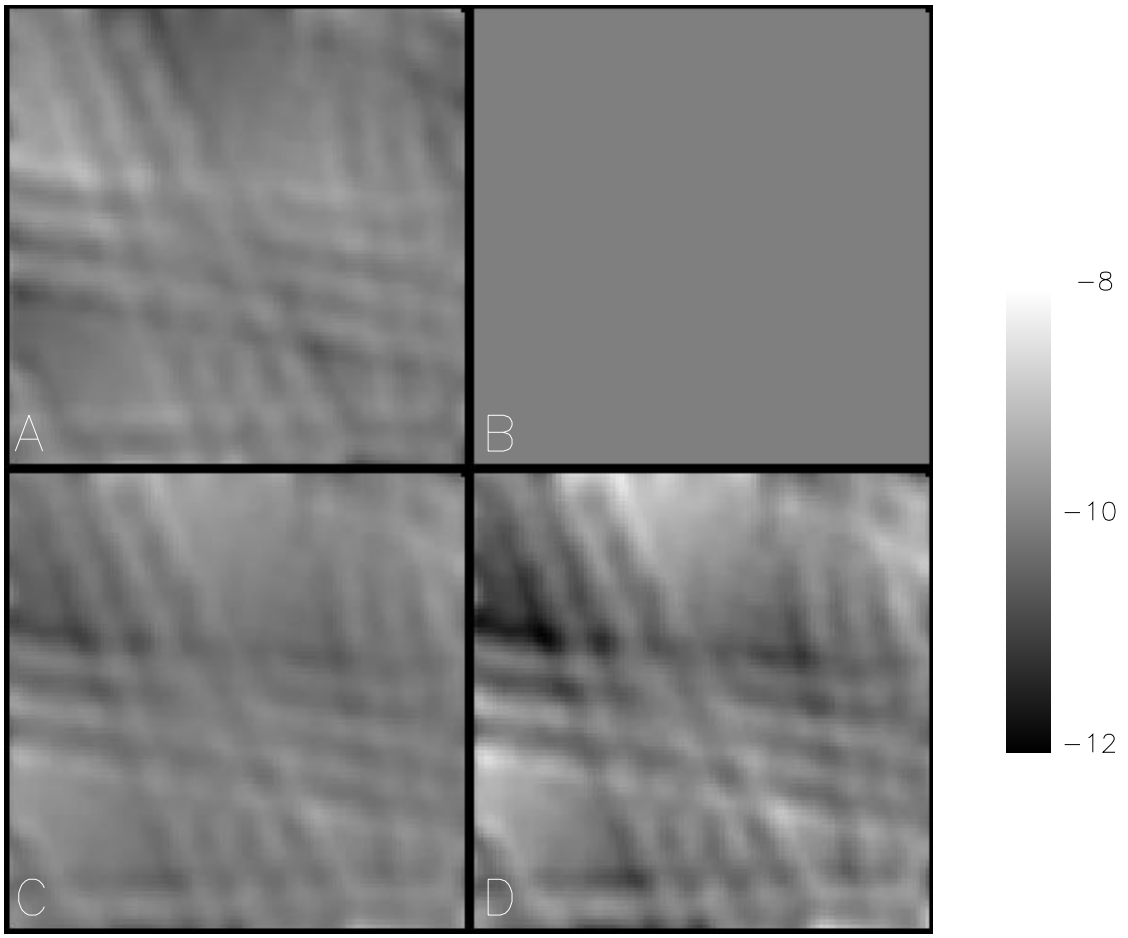


Figure 4.20: A comparison of SIR $\mathcal{A}$ images with various initialization values for the true image set ($\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1$). For this image, the $\mathcal{A}$ value was initialized to the true value of $\mathcal{A} = -10\text{dB}$. The images correspond to various initial $\mathcal{B}$ values: Frame A has an initial $\mathcal{B}$ value of $\mathcal{B} = 0$, Frame B has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.1$, Frame C has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.2$, Frame D has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.3$. Note that when the initialization value equals the actual image value the distortion is minimal.

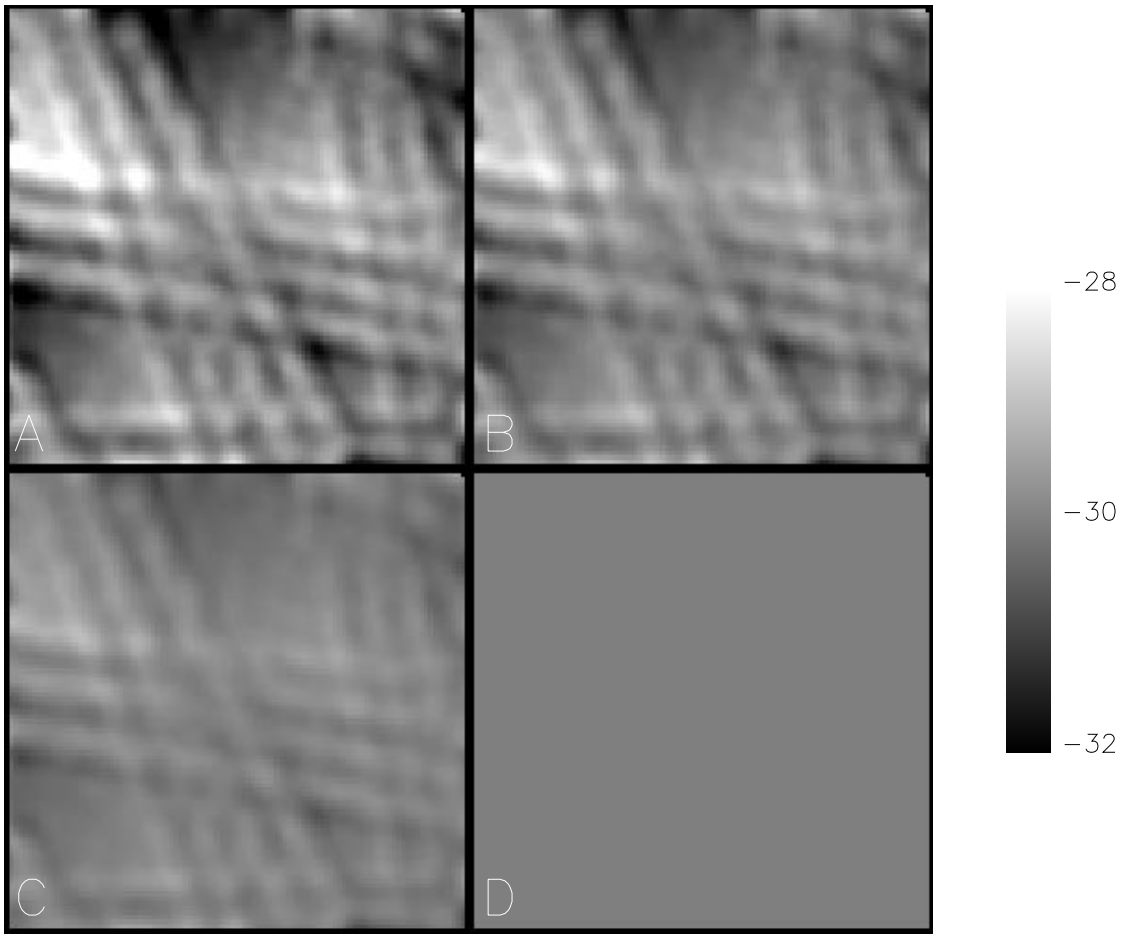


Figure 4.21: A comparison of SIR $\mathcal{A}$ images with various initialization values for the true image set ($\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3$). For this image, the $\mathcal{A}$ value was initialized to the true value of $\mathcal{A} = -10\text{dB}$. The images correspond to various initial $\mathcal{B}$ values: Frame A has an initial $\mathcal{B}$ value of $\mathcal{B} = 0$, Frame B has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.1$, Frame C has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.2$, Frame D has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.3$. Note that when the initialization value equals the actual image value the distortion is minimal.

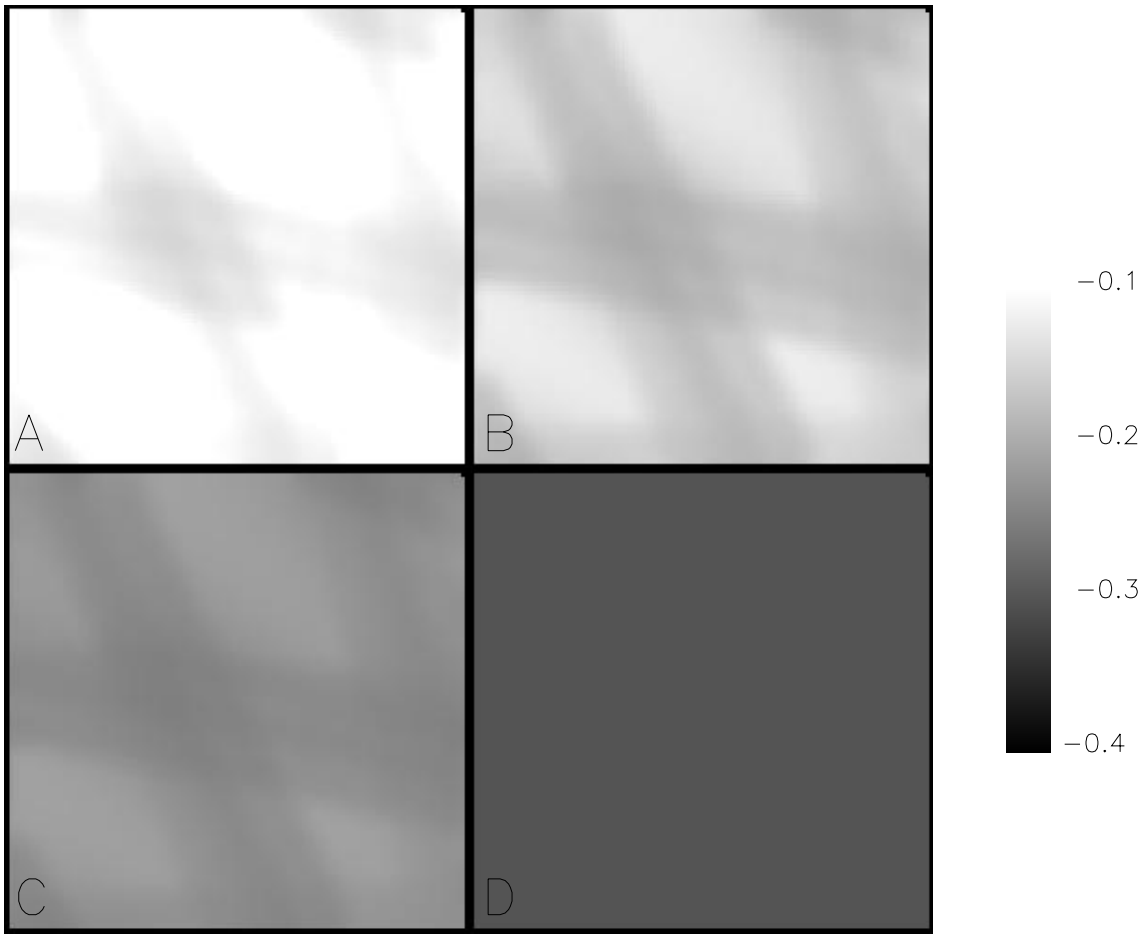


Figure 4.22: This figure shows a comparison of $\mathcal{B}$ value images for a fixed initial $\mathcal{A}$ value ($\mathcal{A} = -30$dB) and various $\mathcal{B}$ values. The true image is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$. Each subimage corresponds to various initial $\mathcal{B}$ values: Frame A has an initial $\mathcal{B}$ value of $\mathcal{B} = 0$, Frame B has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.1$, Frame C has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.2$, Frame D has an initial $\mathcal{B}$ value of $\mathcal{B} = -0.3$. While Figures 4.20 and 4.21 show that the $\mathcal{A}$ value, although high in variance, is still close to the true value for various $\mathcal{B}$ initial values, this image set shows that the $\mathcal{B}$ value fails to reach the actual value of $\mathcal{B}$ in 27 iterations if the initialization is too far from the true value.

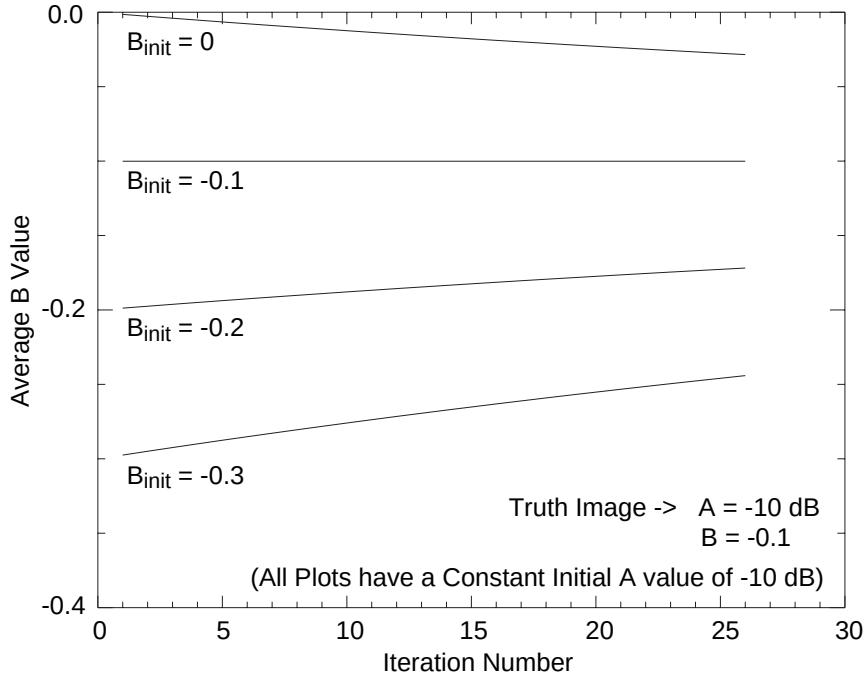


Figure 4.23: Illustration of the convergence of the average $\mathcal{B}$ value for various initial $\mathcal{A}$ values. The true image is $(\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1)$.

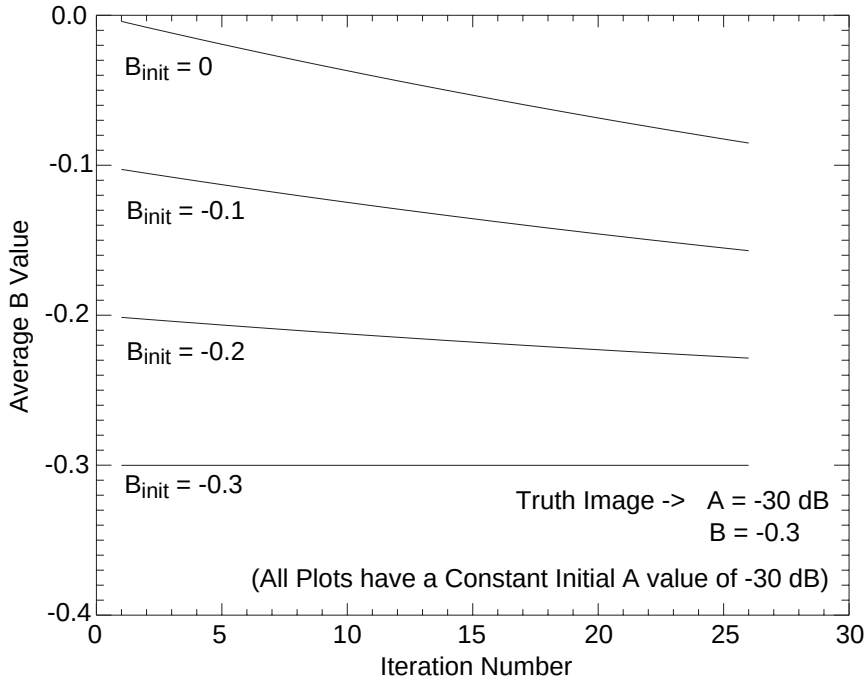


Figure 4.24: Illustration of the convergence of the average $\mathcal{B}$ value for various initial $\mathcal{A}$ values. The true image is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$.

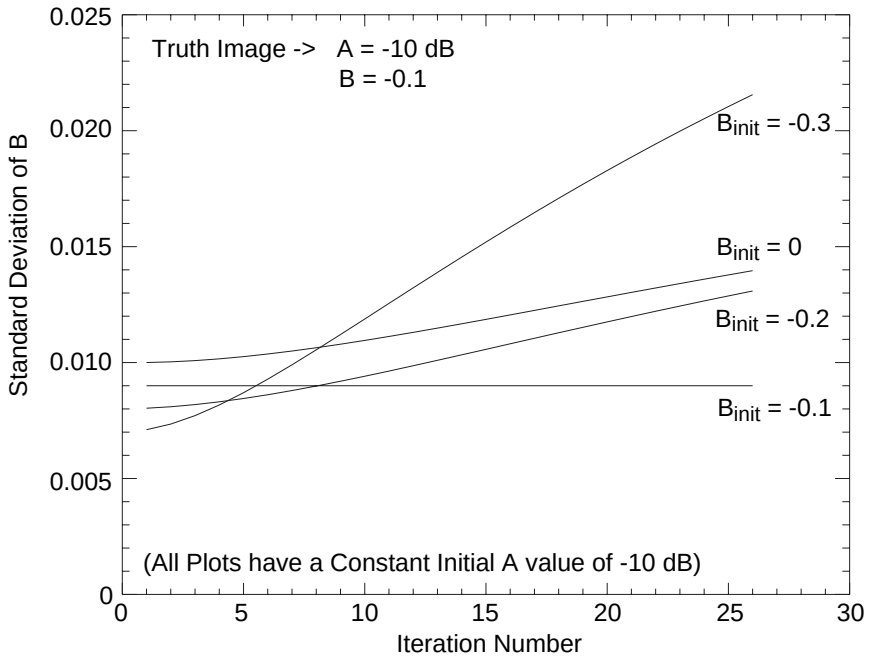


Figure 4.25: Illustration of the standard deviation of the $\mathcal{B}$ value for various initial $\mathcal{A}$ values. The true image is $(\mathcal{A} = -10\mathrm{d}B, \mathcal{B} = -0.1)$. Note how the standard deviation diverges as the iterations increase. This problem is address in the following section.

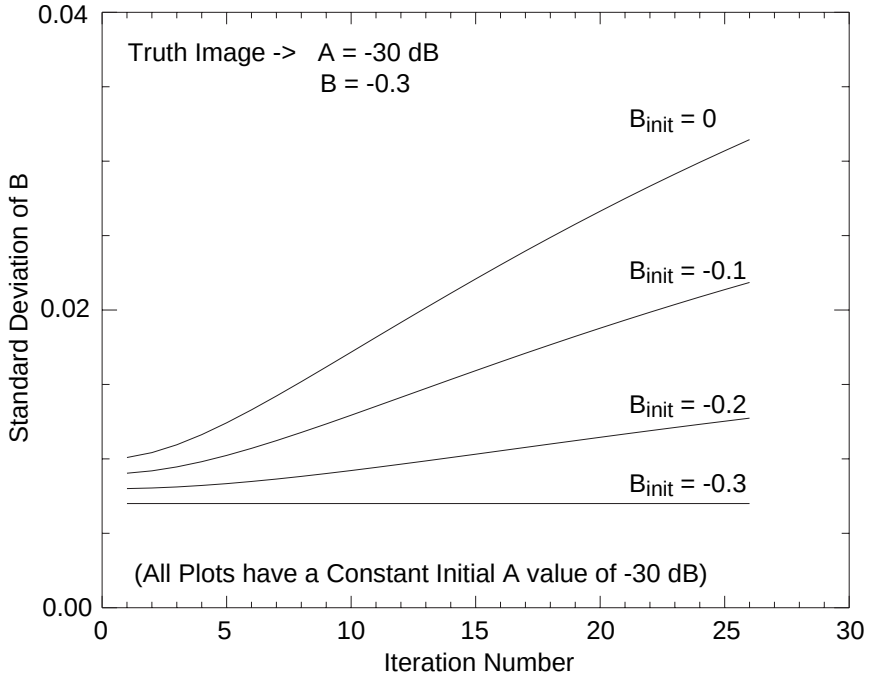


Figure 4.26: Illustration of the convergence of the average $\mathcal{B}$ value for various initial $\mathcal{A}$ values. The true image is $(\mathcal{A} = -30\mathrm{d}B, \mathcal{B} = -0.3)$. Note how the standard deviation actually diverges as the iterations increase.

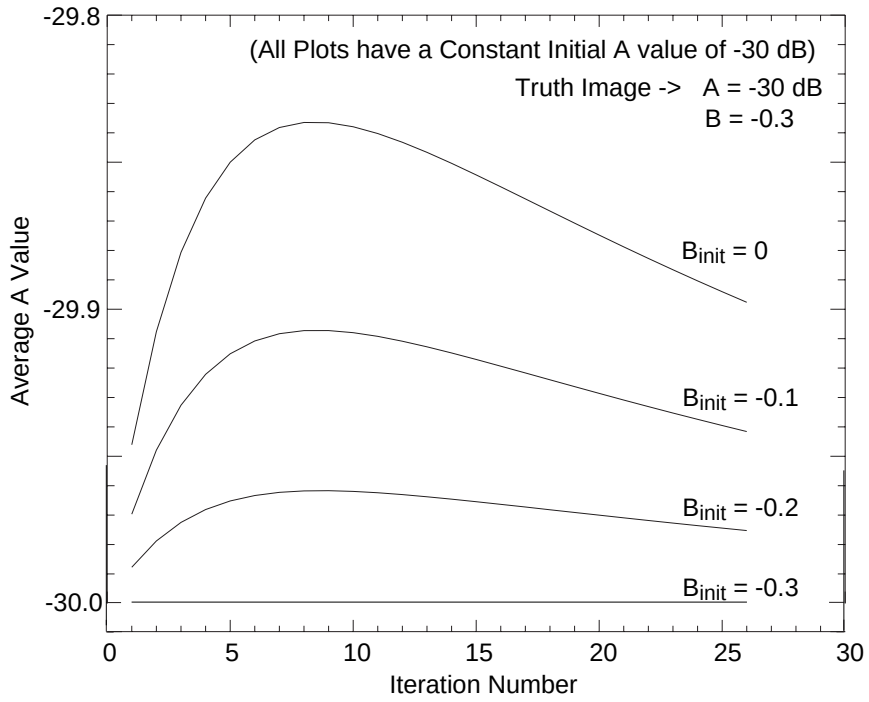


Figure 4.27: An illustration of the lack of convergence of the $\mathcal{A}$ value for various $\mathcal{B}$ initial values. Although the average value is close to the desired value of -30 dB, it is still quite poor compared to the estimate shown in Figures 4.19 where the $\mathcal{B}$ value is initialized to the true value.

angle dependence on SIR imagery. The dependence of σ^o on incidence angle θ can be modeled by a polynomial:

$$\sigma^o = \mathcal{A} + \mathcal{B}\theta + \mathcal{C}\theta^2 + \mathcal{D}\theta^3 + \ldots \tag{4.14}$$

where $\mathcal{A}$ is equivalent to σ^o at an incidence angle of 0 degrees and $\mathcal{B}, \mathcal{C}$ and $\mathcal{D}$ correspond to the linear, quadratic and cubic components of the σ^o vs. incidence angle relationship. A more convenient parameterization of σ^o normalizes the $\mathcal{A}$ value to 40 degrees:

$$\sigma^o = \mathcal{A} + \mathcal{B}(\theta - 40) + \mathcal{C}(\theta - 40)^2 + \mathcal{D}(\theta - 40)^3 + \ldots . \tag{4.15}$$

The $\mathcal{A}$ value is thus normalized to 40 degrees. Note that 40 degrees is the mean scatterometer incidence angle for ERS-1, thus minimizing the higher order terms (i.e. a measurement with an incidence angle of 40 degrees has no higher order terms and thus $\sigma^o = \mathcal{A}$).

While higher order polynomials can be used to approximate σ^o, the linear equation $\sigma^o = \mathcal{A}+\mathcal{B}\theta$ is usually sufficient to give a good approximation of σ^o behavior. However, the linear approximation will have a larger error at the extremes of the scatterometer beam where incidence angles get large (> 55 degrees) or small (< 25 degrees). Figure 4.28 shows an qualitative example of the σ^o vs. incidence angle relationship. The nearly linear region in the center (corresponding to incidence angles of about 25-55 degrees) is accurately modeled by the $\mathcal{A}$ and $\mathcal{B}$ parameters. The tail regions (incidence angles below 25 degrees and above 55 degrees) begin to curve (the curve is exaggerated in the figure), and higher order polynomials like Eq. (4.14) are required to accurately model the behavior of σ^o with large and small incidence angles.

The limitations of the linear model can be seen in Figure 4.2 where the cross hatches exist in the ocean and in the periphery of the sea ice. Over the ocean, we expect to see the cross hatches because of the high dependence of ocean scattering on incidence angle, the azimuthal modulation in the σ^o response, the dependence of σ^o on wind speed and the high temporal variability of the wind speed. Since SIR images are created using 3-7 days of data, the wind can change over this period,

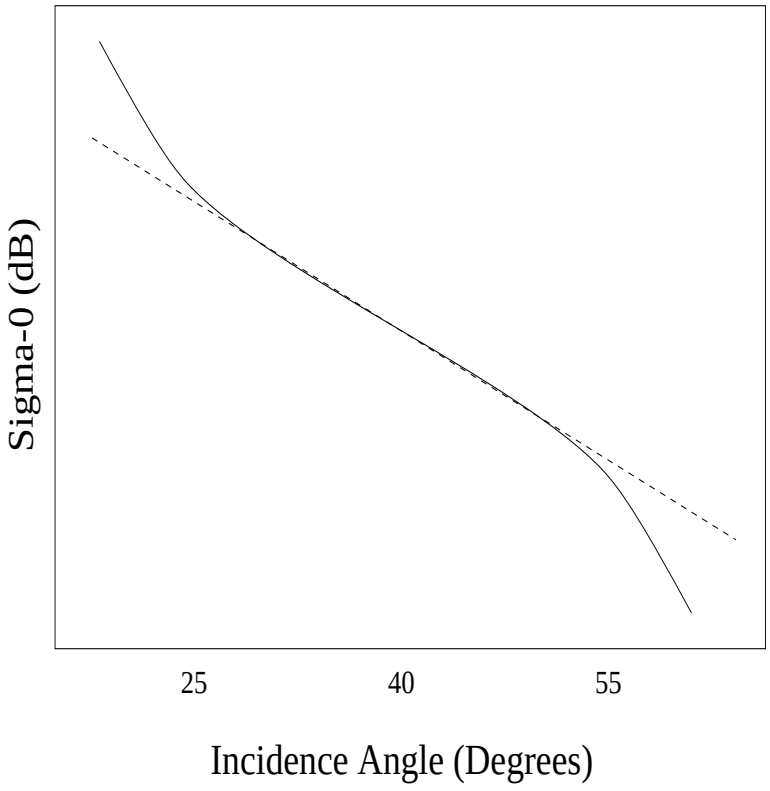


Figure 4.28: A qualitative example of the σ^o vs. incidence angle relationship. Note how the center region is very linear (as illustrated by the dashed line), while the tails begin to show non-linear behavior. The curve is exaggerated in the figure to illustrate the non-linearity. A linear approximation of the σ^o vs. incidence angle relationship is usually sufficient for incidence angles in the range 25-55 degrees.

causing significant changes in σ^o over the open ocean. Since SIR assumes that the surface is constant and isotropic in azimuth and open ocean meets neither of these requirements, SIR imagery of open ocean is of little scientific value.

Over sea ice, there is no significant azimuthal modulation [43] and it is reasonable to assume that the surface scattering characteristics of sea ice remains unchanged over several days. However, the cross hatches are still quite visible in the sea ice pack in Figure 4.2 in both the $\mathcal{A}$ and $\mathcal{B}$ images due to the limitations of the linear model of σ^o vs. incidence angle and inaccurate $\mathcal{B}$ image values as illustrated in the previous section. Although much less pronounced than the cross hatches in the ocean, the artifacts in the sea ice pack present a challenge for using SIR imagery in any classification scheme.

The most significant cross hatching occurs near the ends of the beam. Note in Figure 4.2 how the area in between the cross hatches is relatively consistent throughout the ice pack, while the major differences occur along the edge of each swath. SIR uses a linear approximation in developing the $\mathcal{B}$ value from the $\mathcal{A}$ value. Because the linear approximation begins to fail near the beam edges, it is these areas in the imagery that show the greatest error.

The quadratic and cubic terms $\mathcal{C}$ and $\mathcal{D}$ can be used to reduce the cross hatching in the sea ice regions by increasing the accuracy of the $\mathcal{A}$ and $\mathcal{B}$ estimate from σ^o near the edges of the beam. Previous efforts have demonstrated that SIR is not effective for estimating $\mathcal{C}$ and $\mathcal{D}$ [44], so we opt to use a fixed value of $\mathcal{C}$ and $\mathcal{D}$ based on an examination of values of $\mathcal{C}$ and $\mathcal{D}$ from several study regions over the sea ice. Since we do not care to correct cross hatching over the open ocean, no ocean regions are considered. The study regions and some of their statistics are listed in Table 4.1. The global values chosen are $\mathcal{C} = 0.0015$ and $\mathcal{D} = -0.00015$, which are actually slightly less than the average values. The smaller values were chosen to minimize enhancing the cubic error rather than correcting it. With these global values, σ^o is "corrected" for the quadratic and cubic terms and this new corrected value is used by SIR to make estimates of $\mathcal{A}$ and $\mathcal{B}$. The corrected σ^o value is given

			Linear		Cubic			
Reg	Corner	lon x lat (degrees)	$\hat{\mathcal{A}}$	$\hat{\mathcal{B}}$	$\hat{\mathcal{A}}$	$\hat{\mathcal{B}}$	C x10^{-4}	D x10^{-5}
Glacial 1	(110,-76)	10 x 2	-10.12	-0.188	-10.23	-0.218	29.6	.13
Glacial 2	(140,-74)	10 x 2	-11.01	-0.192	-10.91	-0.206	28.6	-11.0
Glacial 3	(140,-78)	10 x 2	-8.24	-0.130	-8.24	-0.103	20.6	-26.3
Glacial 4	(40,-78)	10 x 2	-11.93	-0.169	-11.84	-0.163	13.0	-22.1
Glacial 5	(-110,-78)	10 x 2	-19.38	-0.275	-19.45	-0.273	15.9	-17.3
Sea Ice 1	(-20,-64)	10 x 2	-14.21	-0.174	-14.06	-0.185	.053	-11.2
Sea Ice 2	(-30,-66)	6 x 4	-16.96	-0.177	-16.82	-0.192	2.49	-9.38
Sea Ice 3	(-170,-74)	10 x 2	-16.98	-0.167	-16.87	-0.182	4.36	-8.17
Sea Ice 4	(-160,-70)	10 x 2	-14.66	-0.181	-14.60	-0.201	12.5	-5.28
Sea Ice 5	(-54,-74)	4 x 2	-16.66	-0.189	-16.57	-0.207	6.55	-6.53

Table 4.1: Study region statistics. The study regions were chosen to be relatively homogeneous regions of the sea ice pack, and their locations and size are given in the first three columns. The 'Linear' columns are linear regression estimates from the raw σ^o of $\mathcal{A}$ and $\mathcal{B}$ with quadratic and cubic terms removed as explained in the text. The 'Cubic' columns are full cubic regression fits with the quadratic and cubic terms C and D estimated from the raw data. Note that there is little difference is most cases between the Linear and Cubic regression estimates of $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$.

by

$$\sigma^o_{corr} = \sigma^o - \mathcal{C}(\theta - 40)^2 - \mathcal{D}(\theta - 40)^3 = \mathcal{A} + \mathcal{B}(\theta - 40). \tag{4.16}$$

The results are mixed. The study regions show a high variability in $\mathcal{C}$ and $\mathcal{D}$ values as illustrated in Table 4.1, so a single global value for $\mathcal{C}$ or $\mathcal{D}$ does not adequately represent the conditions over the entire sea ice pack. In general, there is at most a small reduction in the cross hatches in some images. The effect of the correction varied from image to image. Figure 4.29 shows a surface plot of the difference between an image with higher order correction and the same image without higher order correction. The pattern in the surface plot corresponds with the cross hatch pattern visible in the original images. However, cross hatches are still visible in the SIR images and the maximum correction in Figure 4.29 is less than 0.5 dB.

A comparison of $\mathcal{B}$ images for various values of $\mathcal{C}$ and $\mathcal{D}$ shows that there is an overcorrection near the beam edges for values of $\mathcal{C}$ and $\mathcal{D}$ which are larger than

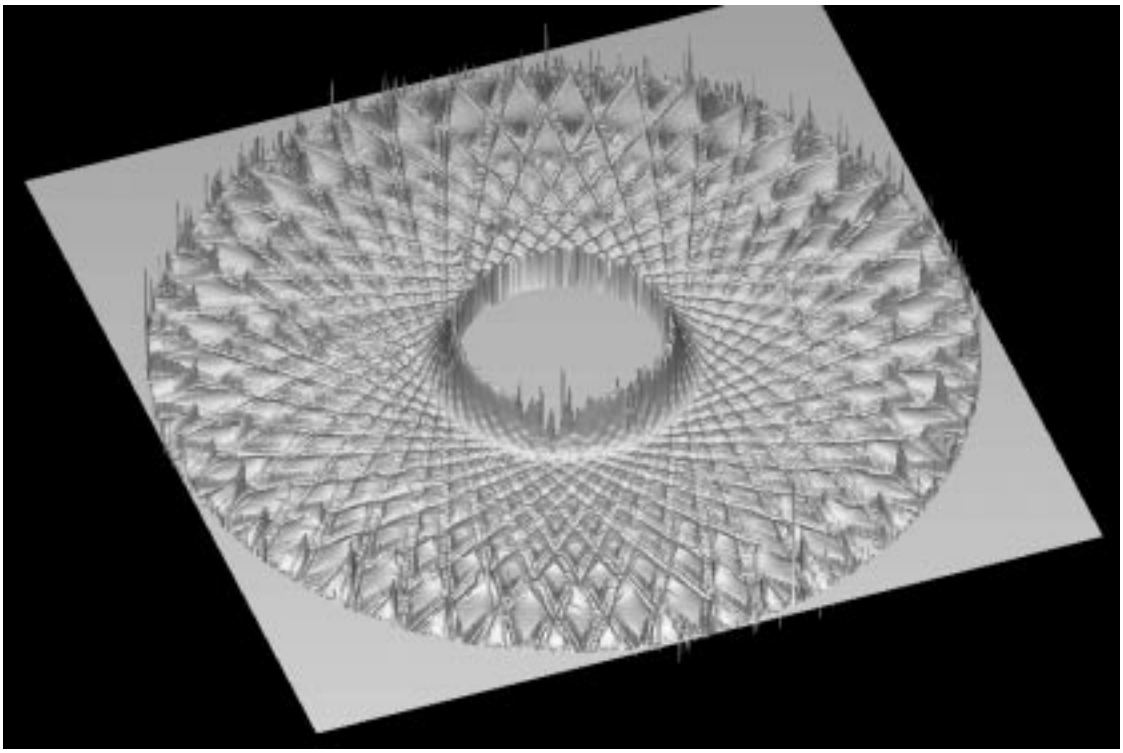


Figure 4.29: Surface plot of the difference between two SIR $\mathcal{A}$ images, one with higher order correction terms and one without. The original image is JD 150 1994 and $\mathcal{C} = 0.0015$ and $\mathcal{D} = -0.00015$. Note how the pattern in the difference image corresponds with the cross hatch pattern.

the true values. This indicates that if used, the global values for $\mathcal{C}$ and $\mathcal{D}$ need to be carefully evaluated. For the regions reviewed in this study, the standard deviation of the $\mathcal{C}$ and $\mathcal{D}$ parameters from several study regions is quite large. A quick review of Figures 4.32 and 4.33 in the last section shows that there are fairly large differences in the non-linear trends in the sea ice data making it difficult to create consistent $\mathcal{C}$ and $\mathcal{D}$ estimates using linear regression. Also, because ERS-1 operates at C-band, we expect less volume scattering content in σ^o compared to instruments operating at Ku-band (SEASAT and NSCAT) and it is the volume scattering component that contributes the most to the non-linear effects in the extreme incidence angle ranges. So a change in volume scattering translates roughly to a change in the non-linearity of the σ^o vs. incidence angle relationship.

The conclusion is that if they are to be used, $\mathcal{C}$ and $\mathcal{D}$ should be chosen smaller than the average calculated values of $\mathcal{C}$ and $\mathcal{D}$ to maximize the benefits of including the non-linear fit but minimize the possibility of over correcting near the beam edges. Unfortunately, because the surface conditions vary significantly in space, it is not possible to pick an ideal $\mathcal{C}$ and $\mathcal{D}$ value that will optimize the SIR algorithm for all areas, and we therefore choose not to implement this correction at this time.

4.4.4 $\mathcal{B}$ Iterative Weighting

Another source error is an incorrect $\mathcal{B}$ value. In each successive SIR iteration, the $\mathcal{A}$ and $\mathcal{B}$ estimates are updated based on information from the raw data. The $\mathcal{A}$ and $\mathcal{B}$ values progress towards some limiting value with each pass of the data. As discussed in the first section, the rate at which the $\mathcal{A}$ value develops must remain fixed to preserve the stability of the algorithm. However, the $\mathcal{B}$ development can be modified by changing the weights used in the iterative update (see Eq. (2.18)).

For this phase of the experiment, the $\mathcal{B}$ update weight is modified by a multiplicative factor. Weights of 10, 25, 50, 100 and 150 times the normal update weight in Eq. (2.18) are used to test the effectiveness of increasing the rate of $\mathcal{B}$ update and the impact this has on the $\mathcal{A}$ image. This effectively emphasizes the linear regression estimate of the $\mathcal{B}$ value as given in Eq. (2.19). A comparison is also made

between the regions in the images with increased $\mathcal{B}$ weighting and linear estimates of $\mathcal{A}$ and $\mathcal{B}$ over the same regions using raw σ^o data and a linear regression algorithm. This comparison is a measure of how well SIR estimates match the information in the raw data.

Test Images and Regions

A single test image is used in this study. The image chosen was JD 150 1994. A total of 10 study regions were chosen to evaluate the effects of accelerated $\mathcal{B}$ development. These regions are listed in Table 4.1.

A valid comparison between the $\mathcal{A}$ and $\mathcal{B}$ values developed by SIR (which is based on a linear model of σ^o vs. incidence angle) and the linear estimates of $\mathcal{A}$ and $\mathcal{B}$ obtained by linear regression of the raw data requires that the quadratic and cubic terms be removed from the raw data prior to the linear regression. This makes the linear estimate of $\mathcal{A}$ and $\mathcal{B}$ from SIR nearly equivalent to the linear regression of the raw data. Note that if we limit Eq. (4.14) to only the cubic, quadratic and linear terms, we get

$$\sigma^o_{corr} = \sigma^o - \mathcal{C}\theta^2 - \mathcal{D}\theta^3 = \mathcal{A} + \mathcal{B}\theta. \tag{4.17}$$

Using a fixed $\mathcal{C}$ and $\mathcal{D}$ values, their effect on σ^o can be removed prior to processing of the data with SIR. Estimates of $\mathcal{C}$ and $\mathcal{D}$ are based on linear regression analysis of various surface areas. A linear regression of SIR data created with this modified data stream is now not biased by the non-linear behavior at the incidence angle extremes. The linear $\mathcal{A}$ and $\mathcal{B}$ values given in Table 4.1 are made after the higher order terms are removed from the raw data. The cubic terms in the table are from a cubic regression using the raw data. Note that there is very little difference between the linear and cubic estimates of $\mathcal{A}$ and $\mathcal{B}$.

The addition of an increased $\mathcal{B}$ weight has dramatic results. Figures 4.30 and 4.31 show comparison scatter plots of $\mathcal{A}$ vs. $\mathcal{B}$ for SIR imagery with and without the additional weighting factor. Each scatter plot corresponds to a study region listed in Table 4.1. The cloud of points represents the $\mathcal{A}$ and $\mathcal{B}$ values for all pixels of the

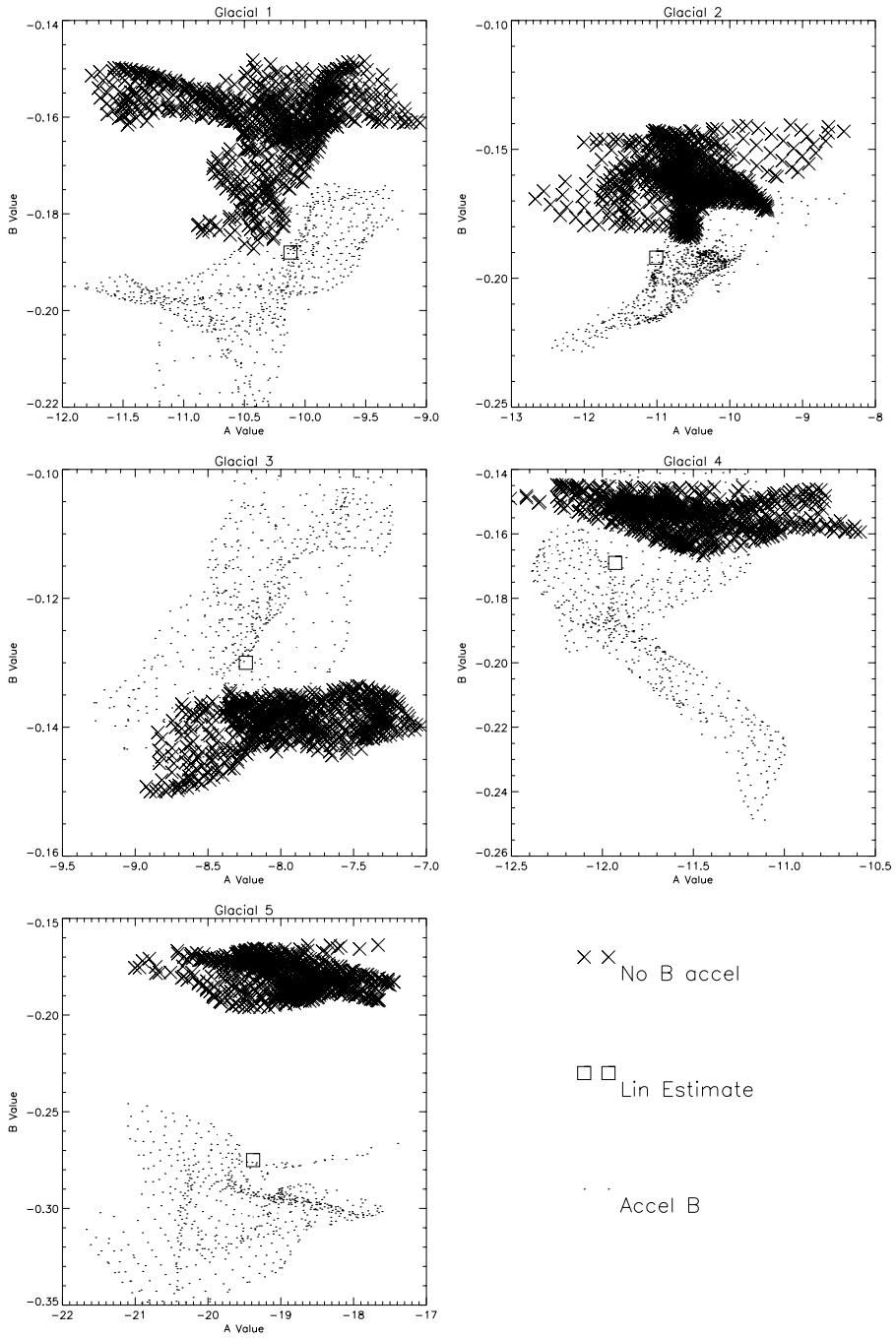


Figure 4.30: Scatter plots of Glacial Ice test regions. The information is from JD 150 1994 with a $\mathcal{B}$ acceleration factor of 50. Note how the $\mathcal{A}$ values remain relatively constant with accelerated $\mathcal{B}$ development, but the $\mathcal{B}$ values change significantly. The increased vertical spread of the $\mathcal{B}$ value for accelerated images better reflects actual surface conditions since the $\mathcal{B}$ value is expected to vary over a large surface area.

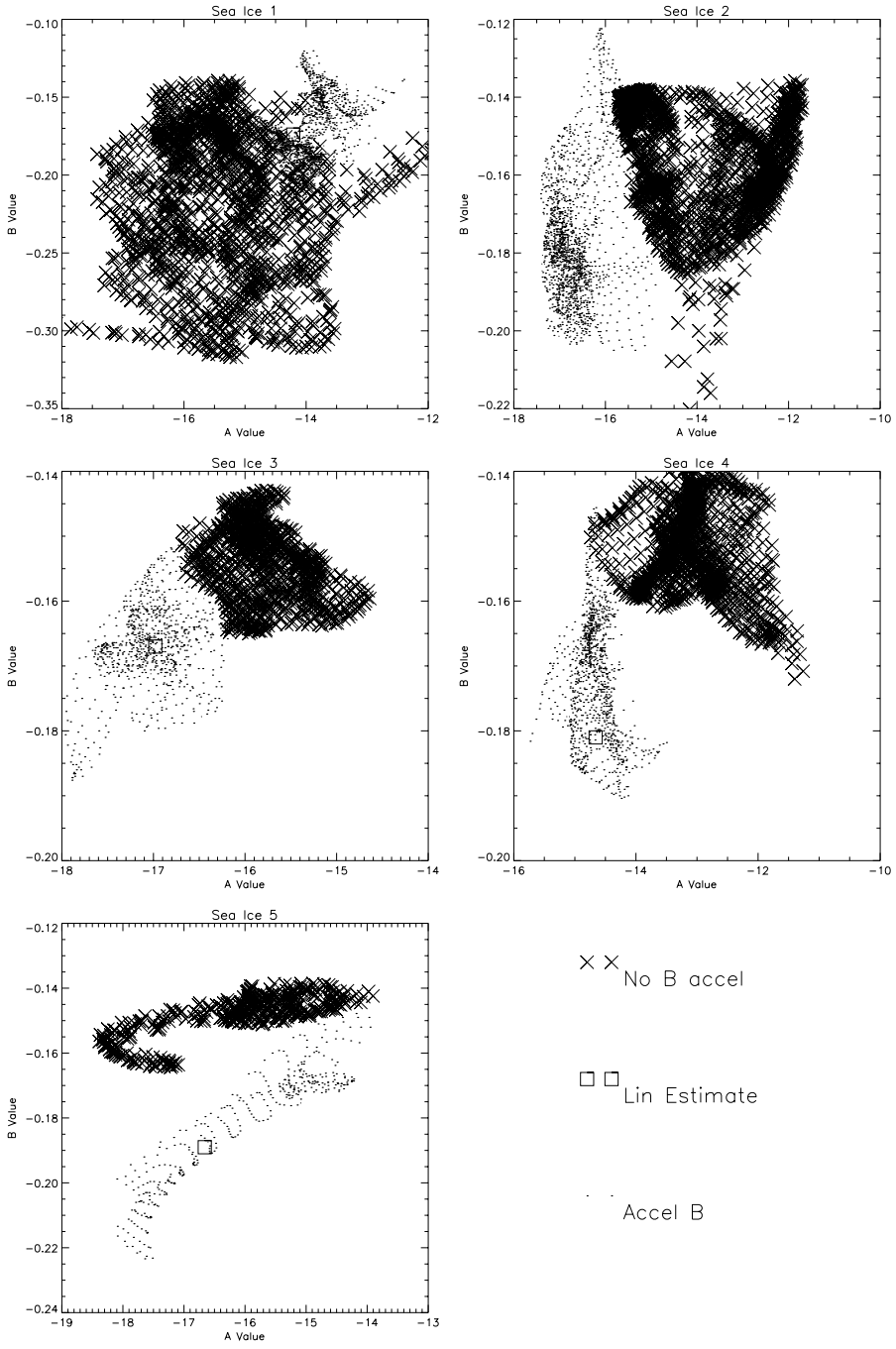


Figure 4.31: Scatter plots of sea ice test regions. The information is from JD 150 1994 with a $\mathcal{B}$ acceleration factor of 50. Both the $\mathcal{A}$ and $\mathcal{B}$ values are affected by increasing the $\mathcal{B}$ weighting.

SIR images that fall within the study region defined in the table. One cloud is the SIR imagery without extra $\mathcal{B}$ weighting and one is with the extra weighting. For comparison, the linear estimate of $\mathcal{A}$ and $\mathcal{B}$ from the raw data in the study region is plotted in each case as a square box. The raw data corresponds to the same time period (JD 150-155) as the image and the same spatial location as noted in Table 4.1. In each case, the weighted development is much closer to the linear estimates for the region.

First, Figure 4.30 shows glacial ice regions. In the each figure, the $\mathcal{A}$ and $\mathcal{B}$ parameters undergo significant change with the increased $\mathcal{B}$ weighting. In each case, the accelerated $\mathcal{B}$ data is clustered more closely to the linear regression $\mathcal{A}$ and $\mathcal{B}$ values. This would indicate that the SIR algorithm imagery from ERS-1 data is more accurate when the $\mathcal{B}$ weighting is higher than the nominal SEASAT weighting. The use of various weights demonstrated a significant change in the resulting $\mathcal{A}$ and $\mathcal{B}$ values as the weight was modified from 10 to 50, but less significant changes as the weight was increased from 50 to 150.

Figure 4.31 shows similar plots for the sea ice study regions. What is significant to note here is the how $\mathcal{A}$ and $\mathcal{B}$ change. For glacial regions in general, the $\mathcal{B}$ value moved more than the $\mathcal{A}$ value, as expected. The $\mathcal{B}$ is more sensitive to the initial value so may have farther to move, and the spread of the $\mathcal{A}$ values in the image remained relatively constant. For the sea ice regions, both the $\mathcal{A}$ and $\mathcal{B}$ values undergo significant changes which may indicate a poor choice for the initial value of $\mathcal{A}$. In all cases the resulting cluster was grouped nearer the linear regression estimates of $\mathcal{A}$ and $\mathcal{B}$.

It is significant to note that for the examples shown, the $\mathcal{A}$ range collapses to a narrow 2-4 dB range similar to the range of σ^o in the raw data shown in Figures 4.32 and 4.33. In the case of sea ice regions 2, 3, and 4, the unaccelerated $\mathcal{B}$ images have $\mathcal{A}$ values with a mean too high to match the trend of the actual raw data.

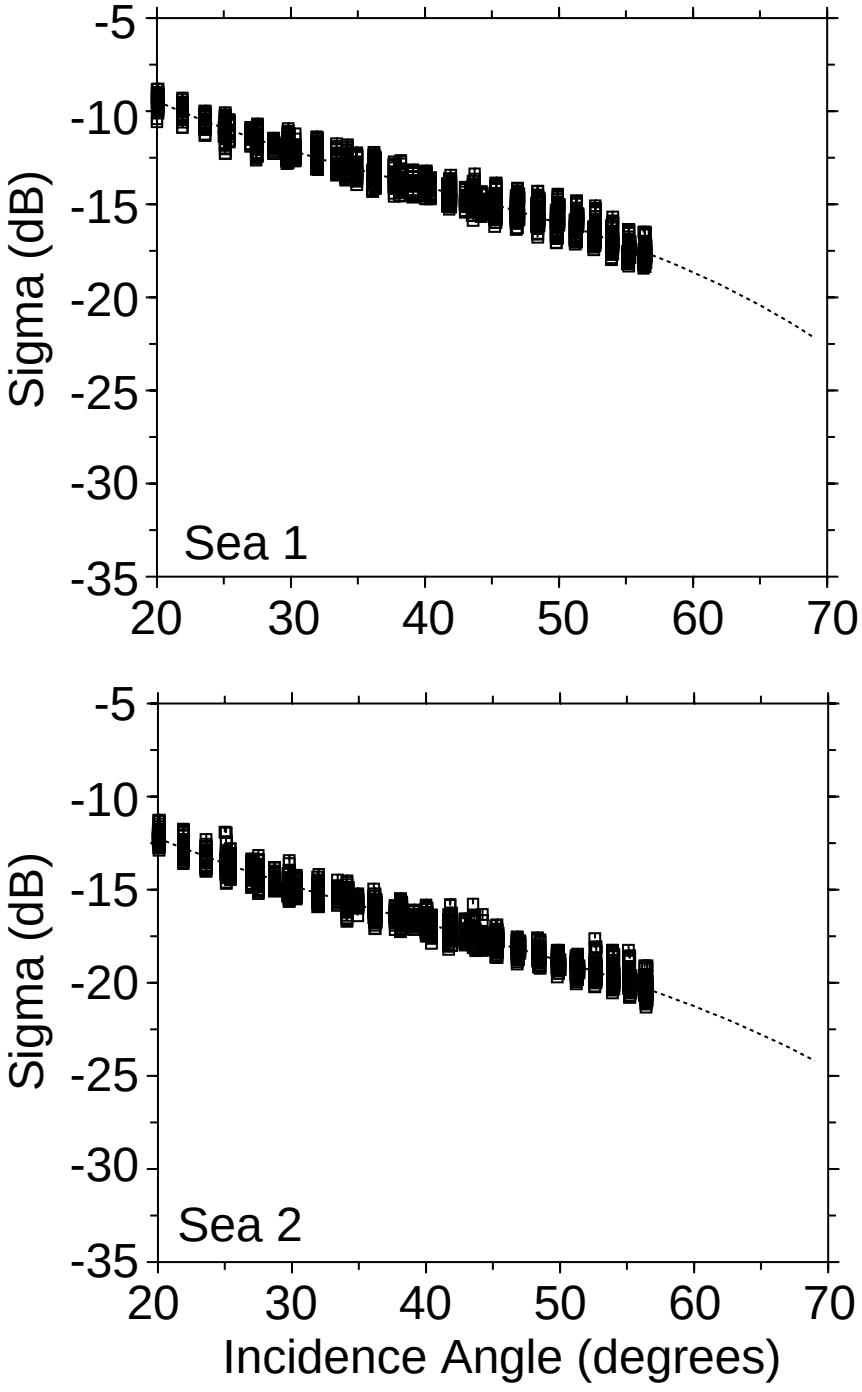


Figure 4.32: σ^o vs incidence angle plots of raw data from sea ice regions 1 and 2 (see Table 4.1).

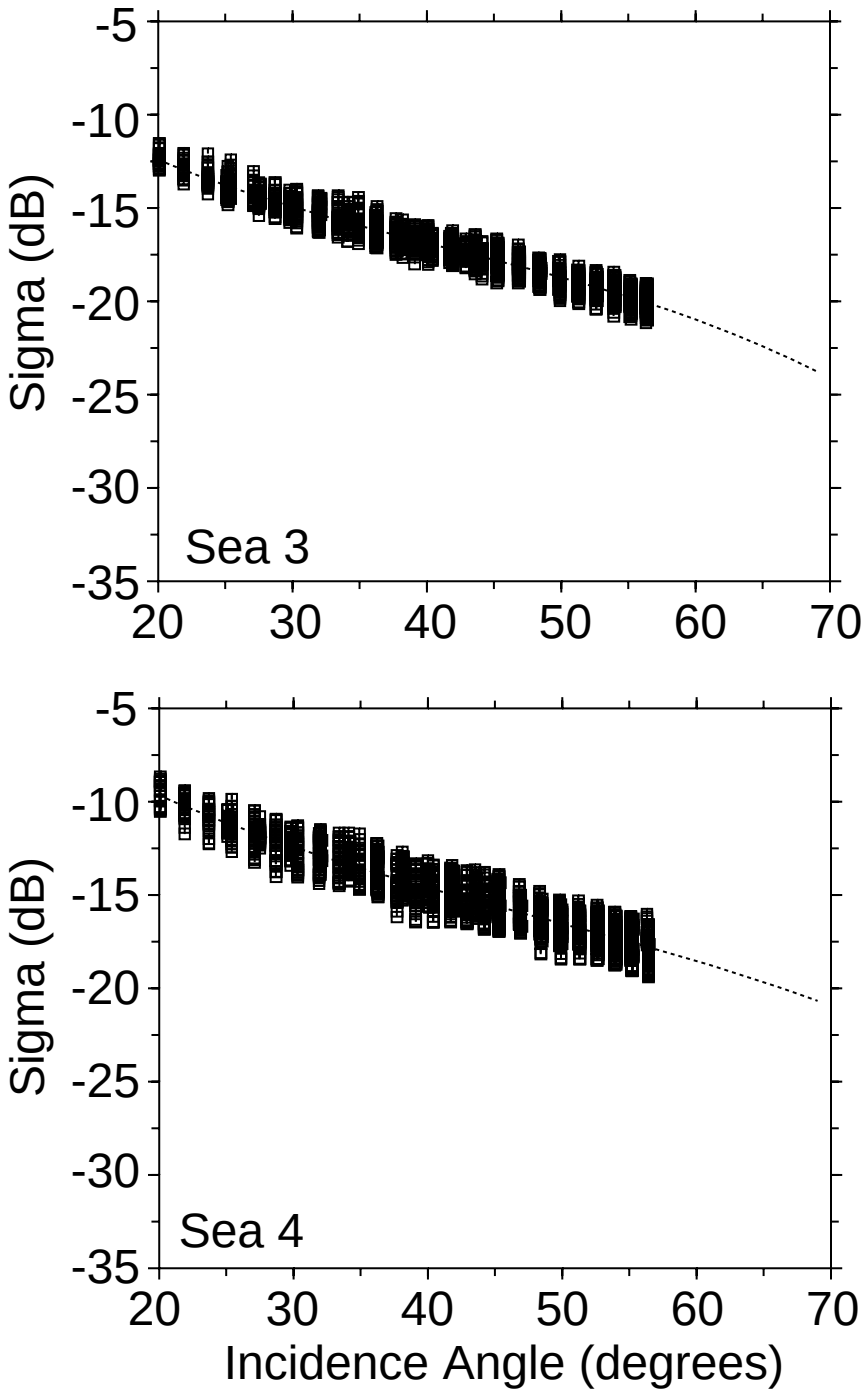


Figure 4.33: σ^o vs incidence angle plots of raw data from sea ice regions 3 and 4 (see Table 4.1).

4.4.5 $\mathcal{B}$ Weighting and Noise

The original argument for damping the $\mathcal{B}$ value development was to reduce the amplification of noise in the resulting imagery [13]. For SASS, the resulting imagery was quite good and the $\mathcal{B}$ images had sufficient dynamic range to be useful. As has been indicated above, accelerating the $\mathcal{B}$ value development did improve the overall accuracy of the ERS-1 test images, and when applied to actual ERS-1 data, the resulting $\mathcal{B}$ images had a usable dynamic range. In a subjective review of the ERS-1 SIR images, there was no evidence that the accelerated $\mathcal{B}$ value had significantly amplified noise.

4.5 Cumulative Effects of the SIR Algorithm Enhancements

In Section 4.4.2, it was demonstrated that the $\mathcal{A}$ value appeared to be relatively immune to initialization values if the $\mathcal{B}$ value was accurate. The $\mathcal{B}$ value, however, showed a diverging standard deviation and the $\mathcal{A}$ value failed to converge for cases where the $\mathcal{B}$ value was not initialized close to the actual value (see Figures 4.18 and 4.19). It is unlikely that initial $\mathcal{B}$ value will be identical to the true value for every pixel in the reconstruction, so now we analyze the performance of SIR for various initial values with the accelerated $\mathcal{B}$ value development outlined in the previous section. Only initialization values for $\mathcal{B}$ are tested on the range -0.3 and 0 since the $\mathcal{A}$ value converged even for poor estimates of $\mathcal{B}$.

In the first set of four images, the true image pair is $(\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1)$. Note here how the $\mathcal{A}$ and $\mathcal{B}$ values both converge in average and standard deviation to the desired value. As expected, a poor estimate of $\mathcal{B}$ requires more iterations to converge than a good estimate, but when compared to the results in Section 4.4.2, this is a significant improvement. Note how the standard deviation of the $\mathcal{B}$ value converges to the same level for the accelerated $\mathcal{B}$ development or unaccelerated $\mathcal{B}$ development (Compare the graphs in Figures 4.37 and 4.25). In the next set of four images, the true image pair is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$. The results are comparable to the previous ones: the convergence of the average and standard deviations of both $\mathcal{A}$ and $\mathcal{B}$ are much improved. The one anomaly is the

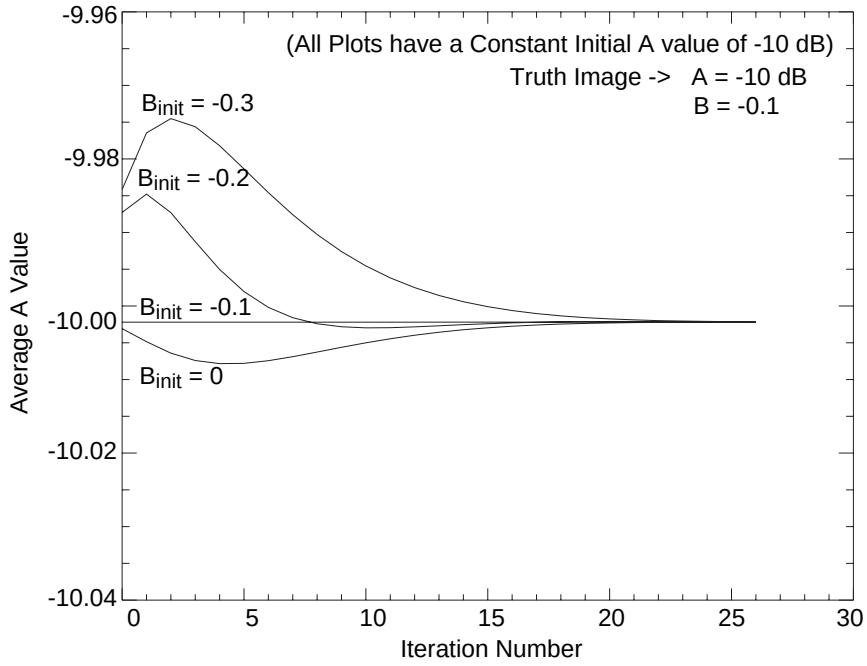


Figure 4.34: The average $\mathcal{A}$ value for accelerated $\mathcal{B}$ development. The true image is $(\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -10\text{d}B$. Note how the $\mathcal{A}$ value converges quickly even for very wrong initial $\mathcal{B}$ values.

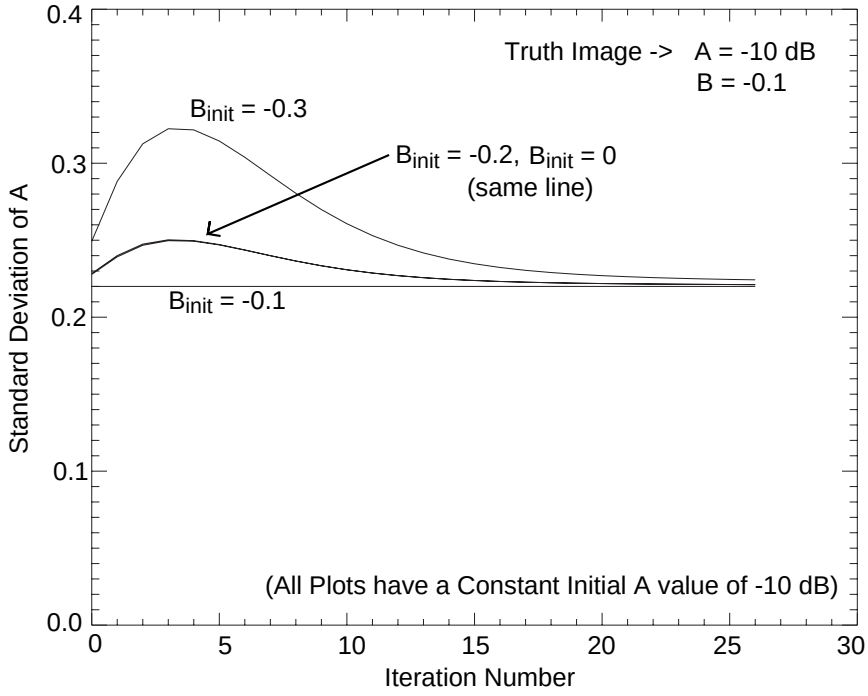


Figure 4.35: The standard deviation of the $\mathcal{A}$ value for the accelerated $\mathcal{B}$ development. The true image is $(\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -10\text{dB}$.

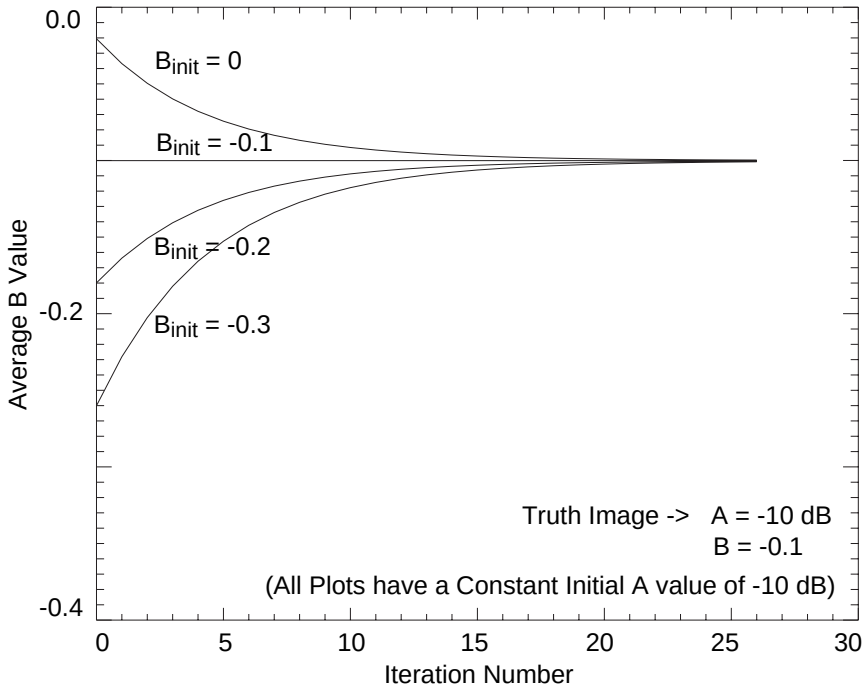


Figure 4.36: The average $\mathcal{B}$ value for accelerated $\mathcal{B}$ development. The true image is $(\mathcal{A} = -10\text{dB}, \mathcal{B} = -0.1)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -10\text{d}B$. Note how the $\mathcal{B}$ value converges quickly even for very wrong initial $\mathcal{B}$ values. (Compare with Figure 4.23).

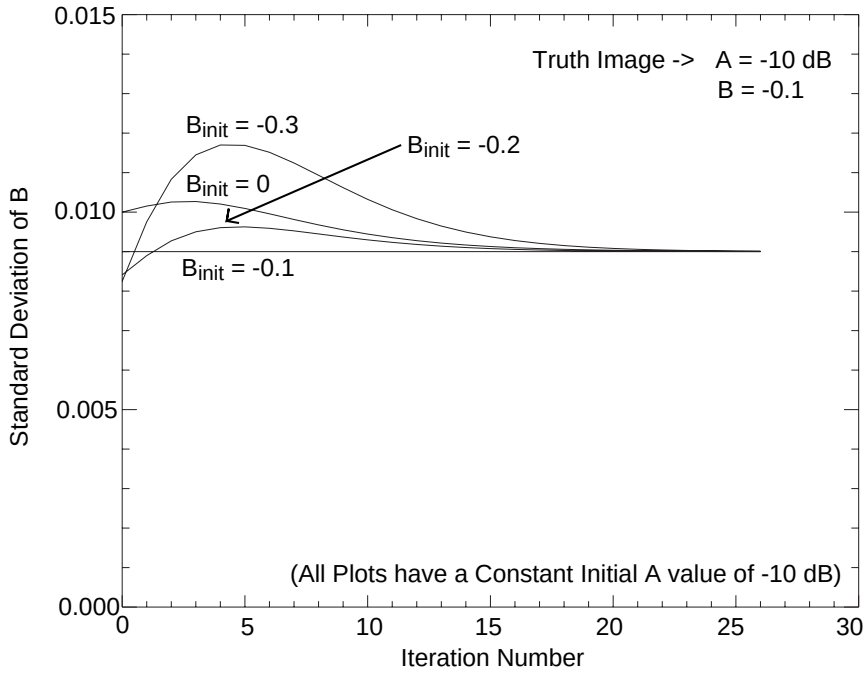


Figure 4.37: The standard deviation of the $\mathcal{B}$ value for accelerated $\mathcal{B}$ development. The true image is $(\mathcal{A} = -10\text{d}B, \mathcal{B} = -0.1)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -10$dB. Note how the standard deviation converges quickly even for very wrong initial $\mathcal{B}$ values. (Compare with Figure 4.25).

improved, but still non-zero, standard deviation of the $\mathcal{A}$ value (Figure 4.39). This is caused by the non-linear nature of the scatterometer measurements. SIR does several calculations in real space (vs. log space), which will result in some differences in the output when compared to similar image configurations but different absolute dB levels in those images. Note that the standard deviation in Figure 4.35 is much higher than the level in Figure 4.39, another result of this mathematical artifact. In both cases, however, the results are acceptable and the accelerated $\mathcal{B}$ value development has been demonstrated to significantly improve the estimate of both $\mathcal{A}$ and $\mathcal{B}$ values.

4.6 Summary

In this Chapter, we demonstrated how multiple overlapping measurements from the scatterometer can be combined to give a more dense sampling of the surface

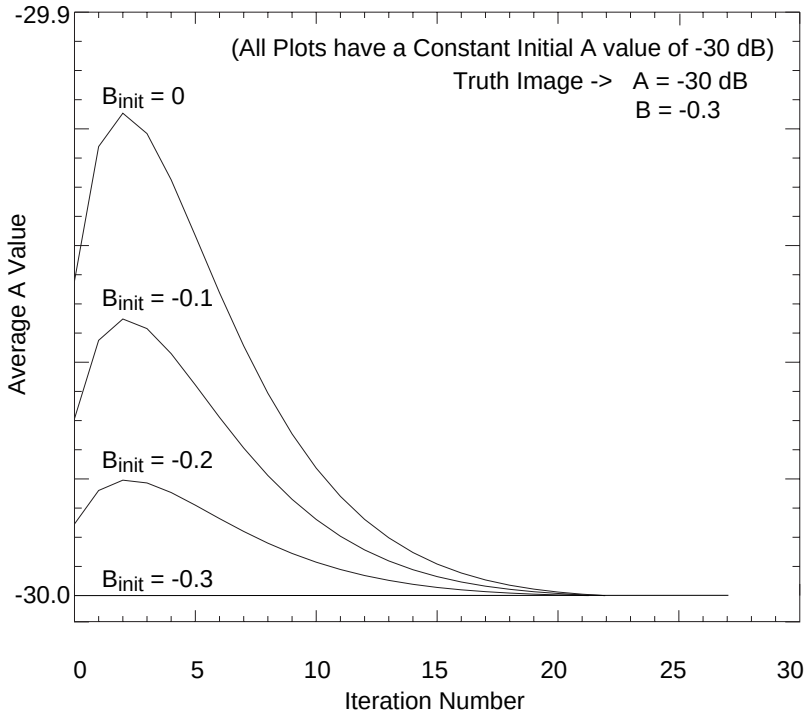


Figure 4.38: The average $\mathcal{A}$ value for accelerated $\mathcal{B}$ development. The true image is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -30\text{d}B$. Note how the $\mathcal{B}$ value converges quickly even for very wrong initial $\mathcal{B}$ values. (Compare with Figure 4.27).

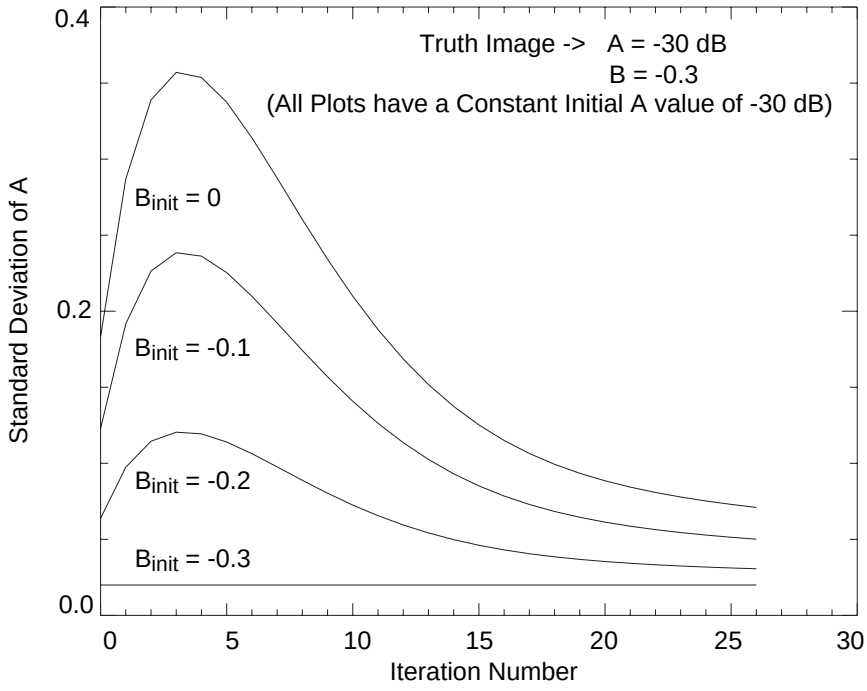


Figure 4.39: The standard deviation of the $\mathcal{A}$ value for accelerated $\mathcal{B}$ development The true image is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -30\text{dB}$.

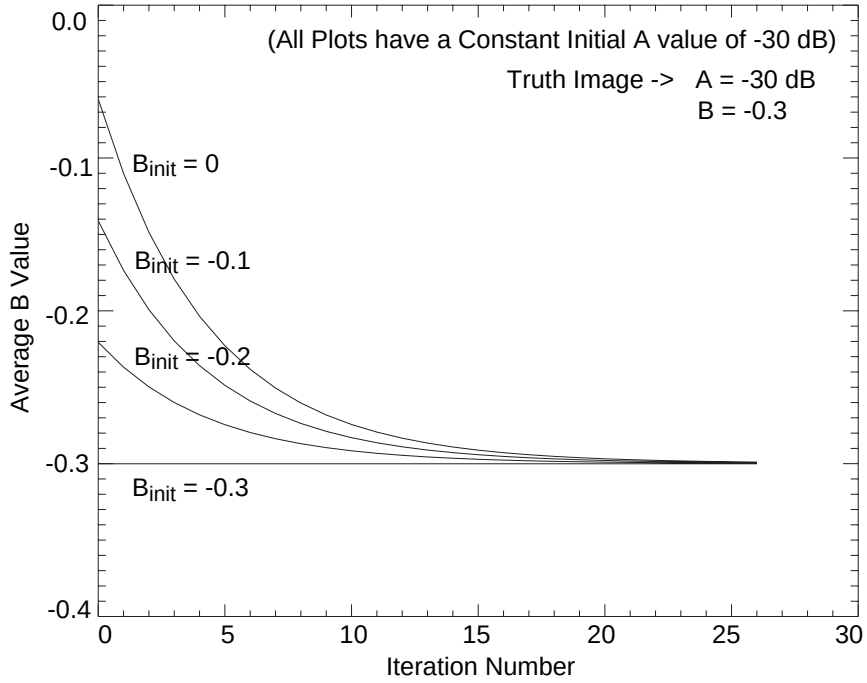


Figure 4.40: The average $\mathcal{B}$ value for accelerated $\mathcal{B}$ development. The true image is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -30\text{d}B$. Note how the $\mathcal{B}$ value converges quickly even for very wrong initial $\mathcal{B}$ values. (Compare with Figure 4.24).

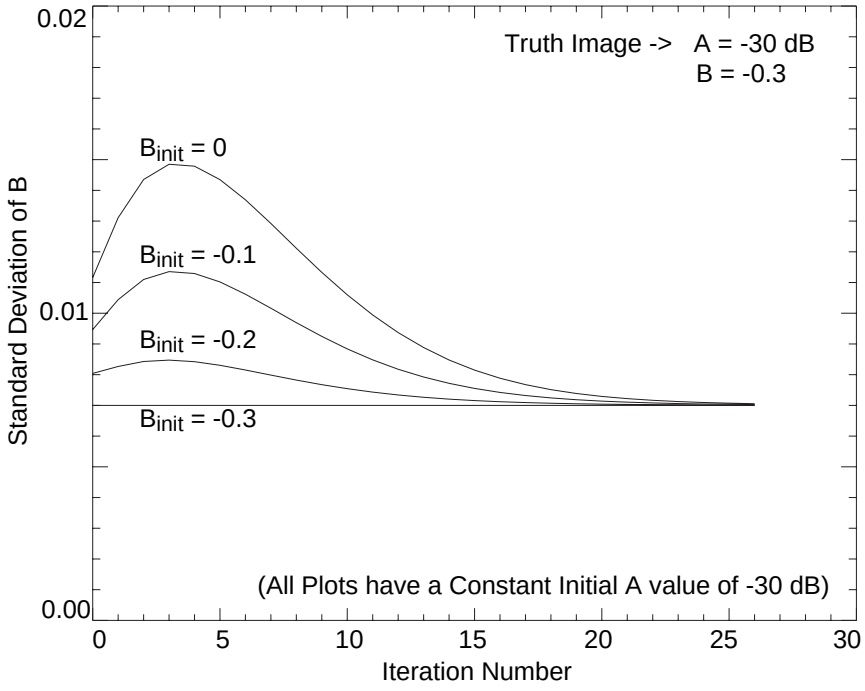


Figure 4.41: The standard deviation of the $\mathcal{B}$ value for accelerated $\mathcal{B}$ development. The true image is $(\mathcal{A} = -30\text{dB}, \mathcal{B} = -0.3)$, and the initial $\mathcal{A}$ value is $\mathcal{A} = -30\text{dB}$ (Compare with Figure 4.26).

than a single pass of the scatterometer, demonstrating the applicability of the previous Chapter's theoretical discussion to the ERS-1 scatterometer. We concluded this section with a brief discussion of the achievable resolution using ERS-1 scatterometer data. Next, several experiments using the SIR algorithm were performed to tune the algorithm for the ERS-1 data. These included reviews of $\mathcal{A}$ and $\mathcal{B}$ value initialization, addition of cubic parameterization to the backscatter model, and acceleration of the $\mathcal{B}$ value development. All of these issues are concerned with improving the accuracy of the SIR images. The results are as follows:

- Although the values will still converge from poor initial values, the initial $\mathcal{A}$ and $\mathcal{B}$ value should be made as close to the true value as possible. In most cases, this entails using the average $\mathcal{A}$ and $\mathcal{B}$ values, computed from the raw measurements, for the region being imaged. Minimizing the difference between true and estimate values reduces the standard deviation of the $\mathcal{A}$ and $\mathcal{B}$ values.

- Cubic parameterization using constant $\mathcal{C}$ and $\mathcal{D}$ values in the backscatter model is shown to be ineffective in a global sense. Spatial variation of the $\mathcal{C}$ and $\mathcal{D}$ parameters is too high even in small regions to make a global value effective in reducing cross hatching in the imagery. Further research into the development of a $\mathcal{C}$ or $\mathcal{D}$ image may be useful.

- Acceleration of the $\mathcal{B}$ value leads to more accurate images, especially the $\mathcal{B}$ images. Even for very inaccurate initial $\mathcal{B}$ values, the SIR algorithm converges to the true average and a minimal standard deviation in all cases.

Chapter 5

Azimuthal Modulation of C-Band Scatterometer σ^o Over Southern Ocean Sea Ice

5.1 Introduction

In previous chapters, we have outlined the theory behind reconstruction from irregular samples, shown the relationship between SIR, MART, AART and Gröchening's reconstruction algorithm, and fine tuned the SIR algorithm for the ERS-1 scatterometer. One issue remains to be addressed before applying SIR to a scientific study: azimuthal modulation. The basis of the SIR algorithm is the ability to combine measurements made at different times on different passes of the satellite over a region.

Because measurements from different passes of the satellite may have different azimuth angles, any modulation in azimuth of the backscatter from the surface could create errors in a reconstruction using SIR because our backscatter model is not parameterized with an azimuth angle (see Eq. (2.2)). In this chapter, we study azimuthal modulation of the C-band microwave signature of Southern Hemisphere sea ice and compare this with azimuthal modulation in the microwave signature of the Antarctic ice sheet in ERS-1 C-band scatterometer data. Azimuthal modulation has previously been observed over the Antarctic ice sheet using the SEASAT Ku-band scatterometer and is generated by wind induced ripples on the ice sheet surface [45]. However, given the dynamic nature of the Antarctic sea-ice pack during the year, we do not expect significant azimuth modulation over Antarctic sea ice at the scale of a

spaceborne scatterometer. Such an assumption is consistent with previous observations of little or no azimuthal modulation in airborne SAR and scatterometer data from Arctic sea ice in the Labrador Sea [17].

5.2 Southern Ocean Ice Characteristics

In order to understand the azimuthal modulation of σ^o over Antarctic sea ice at large scale (50km), we require an understanding of the general surface characteristics of sea ice in that region. Although the volume of *in situ* and radar measurements in the Antarctic is much smaller than similar Arctic measurements, there is sufficient data to make a large scale characterization of Antarctic sea ice [46] [47].

5.2.1 Surface Characteristics

In general, the Antarctic ice pack can be divided into two regimes: an outer ice pack and an inner ice pack, and for this study we use the definitions of these regimes as presented in [48]. Each regime has distinct physical properties that modulate microwave signatures of the ice as described below. The outer ice regime has two distinct phases: one during the winter freeze up and another during the summer melt.

The outer ice regime consists of the Marginal Ice Zone (MIZ) which is the extreme edge of the sea ice pack with a seasonally dependent makeup consisting of sea ice floes (up to several meters) surrounded by open water or slush [48]. The first phase of the outer ice regime occurs during early winter through early spring, when thermodynamic growth causes a rapid advance of the sea ice pack. The outer ice pack, and particularly the MIZ, are by definition regions of unsolidified or uncoalesced ice during winter freeze up, so wave action in this region makes pancake ice predominant in early winter [49] [50]. A photograph of pancake ice taken at the edge of the ice pack is shown in Fig. 1.1 and illustrates the development of pancake ice in the outer ice regime. Oscillatory wave action pushes grease ice, new ice and slush together and eventually thermodynamic cooling causes the pancakes to solidify. Before pancake ice

fields coalesce, the spaces surrounding the pancakes are either open water, frazil or grease ice. Thermodynamic effects will eventually cause the pancake field to coalesce into a solid ice pack. The second phase of the outer ice zone occurs during the spring and summer melt and break up of the sea ice pack. With the spring and summer warming, the pack ice begins to break up and melt, resulting in the MIZ containing large volumes of small, broken floes and brash ice.

The inner ice pack is typically thin to thick first year ice. Evidence from passive microwave systems shows that multiyear ice can survive in the Antarctic and it tends to be concentrated in the western Weddell Sea along the eastern edge of the Antarctic Peninsula [8]. Ridging, a major contributor to large scale deformation in Arctic sea ice, is in general much less intense in the Antarctic than in the Arctic with a lower average ridge height and lower frequency in the main body of the sea-ice pack [51] [52]. Also, as the ice ages, floes in the pack can be laden with snow to cause a negative freeboard condition, flooding the snow-ice interface. The existence of this wet slush layer changes the microwave properties of the sea ice, as does the subsequent re-freezing of this slush layer [46].

5.2.2 Azimuthal Modulation in the Antarctic

Azimuthal modulation of σ^o has been observed over the Antarctic ice sheet. Using Ku-Band SEASAT scatterometer data, *Remy et al.* [45] demonstrated that observed azimuthal modulation over the Antarctic ice sheet is related to the katabatic winds on the continent. Further, any oriented scatterers, including sastrugi, wind oriented drifts and crevasse fields, may create azimuthal modulation in the satellite data. Ice sheets, even without significant oriented scatterers on the surface, can create azimuth modulation if there is a significant surface slope (e.g. a glacier in a mountain valley).

Sea ice, on the other hand, has much different surface characteristics than land ice sheets. Small scale waves such as millimeter, capillary or gravity waves are absent in the outer ice regime and the rest of the sea-ice pack because of the presence on the ocean surface of either solidified pack ice, pancake ice in its various forms

or grease ice which prevent the formation of these small scale waves, eliminating one source of oriented scatterers in the ice pack. Also, the presence of water in the upper snow layer in some areas of the sea ice pack will change the structure of wind-etched surface features such as sastrugi. To further reduce the effects of any oriented scatterers that do develop on the sea ice surface, the dynamic motion of the ice surface causes a randomization of the scatterers over a large scale reducing the cumulative affect of scatterers on the return signal.

Since the sea ice floats on the surface of the ocean, we expect no inherent large scale surface slope associated with sea ice that would induce azimuthal modulation. However, because the ice in the outer ice regime is defined as uncoalesced ice, long wavelength swell-waves are capable of traveling through these outer regions of the sea ice pack [53] and inducing some surface slope. In the Southern Ocean, long wavelength swell waves, with wavelengths of several hundred meters and amplitudes of up to several meters are capable of traveling hundreds of kilometers into the sea-ice pack through pancake ice regions [54]. Once the pancakes have begun to coalesce and solidify however, the waves are quickly damped out by the increasingly rigid sea-ice pack.

In the absence of significant wave action, any significant slope in the sea ice must result from ridging or stacking of ice floes. However, the divergent nature of the sea-ice pack causes break up, rotation and refreezing of sections of the ice which effectively randomizes small scale ridges and other oriented scatterers may form on the surface of the sea ice. This study concentrates on microwave scattering characteristics of Antarctic sea ice on the scale of the ERS-1 AMI scatterometer (50km), and we postulate that over the majority of the sea ice pack, relatively small structure variations in the sea-ice surface will not introduce substantial azimuthal variation in the scatterometer data due to the randomizing effects of the sea-ice pack motion. We shall also establish that long wavelength swell wave penetration in the MIZ will not introduce substantial azimuthal modulation in the scatterometer signature.

5.3 Procedures

Sea-ice surface characteristics have significant spatial and temporal variation over a basin-wide area [36]. As a result, small study regions are used. Also, there is a dependence of σ^o on incidence angle which must be accounted for. Further, correlation between azimuth and incidence angles for a given cell resulting from the orbit geometry can bias azimuthal modulation evaluation. In the following section, σ^o dependence on incidence angle is discussed and a method for removing the dependence based on a linear model is presented. This is followed by a discussion of study region selection methodology that addresses the issues of spatial and temporal variation of the surface.

5.3.1 Removal of Incidence Angle Dependence

As noted in Chapter 2, Section 2.1, the radar measurements may have both incidence angle and azimuth angle dependence. The separation of any incidence angle dependence from the data is crucial for proper interpretation of any azimuthal modulation observed in the plots. Data collected over several days may have many different incidence angles and direct comparison of all measurements in a data set necessitates the removal of incidence angle dependence for some analyzes.

To aid in this removal we use a simple model relating backscatter to incidence angle, first given as Eq. 2.2 and repeated here for easy reference:

$$\sigma^o_{dB} \quad = \quad \mathcal{A} + \mathcal{B}(\theta - 40^\circ) \tag{5.1}$$

Note that σ^o_{dB} is the received backscatter in dB and θ is the incidence angle of the measurement. Recall that $\mathcal{A}$ is the normalized radar backscatter (normalized to 40° in this case) and $\mathcal{B}$ represents the slope of the data with respect to the incidence angle θ. An estimate of the parameter $\mathcal{B}$, denoted $\hat{\mathcal{B}}$, is determined from a linear regression of the σ^o measurements for each region reviewed in this study. With a $\hat{\mathcal{B}}$ estimated for a given study region, the estimate $\widehat{\mathcal{A}}_i$ for each backscatter measurement σ^o_i in the study region is given by

$$\widehat{\mathcal{A}}_i \quad = \quad \sigma^o_i - \widehat{\mathcal{B}}(\theta_i - 40^\circ). \tag{5.2}$$

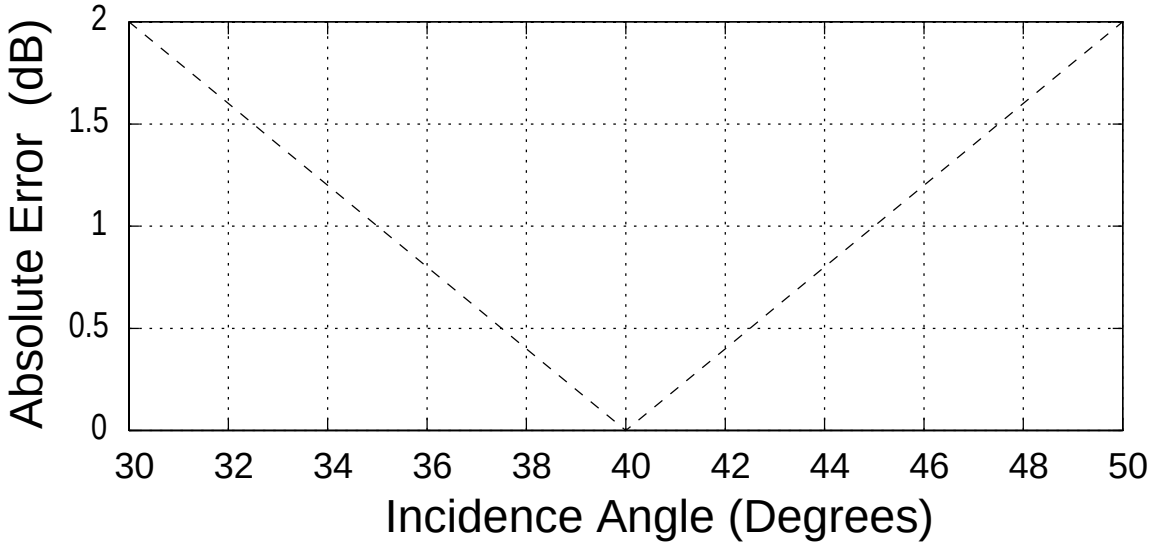


Figure 5.1: Plot of the maximum error in $\hat{\mathcal{A}}$ caused by a worst case $\hat{\mathcal{B}}$ error. This graph shows an example where the measurements are normalized to 40 degrees. Note that the error for the worst case is less than 0.5 dB for an incidence angle range of $\pm 2^{\circ}$ around the normalization angle. It is expected that the error will be considerably less in practice than the worst case error.

The resulting $\widehat{\mathcal{A}}_i$ values represent incidence angle normalized backscatter values, i.e. the value of σ_i^o at $\theta = 40^o$.

5.3.2 $\hat{\mathcal{B}}$ Error

The use of an inaccurate $\mathcal{B}$ value will introduce error into the algorithm. Note that each σ^o measurement in a study region data set represents a unique backscatter measurement from a single radar footprint. Because the surface area of the footprint is smaller than the total area in each study region, we can reasonably assume that due to spatial variability within the study region, each σ^o measurement may have a unique $\mathcal{B}$ value. Error is introduced by using a single estimate of the $\mathcal{B}$ parameter to determine all of the incidence angle normalized backscatter estimates $\mathcal{A}_i$ in a given study region. And although the linear model in (5.1) is a good approximation of the incidence angle dependence, some error is realized from using the linear model.

Suppose that for a given measurement σ_i^o, the true value of $\mathcal{A}$, $\mathcal{A}_t$ is given by

$$\mathcal{A}_t \quad = \quad \sigma_i^o - \mathcal{B}_t(\theta_i - 40^\circ) \tag{5.3}$$

where $\mathcal{B}_t$ represents the true $\mathcal{B}$ value for the i^{th} measurement σ_i^o . If $\widehat{\mathcal{B}}$ is not exactly equal to the true value $\mathcal{B}_t$, the error in the estimate $\widehat{\mathcal{A}}_i$ is given by

$$\Delta\mathcal{A} = \quad \mathcal{A}_t - \widehat{\mathcal{A}}_i \quad = (\widehat{\mathcal{B}} - \mathcal{B}_t)(\theta_i - 40^\circ) \tag{5.4}$$

Introducing similar notation for the error in $\widehat{\mathcal{B}}_i$, let $\Delta\mathcal{B} = \widehat{\mathcal{B}}_i - \mathcal{B}_t$. If we assume that $\Delta\mathcal{B}$ is bounded by some maximum value, the maximum error in $\widehat{\mathcal{A}}_i$ becomes

$$\Delta\mathcal{A}_{max} \quad = \quad \pm\Delta\mathcal{B}_{max}(\theta - 40^\circ). \tag{5.5}$$

For the purposes of evaluating error we can assume, based on the average $\mathcal{B}$ in Tables 5.1 and 5.2 for the regions evaluated in this study, that the mean $\mathcal{B}$ is approximately -0.2 with a worst case range of 0.0 to -0.4, making $\Delta\mathcal{B}_{max} = \pm 0.2$. $\Delta\mathcal{A}_{max}$, is plotted vs. incidence angle in Fig. 5.1.

Because the data can be normalized to any angle, the error introduced by removing incidence angle dependence can be minimized by normalizing the data to an angle in the middle of the incidence range of the data. Note that the graph in Fig. 5.1 shows the worst case error for the assumptions in the previous paragraph; in practice the error will be much smaller. We conclude that the error introduced by faulty $\mathcal{B}$ estimates is negligible when the normalization angle is in the middle of a narrow range of incidence angles. Assuming a narrow incidence angle range (less than 6°), the error will not adversely affect evaluation of azimuthal modulation of 1 dB or more.

5.3.3 Study Region Selection

As part of an evaluation of basin-wide characteristics, study regions in several areas of the Antarctic sea-ice pack are used. Relatively small study regions and short study intervals are used so that areas of relatively constant temporal and

Reg	lon x lat	1993 JD	$\hat{A}$	$\hat{B}$	σ^o Std Dev
I1	8° x 4°	23- 56	-14.41	-0.231	2.69
I2	10° x 4°	119-140	-14.68	-0.199	2.10
I3	14° x 3°	215-242	-14.85	-0.193	3.25
I5	8° x 4°	218-239	-17.26	-0.218	2.28
I6	8° x 3°	248-290	-14.63	-0.230	2.61
I7	18° x 5°	326-350	-15.14	-0.254	2.84
I8	10° x 4°	311-332	-16.80	-0.235	2.42
I10	8° x 3°	200-225	-16.56	-0.201	2.14
I11	8° x 3°	200-225	-16.05	-0.205	2.46
I12	8° x 3°	200-225	-15.81	-0.208	2.43
I13	8° x 3°	200-225	-16.44	-0.218	2.36
I14	8° x 3°	200-225	-16.33	-0.225	2.48
I15	8° x 3°	200-225	-15.65	-0.223	2.88
I16	8° x 3°	200-225	-16.94	-0.208	2.11
I17	8° x 3°	200-225	-17.12	-0.213	2.21
I18	8° x 3°	200-225	-16.92	-0.211	2.24
I19	8° x 3°	200-225	-16.86	-0.210	2.15
I20	8° x 3°	200-225	-16.80	-0.213	2.13
I21	8° x 3°	200-225	-16.60	-0.219	2.25
G1	10° x 2°	109-227	-8.07	-0.174	2.13
G2	10° x 2°	80-117	-15.42	-0.232	3.62
G3	10° x 2°	109-227	-10.56	-0.209	3.09
G4	10° x 2°	80-117	-11.59	-0.188	2.04
G5	10° x 2°	109-227	-15.01	-0.174	4.49

Table 5.1: Study Region Locations and Statistics.

spatial variation can be studied. Selecting areas that are spatially and temporally homogeneous avoids creating study regions with many different ice surfaces that might skew any azimuthal modulation evaluation. Study regions from several areas in the Southern Ocean are used to evaluate azimuthal modulation over many different sea-ice surfaces. The scatterometer requires 3 to 5 days to collect enough readings for each study region to have good azimuth angle diversity. The surface conditions in each study region are assumed constant over the data collection interval, a fair assumption for the scale of observation of the scatterometer.

Study regions are chosen such that the regions are homogeneous in time and space over the data collection interval. To aid in the selection of homogeneous

Reg	lon x lat	1993 JD	$\hat{\mathcal{A}}$	$\hat{\mathcal{B}}$	σ^o StdDev
M1	6° x 2°	102-120	-12.68	-0.197	2.15
M2	10° x 2°	102-120	-12.77	-0.221	2.57
M3	10° x 2°	105-120	-12.66	-0.191	2.07
M4	8° x 2°	128-146	-12.79	-0.211	2.44
M5	6° x 2°	128-146	-12.32	-0.400	5.10
M6	10° x 2°	141-161	-12.42	-0.169	1.96
M7	10° x 2°	142-161	-13.51	-0.207	2.53
M8	10° x 2°	141-161	-15.84	-0.217	2.29
M9	10° x 2°	144-164	-13.16	-0.340	4.29
M10	10° x 2°	150-170	-12.17	-0.179	2.00
M11	10° x 2°	153-173	-13.04	-0.204	2.20
M12	6° x 2°	165-181	-12.67	-0.183	2.06
M13	10° x 2°	165-181	-13.32	-0.217	2.42
M14	10° x 2°	174-181	-12.48	-0.194	2.04

Table 5.2: Marginal Ice Zone Study Region Locations and Statistics.

regions, we use a time series of enhanced resolution images of Antarctic land and sea ice [36] to identify the largest possible regions where the spatial surface response is visually homogeneous. The spatial homogeneity of the σ^o measurements in a study region reduces the variance of the measurements, ensuring accurate assessment of low level (< 1 dB) azimuthal modulation for a given sea ice surface. Study regions that represent several different types of sea ice are selected in order to evaluate azimuthal modulation over different sea-ice surfaces. An example of an enhanced resolution Antarctic image is shown in Fig. 5.2. An explanation of the SIRF algorithm for generating the enhanced resolution images is found in [37].

An additional criterion for selecting viable study regions is adequate diversity of azimuth angles. Azimuth angle diversity is required in order to properly evaluate azimuthal modulation and this diversity is affected by the location and size of a study region and the number of days in the study interval. The scatterometer requires several days of data to generate data with good azimuthal angle diversity. If too many days of data are included, the dynamic nature of the Antarctic sea ice

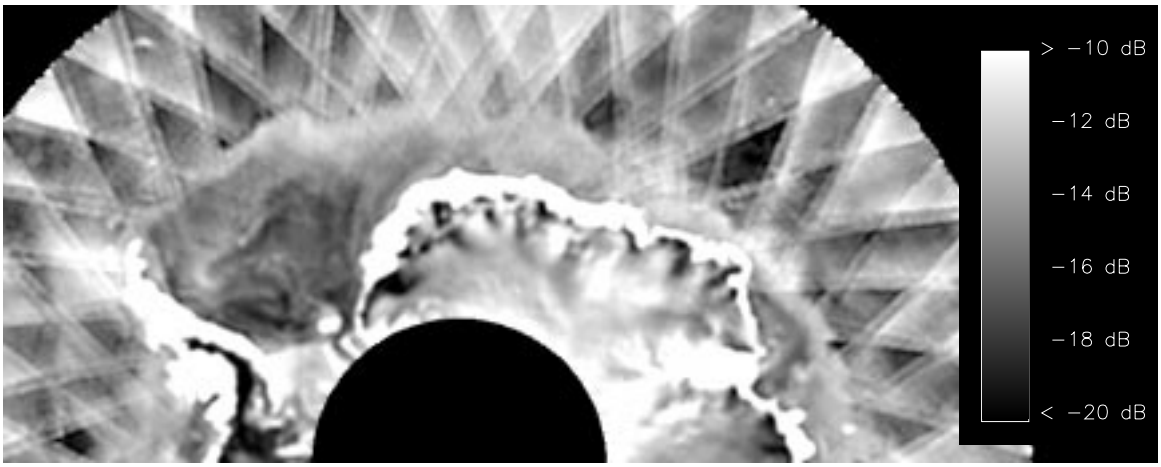


Figure 5.2: A polar stereographic projection image of Antarctica. The image is generated from 6 days of ERS-1 data from JD 126 to JD 131 1993. The hash marks in the periphery of the image are a result of the rapidly changing azimuthal response of the open ocean surface.

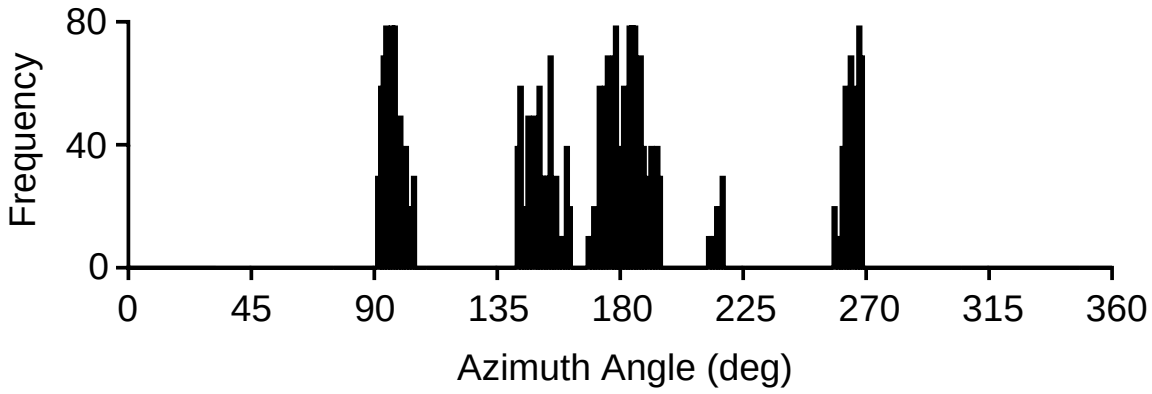


Figure 5.3: Example histogram of azimuthal angles over Antarctic sea ice (Study Region I10). Note that the gaps are caused by the instrument geometry.

increases the probability that the surface will change within the study interval. However, if too few days of data are used, there will not be enough measurements to yield sufficient azimuthal diversity for modulation assessment or sufficient incidence angles to properly estimate the incidence angle dependence of the data in the study region.

The successful tradeoff between study region size (spatial homogeneity) and the number of days (temporal homogeneity) in the data set is evaluated by manually examining data in each study region. An evaluation of each study region is made to determine whether the data is spatially and temporally homogeneous and whether it has sufficient azimuth angle diversity to show azimuthal modulation. The evaluation for homogeneity is done by plotting the σ^o values versus incidence angle and evaluating the data visually for σ^o spread and variance in the σ^o vs. incidence angle plot, plotting the evolution of the σ^o values versus time (to evaluate temporal stability) and by examining the statistics of the data. The data statistics are illustrated in Tables 5.1 and 5.2.

To evaluate the data for azimuth angle diversity, a histogram of azimuth angles for several incidence angle ranges is plotted. An example azimuth angle histogram of ERS-1 scatterometer data over Antarctic sea ice is given in Fig. 5.3 for the incidence angle range 40 to 45 degrees. The ERS-1 data over sea ice shows a limited range of azimuth angles. Notice the groupings of azimuth angles in the example histogram in Fig. 5.3. These groupings are evident in all the data used in this study and are a consequence of the ERS-1 instrument geometry and orbit. Also, ERS-1 instrument geometry and orbit provide very few readings at azimuth angles above 270 degrees or below 90 degrees. Histograms for each incidence angle range are evaluated for each study region for adequate azimuthal diversity. Incidence angle ranges of $\pm 2.5^\circ$ and $\pm 1.5^\circ$ are examined for each study region over the incidence angles from 25° to 55°. Ranges of $\pm 1.5^\circ$ around the average incidence angle of 40° provide good azimuth angle diversity with an acceptably narrow range of incidence angles.

A total of 14 study regions near the edge of the sea-ice pack were selected to study azimuthal modulation in the marginal ice zone during the winter freeze up. The penetration of long wavelength swell waves into the uncoalesced MIZ sea ice may

result in enough surface slope to induce azimuthal modulation which is not coupled to sea ice surface characteristics. The MIZ study regions were selected based on enhanced resolution imagery and are selected to include regions where wave penetration is likely. Note in Fig. 5.2 the bright area near the ice pack edge. The subresolution scatterers in the MIZ cause a brighter microwave signature than other parts of the sea ice pack, and have been observed in the Labrador Sea MIZ [17].

Although the manual evaluations provide only a crude consistency check of the data, the check is sufficient for reviewing time periods and parameter ranges suitable for evaluating azimuthal modulation. A total of 19 study regions in the Antarctic sea-ice pack at various times of the year in 1993 are selected and used in this study, as well as 14 additional regions selected near the ice edge. The study regions are illustrated in Fig. 5.4, and statistics for each region are given in Table 5.1. The 14 MIZ areas are illustrated in Fig. 5.5, with statistics in Table 5.2. The study regions in the Weddell sea are large and overlapping to provide a better picture of the characteristics of the sea-ice pack in this highly dynamic region. The Julian Day (JD) in Tables 5.1 and 5.2 reflect the time period for which data was extracted in each region. In practice, smaller day ranges are used in evaluating azimuthal modulation to better approximate constant surface conditions.

5.4 Analysis

The data is analyzed using two different methods. First, by assuming that the data in each study region is representative of a single type of sea ice, the diversity of azimuth angles in each study region data set leads to a natural test for azimuthal modulation: plotting $\hat{\mathcal{A}}$ versus azimuth angle. If the σ^o spread in the data is low any azimuthal modulation should be apparent in these plots. Using small incidence angle ranges reduces error introduced by the necessary correction for incidence angle dependence.

Second, the fore-aft beam difference is examined. Because the fore and aft beams are 90° apart in azimuth, any significant difference in azimuth response is likely to appear as a difference between the fore and aft beam measurements. Also,

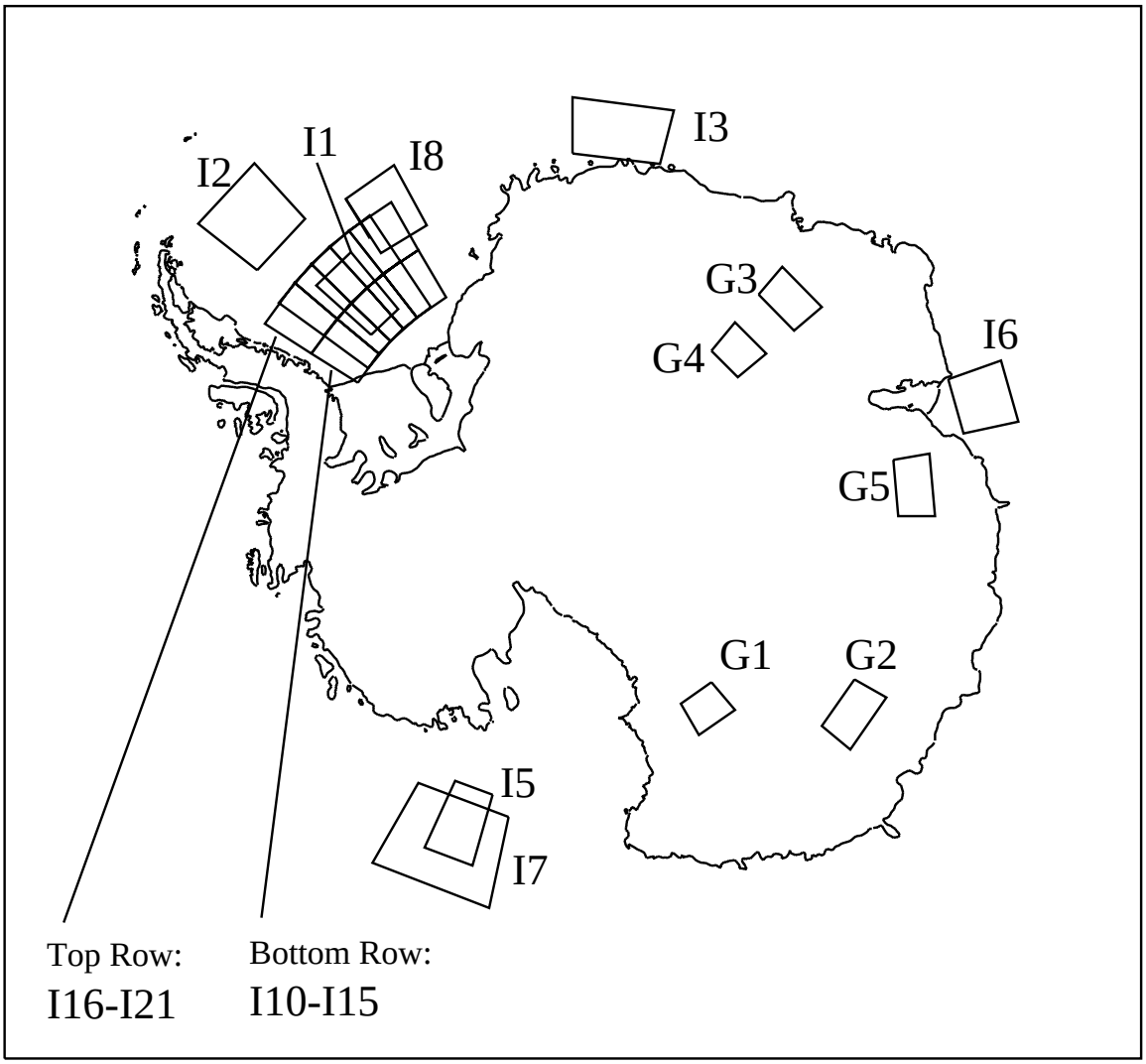


Figure 5.4: Study Regions used with the ERS-1 data. The boxes indicate the location of the data. Different time periods are used for the various boxes. All data is from 1993. Regions I10 through I21 are overlapping regions in the Weddell sea.

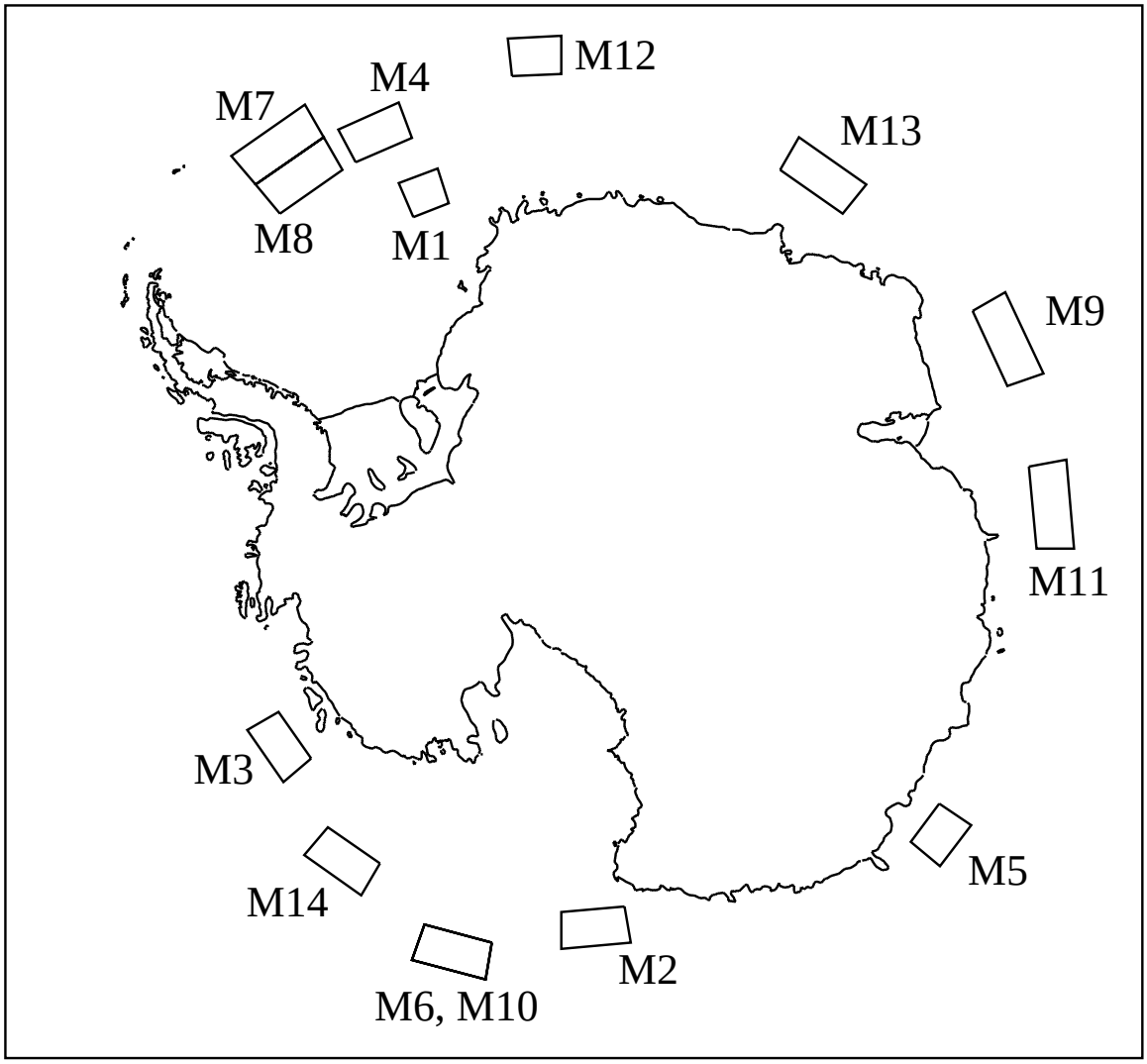


Figure 5.5: MIZ Study Regions used with the ERS-1 data. The boxes indicate the location of the data. Different time periods are used for the various boxes. All data is from 1993. These regions were selected to be near the sea-ice pack edge.

because the fore and aft beams have identical incidence angles, no correction for incidence angle dependence is necessary, thus eliminating a potential source of error in the analysis.

5.4.1 $\hat{\mathcal{A}}$ versus Azimuth

Figs. 5.6 and 5.7 show representative plots of $\hat{\mathcal{A}}$ versus azimuth angle for small incidence angle ranges (37° to 39°) over sea ice. For comparison, Fig. 5.8 shows a representative plot of $\hat{\mathcal{A}}$ versus azimuth angle for a small incidence angle range over the Antarctic ice sheet. The range of $\hat{\mathcal{A}}$ is relatively high in the ice sheet regions but is comparable to the spread found in plots of ice sheet response in *Remy, et al.* [45]. These plots are representative of the graphs produced in this study for all land and sea-ice regions.

Based on the discussion in Section 5.3.2, we can ignore incidence angle dependence when the data is taken over a 3-4 degree range. A comparison of corrected and uncorrected plots over small incidence angle ranges shows little difference if the normalization angle is chosen to be within the incidence angle range

An examination of Figs. 5.6 and 5.7 shows negligible azimuth angle modulation over sea ice. In all sea-ice regions studied, the observed variation in azimuth angle of σ^o was less than 1 dB. Note that the plot in Fig. 5.8 shows significant modulation in azimuth in microwave signatures over land ice sheets.

5.4.2 Fore-Aft Pair Analysis

Because the fore and aft beams have nearly identical incidence angles and azimuth angles 90 degrees apart, it is probable that azimuthal modulation will be displayed in the difference between the fore and aft beam measurements. In this analysis, the simple difference between fore and aft beam measurement pairs are calculated and bulk statistics for each study region are examined. Also, the fore-aft difference over a limited azimuth range is examined over sea ice and land ice.

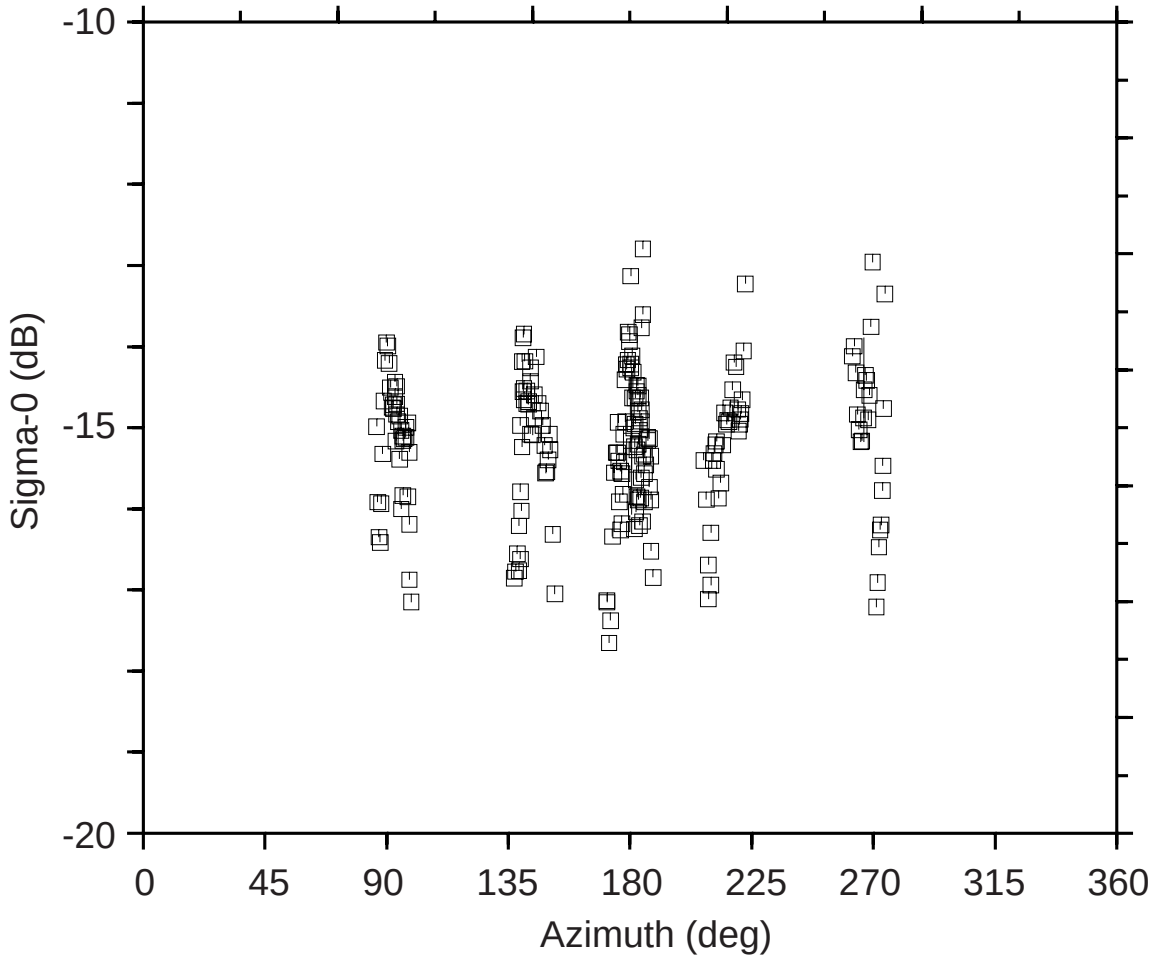


Figure 5.6: Representative σ^o versus azimuth angle plot for sea ice. This region is in the Weddell Sea, region I1 in Fig. 5.4. $\hat{\mathcal{A}}$, the incidence angle normalized σ^o, is plotted.

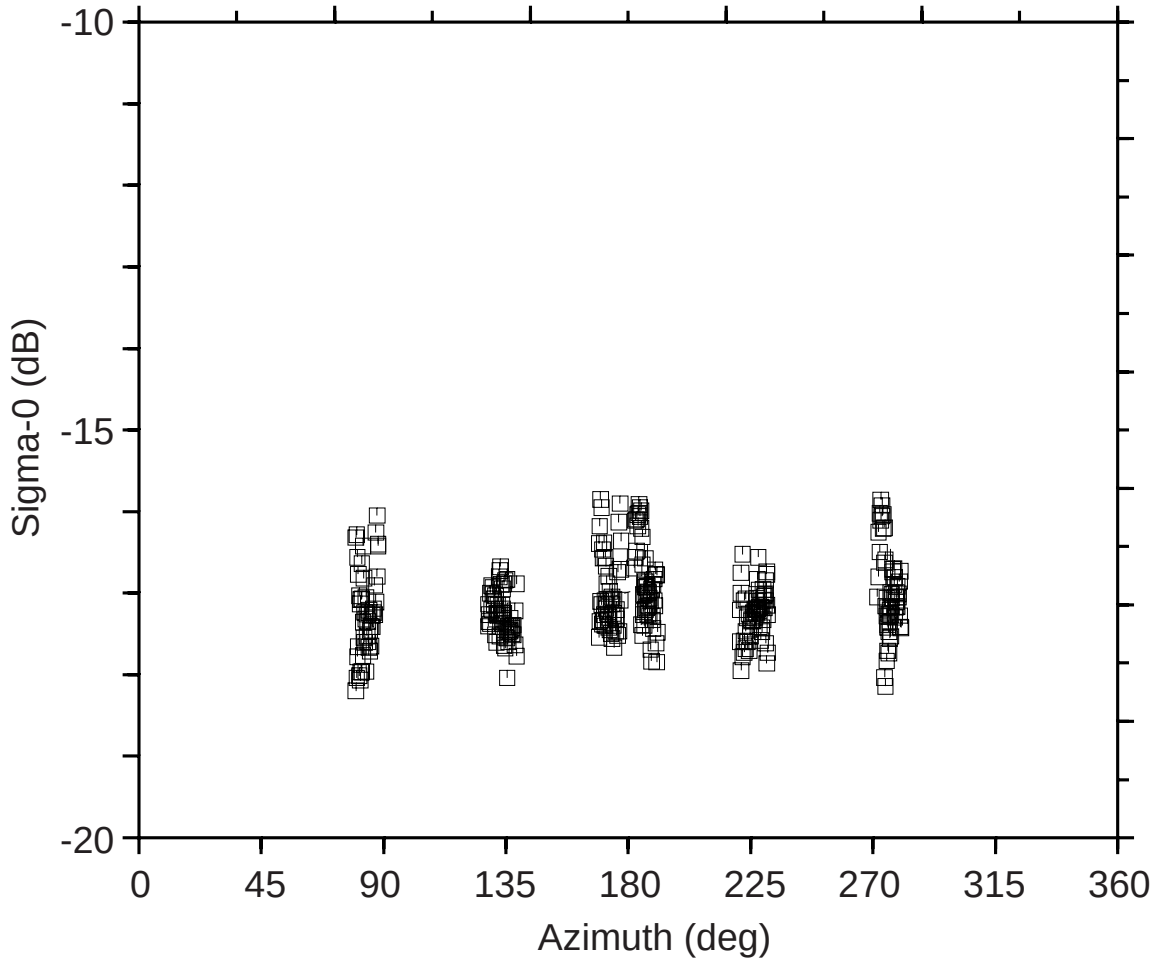


Figure 5.7: Representative σ^o versus azimuth angle plot for sea ice. This region is in the Weddell Sea, region I8 in Fig. 5.4. $\hat{\mathcal{A}}$, the incidence angle normalized σ^o, is plotted.

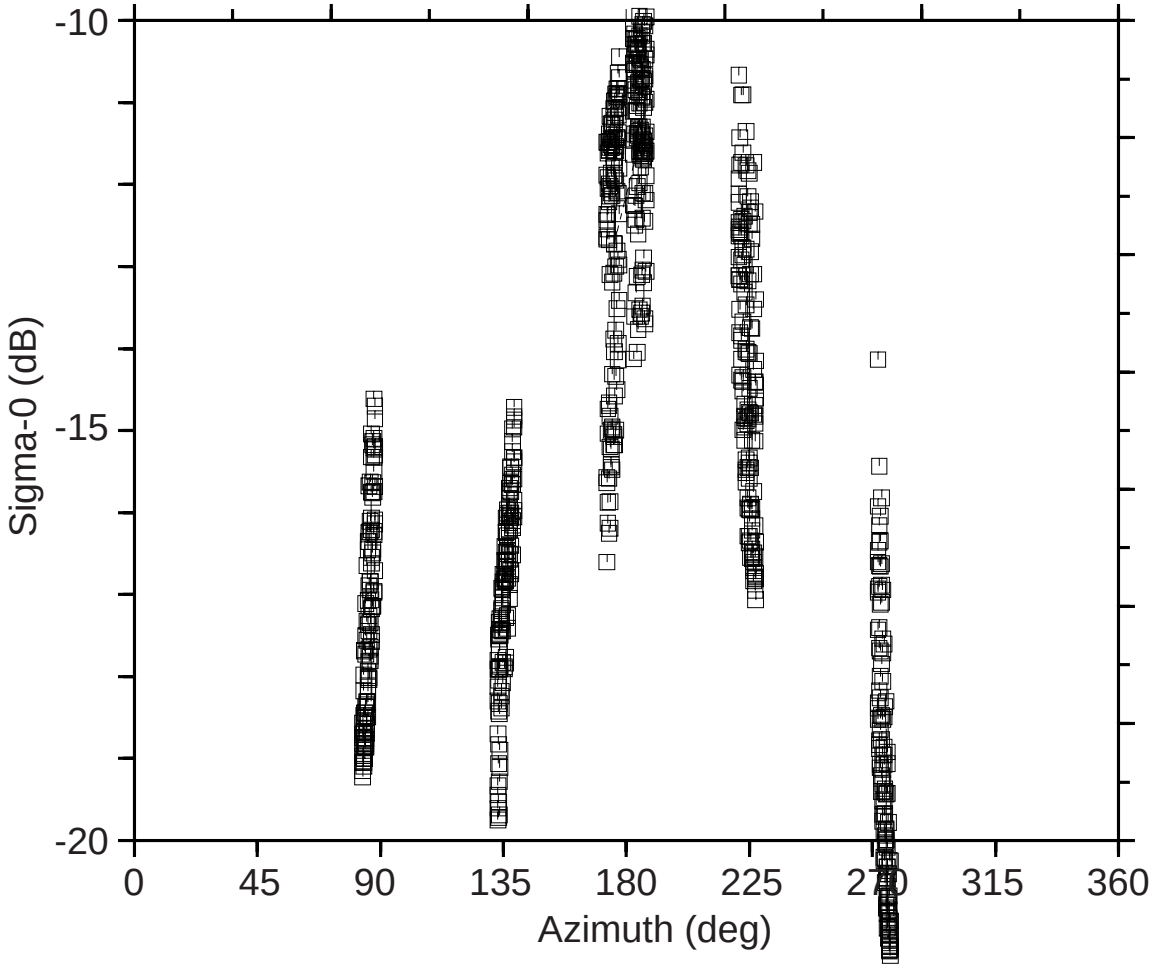


Figure 5.8: Representative σ^o versus azimuth angle plot for Antarctic glacial ice. $\hat{\mathcal{A}}$, the incidence angle normalized σ^o, is plotted.

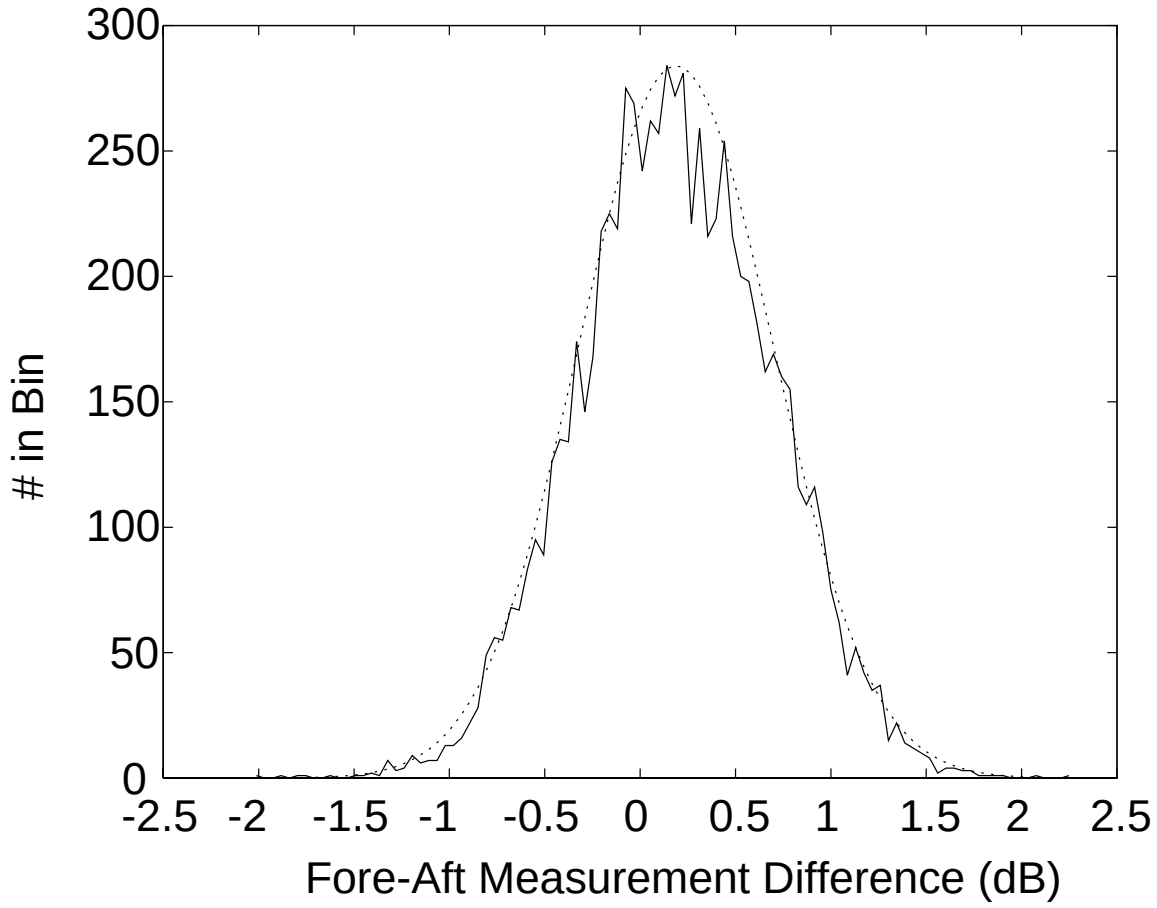


Figure 5.9: Histogram of the difference of the fore and aft beams from the raw scatterometer data for region I10. The dotted line is a Gaussian curve based on the mean and standard deviation of the actual histogram.

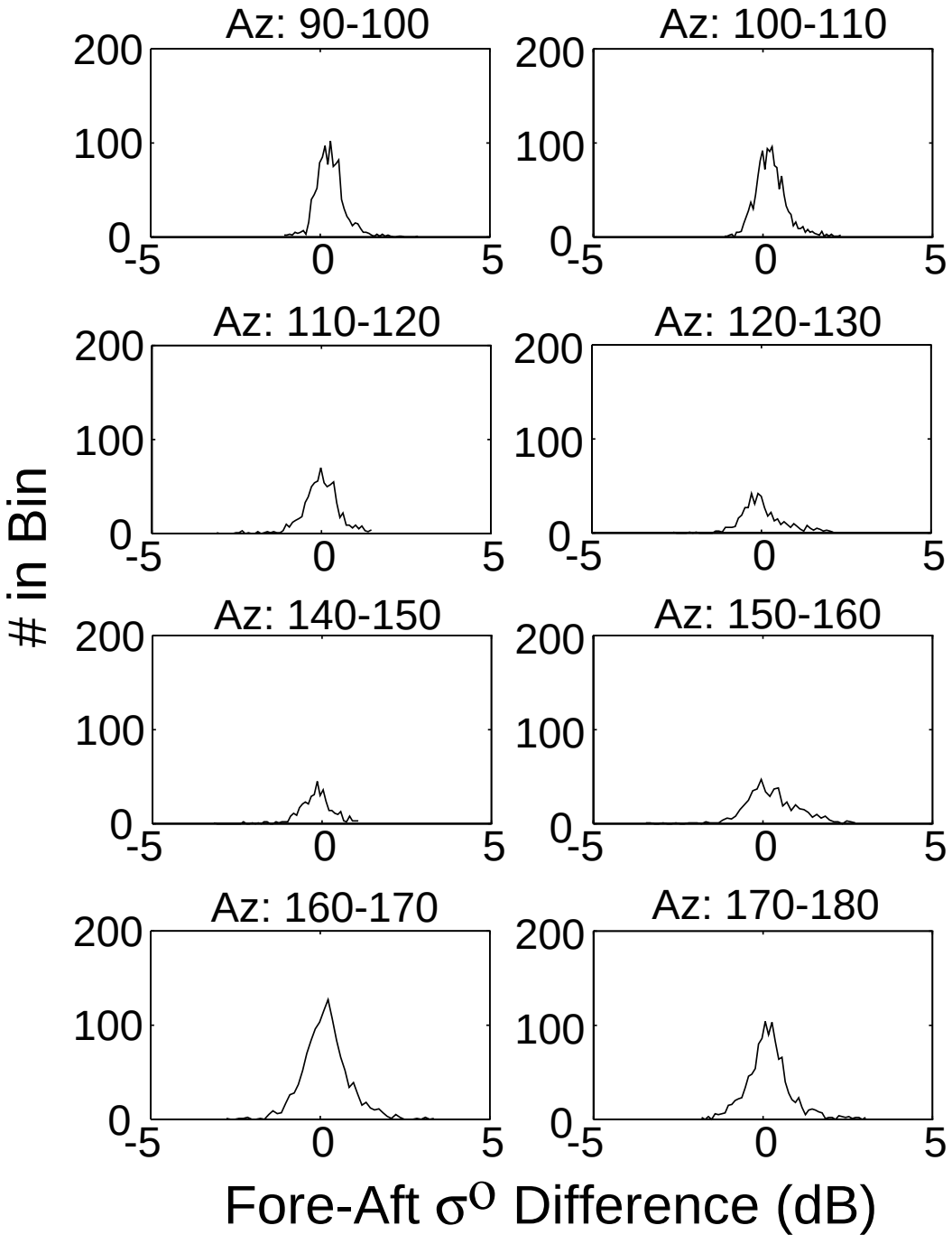


Figure 5.10: Histogram of the difference of the fore and aft beams from the sea ice region I12. These graphs are over small azimuth angle ranges as noted over each graph and illustrate the stability of the histogram mean in azimuth. If the data set does not have measurements in an azimuth range, that azimuth range is not illustrated in the figure.

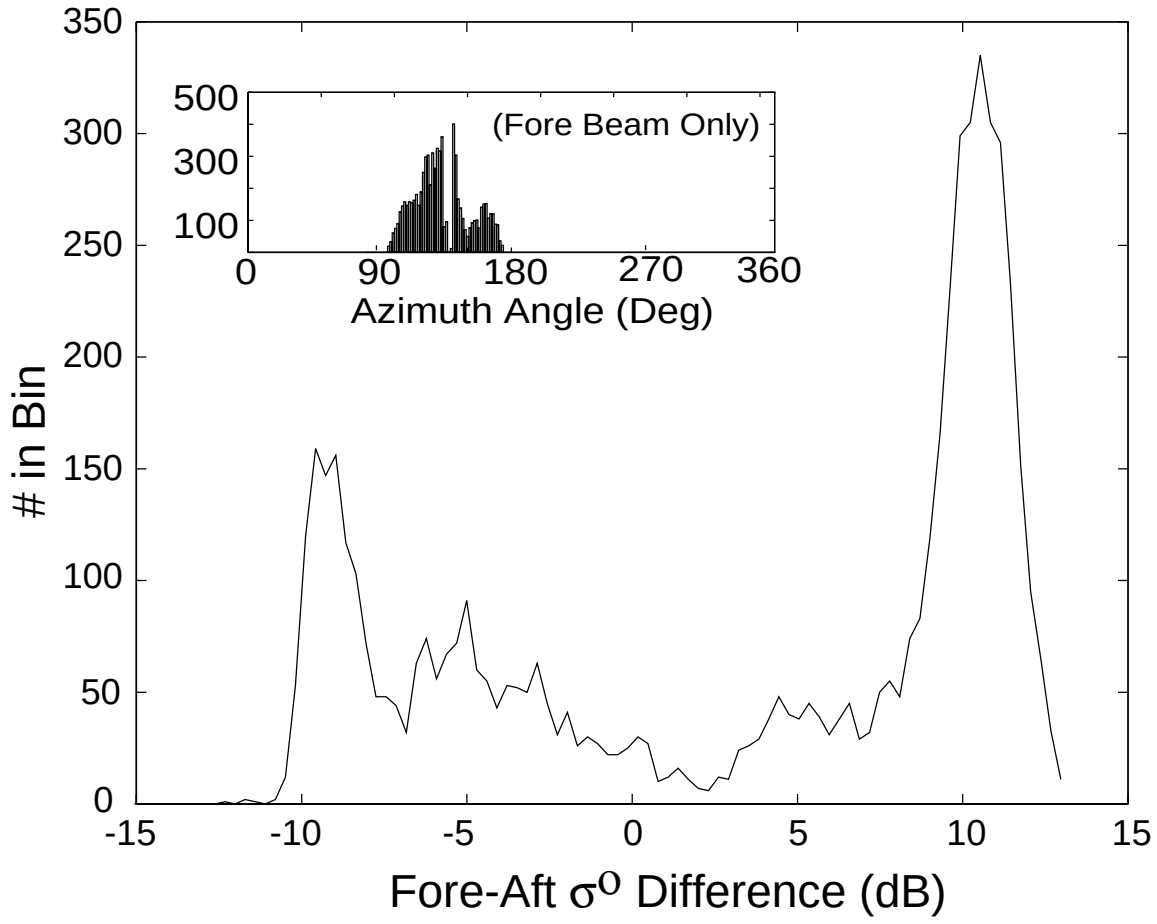


Figure 5.11: Histogram of the difference of the fore and aft beams for scatterometer data from glacial region G2. The inset is a histogram of azimuth angles from the fore beam only to illustrate the diversity of azimuth angles in this study region.

The difference between the fore and aft beam measurements may be modeled by:

$$D = (\sigma_F^o + N_F) - (\sigma_A^o + N_A) \tag{5.6}$$

where N_F and N_A are independent Gaussian noise terms associated with the fore and aft beam measurements, respectively. We can predict D over an azimuthally isotropic medium. For an azimuthally isotropic medium, $\sigma_F^o - \sigma_A^o = 0$ since the incidence angles for each measurement are equal, and D becomes the difference of the noise terms:

$$D = N_F - N_A. \tag{5.7}$$

Since the sum of two independent Gaussian random variables is a Gaussian random variable, we expect the fore-aft beam measurement difference to be a Gaussian random variable. Assuming the noise terms are zero mean, histogram of fore-aft beam measurement differences will be a zero mean Gaussian distribution if the observed surface is isotropic in azimuth.

Fig. 5.9 shows an example histogram of the sea-ice regions studied. The histogram is over all available azimuth angles. The dotted Gaussian curve fitted to the example sea-ice azimuth angle histogram is based on the mean and standard deviation of the data in the histogram. As predicted, the data is Gaussian with a nearly zero mean (For all the regions studied, the mean is less than 0.2 dB in every case.) A Chi-Squared statistical test was applied to the data to verify the Gaussian nature of the curve. Because the number of data points was small for each curve, the probability output from the Chi-Squared test was very small and the test was inconclusive. However, even though the probability was small, it was large compared to the probability that the skewed curves were Gaussian, so we accept this test as indicative of the Gaussian nature of the isotropic surface response. Fig. 5.10 shows similar histograms for the same region but over 10° azimuth angle ranges. Note that the mean remains constant in each azimuth angle bin, which is not the case over land ice as illustrated in Fig. 5.11 which shows a histogram from a land ice region. Over the bulk of the data in this land ice region, there is a non-zero difference between the

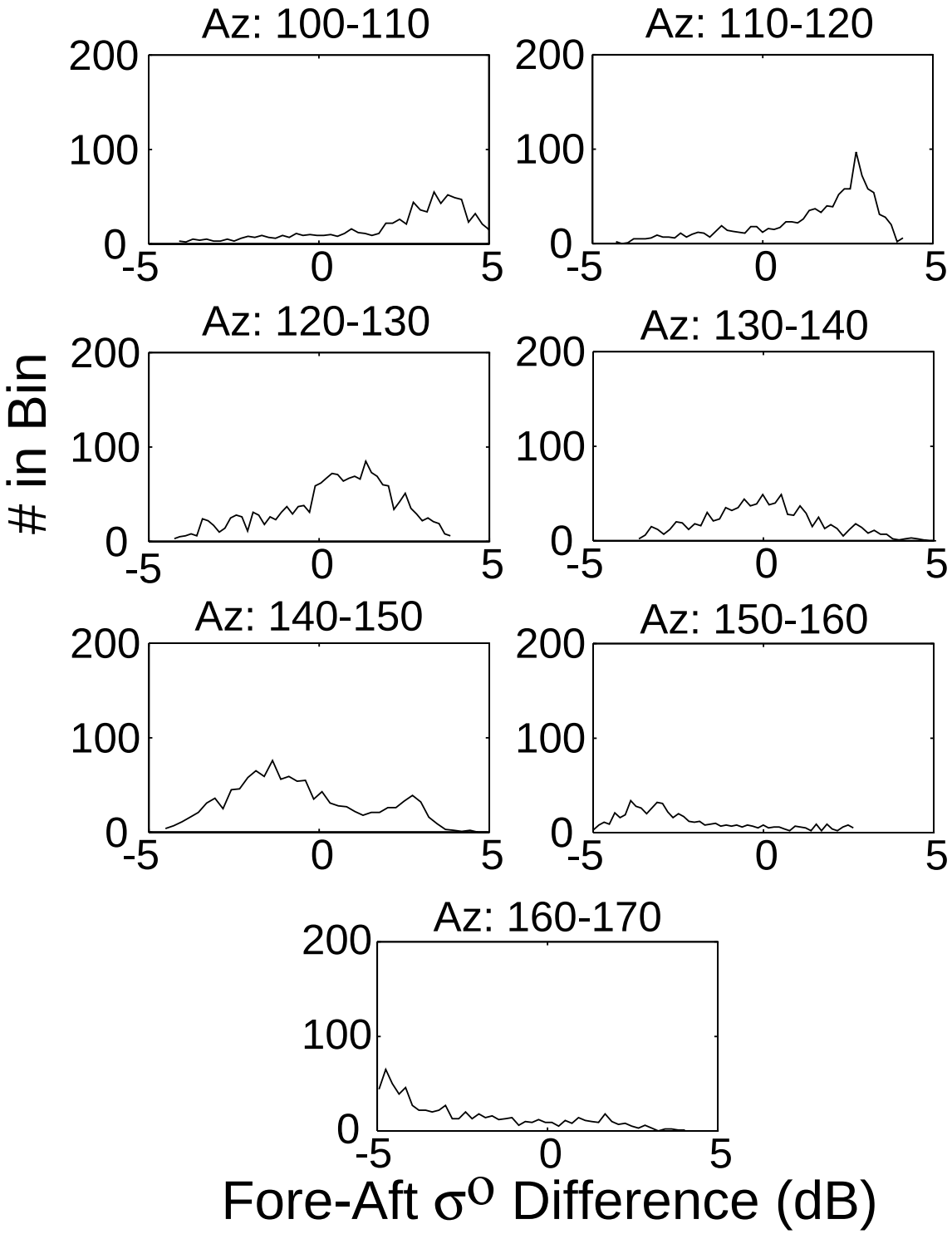


Figure 5.12: Histogram of the difference of the fore and aft beams for scatterometer data from glacial region G2 over 10 degree azimuth ranges from 80° to 180° (fore beam only). Note the progression of the mean from positive to negative as azimuth increases.

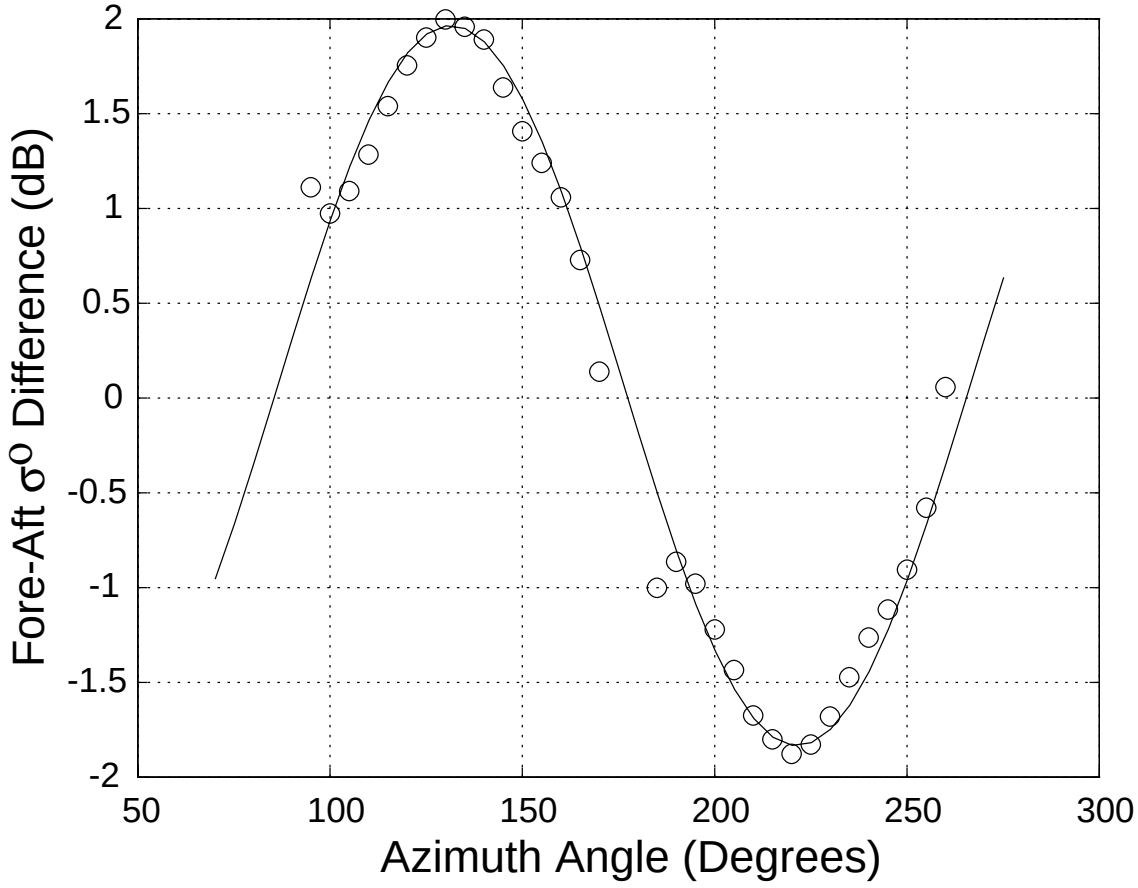


Figure 5.13: A plot of the fore-aft beam measurement difference over glacial ice region G2. Note the double sinusoidal function of the difference in azimuth angle, very similar to the geophysical model function used for predicting near-surface ocean winds from scatterometer data.

fore and aft beams, suggesting that there is modulation in the ice sheet microwave signature. Fig. 5.12 shows the same data but with histograms over 10° incidence angle ranges (Compare with Fig. 5.10). Note the progression of mean from positive at lower azimuth to negative at higher azimuth. A plot of the mean fore-aft difference for 5° azimuth bins versus azimuth angle is shown in Fig. 5.13. The double sinusoid plot is similar to the Ku-band results of *Remy et al.* [45], and is very similar to the geophysical model function used for retrieving near surface ocean winds illustrated in [55].

5.4.3 Marginal Ice Zone

Of special concern are regions of sea ice near the periphery of the sea-ice pack. Unlike areas of open ocean, we do not expect gravity or capillary waves to form and create azimuthal modulation in the microwave signature. But because long wavelength swell waves can propagate deep into the sea-ice pack, a surface slope may be created that is sufficient to induce azimuthal modulation in σ^o. However, given that these swell waves can have wavelengths on the order of several kilometers or more and amplitudes on the order of a few meters, the change in surface slope is quite small. Fourteen regions near the edge of the sea-ice pack were selected between JD 102 and JD 204 1993 and reviewed for azimuthal modulation using a fore-aft difference analysis. Using the previously described methodology it is determined that azimuth modulation in all regions studied was less than 1 dB. Further, in the majority (11 of 14) the azimuth modulation is less than 0.2 dB. We conclude that on the scale of the scatterometer measurements, no azimuthal modulation is visible over Antarctic sea ice.

5.5 Summary

A detailed analysis of C-band ERS-1 scatterometer data reveals that there is no significant azimuthal modulation (less than 1 dB) evident in data taken over Antarctic sea ice at the scale of the ERS-1 scatterometer measurements (nominally 50km). The consistency of the analysis methods used in this study was established by

comparing sea ice results with land ice sheet results. Using the same methodologies for both land and sea ice, azimuthal modulation is shown to be negligible in the sea ice regions studied. In contrast, land ice study regions exhibit significant azimuthal modulation. This result is consistent with the results of previous studies of azimuthal modulation over land ice sheets. Areas in the Marginal Ice Zone, where long wavelength swell waves can penetrate deep into the ice pack, also displayed negligible levels of azimuthal modulation in the ERS-1 scatterometer measurements.

Chapter 6

Application of Enhanced Resolution Imagery to Southern Ocean Sea Ice

6.1 Introduction

The power of the scatterometer for polar remote sensing is two fold: First, the scatterometer provides rapid, repeat coverage of the earth. With its wide 500 km swath and on-board data storage capabilities, the scatterometer is capable of continuously collecting data over the whole globe, which means frequent repeat coverage of individual areas. Second, the scatterometer takes measurements at a variety of incidence angles. This provides additional data which may be used to develop a relationship between the backscatter power and incidence angle, which in turn reveals information about the surface. As will be shown, this allows the separation of ice classes that would not otherwise be distinguishable.

As established in the previous chapters, scatterometer data can be enhanced with the SIR algorithm to a medium scale resolution that makes it useful for scientific studies. Coupled with the higher resolution SAR imagery, processed scatterometer imagery provides a large scale context for interpreting the individual SAR images. SIR images also offer insight into sea ice characteristics and interannual variability of the sea ice. SIR imagery also provides a time series of images illustrating basin wide temporal change and evolution of the polar sea ice pack.

In this Chapter, we take advantage of one of the unique properties of the scatterometer data by developing the measure of azimuthal anisotropy, described in detail in the first section. This new measure is then used to separate ice classes

otherwise inseparable using $\mathcal{A}$ and $\mathcal{B}$ images alone. Six sea ice types are defined and described, and these classes, combined with *in situ* measurements of the $\mathcal{A}$, $\mathcal{B}$ and azimuthal modulation characteristics of each class are used to develop a SIR image classifier. SIR is then applied to ERS-1 scatterometer data to classify Southern Ocean sea ice. The time evolution of the sea ice pack over a three year time period is shown. The following introductory sections provide background on the scattering characteristics of sea ice and introduce the new measure of the surface conditions based on the minimal anisotropic response of the sea ice surface.

6.2 A Measure of Azimuthal Anisotropy

In Chapter 5, we concluded that the anisotropic response of the sea ice pack was negligible for purposes of image reconstruction. However, in some regions of the ice pack, rapid changes in space and time may result in a low level anisotropic backscatter response in azimuth. While these regions have a minimal affect on SIR image reconstruction, they reveal some important characteristics about the surface. For instance, wave action in the marginal ice zone (MIZ) contributes to a small anisotropic azimuth response from the surface. As will be shown, this anisotropic response delineates MIZ ice and multiyear ice, two ice types otherwise indistinguishable from backscatter alone.

We now introduce a measure of the anisotropic azimuth response of the surface originally developed in [56] for ERS-1, the normalized measure of anisotropy, given by

$$\frac{|\sigma_F^o - \sigma_A^o|}{|\sigma_F^o + \sigma_A^o|} \tag{6.1}$$

where σ_F^o is the backscatter received at the fore beam, and σ_A^o is the backscatter response from the same area received at the aft beam. Recall that the fore and aft beams for ERS-1 are 90 degrees apart, and that as noted in Chapter 5, the azimuthal anisotropy has a double sinusoidal response (see Figure 5.13). Azimuthal modulation will be evident in the difference between the fore and aft beams of the ERS-1 scatterometer. Thus the greater the value of Eq. (6.1), the greater the azimuthal

modulation. Because open ocean measurements typically have a large normalized measures of anisotropy and measurements over sea ice typically do not, Eq. (6.1) is used in [56] to separate open ocean measurements from sea ice measurements to generate near surface wind maps from only open ocean measurements.

It is possible that one particular fore-aft measurement pair may not reveal azimuthal modulation even if it is present. Refering to Figure 5.13, note that it is possible to have two measurements 90 degrees apart over an azimuthally anisotropic surface and get the same value at each measurement. To ensure that the surface anisotropy is correctly evaluated, multiple fore-aft measurements are made at various azimuth angles over the same area. For a surface that is azimuthally anisotropic, there will be a range of values calculated from Eq. (6.1), while for an isotropic surface, there will be little variance in the values. The standard deviation of all values from Eq. (6.1) for a particular area constitute a measure of azimuthal anisotropy.

We now adopt the standard deviation of the normalized measure of anisotropy of sea ice defined in Eq. (6.1) as an additional measure of sea ice surface conditions and create an image similar in pixel size and scale to the $\mathcal{A}$ and $\mathcal{B}$ images created by SIR for ERS-1. We call this new image the STD image, in reference to it being derived from the standard deviation of the azimuthal anisotropy. For brevity, references to the "standard deviation of anisotropy" are simply be referred to as the "standard deviation" or STD. For the current application, an image of the normalized standard deviation of the anisotropy is created from the scatterometer measurements taken over the course of 6 days, the imaging interval of the SIR algorithm. For each fore-aft measurement pair, the normalized measure of anisotropy in Eq. (6.1) is back projected onto a pixel grid matching the pixel grid used for the SIR images (for ERS-1, each pixel represents an 8.9 km by 8.9 km square and the image is 960 pixels square), and statistics of the normalized anisotropy measure are kept for each pixel the measure touches. The normalization in Eq. (6.1) eliminates the need to consider the aperture weighting because the aperture weight for the pixel factors out in the normalization. However, the size and shape of the aperture is used to generate the STD image. This process is repeated for each measurement pair in the data set.

The standard deviation of the normalized measures of anisotropy for each pixel is computed using the standard equation:

$$\text{stddev} = \sqrt{\frac{1}{N-1}\left(K - \frac{S^2}{N}\right)} \tag{6.2}$$

where N is the count (number of fore-aft pair measurements that touch a given pixel), K is the sum of the squares of the normalized measure of anisotropy and S is the sum of the normalized measure of anisotropy. The result is an image that reflects the average anisotropic response of the sea ice surface over the imaging interval. We now include the STD image with the $\mathcal{A}$ and $\mathcal{B}$ SIR images for use with a classification algorithm. Note that while the $\mathcal{A}$ and $\mathcal{B}$ images are actually resolution enhanced as shown in previous chapters, the simple method of projecting the standard deviation of anisotropy measurements onto a finer grid does not give greatly improved resolution. In fact, the resolution is the same as the AVE algorithm introduced in [13], which produces images from a simple averaging process. However, it provides adequate results for a preliminary classification algorithm. Further development of the STD image is a topic for more extended research.

Figure 6.1 shows an example STD image from JD 126 (5 May) 1992. In this and subsequent images, the $\mathcal{A}$,$\mathcal{B}$ and STD images are masked with an ocean mask that removes the open ocean pixels and a land mask that removes the Antarctic continent from the images. Both masks are based on Special Sensor Microwave Imager (SSM/I) passive microwave data. Using the zero ice concentration contour in an SSM/I image created using SSM/I data from the same days as the SIR/STD image set, a sea ice mask is created [57]. The mask used to remove the land pixels from the imagery is the standard 25 km land mask provided with the SSM/I data set.

The example STD image in Figure 6.1 is annotated indicating two regions displaying higher standard deviation than the surrounding areas. Region A is an area in the Weddell Sea where warm water upwelling can cause large open ocean polynyas during the early winter expansion of the ice pack [58]. This higher than normal standard deviation is caused by the warm water changing the sea ice surface and ice formation characteristics in this region. Region B is a known coastal polynya

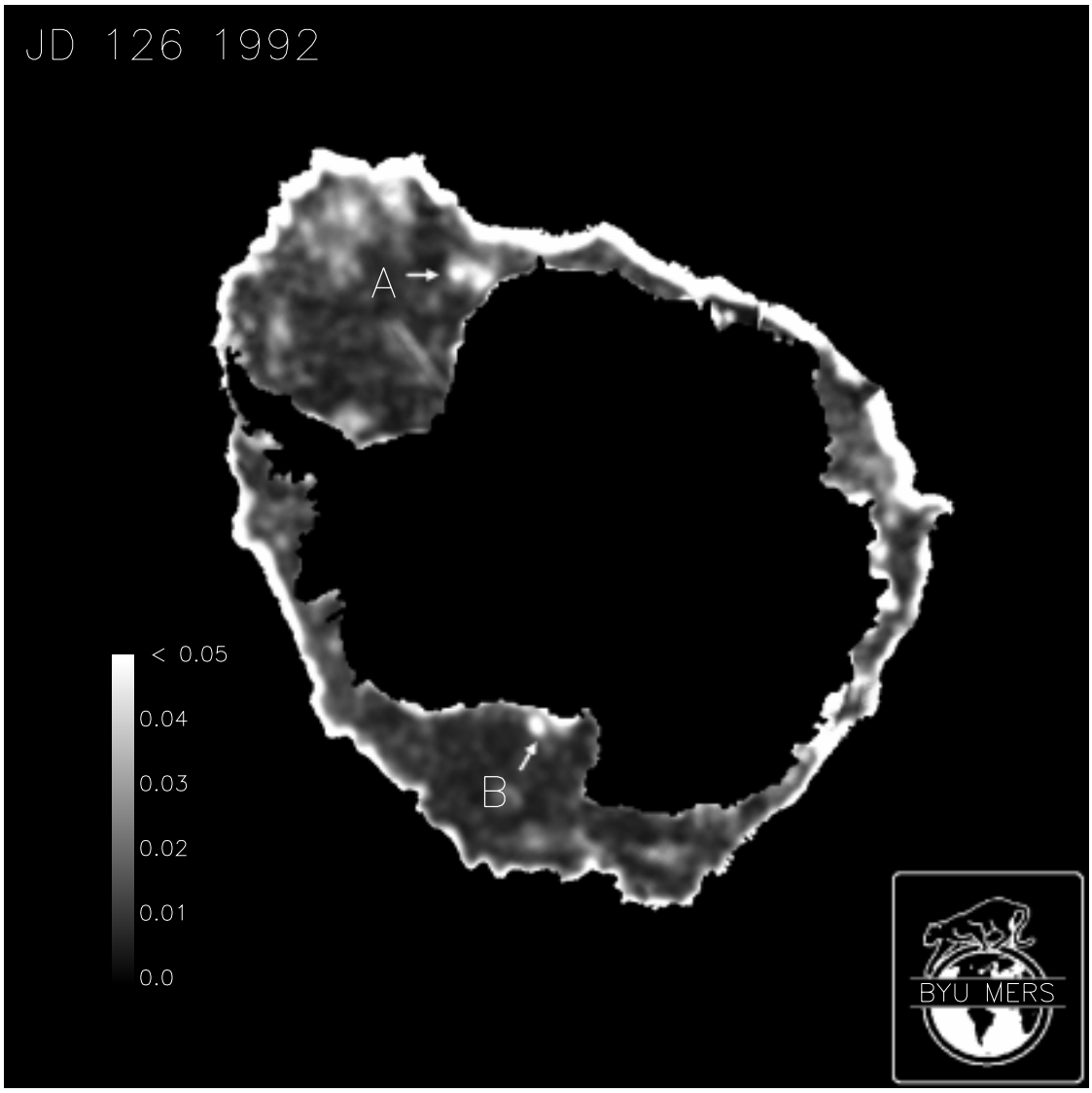


Figure 6.1: An example STD image from JD 126 1992. The two annotated regions display higher than normal standard deviation. The first region, A, indicates a known open ocean polynya caused by warm water upwelling. The second region, B, indicates the regions off the Ross ice shelf where the ice pack has pulled away from the continent and allowed new ice to form in the gap.

off the Ross ice shelf. Here, the ice pack is drawn away from the continent by dynamic forcing and new ice is allowed to form in the exposed open water. Where the previous anisotropic response was due to warm water upwelling, coastal polynyas are scientifically interesting because they produce cold water downwellings: Super cooled, highly saline water is ejected into the ocean as ice forms, and eventually this cold water mass falls to the ocean floor, giving rise to cold bottom currents and trapping and storing greenhouse gases in deep ocean currents [40].

In the following sections, we examine the coupling of STD images to the $\mathcal{A}$ and $\mathcal{B}$ images to create more accurate sea ice classifiers. In the next section, we discuss ice types and the expected backscatter response from the various types. Following this, an overview of work done on a one dimensional $\mathcal{A}$ image classifier is presented, and finally the fusion of $\mathcal{A}$, $\mathcal{B}$ and STD images to produce classified imagery is presented.

6.3 Sea Ice Type Definitions and Scattering Characteristics

Every year, the Antarctic seasonal ice pack undergoes a regular cycle of growth and decay. We now define six ice classes used to characterize the Southern Ocean sea ice through this annual cycle. For each class, a summary of the general characteristics of the $\mathcal{A}$, $\mathcal{B}$ and STD values is given.

Multiyear Sea Ice

Multiyear ice is ice that survives the summer melt, so this class of sea ice is typically very thick. Older ice is less saline than younger ice types because of brine drainage from the freeboard portion of the ice over time, significantly changing the scattering characteristics of the sea ice. Also, snow accumulation and the seasonal melt/freeze cycle that the surface of the ice undergoes in the transition from summer into winter both have a substantial affect on the backscatter characteristics of this class of ice. The layering effects of the accumulated snow and ice lenses formed by refreezing melt ponds increase the volume (non-specular) scattering from the ice and these cause a generally high backscatter return at all incidence angles [57]. Thus, the

$\mathcal{A}$ value is high. These same affects also cause a strong dependence on incidence angle, and we expect the $\mathcal{B}$ value to be very low for this ice class. Due to the smoothing of potentially large scale roughness of multiyear ice by snow layering, the standard deviation of anisotropy is low.

Nilas

Sea ice generally forms first as small crystals that form in the open ocean as the surface layer cools [1]. If the conditions are calm, the crystals begin to coalesce forming a thin (< 10 cm thick), elastic layer called nilas [1]. During sea ice pack expansion in the winter, the areas of largest ice production are at the sea ice edge and near the large coastal polynyas. Near the sea ice edge, wave action usually does not allow for calm conditions that favor the formation of nilas, so large expanses of nilas have not been observed since most expeditions concentrate on the highly dynamic marginal ice zone (MIZ) [57]. Nilas most likely forms in areas of long term divergence, typical of the coastal polynyas where the pack continually pulls out from the continent. Even then, the nilas is short lived, quickly thickening and hardening into young first year ice. Nilas is characterized by a highly wet, saline surface. Little if any volume scattering contributes to the total backscatter return. The $\mathcal{A}$ value is expected to be very low, with a strong dependence on incidence angle. The $\mathcal{B}$ value is low, and the relatively smooth, unbroken surface indicates a low standard deviation value.

MIZ/Pancakes

In most cases, new ice is forming in areas where wind and ocean waves push the ice crystals together. This action forms pancake ice (Illustrated in Figure 1.1) as the ice crystals are pushed into groups. The pancakes are characterized by the pronounced lip caused by wind driving the pancakes into each other, the force pushing the soft ice up into a ridge. Eventually, the cooling temperatures cause the pancake formations to congeal into a solid pack of ice. Meanwhile, ocean waves are capable of traveling in the ice pack in these regions. This type of ice is prevalent at

the edge of the ice pack, making this type of ice characteristic in the MIZ. *In situ* measurements have shown that regions of pancake ice have a very high backscatter return, similar to multiyear ice [50,57]. However, unlike multiyear ice, pancake ice has a much higher $\mathcal{B}$ value and, due to the wave actions and small scale ridges prevalent in these regions, a much higher standard deviation of anisotropy.

Smooth First-year Sea Ice

As pancake or nilas ice congeals and thickens, it evolves into smooth (undeformed) first year ice, with a thickness for this class of over 50 cm. The backscatter from this ice class is typically high because of the increased contribution of volume scattering to the return as brine begins to drain from the free board portion of the ice. The return is smaller than multiyear ice due to the less complex layering and ice lenses. The scattering is still specular, however, because of the relatively smooth surface, and the $\mathcal{B}$ value is quite low. The large-scale uniformity of the roughness in these regions indicates a very isotropic response, so the standard deviation of anisotropy is small.

Rough First-year Sea Ice

Because the sea ice pack is in motion due to wind and ocean currents pushing the ice on the ocean surface, ridging and deformation of the ice pack is inevitable. Rough first year ice is deformed and generally thicker than smooth first year ice due to piling of the ice pack during deformation. The rougher surface contributes to more specular scattering loss and the backscatter from this class is lower than smooth first year ice. However, the increased surface roughness and ridging increases the specular scattering components to increase at various incidence angles making the $\mathcal{B}$ value lower. Again, the large scale uniformity of the surface roughness indicates a low standard deviation of anisotropy.

Ice Type	$\mathcal{A}$ Value	$\mathcal{B}$ Value	STD Value
Icebergs	Very High	Mixed	Mixed
Multiyear	High	Low	Low
MIZ/Pancakes	High	High	High
Rough First-year	Mid	Mid	Mid
Smooth First-year	Mid	High	Low
Nilas	Low	High	Low

Table 6.1: A summary of $\mathcal{A}$, $\mathcal{B}$ and STD values for the six sea ice classes. Note that multiyear ice and MIZ/pancake ice have different $\mathcal{B}$ and STD characteristics compared to the identical $\mathcal{A}$ value ranges for the two classes.

Icebergs

Icebergs constitute the last ice class. Iceberg are large, thick glacial ice that has broken or calved off the Antarctic glacial ice sheet. These large ice floes have a much lower salinity than even old multiyear ice, and have a much larger backscatter coefficient. This class has the brightest $\mathcal{A}$ value. The $\mathcal{B}$ value depend on the orientation of the iceberg surface. Tabular bergs are isotropic simply because the surface has no slope. Most bergs have some surface slope so the backscatter depends on incidence angle and azimuth as well, resulting a large variance and higher mean in the $\mathcal{B}$ and STD values.

Table 6.1 shows a summary of the expected $\mathcal{A}$, $\mathcal{B}$ and STD values for the various ice types outlined above. In the following section, more detailed values are added to this table to create a classification scheme for SIR imagery.

6.3.1 Sea Ice Field Measurements

Figure 6.2 shows the C-band (4.3 GHz) scatterometer measurements made during a 1992 expedition to the Southern Ocean by Dr. Mark Drinkwater of the Jet Propulsion Laboratory (JPL) (Figure courtesy of Dr. Mark Drinkwater). Graph (a) in Figure 6.2 is of white ice, a category not described above, is similar in some respects to nilas but is ignored for this study. In each graph, σ^o versus incidence angle is plotted for both VV and VH polarization with the bar indicating the standard deviation of

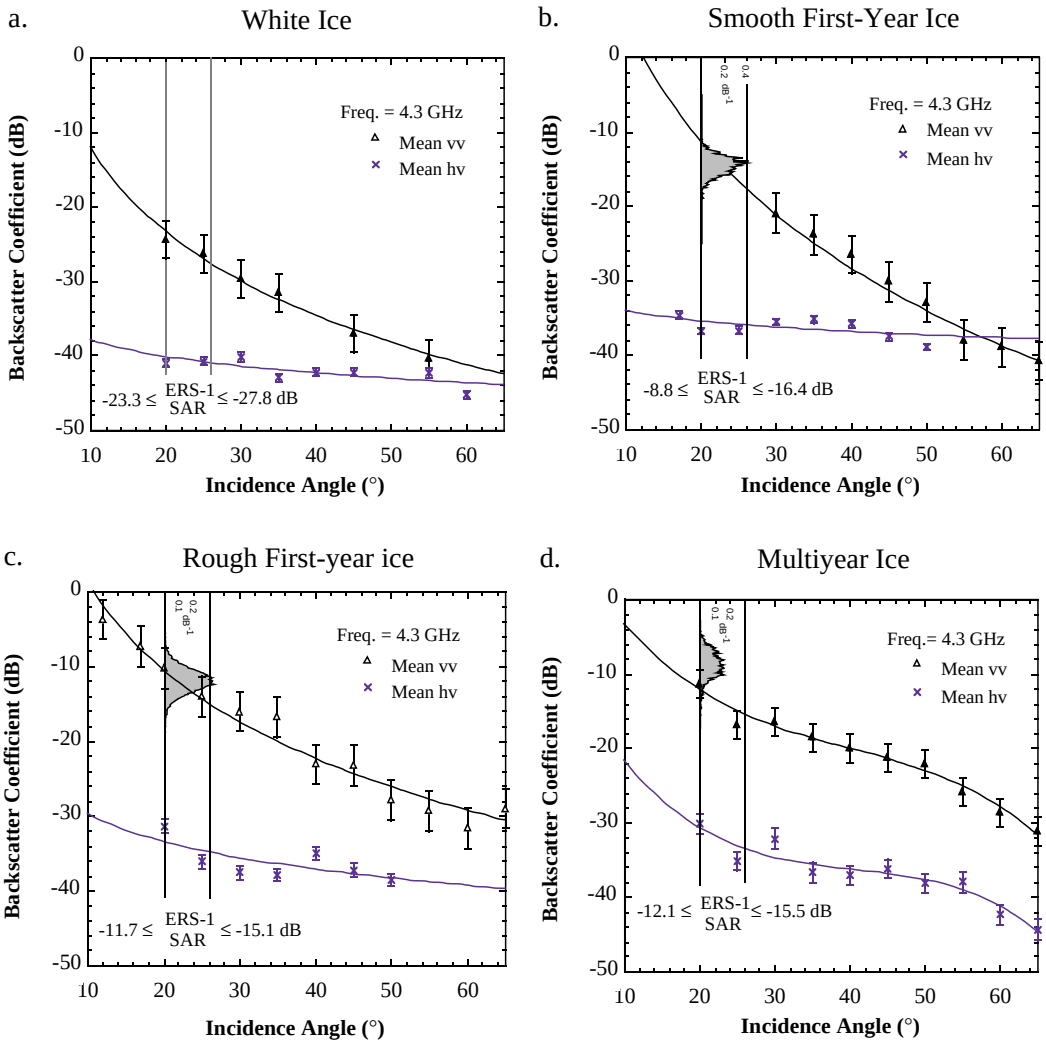


Figure 6.2: Each plot in this figure shows the σ^o versus incidence angle relationship for three different ice types considered in this study. The fourth plot, (a) White Ice, is not discussed here but is similar to nilas. Note the varying levels of backscatter return and the varying dependence on incidence angle for each ice type. This figure courtesy of Dr. Mark Drinkwater, JPL [57].

the measurements made at that incidence angle. The ERS-1 scatterometer is limited to VV polarization, so we ignore the VH plot. The shaded plots turned 90 degrees clockwise represent a distribution of σ^o for the ERS-1 SAR measurements for that ice type. The ERS-1 SAR is limited to a narrow range of incidence angles (20° to 26° , approximately) and so the histogram of SAR measurements is plotted at this point on the incidence angle axis. Notice how in each graph the distribution of SAR σ^o generally matches the curve fit to the scatterometer data.

Now compare the graphs in Fig 6.2 to Table 6.1 showing the characteristic $\mathcal{A}$, $\mathcal{B}$ and STD values for the various sea ice types. For Smooth First Year ice in Figure 6.2(b), the backscatter is relatively low compared to the other types. Also, the variation with incidence angle is steep, indicating a high $\mathcal{B}$ value. Rough First Year ice in Figure 6.2(c) is characterized by a higher backscatter, but less steep dependence on incidence angle. Finally, Multiyear ice in Figure 6.2(d) shows the highest backscatter and the least dependence on incidence angles. This last graph also clearly shows the higher order nature of the incidence angle dependence discussed in Section 4.4.3. Notice how the graph is not actually linear near the extreme incidence angle values, but that it is linear in the mid range incidence angles where the majority of measurements are made. Also notice how each class has a different level of cubic dependence, a fact which prevented a global application of higher order parameters to the data model in Section 4.4.3.

6.4 $\mathcal{A}$ Image Classification

As part of ongoing research on the utility of ERS-1 SIR imagery, a simple, one-dimensional classifier has been developed using only normalized backscatter values [57]. The graph in Figure 6.3 summarizes the SAR, ship borne scatterometer and ERS-1 scatterometer data used to develop this classifier. SAR imagery was taken where the surface conditions were known. This was coupled with *in situ* measurements taken from a ship mounted C-band scatterometer and corresponding ERS-1 scatterometer SIR $\mathcal{A}$ image values (corresponding backscatter at an incidence angle of 40°) to create the graph in Figure 6.3. The five main ice types (excluding nilas)

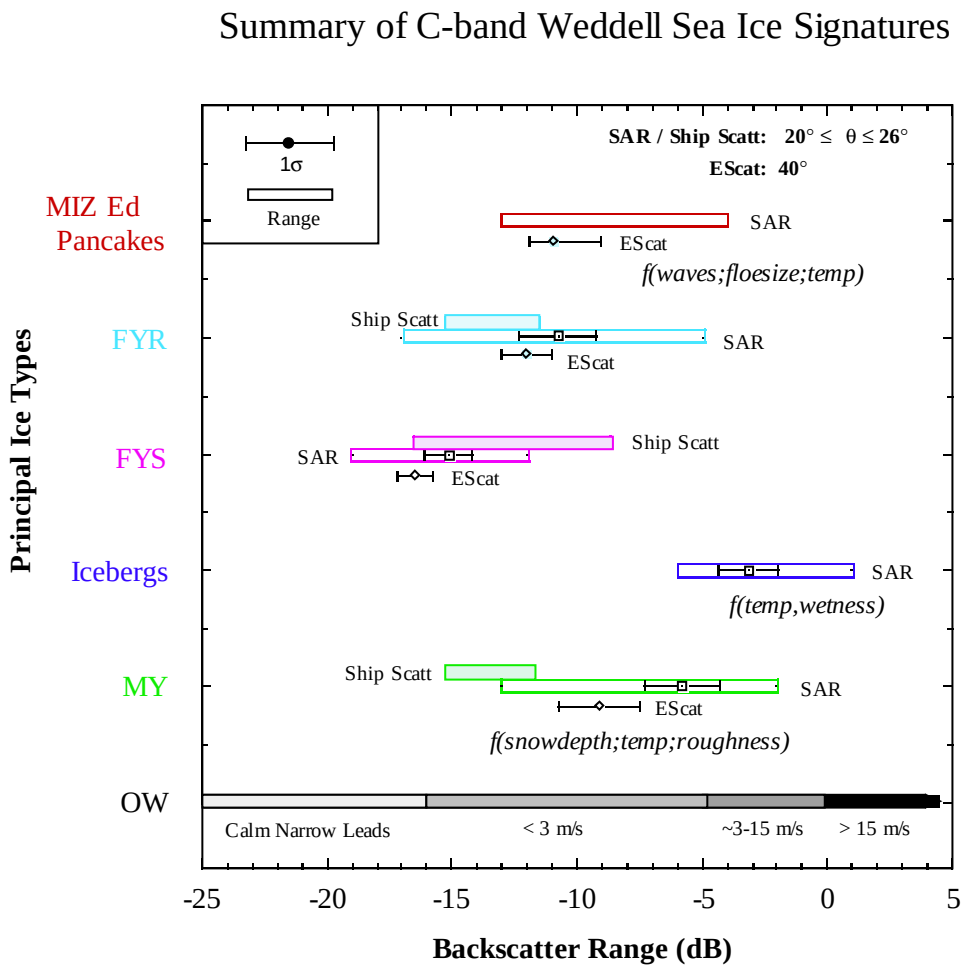


Figure 6.3: This figure illustrates the relationship between the ship-borne scatterometer (Ship Scatt), the ERS-1 SAR (SAR) and ERS-1 scatterometer (EScat) backscatter characteristics for the 5 ice types considered. Open Water (OW) characteristics are also shown. Note that the MIZ/Pancake ice class overlaps the Multi-Year (MY) ice class, making separation of these classes by backscatter alone difficult. For several ice types, the factors influencing the surface backscatter are shown as functionals: $f(\cdot)$. Figure courtesy of Dr. Mark Drinkwater, JPL [57].

Ice Type	Backscatter Range
Icebergs	$-6.0 \leq \sigma^o < 0.0$
Multiyear	$-11.0 \leq \sigma^o < -6.0$
MIZ/Pancakes	$-11.0 \leq \sigma^o < -6.0$
Rough First-year	$-14.0 \leq \sigma^o < -11.0$
Smooth First-year	$-20.0 \leq \sigma^o < -14.0$
Nilas	$-32.0 \leq \sigma^o < -20.0$

Table 6.2: A summary of winter $\mathcal{A}$ value classes. Note that multiyear ice and MIZ/pancake ice have the same dB range, making them inseparable with only the $\mathcal{A}$ image. The ice types are summarized in Section 6.3. From [57].

discussed in the previous sections are shown along with an open water (OW) class showing the relationship between backscatter from open water and wind speed.

Note in Figure 6.3 how Multiyear ice (MY) has almost the same signature as MIZ/Pancake ice, indicating that from backscatter alone it is not be possible to distinguish between Multiyear and MIZ/Pancake ice [57]. However, the other classes, First Year Smooth (FYS), First Year Rough (FYR) and Icebergs are generally distinguishable using normalized backscatter alone. Based on the *in situ* data available, a simple threshold classifier from [57] is presented in Table 6.2. The Multiyear and MIZ classes in this classifier have identical backscatter ranges, making distinguishing these two classes from normalized backscatter alone impossible.

6.5 Separation of MY and MIZ Classes

We desire to separate the Multiyear (MY) and MIZ/Pancake (MIZ) classes. Recall that summarized in Table 6.1 are general $\mathcal{A}$, $\mathcal{B}$ image and STD image characteristics of the various sea ice types, and that in Table 6.2 specific ranges for the $\mathcal{A}$ values for each class are given. Because all but MY and MIZ ice are separable by $\mathcal{A}$ value, the $\mathcal{B}$ value and STD value are only used at this point for the case of separating MY ice from MIZ ice. Since the $\mathcal{B}$ and STD values for MIZ are higher than for MY ice, we determine a threshold for an extended classification algorithm from histograms of $\mathcal{B}$ and STD values shown in Figure 6.4 and 6.5 respectively. These

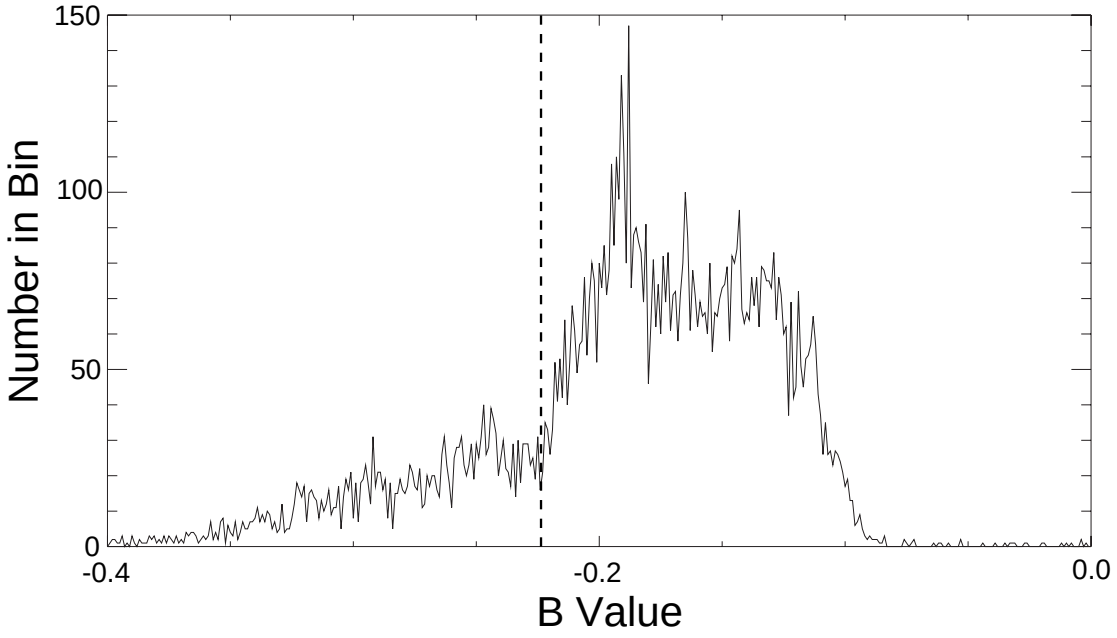


Figure 6.4: A histogram of the $\mathcal{B}$ values for the MY/MIZ ice class from JD 168 1992. The dotted line indicates the natural split in the data and provides a threshold between two populations representing MY (low $\mathcal{B}$) and MIZ ice (high $\mathcal{B}$).

histograms are made using all the $\mathcal{B}$ and STD pixels in the MY/MIZ classes created with the one dimensional classification algorithm given in Table 6.2.

In Figure 6.4, the histogram shows two populations divided by the dotted line. Per Table 6.1, the lower $\mathcal{B}$ values correspond to the Multiyear ice, with the expectation of less MY ice than MIZ ice confirmed because the distribution favors the higher $\mathcal{B}$ values of MIZ ice. Figure 6.5 shows a histogram of the STD values for the same data set. This graph has a less well defined split in the populations, but the dotted line on the plot indicates a likely split in the data. There are also values of higher (> 0.1) normalized standard deviation that extend in a tail beyond the graph. These high values correspond to the MIZ ice. The lack of separation in this data is

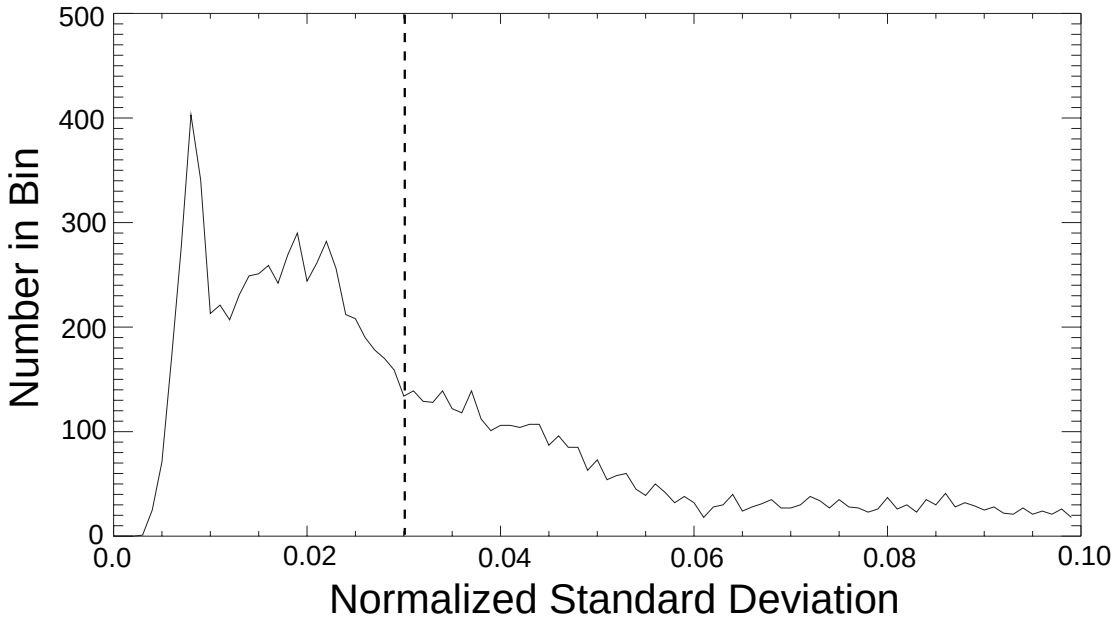


Figure 6.5: A histogram of the normalized standard deviation (STD) values for JD 168 1992. Unlike Figure 6.4, the populations in this plot are less separated, but the dotted line indicates a probable split in the two populations.

likely the result of the sensitivity of the STD values to either data dropout or too few measurements for a given pixel.

Using the histograms, the threshold for the $\mathcal{B}$ image data is set at -0.2. Values above this threshold are considered MIZ, values below are MY. The threshold for the normalized standard deviation value is set at 0.03, with higher values indication MIZ, lower values indicating MY sea ice. We now apply this modified algorithm to a specific time period over the Southern Ocean for validation.

6.6 Validation Using 1992 *Polarstern* Data

For the Winter Weddell Gyre Study (WWGS) of 1992, the research vessel *Polarstern* was dispatched to enter the sea ice pack and study the surface conditions [50, 59]. During this expedition, a ship borne C-band scatterometer made measurements of the surface conditions of homogeneous ice types, and visual observations of surface conditions were made. This expedition provides some of the few *in situ* observations available for validation of the SIR scatterometer imagery. There are two very important surface observations made during this mission that are used to validate the modified algorithm used in this study. First, just after the ship entered the sea ice pack, a large, persistent field of pancake ice was encountered. This is where the photograph in Figure 1.1 was taken. Second, the ship later skirted around a large patch of rough first-year ice in the middle of the Weddell Sea.

Figure 6.6 shows a SIR image for JD 168 1992 where the simple one dimensional classification using $\mathcal{A}$ values has been applied. This date corresponds to the time that *Polarstern* encountered the persistent pancake ice field in the MIZ. The figure is oriented with the 0 degree longitude at the top of the picture. Note there is no MIZ class represented in the image. All MIZ is classified as MY ice per the simple $\mathcal{A}$ image classifier in Table 6.2. The large red grouping of MY ice at 0 degrees longitude at the ice edge is in the region where *Polarstern* observed the large pancake field just after entering the ice pack [50, 59].

We now apply the multi-dimensional classifier using the $\mathcal{B}$ and STD values to separate the MY and MIZ ice classes. Figure 6.7 shows the resulting classified

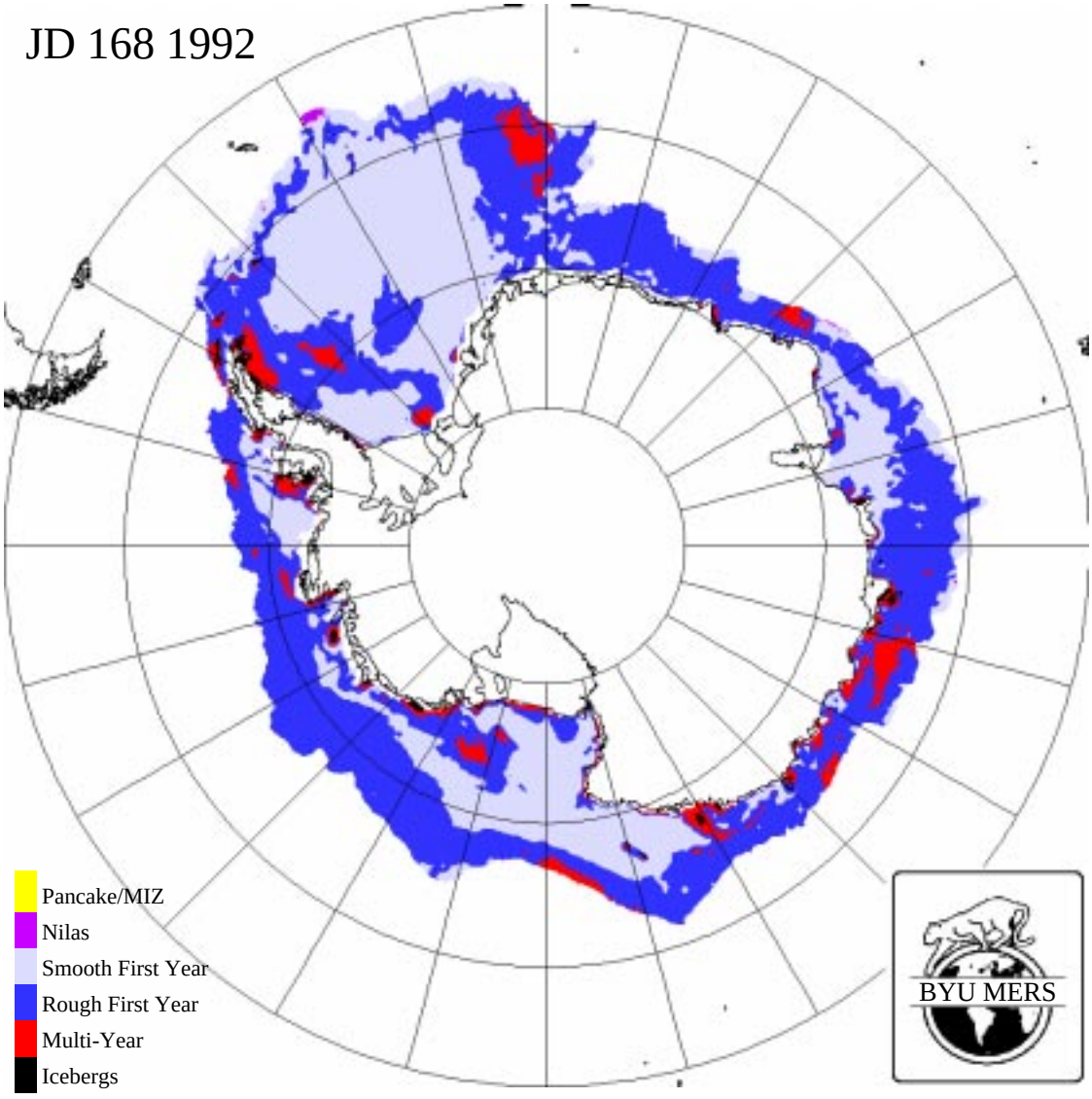


Figure 6.6: A classified image based on the SIR $\mathcal{A}$ image only.

image. Note how the same region near 0 degrees longitude referenced above is now classed as MIZ/Pancake, not MY as in the previous image. This corresponds to the observed region in 1992 that consisted of pancake ice [59]. Other regions in and around the sea ice pack have also been classed as MIZ/Pancake ice, most notably the regions at the ice edge at 45 degrees west and 135 degrees west where we would expect the MIZ/Pancakes because of wave action at the ice edge. The confidence developed in the MY-MIZ class separation because of the *in situ* data available supports the classification of these other regions of the sea ice pack. Also classed as MIZ/Pancakes are the ice production regions off the Ronne-Filchner ice shelf in the Weddell Sea and the ice production area off the Ross ice shelf, where winds off the continental shelf produce MIZ-like conditions in the newly formed sea ice.

While the addition of $\mathcal{B}$ and STD information enhances the one dimensional $\mathcal{A}$ classification, there are indications that some misclassification has occured. Note the small amount of yellow in the typically old ice mass at the tip of the Antarctic Peninsula. Review of images prior to and after this image show that little if any change occurs in the size and shape of this region confirming the supposition that the area is MY ice: MIZ ice tends to rapidly change size and shape as the surface congeals and solidifies into first-year smooth ice while MY sea ice tends to change very slowly. Also, there are still small areas in the largely MIZ ice class at 0 degrees longitude that remain classified as MY ice. Since previous to this multi-dimensional algorithm *all* MIZ ice was classified as MY, the separation of the majority of pixels in these two classes is considered a major improvement over the one dimensional algorithm. Some misclassification is expected due to the significant overlap of the populations in the $\mathcal{B}$ and STD value histograms shown in Figures 6.4 and 6.5.

Another validation of the algorithm concerns the large dark blue mass near the center of the Weddell Sea in Figure 6.7. Records from the 1992 *Polarstern* expedition show that the ship course was modified to the north to avoid a large mass of rough first-year ice in the center of the Weddell Sea [57] [59]. This feature of the ice pack is correctly classified in the SIR imagery.

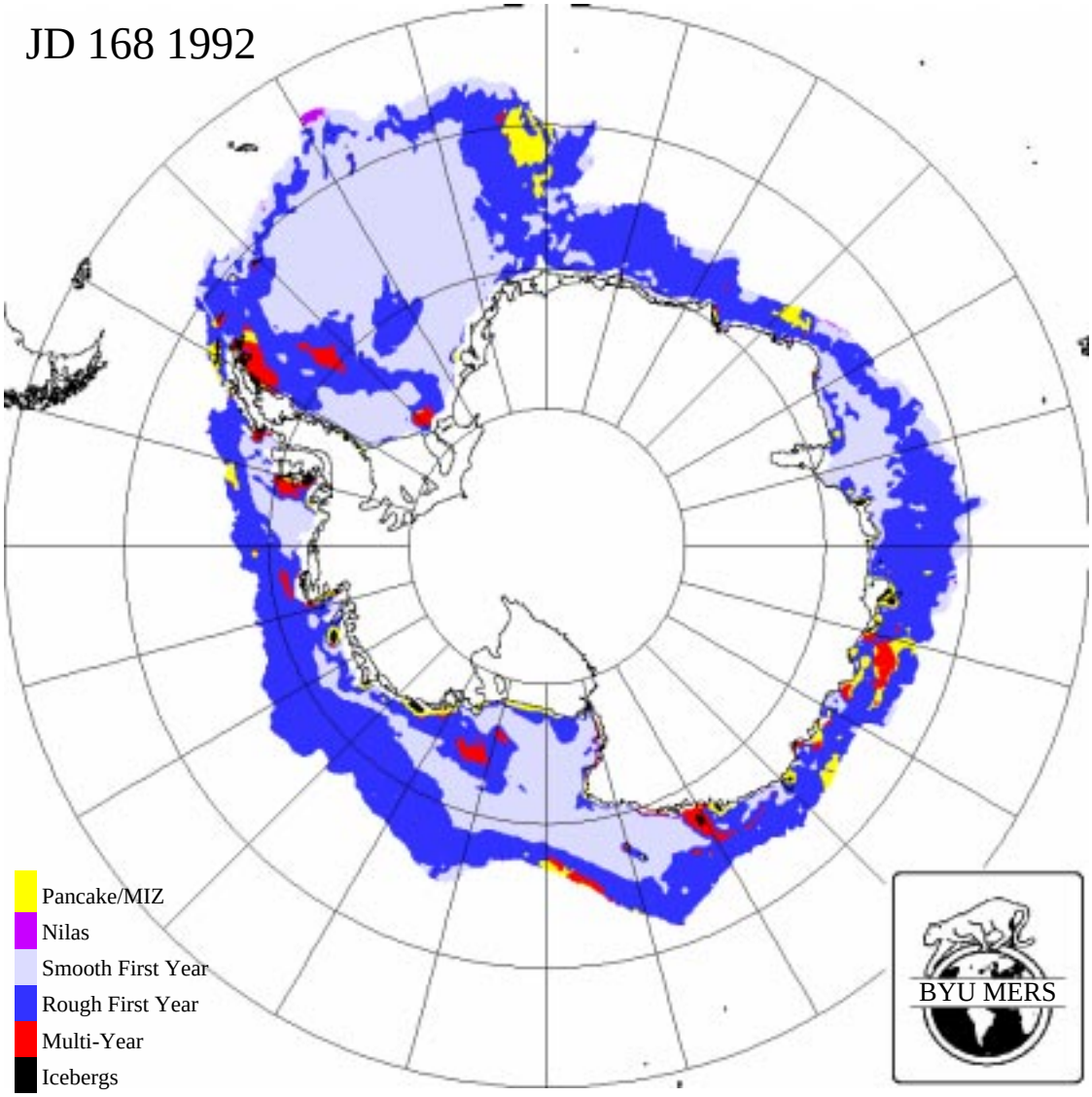


Figure 6.7: A classified image based on the SIR $\mathcal{A}$, $\mathcal{B}$ and STD images. Note the difference between this and the image in Figure 6.6 in that the MIZ is a separate class.

6.6.1 Consistency Check: The Annual Sea Ice Cycle 1995

Figures 6.8 through 6.13 show a typical annual cycle of sea ice growth and decay for 1995. Each year the cycle has differences from the year before, but here we wish to establish the consistency of the classes used in the classification algorithm. Differences from year to year of temperature, wind and ocean circulation and precipitation causes significant interannual variation of the amount of each ice type and the extent of the sea ice sheet. However, we may expect some basic consistency assuming the weather variations are not abnormally large. Also, we should expect that the distribution of ice type throughout a year long cycle to be more or less continuous throughout the year. In this section, we check the consistency of the classification of SIR imagery over the course of one year.

Figure 6.8 shows an example of the early expansion of the sea ice sheet. The prominent feature is the presence of relatively large quantities of MY ice both in the Weddell Sea and along many parts of the continental coast line. In the Weddell Sea, a band of rough first year ice surrounds an inner core of smooth first year ice. This is expected because the ice production begins near the continent, and as the ice ages, deforms and thickens it moves out to make room for newer ice to form off the Ronne-Filchner ice shelf.

Figures 6.9 and 6.10 show the expanding sea ice pack and the gradual break up of the older ice masses. Notice the substantial reduction in the MY sea ice compared to the first image. Figure 6.11 shows a good example of a typical MIZ ice class near the ice edge at the top of the figure. Figure 6.13 shows the sea ice pack near its maximal extent.

In general, examination of the set of images from 1995 shows that the various ice classes outlined in Section 6.3 display a spatial and temporal stability over the course of a one year sea ice cycle. The relatively small quantity of the MIZ class throughout the series is typical in that pancake ice and areas where wave action is prevalent are scattered and short lived [57]. The consistency of the ice classes from image to image further supports the definition of the classes and the spatial stability indicates that the classes are useful for studying dynamic motion in the sea ice pack.

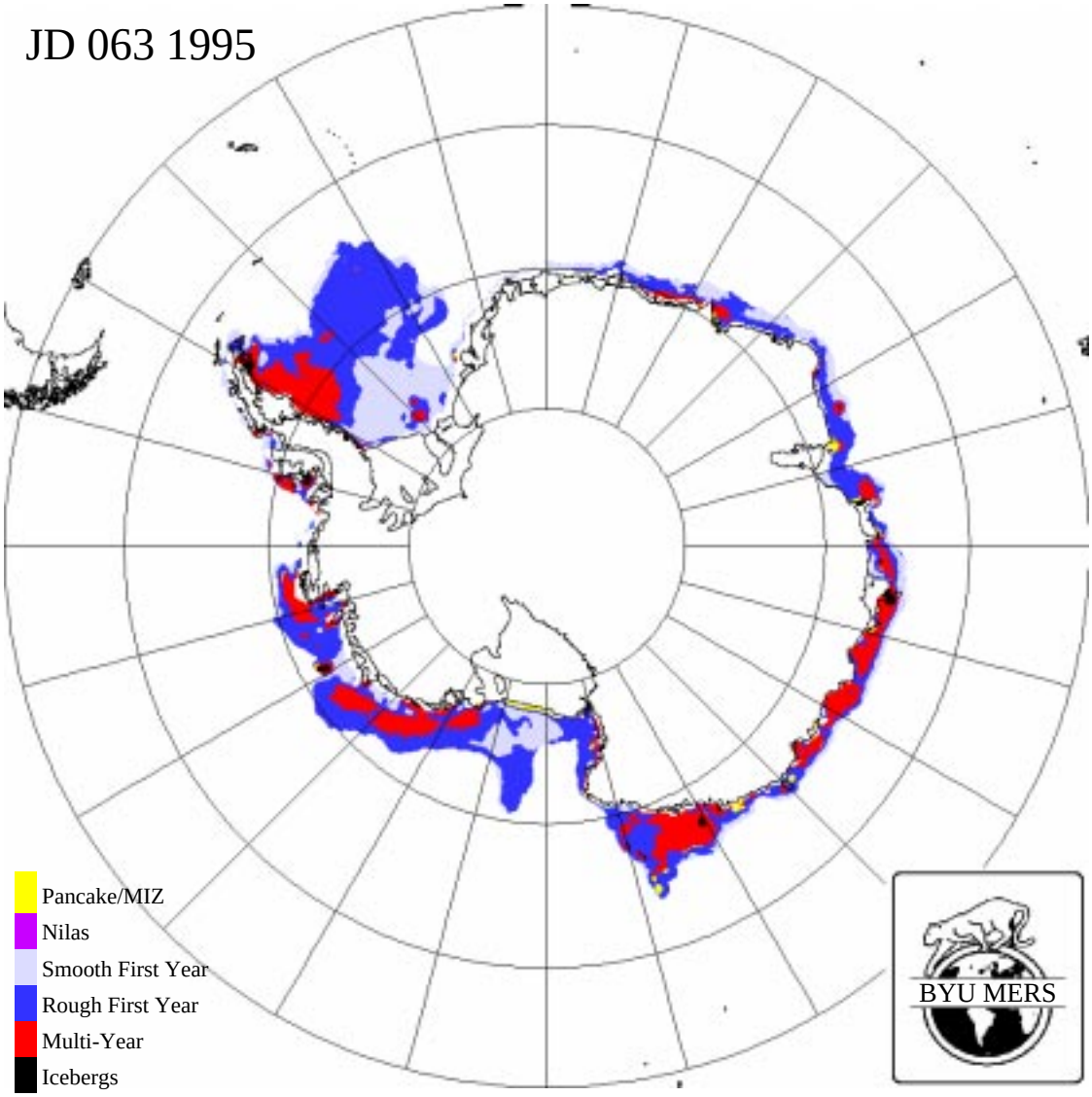


Figure 6.8: A classified image based on the SIR $\mathcal{A}$, $\mathcal{B}$ and STD images of JD 063 1995.

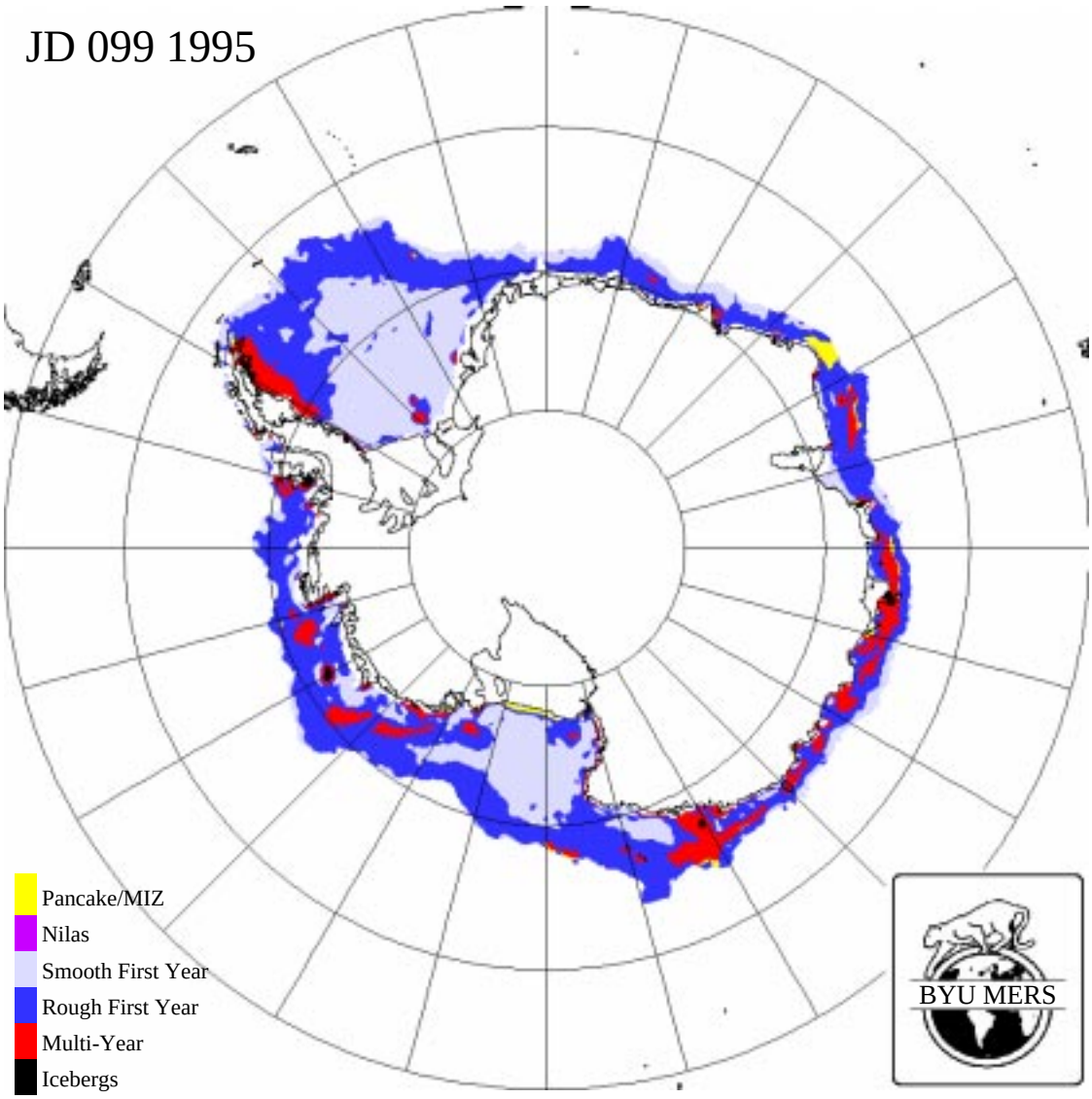


Figure 6.9: A classified image based on the SIR $\mathcal{A}$, $\mathcal{B}$ and STD images of JD 099 1995.

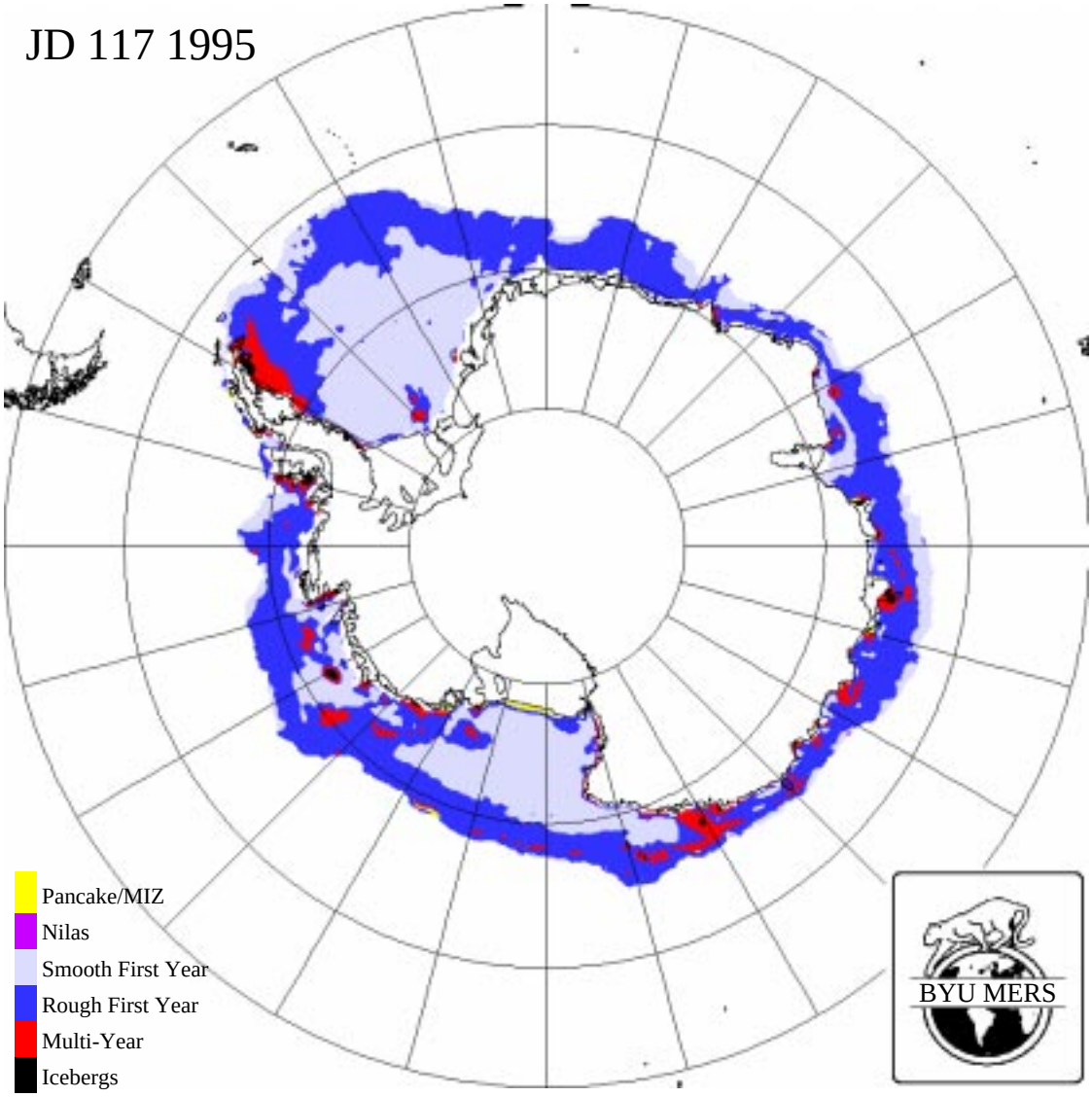


Figure 6.10: A classified image based on the SIR $\mathcal{A}$, $\mathcal{B}$ and STD images of JD 117 1995.

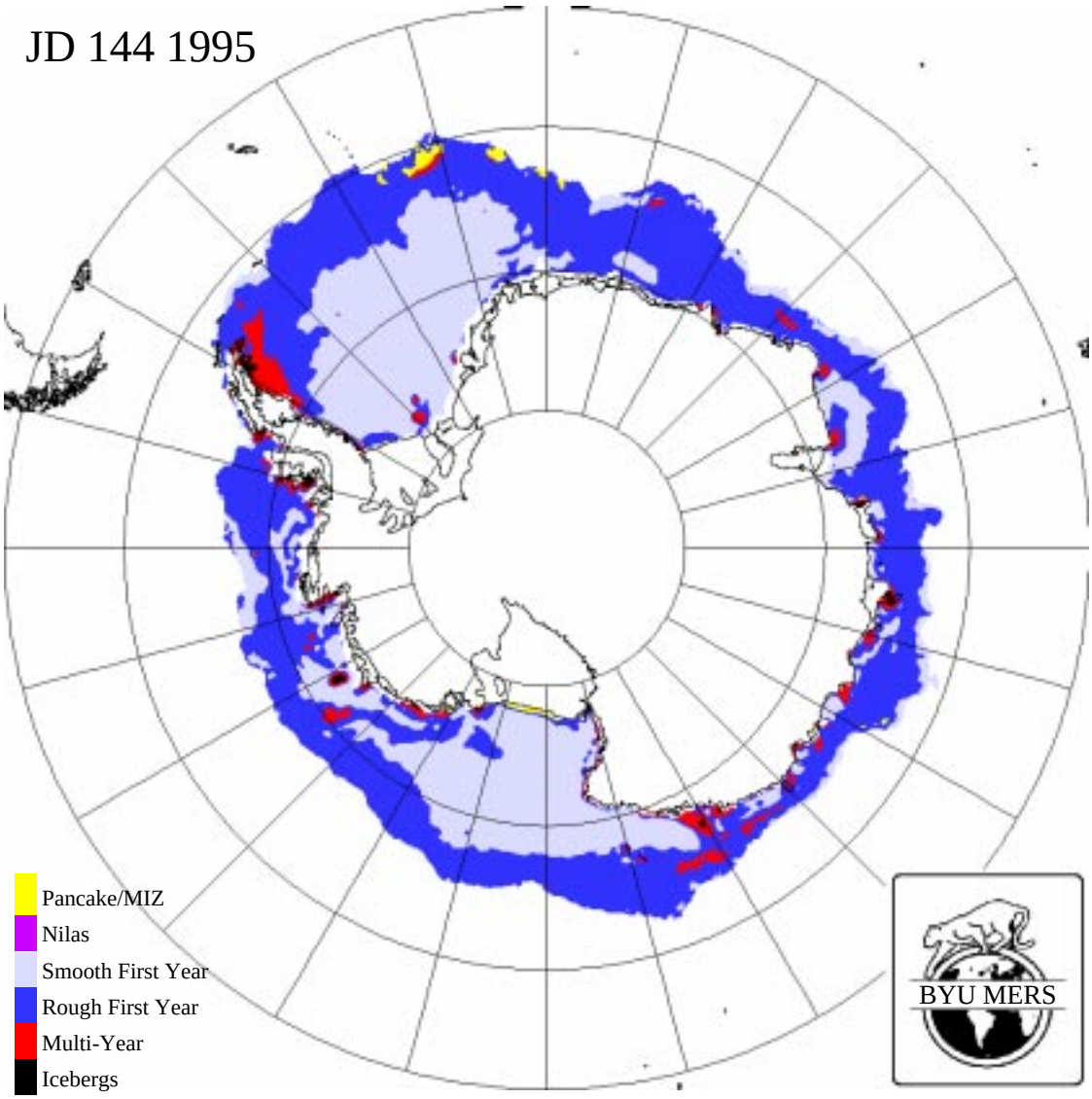


Figure 6.11: A classified image based on the SIR $\mathcal{A}$, $\mathcal{B}$ and STD images of JD 144 1995.

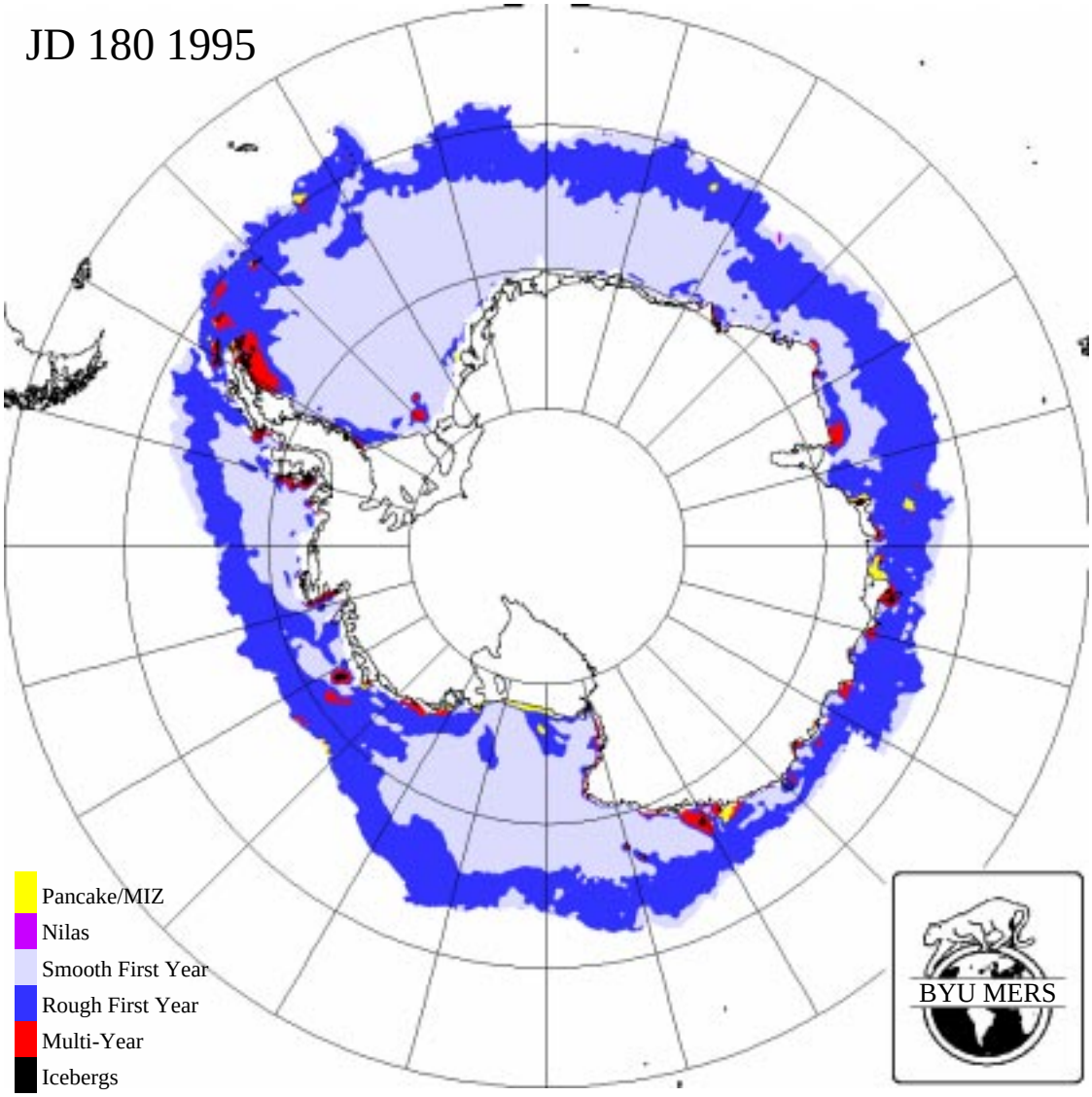


Figure 6.12: A classified image based on the SIR $\mathcal{A}$, $\mathcal{B}$ and STD images of JD 180 1995.

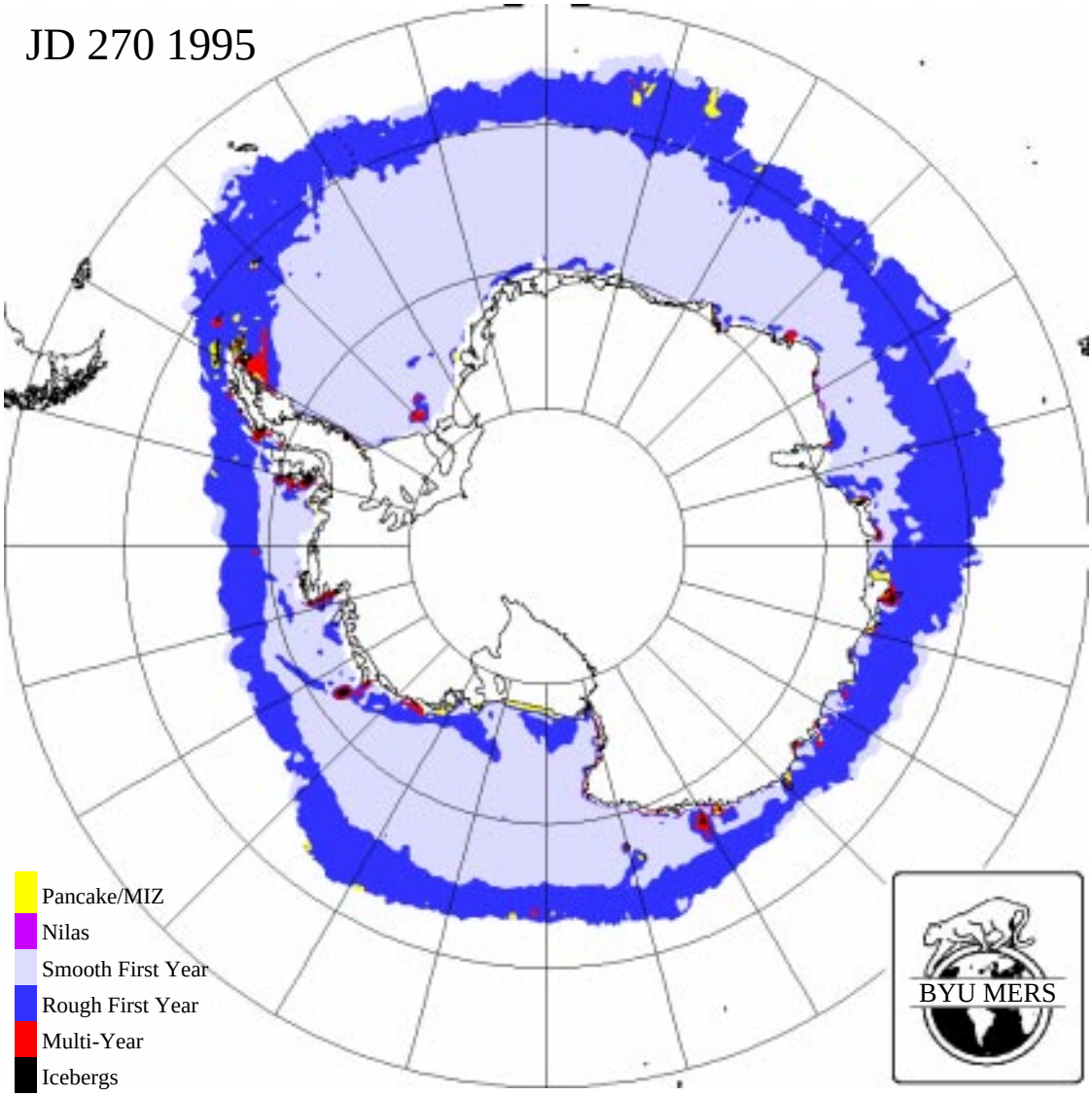


Figure 6.13: A classified image based on the SIR $\mathcal{A}$, $\mathcal{B}$ and STD images of JD 270 1995.

In summary, the consistency of the classification algorithm has been demonstrated first by a comparison of classified imagery to the few *in situ* observations available, where the MIZ and rough first year sea ice were correctly classified. Second, the temporal and spatial consistency of the classification throughout a single seasonal cycle has been demonstrated. This leads to a high confidence level in the sea ice classification algorithm. As more *in situ* observations become available, the algorithm may be refined to further reduce any classification errors.

6.7 Interannual Ice Type Variations 1992 to 1994

An interannual comparison of the changing areal values of each ice type is the first application of this new classifier to a study of Southern Ocean sea ice. The interest here is two-fold: first, to examine ice types individually, second to examine the overall annual sea ice cycle over the course of three years. The areal comparisons made in this chapter use the correction factor needed for the polar stereographic projection as derived in Appendix C.

6.7.1 Individual Sea Ice Type Interannual Variation

In Figures 6.14 to 6.18, the total area classified as a given sea ice type is plotted versus date. The y-axis is in square kilometers, the x-axis in months. Except for Figure 6.16 which shows first year rough and smooth sea ice classes from 1993 on the same plot, data from 1992 to 1994 is plotted together on a single plot, with different years having different line types as noted in the legends.

Figure 6.14 shows the annual evolution of first year smooth (FYS) sea ice. Around February/March, the area of FYS ice reaches a minimum and begins to increase as the fall freeze up commences. Note the consistently smooth transition in all three years from summer to fall, characterized by a gentle change in curvature around JD 50 of each year. The growth rate in all three years plotted is roughly linear and identical year to year at 64000 km^2 per day. In 1992 the minimum was lower and overall growth consistently lagged the other two years by about 1-2 weeks, although the growth rate was identical to the two following years. All three years reach a

peak around August/September at about 12.5 million square kilometers of FYS ice. The peak is short lived, as the rapid spring melt begins around September/October. Continual growth of this ice type is expected prior to the summer melt, since new ice formation at the periphery of the sea ice sheet and at interior polynyas continues right up to beginning of the summer melt. As temperatures begin to change, the area peaks, then falls rapidly as the ice sheet melts and breaks up.

Figure 6.15 shows the annual evolution of first year rough (FYR) sea ice. Recall that this type of ice is primarily made up of FYS ice that has been deformed through the natural motion of the sea ice sheet. The graph reveals several things. In February/March of each year, the area of FYR ice reaches a minimum just as it did for FYS ice. The growth rate in each year is again linear and identical year to year at 59000 km^2 per day, very similar to the rate of growth of FYS ice. The lag displayed in 1992 for FYS ice is not apparent in the FYR class. FYR ice also reaches a peak three months earlier than FYS ice, peaking around May/June. The peak area is also much less than for FYS, with FYR rough ice peaking at about 8 million square kilometers.

An interesting comparison is made in Figure 6.16 of the FYS and FYR classes from 1993 data. It shows that the growth rate for each class is nearly concurrent up to the time that FYR ice peaks in May/June. It also shows that the consistently smooth transition from summer to fall displayed by FYS data is not the same as FYR data, where the transition from summer to fall is relatively abrupt for FYR ice (compare Figures 6.14 and 6.15).

Figure 6.17 shows the interannual variation of multiyear (MY) ice. Through the summer melt up to the beginning of the fall freeze up (September through January), the volume of MY ice in the graphs progresses to a minimum value, which is unexpected since we expect a maximum value at the end of the summer melt. This is likely a result of the surface melting and pooling of water on the ice surface and the significant change this has on the backscatter from the old sea ice. Surface water significantly reduces the $\mathcal{A}$ value from MY sea ice, causing some misclassification in the late melt period by the one dimensional classification algorithm. Also, break up of the older ice and integration of these smaller pieces of old ice into the body of the

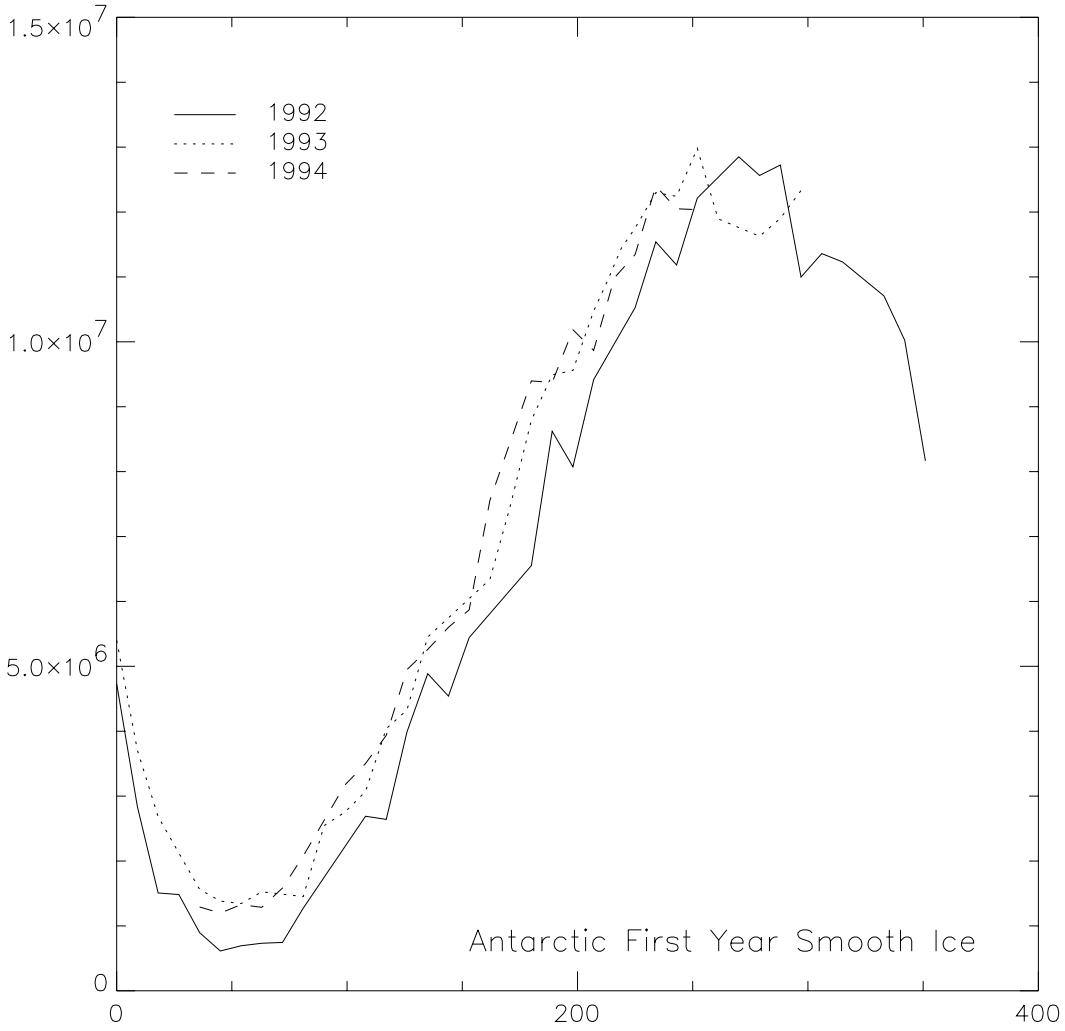


Figure 6.14: This plot shows a comparison of the total surface area of smooth first year ice using the classification algorithm proposed in the text. Note that just as for rough first year ice, the increase in surface area for the type during the fall freeze up is consistent for all three years.

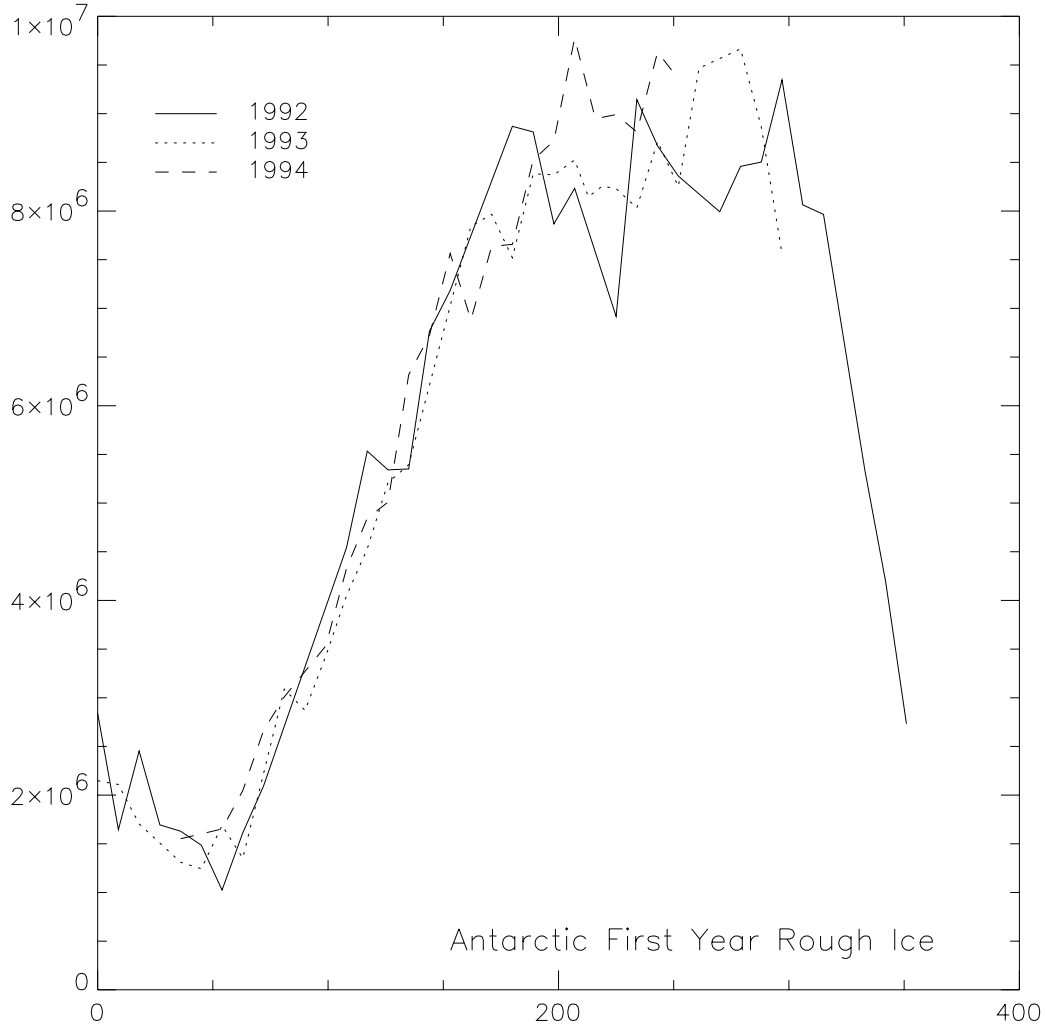


Figure 6.15: This plot shows a comparison of the total surface area of rough first year ice using the classification algorithm proposed in the text. Note that for each year, the increase in surface area for the type during the fall freeze up is consistent for all three years.

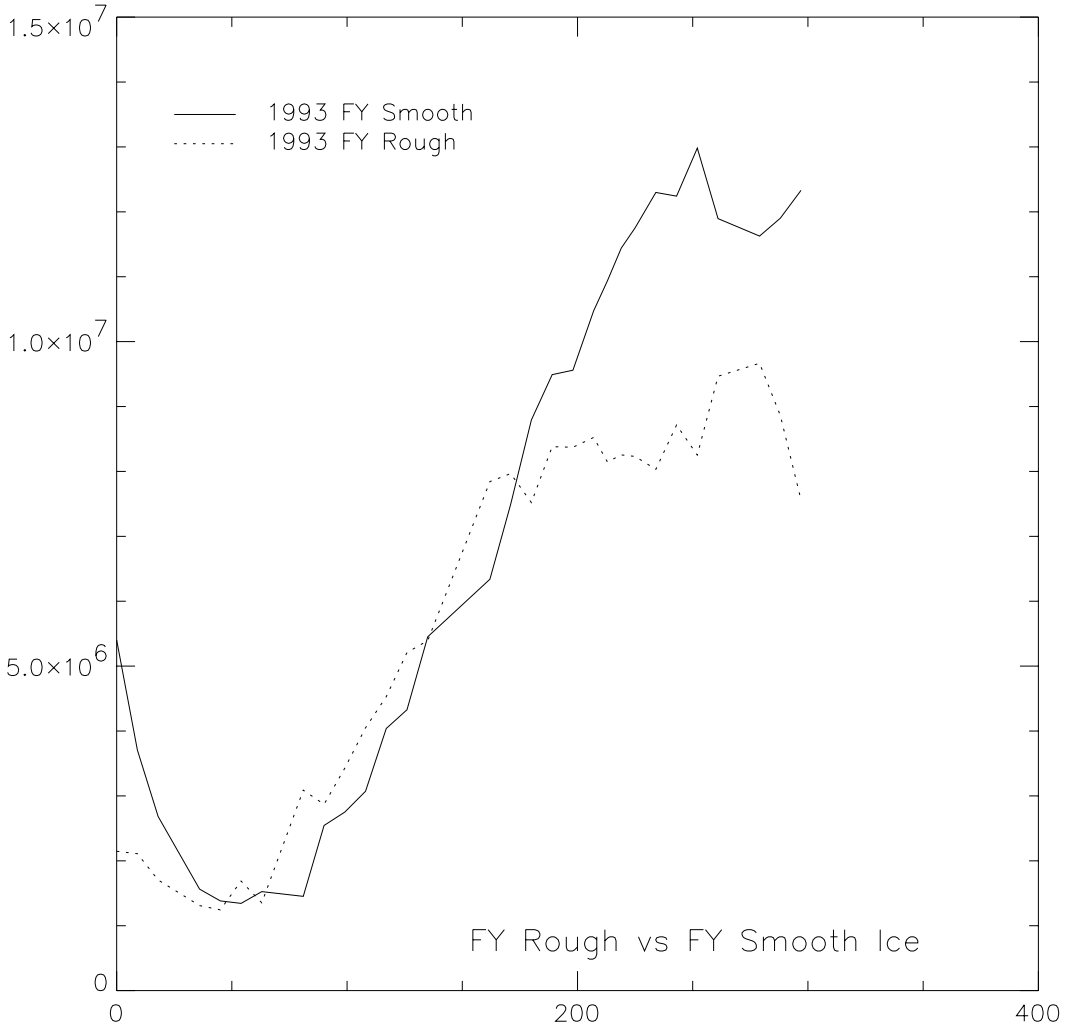


Figure 6.16: This plot shows a comparison of the total surface area of smooth first year ice and rough first year ice in 1993. It is interesting to see that rough first year ice has a larger areal coverage than smooth first year ice, which plateaus earlier in the freeze up cycle. This is consistent with the idea that rough ice forms during deformation, which continues throughout the ice growth period.

ice sheet reduce the area of MY ice over the course of the annual cycle. One other interesting thing to note is the nearly 50% reduction in MY ice area in 1993 compared to 1992 and 1994.

Figure 6.18 shows the plots for icebergs, the final ice type considered. This figure only shows 1993 and 1994 data due to heavy SAR use in 1992 which caused a great deal of dropout errors in the scatterometer data in the areas where bergs are most common. The high areal peak around JD 250 in 1993 is a result of a similar drop out problem where the dropout caused a number of pixels to be classified incorrectly as icebergs. Current processing methods do not account for drop out errors, and the result is the insertion of drop out errors into the iceberg class. Nonetheless, the plots do show that bergs have a relatively constant area in the ice sheet throughout most of the year, with the drop in area at the beginning of the season due to the presence of surface water forming on the bergs as the temperatures increase.

The final plot in this review is an interannual cycle plot composed of the sum of the area of all ice classes plotted versus time for all three years. The plot in Figure 6.19 shows the interannual cycle of the Souther Ocean sea ice sheet. The irregularities in the 1992 cycle are the result of some minor data losses in 1992 due to the frequent use of the SAR for study of Antarctica. In each year, the total sea ice volume peaks between 21 and 22 million square kilometers.

Although there is a general upward trend in total ice area year to year, this three year sample is unlikely to reveal any long term patterns. However, the slight upward trend is consistent with the results of another study of long term ice trends indicating that Southern Ocean sea ice is gradually reaching larger and larger extents [8]. It is interesting to note that if this trend were true, it would indicate global *cooling*, not global warming. These results should be the beginning of a longer term study to establish a statistically significant trend in sea ice extent. This data, combined with other C-band data from ERS-1 and ERS-2, and Ku-band data from the recently launched NSCAT scatterometer allows data to be compiled over the next few years so that an extended study of Southern Ocean sea ice behavior may be made. SIR imagery should prove to be a valuable tool in basin-wide sea ice analysis.

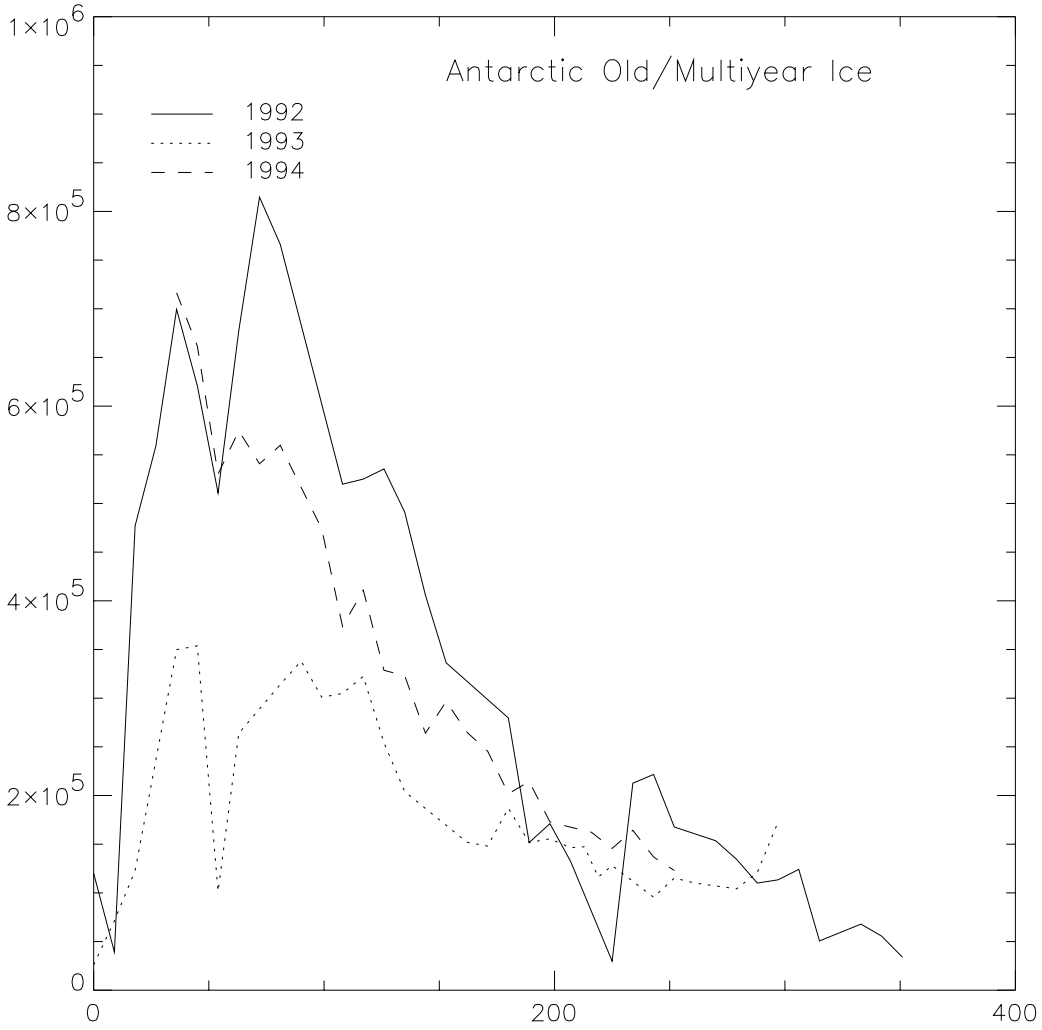


Figure 6.17: This plot shows a comparison of the total surface area of multiyear ice using the classification algorithm proposed in the text.

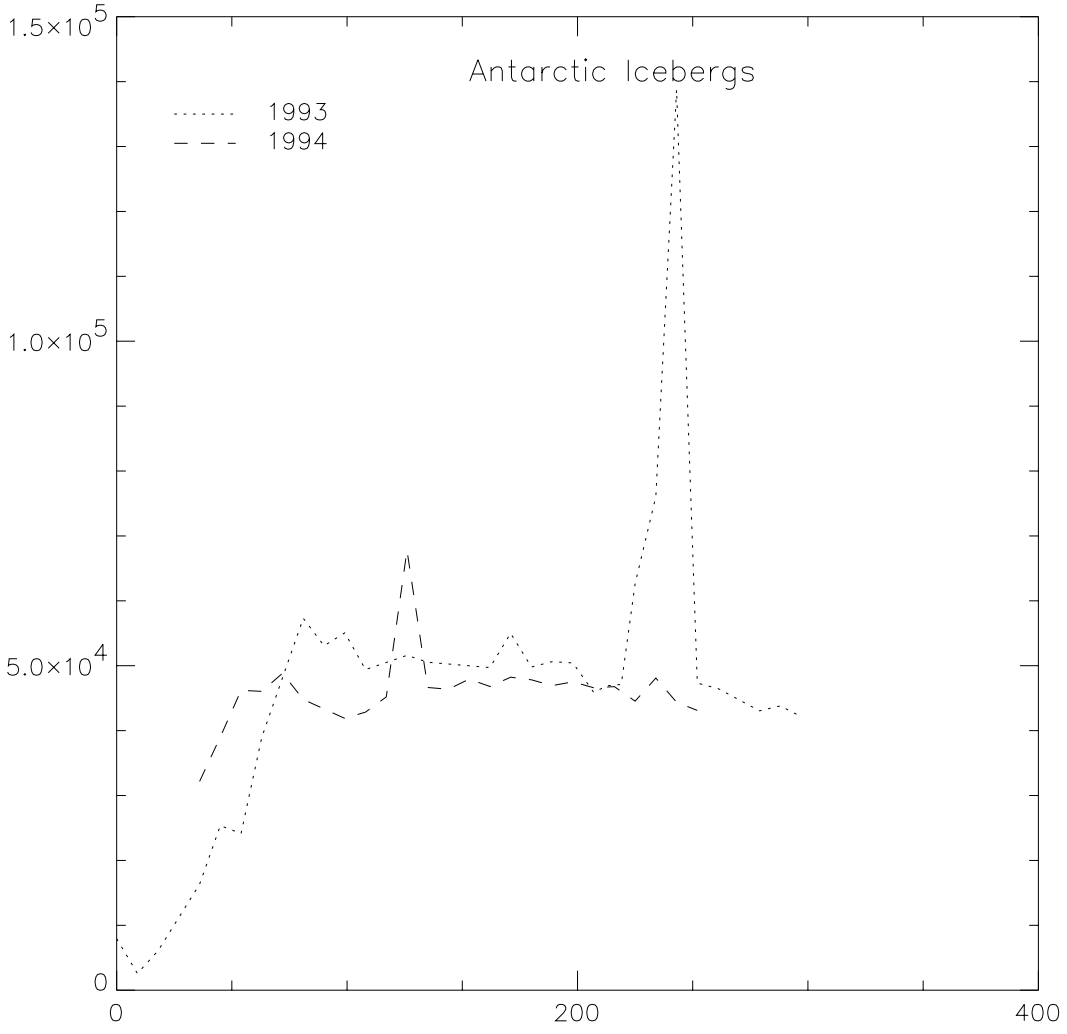


Figure 6.18: This plot shows a comparison of the total surface area of ice berg using the classification algorithm proposed in the text.

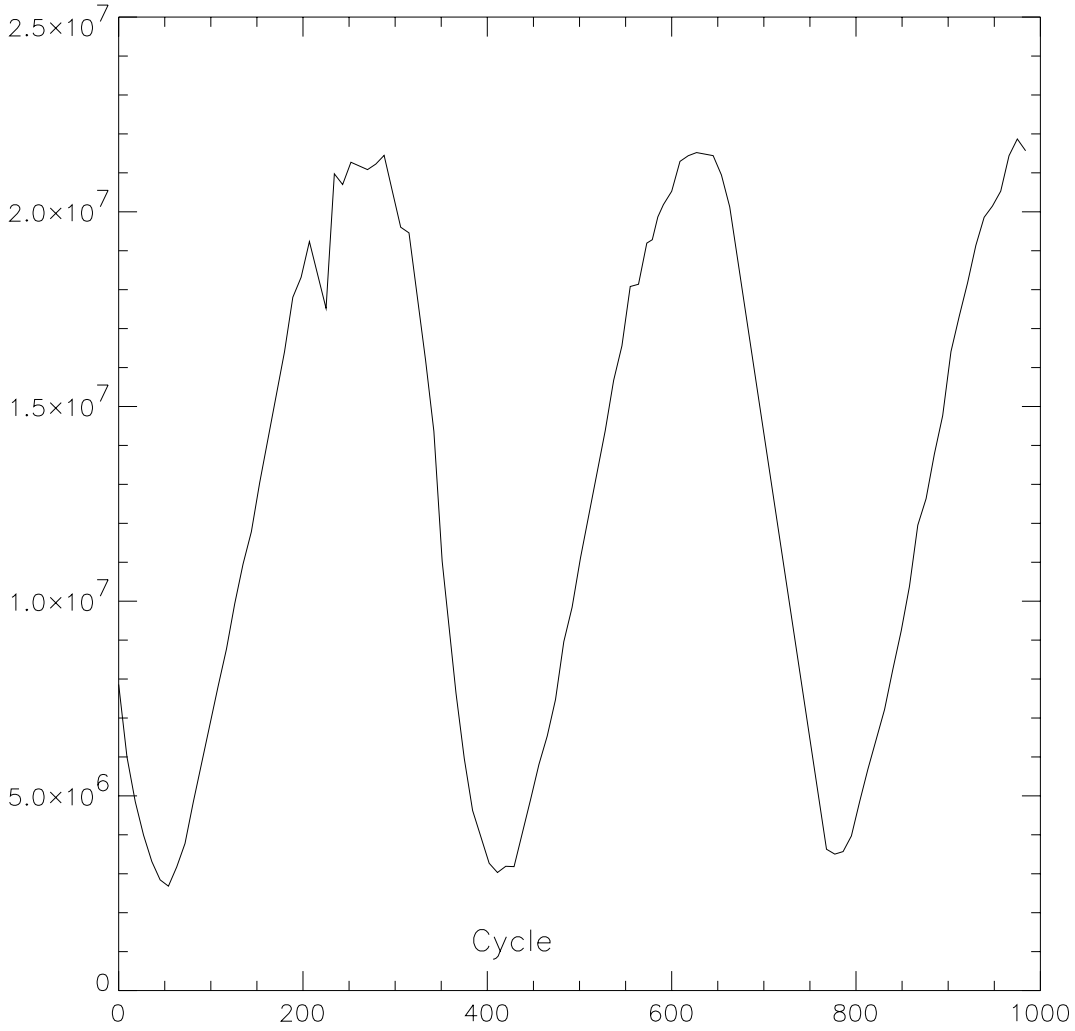


Figure 6.19: Interannual cycle of total ice area for 1992, 1993 and 1994. Note the slight upward trend in ice area, which is consistent with findings in other studies [8].

6.8 Summary

In this chapter, we presented background on ice classes and a simple, previously developed one dimensional sea ice classifier based on SIR $\mathcal{A}$ imagery. This routine was extended to separate MIZ/Pancake ice from MY, two classes indistinguishable using $\mathcal{A}$ values alone, by incorporating $\mathcal{B}$ and STD image information in the classification algorithm.

As validation of this new classification algorithm, *in situ* observations of the sea ice pack made in 1992 are compared with classified imagery. The result is a strong correlation between the observed sea ice conditions and SIR imagery classified with the extended algorithm. An examination of a full year cycle of sea ice growth and decay in 1995 reveals that the various ice classes show strong spatial and temporal stability from the beginning of ice pack expansion through the decay of the sea ice pack. The stability of the classification indicates that the images have value for study of short term motion in the sea ice pack, as well as for study of interannual changes in sea ice type proportions.

Finally, an interannual comparison of areal variation of the various ice classes is made. In each case, the trends of growth and decay match expectations for each ice type, further supporting the accuracy of the SIR images and classification algorithm. The final interannual comparison of total sea ice in the Southern Ocean indicated a slight increase in maximal sea ice extent over three years. As more and more data is collected, current trends, whether towards greater ice extent or lesser ice extent, may be better evaluated and the conditions of global sea ice may be more regularly monitored.

Chapter 7

Conclusion

In the preceding chapters, various aspects of scatterometer image reconstruction were discussed. Following a background of satellite imaging, with an emphasis on polar remote sensing, the mechanics of the SIR algorithm, an image reconstruction algorithm which has been used successfully for various remote sensing applications, was presented. This introduction to the SIR algorithm sets the stage for developing the theory behind the algorithm. In the theory chapter, a theory of reconstruction from irregular samples was combined with the aperture characteristics of the ERS-1 scatterometer to demonstrate the reality of the resolution enhancement of the SIR algorithm. This was combined with synthetic images which quantified the resolution enhancement for the ERS-1 scatterometer and demonstrated empirically the resolution enhancement.

The next chapter presented an analysis of C-band azimuthal modulation over Southern Ocean sea ice. This was the first time such a study was carried out for the C-band scatterometer. The study concluded that the azimuthal modulation of less than 1 dB and is negligible for the purposes of image reconstruction.

The final section presented a first generation sea ice classifier based on the SIR algorithm and a new measurement of surface anisotropy. The results showed strong temporal and spatial consistency, was consistent with available *in situ* measurements, and demonstrated the viability of SIR for tracking sea ice type evolution, sea ice extent and motion within the sea ice pack.

7.1 Contributions

This work represents a significant contribution in the field of scatterometer image reconstruction and resolution enhancement. The contributions can be summarized under three headings as follows:

Additive ART: Complete Reconstruction

The theory presented herein has broad application. The demonstration of additive ART as a complete reconstruction under Gröchenig's lemma validates this particular method of iterative reconstruction as a complete reconstruction in the limit. Also, the relationship between the aperture function and the sampling has implications in general radar application. Depending on the application, it may be possible to increase performance of a radar system by widening the main lobe and sampling more frequently and then using a reconstruction algorithm that meets the criteria of the lemma.

Theory - SIR

Although the SIR algorithm has been in use for some time (See for example [3] published in 1993), the theory presented here in Chapter 3 establishes for the first time the validity of the SIR algorithm approach and qualifies the theoretical limits of the algorithm [60–62]. This work also lays the necessary ground work and justification for hardware modifications that makes future scatterometers more useful as imaging platforms.

Azimuthal Modulation

The azimuthal modulation study published in 1997 was the first of its kind for Antarctic sea ice [39, 43], and while necessary for validating the SIR algorithm, provides a nice stepping stone to other areas of study using a measure of anisotropy. Here, a methodology was established and applied to the scatterometer data which establishes the nature of azimuthal modulation at the scale of the spaceborne scatterometer.

Classification

In the final chapter, a basic classification algorithm developed from available *in situ* data is presented and applied to Antarctic sea ice [35]. Here, through a series of SIR images, the consistency of the classification is checked and the inter-annual variation of the classification is presented for 3 years (1993, 1994 and 1995). While the algorithm may require some additional fine tuning as more *in situ* measurements are gathered, this first cut appears to have good potential for a basin wide monitoring of the various ice types.

There are several publications derived from this work, including more than 10 conference papers and two journal articles: one on the azimuthal modulation material [43], and the other covering the theoretical basis for SIR [62]. A third journal article is being prepared from the classification material in the last chapter.

7.2 Future Work

There are several natural extensions of the current work that may be pursued to enhance and expand the results of this project.

Relationship of side lobe level to iterations

As outlined in the main body of the text, the attenuation of the frequency content of the original surface is irrelevant to the reconstruction in a theoretical sense. However, in a practical environment, finite iterations may limit the ultimate development of the image. A study of the relationship of side lobe level to iterations and quantifying the error associated with that in an ideal case would be instructive in further development of the SIR algorithm.

NSCAT Data

Because of the post-processing techniques used by ESA to produce the various ERS-1 scatterometer products, resolution enhancement is limited. With the recent acquisition of 9 months of data from the NSCAT Ku-band scatterometer, however, there is a new set of data that is capable of significantly better resolution.

Results have demonstrated that the resolution enhancement of NSCAT data is much better than is possible using the ERS-1 data (12 km per pixel versus 33 km per pixel), resulting in even more detailed imagery. The application of SIR to NSCAT requires several preparatory studies:

Theory The NSCAT measurement cells are oblong, not round, and in any given image region there are multiple orientations and sizes of cells. The impact of this on final resolution of the images needs to be examined. A first step would be a reworking of the theory presented in Chapter 3 to account for the new shape and orientation of the measurement cells. One possible approach is an analysis of the size and shape of the cells at certain latitudes and longitudes used to create statistical relationship of the size/shape pattern to the resolution recoverable from the data.

Azimuthal Study Because NSCAT is at Ku-band, a repeat of the azimuthal modulation study presented in Chapter 5 was necessary and has recently been completed. The results showed less than 0.6 dB of azimuthal modulation, with significant modulation over glacial regions, as expected [63]. If nothing else, this study confirms the isotropic azimuthal response of sea ice for a wide range of microwave frequencies.

Classification It is noted in the literature that there is more distinction between sea ice classes at Ku-band compared to C-band [19,46].As such, the NSCAT data has the potential of more and/or better separated sea ice classes. *In situ* measurements similar to the ones used to create the sea ice classes in Chapter 4 need to be assembled and studied to form new ice type classifications based on $\mathcal{A}$, $\mathcal{B}$ and standard deviation of anisotropy measures.

Standard Deviation of Anisotropy

This measure as currently implemented is not truly resolution enhanced when compared to the $\mathcal{A}$ and $\mathcal{B}$ values which have been produced using SIR. While the current methodology is sufficient for a first generation classifier, further development and enhancement of the standard deviation of anisotropy measure would increase the

accuracy of future classification methods. One approach has been to use the resolution enhanced value of $\mathcal{A}$ compared to the actual measurement values to determine an $\mathcal{A}$ STD value, which is being evaluated. Adaptation of SIR to directly calculate the measure or the use of a similar resolution enhancing method would be appropriate. Also, application of this measure to NSCAT data requires some development since the current data format does not have co-located fore and aft measurement cells like the ERS-1 data has. However, this lack of registration may prove to be useful in developing a better measure of surface anisotropy.

Confidence Image

With a first generation classifier under development, a good next step would be a confidence image. This would give some indication of how good the estimate of $\mathcal{A}$ or $\mathcal{B}$ is relative to various data aspects that affect the estimate. In particular, the number of measurements that touch a point in the enhanced resolution image and the standard deviation of the incidence angles touching a pixel could be statistically related to how well the estimate can be made from the data. The result would lend credence to the classified images and help flag areas in the image that may be less than accurate.

Algorithm Refinement

The simple classifier presented in this work is the first generation of a valuable research tool. With SIR, images of Antarctica can be generated about every three days with NSCAT, and since the processing time is relatively short compared to SAR image processing, SIR has the potential to provide very rapid medium scale imagery of remote earth locations. Coupled with a well-developed classifier, and this data would be invaluable in the study of interannual changes and variation. Further development of the classifier by integrating more surface studies and possible validation using *in situ* data from a future field study would be a first step. Investigation and use of other parameters beyond the $\mathcal{A}$, $\mathcal{B}$ and standard deviation of anisotropy could yield more information about the surface.

Refinement of SIR

Currently, the standard deviation parameter is not a resolution enhanced figure. It may be possible to get from SIR other relatively independent parameters that would aid in classification of surface features. Investigation of other forms of the multivariate SIR might yield more surface information.

Inter-Instrument Comparison

Since there are now three scatterometer data sets (SASS, ERS-1/2 and NSCAT) at two different frequencies spanning more than three decades, a study utilizing all these instruments to study interannual variations in sea ice formation can now be performed.

Broad Expansion of Application of SIR

With the establishment of SIR as a valid algorithmic approach, the potential application is large. Previous studies using SIR have focused on regional vegetation studies [64–67] and Greenland ice sheets [38, 68, 69]. Such studies could be, and are in the process of being broadened and refined, and include such things as farmland studies, hydrology, deforestation in various regions [70, 71] and north pole sea ice. Some work is already in progress to expand the application of SIR to other earth surfaces [72]. In another type of expansion, the SIR algorithm approach could be applied to different instruments, such as the SSM/I passive radar system as in [34]. With proper *in situ* measurements and model development, the application of enhanced resolution scatterometer imagery for large scale studies of earth surfaces using multiple spaceborne platforms is quite large: from farming studies to deforestation to polar sea ice.

Appendix A

The ERS-1 Scatterometer: A Functional Summary

A.1 ERS-1 Scatterometer Instrument Geometry

In this dissertation, ERS-1 scatterometer data is used. The ERS-1 scatterometer is part of the Active Microwave Instrument (AMI) on board the ERS-1 satellite. Figure A.1 illustrates the geometry of the scatterometer measurements. Each beam of the scatterometer is a fan beam antenna with the beam pattern oriented as shown in the figure, with the fore, mid and aft beams oriented respectively at 45°, 90° and 135° relative to the flight path. Since the orientation of the flight path relative to north is known, for each track of the satellite an absolute azimuth angle of the readings for a given beam is known. The multiple azimuth measurements are important in making near surface ocean wind measurements. The incidence angle of the measurements range from 18° to 47° for the mid beam and 25° to 59° on the fore and aft beams based on the 500 km wide swath and an 800 km average instrument altitude with small variations due to curvature in the Earth's surface.

There are several data products that are available from the ERS-1 scatterometer, but very little has been released concerning the details of how these products are made. The data used in this study is the lowest level data currently available from the European Space Agency (ESA). The data is organized on a regular 25x25 km grid oriented with the satellite track and there are 19 cells across the swath. It should be noted that as the satellite track changes with each orbit, so does the orientation of the regular grid relative to previous passes. Thus the measurement from one pass over a specific area may be in a slightly different location on the next pass. This lack

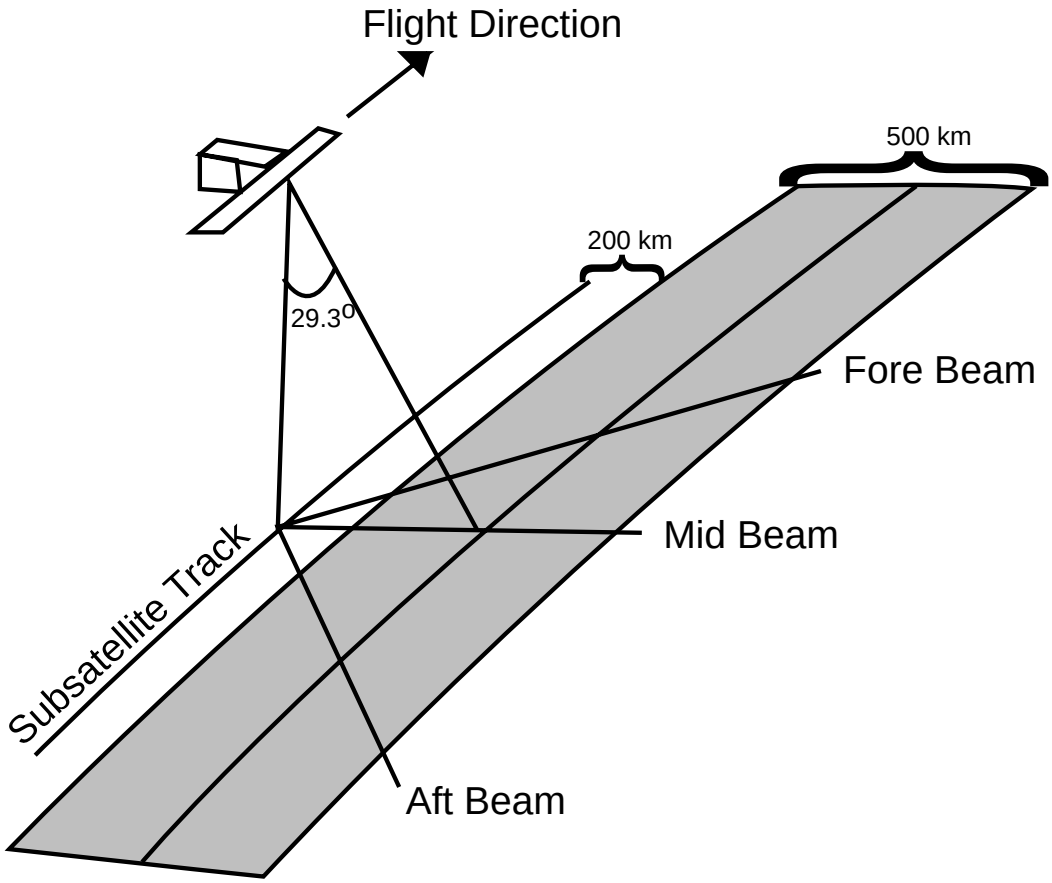


Figure A.1: A graphical representation of the ERS-1 scatterometer geometry. The gray stripe represents the swath, which is 500 km wide. The fan beam antennas of the scatterometer are oriented at 45° , 135° and 180° relative to the satellite flight vector. This provides three different azimuth measurements for each pass. The swath is computationally divided into a regular grid of 19 "nodes" or cells across the swath spaced 25 km apart. The raw scatterometer data is processed so that each cell in the swath has a fore, mid and aft beam measurement.

of measurement registration from pass to pass is illustrated later in Chapter 3 and is the core of the SIR resolution enhancement algorithm, which relies on multiple, overlapping measurements to achieve resolution enhancement.

The actual raw scatterometer measurements are not taken on a regular grid. After the raw data is gathered, it is resampled into the regular 25x25 km grid using a predefined impulse response function applied to the raw measurements. The impulse function was chosen to minimize aliasing by filtering out high frequency components from the raw data. The resampling causes the impulse response function to be nominally the same for every cell, even at the very ends of the beams where normally elongation distorts the cell size. In contrast, the NSCAT scatterometer has a variable cell size and response depending on where the cell is along the beam. At this point, we note again that there is a marked lack of published information about exactly how the ground processing resamples the raw ERS-1 scatterometer data into this 25x25 km grid.

In essence, the cell locations are assigned to a swath and then raw data from the fore beam that falls within the region of support of the impulse function centered at a given cell is used to generate a fore beam measurement. This measurement value is assigned to that cell, and the process is repeated for the mid and aft beams. Thus, for each cell location, there is a fore, mid and aft beam measurement. Further information about the exact application of the impulse function is lacking at this time.

The impulse response used for ERS-1 is a Hamming window. The impulse response used for ERS-1 is assumed to have 50 km region of support. However, the 50 km distance is approximately the 3 dB level of the impulse response [42], and the actual region of support extends to between 85 and 100 km. Figure A.2 shows the cell centers for an example swath. Several of the cell centers in the figure have had a 50 km diameter circle superimposed on them to illustrate the overlap of the various measurements.

The lack of access to the raw scatterometer data and the particular choice of the Hamming window as an aperture function with its very low side lobes hinders

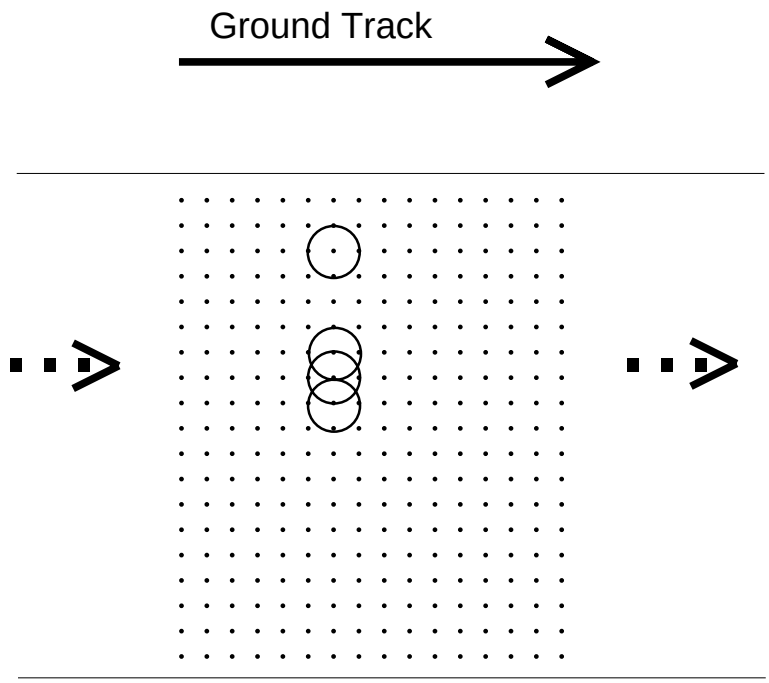


Figure A.2: This figure illustrates the 19 nodes that are oriented across the swath. The cells are spaced 25 km apart in each direction. The impulse response assigned to each cell is circular with approximately a 50 km diameter, so the impulse response of one cell overlaps the responses of its immediate neighbors.

the reconstruction of imagery from the ERS-1 scatterometer (see Chapters 3 and 6). However, the data in its current format can be enhanced and used successfully for land and sea ice studies.

A.2 Derivation of The ERS-1 Aperture Frequency Response

As part of the evaluation of the ERS-1 scatterometer as a land imaging instrument, the effects of the aperture function associated with the scatterometer measurements is necessary. Each measurement cell of the ERS-1 scatterometer has a cos^2 roll off characteristic. The aperture function effectively filters the incoming signal with a low pass filter with a frequency response corresponding to the Fourier transform of the cos^2 roll off function. In this section, the Fourier transform of the ERS-1 scatterometer aperture function is derived and the analyzed.

A.2.1 Derivation

First, define the roll off function:

$$\begin{aligned} f(x,y) &= \cos^2\left(\frac{2\pi}{100}\sqrt{x^2+y^2}\right) \\ 0\,\text{km} \leq & \sqrt{x^2+y^2} \leq 50\,\text{km} \end{aligned} \tag{A.1}$$

where x and y are given in kilometers. The region of support for the aperture is represented by A and is assumed to be a circle of radius 50 km. The Fourier Transform of Eq. (A.2) is given by:

$$F(\xi_1,\xi_2) = \int_{-50}^{50} dx \int_{-50}^{50} dy\, f(x,y) e^{-j2\pi(\xi_1 x+\xi_2 y)}. \tag{A.2}$$

Note that the integration is only over the aperture region of support. Because the aperture is circularly symmetric, a change to cylindrical coordinates is appropriate:

$$\begin{gathered} \int_{-50}^{50} dx \int_{-50}^{50} dy \rightarrow \int_0^{2\pi} d\theta \int_0^{50} \rho d\rho \\ x = \rho\cos(\theta) \;\; y = \rho\sin(\theta) \\ F(\xi_1,\xi_2) = \int_0^{2\pi} d\theta \int_0^{50} d\rho\, \rho \cos^2(\tfrac{2\pi}{100}\rho) e^{-j2\pi(\xi_1\rho\cos(\theta)+\xi_2\rho\sin(\theta))} \end{gathered} \tag{A.3}$$

and for brevity, the following values are defined:

$$\beta = \frac{2\pi}{100} \tag{A.4}$$

$$k = 2\pi\sqrt{(\xi_1^2 + \xi_2^2}. \tag{A.5}$$

Begin by reorganizing the exponent:

$$\begin{aligned} F(\xi_1, \xi_2) &= \int_0^{2\pi} d\theta \int_0^{50} d\rho\, \rho \cos^2(\beta\rho) e^{-j2\pi(\xi_1 \rho\cos(\theta) + \xi_2 \rho \sin(\theta))} \\ &= \int_0^{2\pi} d\theta \int_0^{50} d\rho\, \rho \cos^2(\beta\rho) e^{-j2\pi\rho(\xi_1 \cos(\theta) + \xi_2 \sin(\theta))} \\ &= \int_0^{2\pi} d\theta \int_0^{50} d\rho\, \rho \cos^2(\beta\rho) e^{-j2\pi\rho\sqrt{\xi_1^2+\xi_2^2}\cos(\theta - \phi_k)} \end{aligned} \tag{A.6}$$

and note that because the function is symmetric, the phase term ϕ_k is unimportant in the integration through all values of θ. We can now set $\phi_k = 0$ and then rearranging the integrals results in

$$F(\xi_1, \xi_2) = \int_0^{50} d\rho\, \rho \cos^2(\beta\rho) \int_0^{2\pi} d\theta e^{-j2\pi\rho\sqrt{\xi_1^2+\xi_2^2}\cos(\theta)}. \tag{A.7}$$

Recall that the Bessel function is defined as

$$J_n(x) = \frac{1}{2\pi} \int_0^{2\pi} d\theta e^{-j(x\cos(\theta) - n\theta + n\frac{\pi}{2})} \tag{A.8}$$

so that the integration of θ in Eq. (A.7) becomes the zero-order Bessel Function since $n = 0$ will eliminate the two phase terms in Eq. (A.8). The final result, then, is the Fourier transform equals:

$$F(\xi_1, \xi_2) = 2\pi \int_0^{50} d\rho\, \rho \cos^2(\beta\rho) J_0(k\rho) \tag{A.9}$$

where k is defined in Eq. (A.5).

A.2.2 Evaluation

Because of the Bessel function, the integration in Eq. (A.9) is not trivial. While approximations can be made for the Bessel function for large and small arguments, the approximations give less accurate results near the origin, so the most

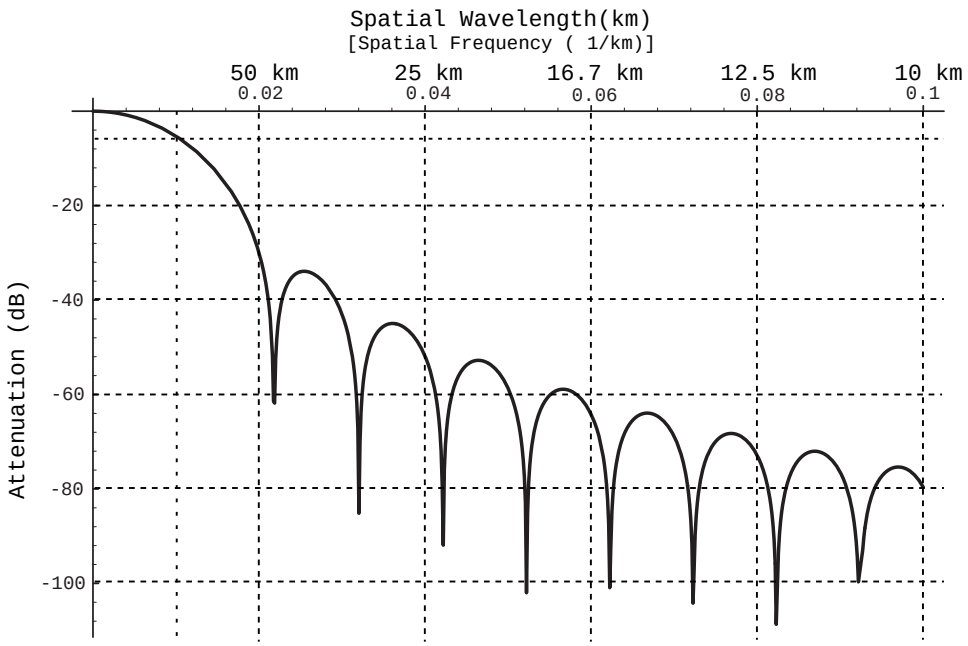


Figure A.3: The result of the numerical evaluation of the integral in Eq. (A.9). This is the frequency response of the ERS-1 aperture function. The plot is $20 \log_1 0(\|F(\xi_1, \xi_2)\|)$.

profitable approach is a numerical evaluation of the integral. For the purposes of this study, we use the final numerical result to characterize the aperture function.

Using available math tools, the integration can be evaluated and the results are illustrated in Figure A.3 which is a plot of $20 \log_1 0(\|F(\xi_1, \xi_2)\|)$. This should be recognized as a plot of the Hanning window frequency response. The aperture acts as a low pass filter with a 3 dB point at a wavelength of 100 km and the first null at about a wavelength 45 km.

Appendix B

Motion Error Characteristics of the SIR Resolution Enhancement Algorithm

B.1 Introduction

SIR attempts to recover as much of the frequency information as possible from the scatterometer data stream. By using several days worth of data, multiple overlapping cells are algebraically recombined in a weighted iterative process. Resolution enhancement is achieved by backprojecting the data onto a higher resolution grid. The achievable enhancement is determined in part by the overlap characteristics of the cells and the cell shape. Seasat data, for instance, can be enhanced to greater resolution than ERS-1 data because the cell shape is oblong, resulting in a higher spatial sampling in one direction. Recombination of several days of data with various cell orientations results in a greater resolution enhancement than is possible with the larger, circular ERS-1 cell shapes. However, SIR does provide resolution enhancement to increase the use and applicability of ERS-1 data over land and ice.

This appendix addresses the issue of radiometric and frequency error in the SIR algorithm for ERS-1 data. Because SIR was originally developed for use over the Amazon where microwave signatures are relatively stable over periods of days, there is a need to characterize the behavior of the algorithm over the more dynamic polar ice regions. For brevity, we consider only the simple characterization of the SIR algorithm as a low pass filter for surfaces in motion and radiometric accuracy of SIR compared to nominal ERS-1 resolution.

B.2 Methodology

Simulated data is used to characterize the SIR imagery. High resolution test scenes are processed using cell location, azimuth and incidence angle information extracted from actual ERS-1 scatterometer data. The cell response shown in Fig. 4.3 is used in calculating the simulated σ^o values. For the purposes of this simulation, we assume that this response is exact and that there is no location error in the latitude and longitude extracted from the actual ERS-1 data.

We model σ^o as a linear function of incidence angle:

$$\sigma^o_{dB} \quad = \quad \mathcal{A} + \mathcal{B}(\theta - 40^\circ) \tag{B.1}$$

where θ has been normalized to 40°, the median incidence angle in the data. Each test scene consists of two colocated images: an $\mathcal{A}$ image and a $\mathcal{B}$ image. In this study, the test scenes are at 1 km/pixel resolution. For any given cell measurement, we assume that the incidence angle is identical for each pixel in the footprint. For each pixel, the σ^o value for that pixel is calculated based on the $\mathcal{A}$ and $\mathcal{B}$ values from the test images and the incidence angle associated with the cell using (B.1). The cos^2 weighting function is applied and the weighted average of the pixels within a footprint represent a linear estimate of σ^o for the test 'surface'.

The simulated data stream is then processed with SIR. Surface plots of sample input and output image are shown in Fig. B.1. The test scenes used in this study are two-level scenes as illustrated in Fig. B.1(A). Step sizes vary from 2 dB to 20 dB, but the results for various step sizes have negligible differences.

A total of 5 days of location, azimuth and incidence data are extracted from the ERS-1 data set. Motion scenes are created using five $\mathcal{A}$ and $\mathcal{B}$ image sets, one for each day. The scenes are considered stationary over each day period. Motions of 2 km/day, 5 km/day, 10 km/day, 15 km/day and 30 km/day are considered, although a sustained speed of 30 km/day in sea ice is unlikely.

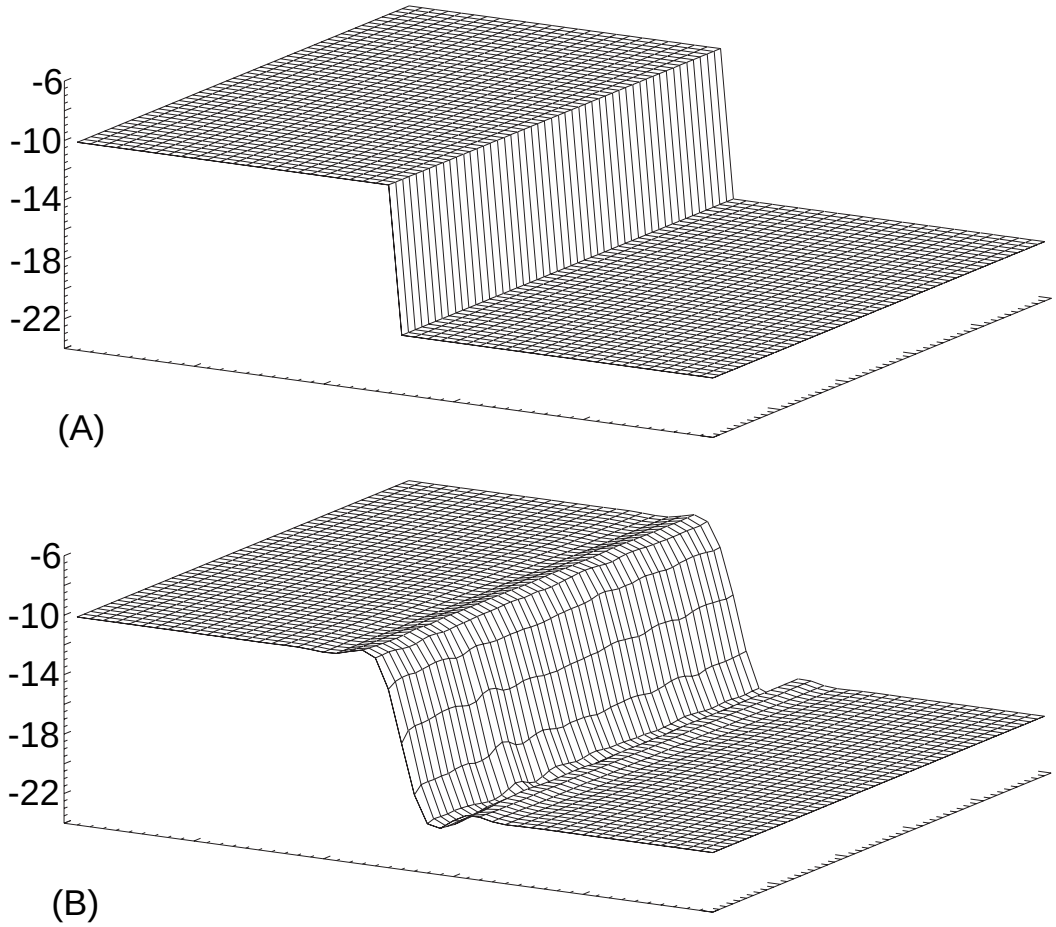


Figure B.1: *Example test image and SIR output. Plot (A) is the original single step image. Plot (B) is the output of SIR.*

B.2.1 Comparison images

Comparison of SIR imagery to low pass filtered versions of the original high resolution imagery is used to assess the SIR algorithm performance. Each test scene in this study is low pass filtered to create a 'truth' image for comparison with the SIR output. The low pass filter used to create comparison 'truth' images is ideal with a spatial cutoff frequency equivalent to the resolution of the SIR imagery. For motion over multiple days, each day's image is low pass filtered and the truth image is created by averaging the individually filtered daily images.

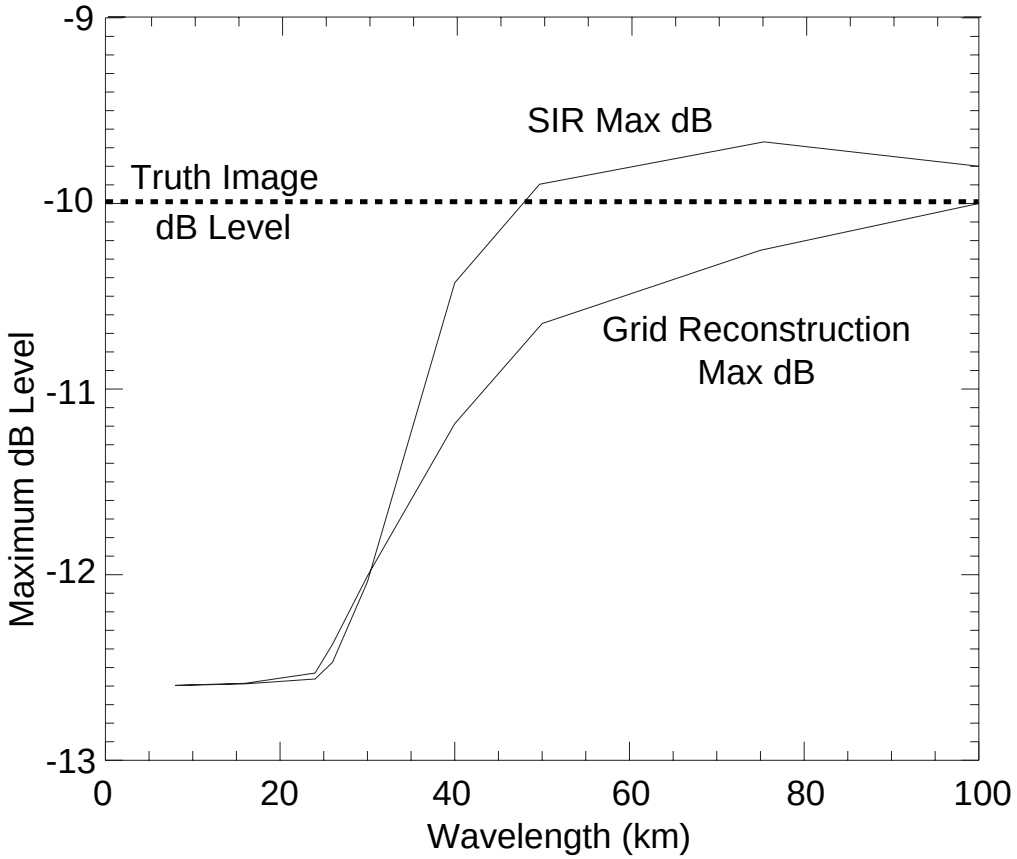


Figure B.2: *Convergence of the maximum dB levels for the SIR algorithm and a simple single pass reconstruction. Note SIR becomes more accurate relative to the true dB level faster than the simple grid reconstruction even with the overshoot.*

B.3 Results

Because the ERS-1 footprint extends beyond the nominal 25km grid spacing (approximately a 100 km diameter circle), significant reduction of the spatial frequency content of the original image is expected at wavelengths less than 100km. The SIR algorithm enhances the frequency content and results in a reconstructed image that is accurate at higher spatial frequencies than the nominal data provides. This is illustrated in Fig. B.2. The SIR image converges to within 0.5dB of the true dB level at wavelengths of around 40km, although there is a slight overshoot. The idealized data does not reach this level of accuracy for wavelengths less than about 60km.

B.3.1 Motion

Fig. B.3 shows profile comparisons of SIR imagery with the low pass 'truth' images. Note that as the magnitude of the motion increases, the overshoot and undershoot of the truth images are reduced by the averaging of the individually filtered day images, but that the transition region between the high and low dB levels is the same for both SIR and the low pass truth images. Also note that the SIR images have less over- and undershoot and ripple than the low pass truth images.

Figs. B.4(A) and B.4(B) show a comparison of 15 km/day motion and 30 km/day motion. The ripples in the 30 km/day transition region are because the scatterometer passes over the regions at various ascending and descending angles which result in unique distributions of overlapping cells in any given scene. Because SIR uses multiple overlapping cells to reconstruct the images, areas in a scene that have limited or reduced cell coverage may not be as accurately reconstructed. In Fig. B.4(A), where the underlying motion is more reasonable, the textural anomalies in the transition region are less significant.

B.4 Summary

The ability of the SIR reconstruction algorithm to create accurate images at higher spatial frequencies is illustrated. SIR provides more accurate dB estimates at lower wavelengths resulting in an approximately 33% resolution gain for errors less than 0.5 dB when compared to simple grid reconstructions even with the severe limitations of the ERS-1 cell response. The SIR algorithm also produces images of stationary scenes and scenes in motion that are similar to low pass filtered versions of the original scenes allowing the algorithm to be characterized as a low pass filter. Because of the non-linearities of the SIR algorithm, the overshoot and undershoot characteristic of low pass filtering are reduced in SIR imagery. This suggests that SIR is more effective than simple construction of imagery from raw ERS-1 scatterometer data.

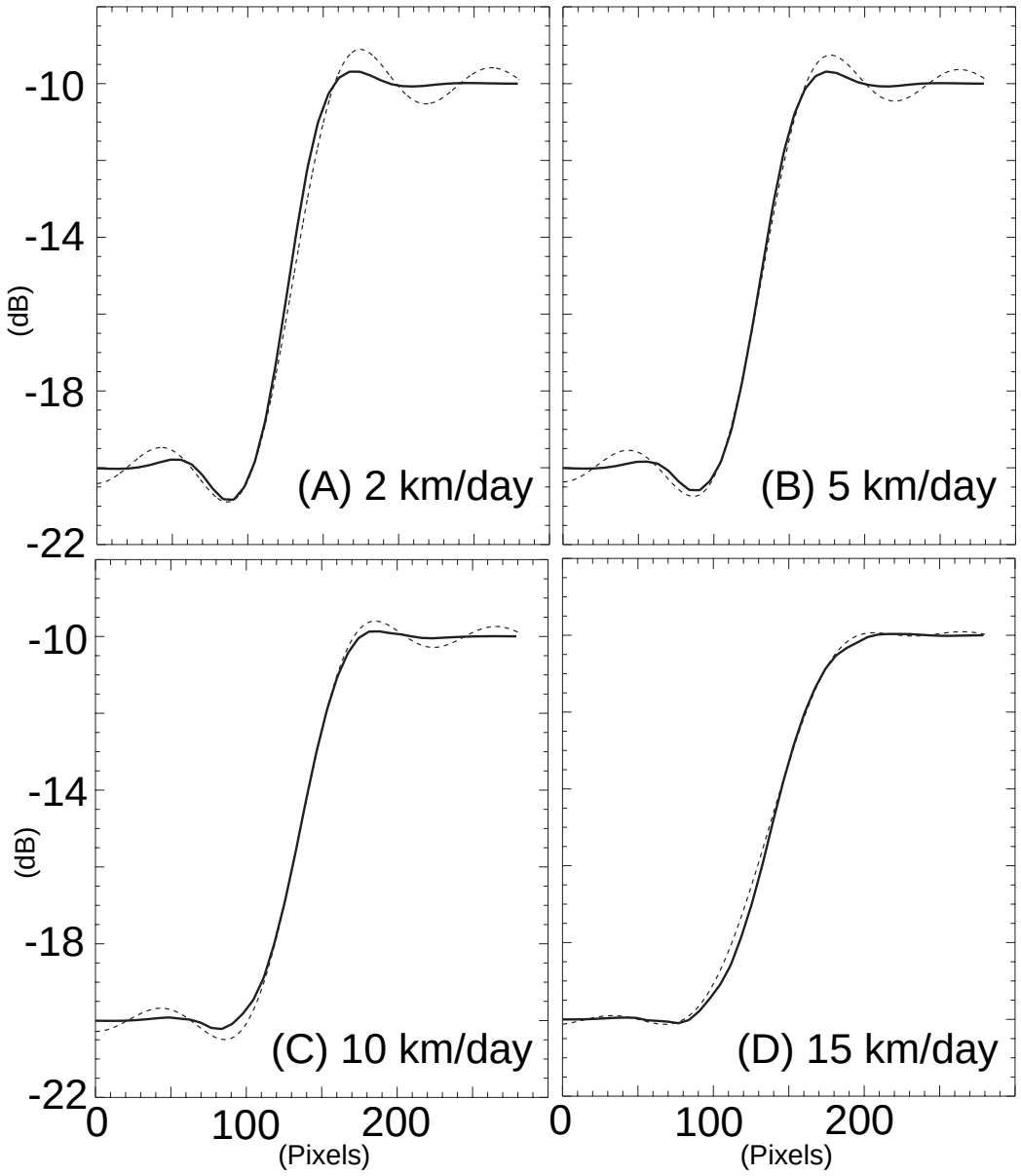


Figure B.3: *Cross sections of several motion examples. The solid lines are SIR image profiles, the dotted lines are low pass filtered truth image profiles.*

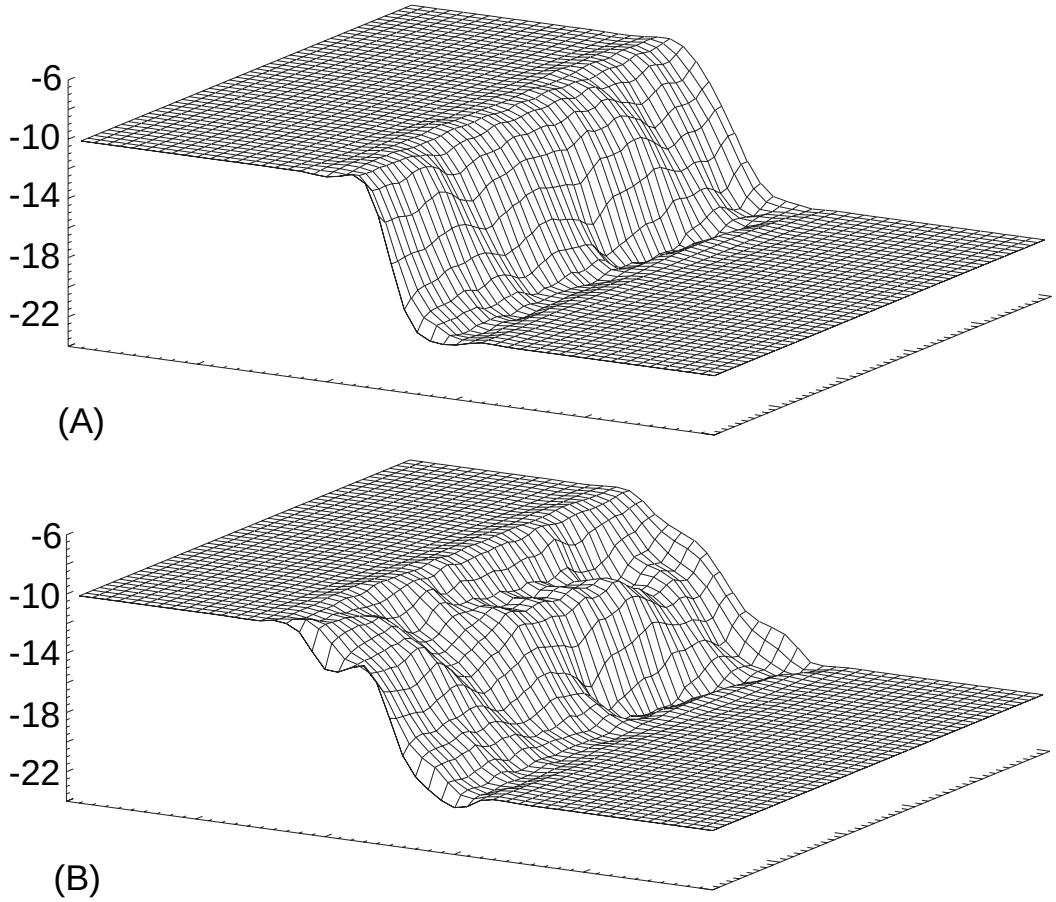


Figure B.4: *Comparison of two motion images. Plot (A) is 15 km per day motion, plot (B) is 30 km per day. Note the textural similarities in the transition regions and the very flat response in the extremities.*

Appendix C

Effects of Projection Selection: Comparison of the Polar Stereographic and Azimuthal Equal Area Projections

C.1 Introduction

The appropriate use of a selected map projection when mapping satellite data is an important step in creating statistically significant data sets. In recent years there has been an increase in availability of high quality medium and large scale satellite data sets covering many geophysically interesting global surfaces. The use of this data for statistical analysis of surface conditions has great potential in increasing our understanding of earth processes, particularly in polar regions.

The statistical analysis of satellite data will center on areal classifications and interannual or interseasonal comparison of areal change. However, the common projection used for polar ice study is the polar stereographic, which introduces significant areal distortion away from the projection origin as explained in the following sections. Errors will arise if large or medium scale data from radiometers or scatterometers is used for statistical analysis of the polar sea ice characteristics and the satellite data is not corrected for areal distortions introduced by map projections. Small scale data, such as SAR imagery, will in general be unaffected by the projection used since the origin of the projection is within the image.

This paper outlines some of the issues involved with selection of map projections and the associated errors. While we concentrate on issues in polar ice study, the ideas are applicable in any field using large scale satellite data. The first section outlines some earth model issues and the errors introduced by using different earth

models. The second section introduces two map projections, the polar stereographic and Lambert equal area projections, and outlines the primary differences between them as they pertain to polar ice study. This is followed by an analysis of the error between these two projections and possible error corrections.

C.2 The Projections

The problem of map projection lies in the placement of a spherical surface on the flat surface of a map. There are many projections available to map a sphere to a flat surface, each having desirable characteristics as well as weaknesses. The task of the map maker is to evaluate the needs of those using the map and choose an appropriate projection.

In Antarctic and Arctic maps, the stereographic projection is commonly used. The stereographic projection is also used for the DMSP SMMR and the NIMBUS-7 ice concentration maps, making it a historically important map projection in the context of decadal remote sensing studies of the polar regions. Work using Ku-Band and C-band scatterometers for imaging of land has used a Lambert Azimuthal Equal-Area projection (CITE US)(see [73] for a detailed description of the Lambertian projection). However, the projection of choice for current historic satellite research is the polar stereographic projection. In this section, both the Lambert Equal Area and Polar Stereographic projections are outlined. Both will be used for the numerical analysis of projection accuracy.

C.2.1 Stereographic Projection

The geometry of the Lambert equal area projection is shown in Figure C.1. The figure illustrates a projection for the North pole, where the origin is the North pole.

In Figure C.1, r, the radial distance from the origin of the transform (in polar aspects, the origin is one of the poles and is the North pole in the figure), is given by $r = 2R \tan \frac{90-\phi}{2}$. ϕ is the angle of latitude and R is the radius of the earth at the origin of the projection, in this case the polar radius of the earth.

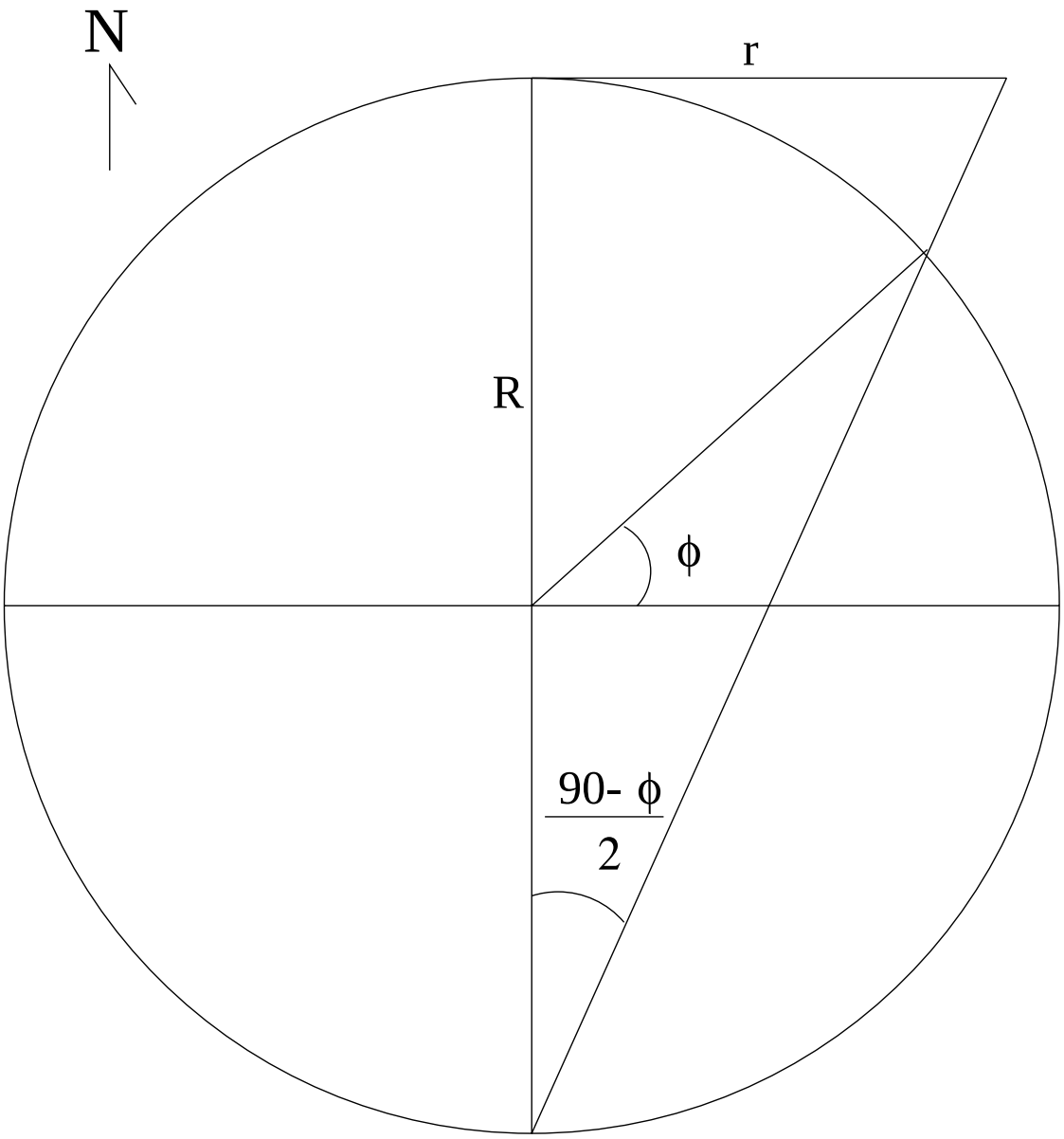


Figure C.1: *Geometry of the stereographic projection.*

The stereographic projection is conformal. Conformal projections maintain the angular separation of lines from the surface to the map. A small intersection at some angle θ on the earth will have the same angle in the projection. Thus, relative shape and angle are maintained through the projection while area is distorted. This type of projection is desirable in maps used for plotting courses or where shape is more important than size. Unfortunately, the projection does not have equal area (see the discussion on Lambert transforms below).

The stereographic projection is an easily visualized geometric projection. As shown in Figure C.1, the geometry of the projection is easily illustrated, and the equations for the projection both forward and backward are simple even for an elliptical earth model, making this projection computationally conservative.

C.2.2 Lambert Equal Area Projection

The geometry of the Lambert equal area projection is shown in Figure C.2. The figure illustrates a North polar aspect. In Figure C.2, r, the radial distance from the origin of the transform (in polar aspects, the origin is one of the poles and is the North pole in the figure), is given by $r = 2R \sin \frac{90-\phi}{2}$.

The Lambert transform is not conformal, thus relative shape is not preserved in the projection, but it is equal area. This means that any shape placed on the map at two different points will cover the same amount of area.

Because the Lambert transform preserves area, it will be used as a point of comparison for numerical analysis of the Polar Stereographic projection. Ice typing and classification will rely on accurate areal measurements, and by using an equal area projection, the need to correct for pixel area distortion for different regions of the map is avoided. However, the Lambert projection is not commonly used because of the difficulty of inverting the projection.

C.2.3 Ellipsoidal Model of the Earth

Projection algorithms are based on a model of the earth, usually a sphere or an oblate spheroid. The earth has a slightly shorter radius at the pole than at

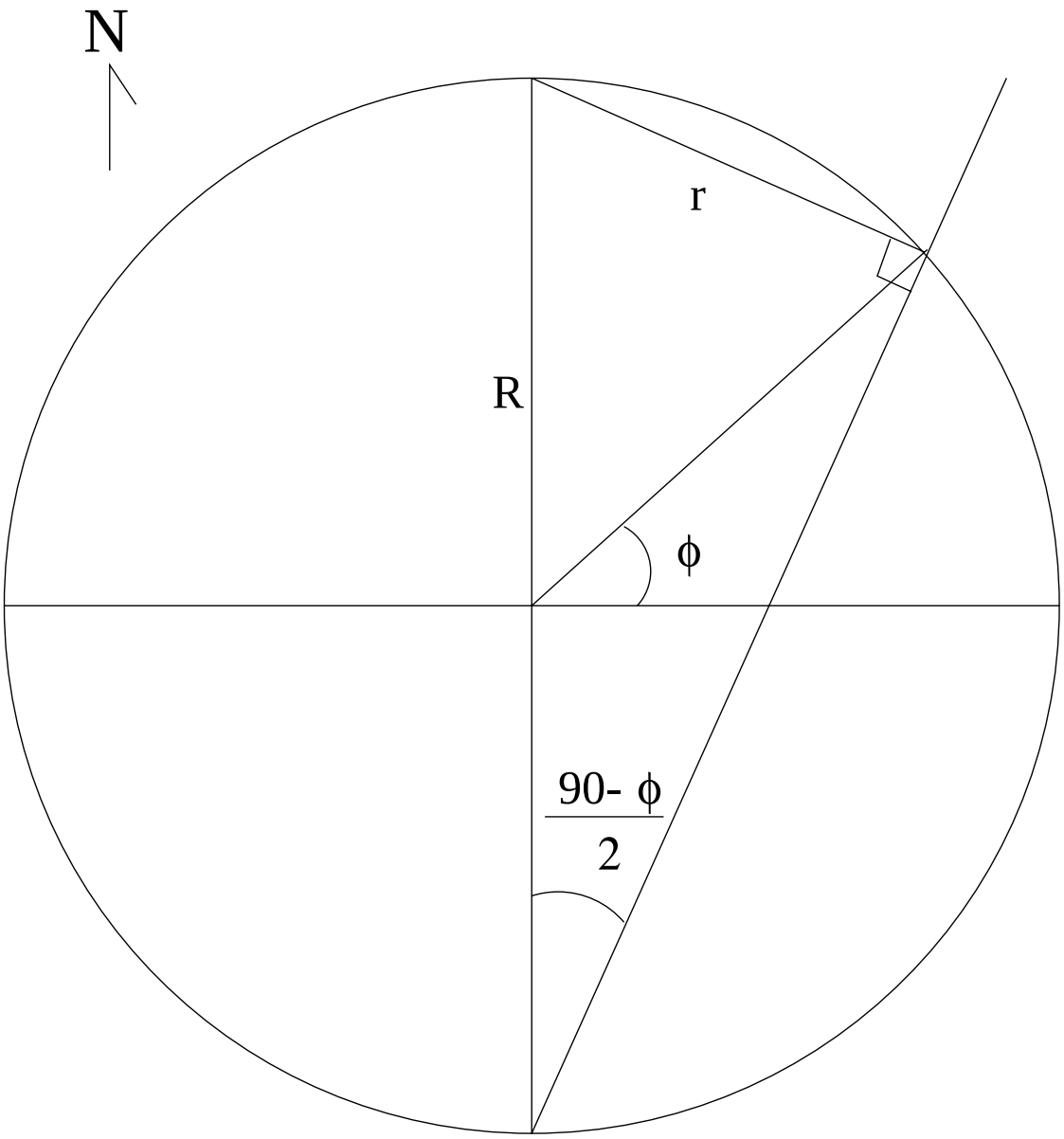


Figure C.2: *Geometry of the Lambert equal area projection.*

the equator making a simple sphere a poor model of the earth. A poor earth model will introduce error into any classification scheme because the data will be distorted by the map projection. Generally, since the oblate spheroid is closer to the actual shape of the earth, projections using the oblate spheroid model will be closer to actual ground truth than a simple spherical model.

Huges Ellipsoid

In a 1980 Hughes Aircraft SSM/I computer program specifications manual, an ellipsoid is defined that is used for SSM/I data and several other satellite systems [74]. No origin for this ellipsoid is available in the literature, and the ellipsoid differs slightly from other published ellipsoids (See Table for a listing of the Hughes and other ellipsoid models), but it is commonly used and is the model of choice for this study. It will be shown later that minor variations in the choice of ellipsoid have little quantitative impact on the resulting projection.

Ellipsoid	Eq. Rad.	Polar Rad.	1/f
Hughes	6378.273	6356.889	298.279
Hough	6378.270	6356.794	297
IUGG 67	6378.160	6356.775	298.25
WGS 72	6378.135	6356.330	298.26
IUGG 75	6378.140	6356.755	298.257
WGS 84	6378.137	6356.752	298.257

C.2.4 The Ellipsoidal Model vs. Spherical

Just to illustrate the magnitude of the error associated with a non-ellipsoidal model of the earth, consider a Lambert projection near the origin, with $\phi = 80^o$ as shown in Figure C.2. Calculating the radial distance r as illustrated in Figure C.2 using both a spherical and Ellipsoidal model we get

$$r_{\text{Spherical}} = 1111.83\text{km} \quad \text{(C.1)}$$

$$r_{\text{Ellipsoidal}} = 1108.03. \quad \text{(C.2)}$$

These numbers reveal an error of about 4km at latitudes close to the origin, but the error increases to over 10km for latitudes over 60° which is near the extent of polar ice in both the Northern and Southern Hemispheres. The error will, of course, vary with the projection used.

C.3 Analysis

The next sections deal with exact analysis of error between the map projections outlined in the previous section and the actual area on the spheroid. No pixelization or quantization error is considered.

C.3.1 Concentric Rings

This part of the study concentrates on changes in area of concentric rings around the pole. Beginning at 80° south and continuing through 50° south, 1° in latitude rings are created centered on the south pole. A true earth area is calculated for each ring and then also for the polar stereographic projection. The first parameter is the percentage error between true and polar stereographic ring areas. The error is given by

$$100 \times (\frac{A_{Pol} - A_{True}}{A_{True}}) \tag{C.3}$$

and is shown in Figure C.3. Note that the error is negative below the polar stereographic origin latitude of 70° . This is consistent with calculations from the previous section showing a contraction of area in this region for the polar stereographic projection. This plot illustrates that above 57° south, there is greater than 10% error between the true and projected areas. Figure C.4 shows how this translates into absolute area. Above 57° south, the absolute area difference is greater than 250 000 square kilometers, quickly rising to 500 000 square kilometers at 52° south, the northernmost extent of the Southern Ocean sea ice sheet.

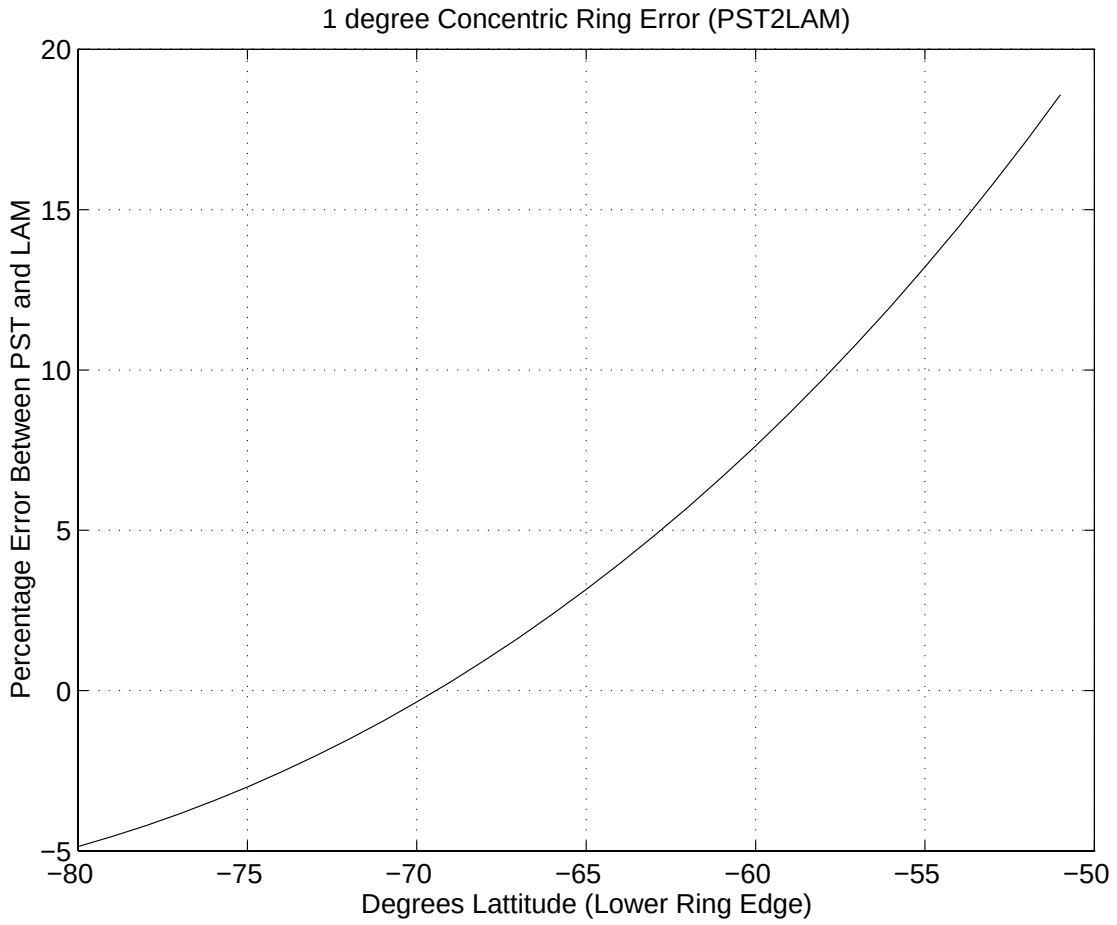


Figure C.3: *Percentage error between a polar stereographic projection and the true area. The area are polar concentric circles for ease of area calculation. The polar stereographic projection has on origin latitude of 70° south.*

C.3.2 Features in Motion

Another concern the monitoring of objects in motion in satellite imagery. In most studies that concentrate on land features, movement is not a concern. However, sea ice at the poles is very dynamic and can move with velocities of several kilometers per day. Because of this, a feature studied and classified during one time period may move significantly during the annual freeze-melt cycle. The same feature at low lattitudes early in the winter may move to higher lattitudes in a matter of a few weeks or days.

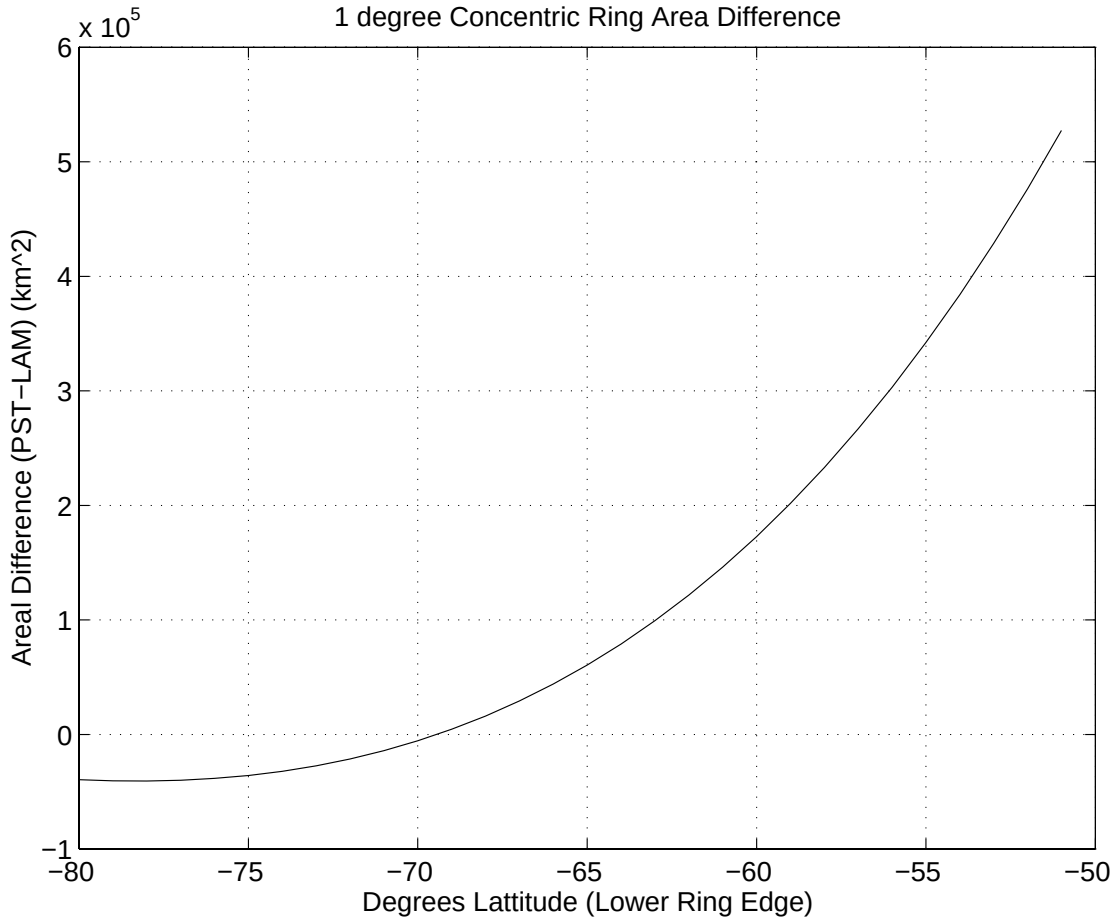


Figure C.4: *Absolute difference between true areas and polar stereographic errors in square kilometers.*

The change in location of an object brings a change in the projected size of that object under the polar stereographic projection. A square, constant area region was moved through a series of 1° steps from the polar stereographic origin of 70° south to 55° south. The original region was chosen to be 4° in latitude by 4° in longitude.

Figure C.5 illustrates how much larger an object can appear under the polar stereographic projection. Near the origin, distortion is minimal, but quickly increases to over 10% at 13 degrees from the origin.

The implications of areal distortion are most vivid for sea ice studies of divergence and convergence. In essence, the uncorrected areal distortion of ice mass

moving away from the projection origin appears to be a divergence of the ice mass. If the pixel area is not corrected, calculations of sea ice divergence or convergence will be highly inaccurate. Sea ice typing will also yield inexact results if the areal distortion is not corrected before percentage changes or type comparisons are made.

C.3.3 Areal Distortion Correction

Because the polar stereographic projection is used in many areas of remote sensing and for many historical data sets, and because the forward and reverse algorithm for both a sphere and spheroidal model of the earth are easily implemented, it will continue to be the transform of choice. Thus, we need to develop a correction factor that can be used to correct the areal distortion introduced by the algorithm. One such factor is introduced in [8], but due to a apparent printing error, there is a significant error in the equation presented. Since no background is provided, we derive a correction factor here. The correction factor will be based on equations and background material from [8] and [73]. For a complete discussion of map projections in general, the reader is refered to [73].

For a conformal projection, the areal distortion m_a is the square of the linear distortion m, given by [73]

$$m = \frac{\rho c}{R \cos \phi} \tag{C.4}$$

$$m_a = m^2 \tag{C.5}$$

$$R = \frac{R_e R_p}{\sqrt{R_p^2 \cos \phi^2 + R_e^2 \sin \phi^2}} \tag{C.6}$$

where ρ is the distance from the pole on the map projection, $c = 1$ for the polar stereographic projection, R is the radius of the sphere or spheroid at the latitude of the point being projected, R_e and R_p are the equatorial and polar radii respectively and ϕ is the latitude. For an unmodified polar stereographic projection, ρ is given by

$$\rho = 2R_p \tan \frac{90^o - \phi}{2}. \tag{C.7}$$

However, the projection we are considering has been modified so that the projection origin is at 70^o latitude, so instead of $2R_p$, we need to use the shorter distance shown

in Figure C.1 as $R_p + R_{70^o} \sin 70^o$. Substituting this value into Eq. (C.7) and then into Eq. (C.5), we get as an areal distortion factor for the modified polar stereographic projection:

$$m_{70^2} = \frac{(R_p + R_{70^o} \sin 70^o)^2 (R_p^2 \cos \phi^2 + R_e^2 \sin \phi^2) \tan \frac{90^o - \phi}{2}}{R_e^2 R_p^2 \cos phi^2} \quad \text{(C.8)}$$

$$R_{70^o} = \frac{R_e R_p}{\sqrt{R_p^2 \cos 70^{o2} + R_e^2 \sin 70^{o2}}}. \quad \text{(C.9)}$$

C.4 Summary

As the research community works to make better use of medium to large scale satellite data, careful consideration of the projection used needs to be made. The current use of the polar stereographic projection requires areal correction based on location relative to the projection origin in order for statistical evaluations to remain accurate. Even with the origin of the projection placed at a mid latitude of polar imagery (70° south), this analysis shows that errors of 5% or more are possible in various areal comparisons or classification. Since changes on the order of a few percent are considered important for global climate modeling, it is imperative that the scientific community be aware of the implications of areal distortion in satellite imagery

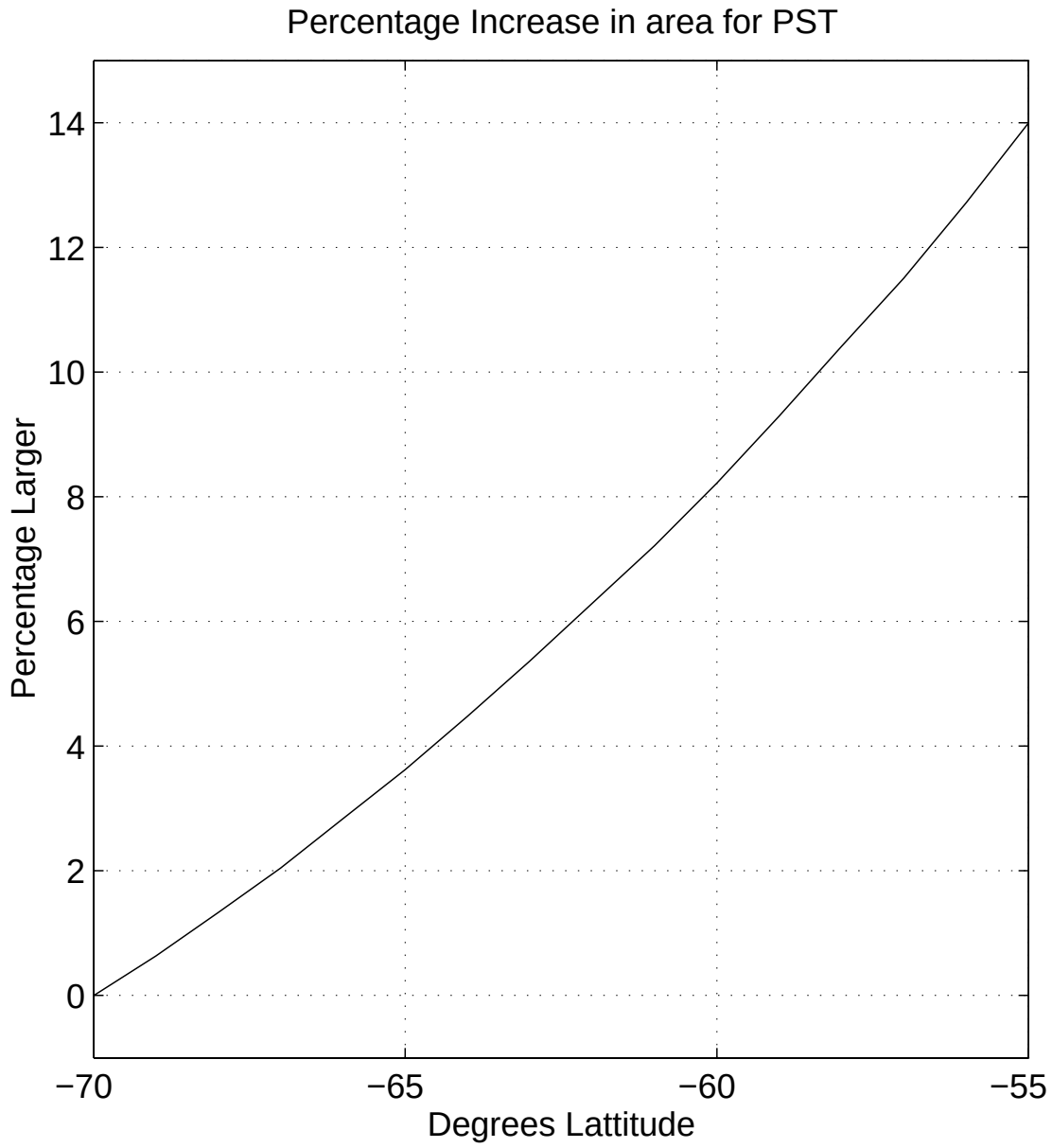


Figure C.5: *Percentage increase of a constant area, square region compared to an 'origin' region at 70° south. The region was moved in 1° steps for each comparison.*

Bibliography

[1] W. F. Weeks and S. F. Ackley, "The Growth, Structure, and Properties of Sea Ice", in *The Geophysics of Sea Ice*, Norbert Untersteiner, Ed., chapter 1. Plenum Press, Washington, DC, 1986.

[2] 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, 1991.

[3] D. Long, P. Hardin, and P. Whiting, "Resolution Enhancement of Spaceborne Scatterometer Data", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, pp. 700–715, 1993.

[4] F. Carsey, "Review and status of remote sensing of sea ice", *IEEE Journal of Oceanic Engineering*, vol. 14, pp. 127–137, 1989.

[5] B. Holt, R. Kwok, and E. Rignot, "Status of the ice classification algorithm in the alaska sar facility geophysical processor system", in *Proceedings of IGARSS 90*. IEEE, 1990, pp. 2221–2224.

[6] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing*, vol. 1, Artech House, Boston, 1981.

[7] J. C. Comiso, "Characteristics of arctic winter sea ice from satellite multispectral microwave observations", *Journal of Geophysical Research*, vol. 91, no. C1, pp. 975–994, 1986.

[8] P. Gloersen, W. J. Campbell, D. J. Cavalieri, J. C. Comiso, C. L. Parkinson, and H. J. Zwally, *Arctic and Antarctic Sea Ice, 1978–1987: Satellite Passive–Microwave Observations and Analysis*, NASA, Washington, D.C., 1992.

[9] P. Lecomte, M. Davison, and A. Cavanie, "Ice Boundary Mapping using ERS-1 Scatterometer Data", *Earth Observation Quarterly*, vol. 40, pp. 8, 1993.

[10] D. Long and M. Drinkwater, "Greenland Ice-Sheet Surface Properties Observed by the Seasat-A Scatterometer at Enhanced Resolution", *Journal of Glaciology*, vol. 40, no. 135, pp. 213–230, 1994.

[11] P. Lecomte, A. Cavanie, and F. Gohin, "Recognition of Sea Ice Zones using ERS-1 Scatterometer Data", in *Proceedings of IGARSS 93*. IEEE, 1993, pp. 855–857.

[12] F. Gohin and A. Cavanie, "A first try at identification of sea ice using the three beam scatterometer of ers-1", *International Journal of Remote Sensing*, vol. 15, pp. 1221–1228, 1994.

[13] P. Whiting, "Resolution Enhancement of SEASAT Scatterometer Data", Master's thesis, Brigham Young University, 1992.

[14] P. Whiting and D. Long, "Resolution Enhancement of Seasat Scatterometer Data", Tech. Rep. TR-L101-92.1, Brigham Young University, 1992.

[15] R. G. Kennett and F. K. Li, "Seasat Over-land Scatterometer Data, Part I: Global Overview of Ku-Band Backscatter Coefficients", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27, pp. 592–605, 1989.

[16] R. K. Moore and A. K. Fung, "Radar Determination of Winds at Sea", *Proceedings of The IEEE*, vol. 65, pp. 1504–1521, 1979.

[17] C. E. Livingstone and M. R. Drinkwater, "Springtime C-Band SAR Backscatter Signatures of Labrador Sea Marginal Ice: Measurements Versus Modeling Predictions", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, no. 1, pp. 29–41, 1991.

[18] M. Drinkwater, D. Long, and D. Early, "ERS-1 Investigations of Southern Ocean Sea Ice Geophysics Using Combined Scatterometer and SAR Images", in *Proceedings of IGARSS 94*. IEEE, 1994, pp. 165–167.

[19] M. Pettersson, J. Grandell, A. Carlström, J. Pallonen, L.M.H. Ulander, and M. Hallikainen, "Analysis of C-Band Backscatter Measurements of Thin Arctic Sea Ice", in *Proceedings of IGARSS 95*. IEEE, 1995, pp. 360–362.

[20] R. Gordon, "A Tutorial on ART", *IEEE Transactions on Nuclear Science*, vol. NS-21, pp. 78–93, 1971.

[21] Y. Censor, "Finite series-expansion reconstruction methods", *Proceedings of the IEEE*, vol. 71, no. 3, pp. 409–419, March 1983.

[22] T. Elfving, "On Some Methods for Entropy Maximization and Matrix Scaling", *Linear Algebra and its Applications*, , no. 34, pp. 321–339, 1980.

[23] H.J. Trussell, "Convergence Criteria for Iterative Restoration Methods", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, pp. 129–136, 1983.

[24] G. T. Herman, A. Lent, and S. W. Rowland, "ART: Mathematics and Applications", *Journal of Theoretical Biology*, pp. 1–32, 1973.

[25] K. A. Dines and R. J. Lytle, "Computerized Geophysical Tomography", *Proceedings of The IEEE*, , no. 67, pp. 1065–1073, 1979.

[26] A. V Oppenheim and A. S. Willsky, *Signals and Systems*, Prentice Hall, Englewood Cliffs, NJ, 1983.

[27] H. Freichtinger and K. Gröchenig, "Iterative Reconstruction of Multivariate Band-limited Function from Irregular Sampling Values", *SIAM Journal of Mathematical Analysis*, vol. 23, no. 1, pp. 244–261, 1992.

[28] K. Gröchenig, "Reconstruction Algorithms in Irregular Sampling", *Mathematics of Computation*, vol. 59, no. 199, pp. 181–194, 1992.

[29] D. S. Early and D. G. Long, "Error and Resolution Characteristics of the SIRF Resolution Enhancement Algorithm", in *Proceedings of IGARSS 96*. IEEE, 1996, pp. 124–126.

[30] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[31] Brian Jeffs, "Private Communication", BYU, 1997.

[32] R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*, Kluwer Academic Publishers, Boston, 1991.

[33] D. Daum, D. Long, and W. Davis, "Reconstructing Enhanced Resolution Images From Spaceborne Microwave Sensors", in *Proceedings of IGARSS 94*. IEEE, 1994, pp. 2231–2233.

[34] D. Long and D. Daum, "Spatial Resolution Enhancement of SSM/I Data", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 2, pp. 407–417, 1997.

[35] D. S. Early and D. G. Long, "A Simple Ice Classification Algorithm", in *Proceedings of IGARSS 97*. IEEE, 1997.

[36] M. Drinkwater, D. Long, and D. Early, "Enhanced Resolution ERS-1 Scatterometer Imaging of Southern Ocean Sea Ice", *ESA Journal*, vol. 17, pp. 307–322, 1993.

[37] D. Long, D. Early, and M. Drinkwater, "Enhanced Resolution ERS-1 Scatterometer Imaging of Southern Hemisphere Polar Ice", in *Proceedings of IGARSS 94*. IEEE, 1994, pp. 156–158.

[38] D. Early, D. Long, and M. Drinkwater, "Comparison of Enhanced Resolution Images of Greenland from the ERS-1 and Seasat Scatterometers", in *Proceedings of IGARSS 94*. IEEE, 1994, pp. 2382–2384.

[39] D. Early and D. Long, "A Study of C-Band σ^o Azimuthal Modulation over Antarctic Sea Ice using ERS-1", in *Proceedings of IGARSS 95*. IEEE, 1995, pp. 156–158.

[40] M. Drinkwater, D. Long, and D. Early, "Comparisons of Variations in Seasonal Sea-Ice Formation in the South-Western Weddell Sea with Seasonal Bottom-Water Outflow Data", in *Proceedings of IGARSS 95*. IEEE, 1995, pp. 402–404.

[41] M. Drinkwater, D. Long, and D. Early, "Enhanced Resolution ERS-1 Scatteromter Imaging of Antarctic Ice", *Earth Observation Quarterly*, vol. 43, pp. 4–6, 1994.

[42] E. Attema, "Private Communication", 1996.

[43] D. Early, D. Long, and M. Drinkwater, "Azimuthal Modulation of C-Band Scatterometer σ^o over Southern Ocean Sea Ice", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 1201–1209, 1997.

[44] D. Long, "Private Communication", BYU, 1997.

[45] F. Remy, Ledroit, and J. F. Minster, "Katabatic Wind Intensity and Direction over Antarctica Derived from Scatterometer Data", *Geophysical Research Letters*, vol. 19, pp. 1021–1024, 1992.

[46] A. R. Hosseinmostafa, V. I. Lytle, K. C. Jezek, S. P. Gogineni, S. F. Ackley, and R. K. Moore, "Comparison of Radar Backscatter from Antarctic and Arctic Sea Ice", *Journal of Electromagnetic Waves and Applications*, vol. 9, no. 3, pp. 421–438, 1995.

[47] M. R. Drinkwater, R. Hosseinmostafa, and W. Dierking, "C-band microwave backscatter of sea ice in the Weddell Sea During the winter of 1992", in *Proceedings of IGARSS 93*. IEEE, 1993, pp. 446–448.

[48] J. Comiso, T. Grenfell, M. Lange, A. Lohanick, R. Moore, and P. Wadhams, "Microwave Remote Sensing of the Southern Ocean Ice Cover", in *Microwave Remote Sensing of Sea Ice*, F. D. Carsey, Ed., chapter 12. American Geophysical Union, Boston, 1992.

[49] M. A. Lange, S. F. Ackley, P. Wadhams, G. S. Diekmann, and H. Eicken, "Development of Sea Ice in the Weddell Sea", *Annals of Glaciology*, , no. 12, pp. 92–96, 1991.

[50] M. R. Drinkwater and C. Haas, "Snow, Sea-ice, and Radar Observations During ANT X/4: Summary Data Report", Tech. Rep. 53, Alfred Wegener Institut, July 1994.

[51] V. I. Lytle and S. F. Ackley, "Sea Ice Ridging in the Eastern Weddell Sea", *Journal of Geophysical Research*, vol. 96, no. 10, pp. 18411–18416, 1991.

[52] W. Dierking, "Laser Profiling of the Ice Surface Topography During the Winter Weddell Gyre Study 1992", *Journal of Geophysical Research*, vol. 100, no. C3, pp. 4807–4820, 1995.

[53] A. K. Liu and S. Häkkinen, "Wave Effects on Ocean-Ice Interaction in the Marginal Ice Zone", *Journal of Geophysical Research*, vol. 98, no. C6, pp. 10025–10036, 1993.

[54] M. Drinkwater, "Private Communication", 1996.

[55] M.H. Freilich and R.S. Dunbar, "A Preliminary C-band Scatterometer Model Function for the ERS-1 AMI Instrument", in *Proceedings of the First ERS-1 Symposium*, Cannes, France, 1993, ESA, pp. 79–84, SP-359.

[56] P. Lecomte, M. Davison, and A. Cavanie, "Ice Boundary Mapping using ERS-1 Scatterometer Data", *Earth Observation Quarterly*, , no. 40, pp. 8, 1993.

[57] M. R. Drinkwater, "Satellite Microwave Radar Observations of Antarctic Sea Ice", in *Recent Advances in the Analysis of SAR for Remote Sensing of the Polar Oceans*, C. Tsatsoulis and R. Kwok, Eds., chapter 8. Springer Verlag, Berlin, 1998.

[58] S. Martin, K. Steffen, J. Comiso, D. Cavalieri, M. Drinkwater, and B. Holt, "Microwave Remote Sensing of Polynyas", in *Microwave Remote Sensing of Sea*

Ice, F. D. Carsey, Ed., chapter 15. American Geophysical Union, Boston, 1992, Geophysical Monograph 68.

[59] M. R. Drinkwater and C. Haas, "Sea-ice conditions during the Winter Weddell Gyre Study 1992 ANT X/4 and RV "Polarstern": Shipboard observations and AVHRR satellite imagery", Tech. Rep. 34, Alfred Wegener Institut, December 1992.

[60] D. G. Long and D. S. Early, "Enhanced Resolution Scatterometer Imaging with Irregular Samples", in *Proceedings of the SPIE 3117*. SPIE, 1997.

[61] D. S. Early and D. G. Long, "Enhanced Resolution Imaging from Irregular Samples", in *Proceedings of IGARSS 97*. IEEE, 1997, pp. 1844–1846.

[62] D. Early and D. Long, "A Theory of Enhanced Resolution Imagery from Scatterometers", *IEEE Transactions on Geoscience and Remote Sensing*, To be Submitted for Publication, Spring 1998.

[63] Q. Remund, "Private Communication", 1997.

[64] D. G. Long and P. J. Hardin, "Vegetation Studies of the Amazon Basin Using Enhanced Resolution Seasat Scatterometer Data", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 449–460, 1994.

[65] D. G. Long, D. Daum, and P. Hardin, "Spatial Resolution Enhancement of SSM/I Data: Vegetation Studies of the Amazon Basin", in *Proceedings of IGARSS 96*. IEEE, 1996, pp. 1606–1608.

[66] P. Hardin, D. Long, and Q. Remund, "Discrimination of Africa's Vegetation Using Reconstructed ERS-1 Imagery", in *Proceedings of IGARSS 96*. IEEE, 1996, pp. 827–829.

[67] P. Hardin, D. Long, Q. Remund, and D. Daum, "A Comparison of Reconstructed Ku-Band Scatterometery, C-Band Scatterometery, and SSM/I Imagery fro Tropical Vegetaion Classification", in *Proceedings of IGARSS 96*. IEEE, 1996, pp. 848–850.

[68] D. Long and M. Drinkwater, "Greenland Ice Sheet Surface Properties Observed by the Seasat-A Scatterometer at Enhanced Resolution", *Journal of Glaciology*, vol. 40, pp. 213–230, 1994.

[69] D. G. Long and M. R. Drinkwater, "Studies of Greenland Using the NASA Scatterometer", in *Proceedings of the SPIE 3117*. SPIE, 1997.

[70] C. J. Wilson III and D. G. Long, "Analysis of the Canadian Boreal Forest Using Enhanced Resolution ERS-1 Scatterometer Imagery", in *"Prodeedings of the Rocky Mountain NASA Space Grant Consortium 1996 Fellowship Student Symposium"*. University of Utah, Salt Lake City, Utah, 1996.

[71] C. J. Wilson III and D. Long, "Analysis of the Canadian Boreal Forest using Enhanced Resolution ERS-1 Scatterometer Imagery", in *Proceedings of IGARSS 96*. IEEE, 1996, pp. 218–220.

[72] P. J. Hardin, D. G. Long, and R. R. Jensen, "Characterizing Earth's Surface Using Moderate Resolution 14 GHz Scatterometer Imagery: Early Results from NSCAT Reconstruction", in *Proceedings of IGARSS 97*. IEEE, 1997, pp. 1835–1837.

[73] F. Pearson, II, *Map Projections: Theory and Application*, CRC Press, Boca Raton, FL, 1990.

[74] NSIDC Distributed Active Archive Center, *DMSP SMM/I Brightness Temperature and Sea Ice Concentration Grids for the Polar Regions: User's Guide*, Second edition, 1996.