JPL Space Science and Technology Series

Joseph H. Yuen, Series Editor

Synthetic Aperture Radar Polarimetry

Jakob Van Zyl and Yunjin Kim

Jet Propulsion Laboratory
California Institute of Technology

JPL SPACE SCIENCE AND TECHNOLOGY SERIES

Joseph H. Yuen, Editor-in-Chief

Published Titles in this Series

Fundamentals of Electric Propulsion: Ion and Hall Thrusters
Dan M. Goebel and Ira Katz

Synthetic Aperture Radar Polarimetry
Jakob Van Zyl and Yunjin Kim

Synthetic Aperture Radar Polarimetry

Synthetic Aperture Radar Polarimetry

Jakob Van Zyl and Yunjin Kim

Jet Propulsion Laboratory
California Institute of Technology

Synthetic Aperture Radar Polarimetry

December 2010

The research described in this publication was carried out at the
Jet Propulsion Laboratory, California Institute of Technology, under
a contract with the National Aeronautics and Space Administration.

Contents

Note from the Series Editor ix
Foreword xi
Preface xiii
Acknowledgments xv
Authors xvii

Chapter 1: Synthetic Aperture Radars (SAR) Imaging Basics 1
1.1 Basic Principles of Radar Imaging 2
1.2 Radar Resolution 6
1.3 Radar Equation 10
1.4 Real Aperture Radar 12
1.5 Synthetic Aperture Radar 14
1.6 Radar Image Artifacts and Noise 17
1.6.1 Range and Azimuth Ambiguities 18
1.6.2 Geometric Effects and Projections 20
1.6.3 Signal Fading and Speckle 22
1.7 Summary 23
References 25

Chapter 2: Basic Principles of SAR Polarimetry 27
2.1 Polarization of Electromagnetic Waves 27
2.2 Mathematical Representations of Scatterers 32
2.3 Implementation of a Radar Polarimeter 39
2.4 Polarization Response 42
2.5 Optimum Polarizations 62
2.5.1 General (Bistatic) Case 63
2.5.2 Backscatter (Monostatic) Case 65
2.5.3 Special Case: Single Scatterer in Backscatter (Monostatic) Case 68
2.5.4 Special Case: Multiple Scatterers with Reflection Symmetry 70
2.5.5 A Numerical Example 73
2.6 Contrast Enhancement 76
2.6.1 Numerical Example 78
2.6.2 Image Example 80
2.7 Summary 82
References 83

Chapter 3: Advanced Polarimetric Concepts 85
3.1 Vector-Matrix Duality of Scatterer Representation 85

3.2 Eigenvalue and Eigenvector-Based Polarimetric Parameters 88
3.2.1 Parameters Used to Describe Randomness in Scattering 89
3.2.2 Alpha Angle 92
3.3 Decomposition of Polarimetric Scattering 98
3.3.1 Scattering Decomposition in the Incoherent Case Using Orthonormal Bases 102
3.3.2 Model Based Scattering Decomposition in the Incoherent Case 107
3.4 Image Classification 139
3.4.1 Supervised Classification 139
3.4.2 Physics-Based Unsupervised Classification 144
3.4.3 Combined Unsupervised and Bayes Classification Algorithms 145
3.5 Polarimetric SAR Interferometry 147
3.6 Summary 154
References and Further Reading 155

Chapter 4: Polarimetric SAR Calibration 159
4.1 Polarimetric Radar System Model 159
4.2 Cross-Talk Estimation and Removal 170
4.3 Co-Polarized Channel Imbalance Calibration 177
4.4 Absolute Radiometric Calibration 183
4.4.1 Effect of Topography on the Scattering Area 187
4.4.2 Effect of Topography on Antenna Pattern Corrections 192
4.4.3 AIRSAR Image Example 194
4.5 Faraday Rotation 195
4.6 Summary 198
References 198

Chapter 5: Applications: Measurement of Surface Soil Moisture 201
5.1 Surface Electrical and Geometrical Properties 202
5.1.1 Geometrical Properties 203
5.1.2 Electrical Properties 208
5.1.3 Penetration Depth 211
5.1.4 Soil Moisture Profile 214
5.2 Scattering from Bare Rough Surfaces 217
5.2.1 First-Order Small Perturbation Model 220
5.2.2 The Integral Equation Model 221
5.3 Example Bare Surface Soil Moisture Inversion Models 222
5.3.1 The First-Order Small Perturbation Model 224
5.3.2 Algorithm Proposed by Oh et al. 226
5.3.3 Algorithm Proposed by Dubois et al. 229
5.3.4 Algorithm Proposed by Shi et al. 231

5.4 Comparison of the Performance of Bare Surface Inversion Models 234
5.5 Parameterizing Scattering Models 239
5.6 Inverting the IEM Model 247
5.6.1 Using a Data Cube 247
5.7 Scattering from Vegetated Terrain 251
5.7.1 Scattering from the Vegetation Layer (Scattering Path 1) 253
5.7.2 Backscatter from the Underlying Ground Surface (Scattering Path 4) 257
5.7.3 Double Reflection Scattering (Scattering Paths 2 and 3) 258
5.8 Simulation Results 265
5.8.1 Effect of the Angle of Incidence 266
5.8.2 Effect of Cylinder Radius 267
5.8.3 Effect of Cylinder Moisture 270
5.8.4 Radar Vegetation Index 270
5.8.5 Effect of Soil Moisture 274
5.8.6 Inverting for Soil Moisture: The Data Cube 277
5.9 Time Series Estimation of Soil Moisture 279
5.10 Summary 286
References 287

Appendices

A: Tilted Small Perturbation Model Details 293
B: Bistatic Scattering Matrix of a Cylinder with Arbitrary Orientation 299
C: Nomenclature 309

Note from the Series Editor

The Jet Propulsion Laboratory (JPL) Space Science and Technology Series broadens the range of the ongoing JPL Deep Space Communications and Navigation Series to include disciplines other than communications and navigation in which JPL has made important contributions. The books are authored by scientists and engineers with many years of experience in their respective fields, and lay a foundation for innovation by communicating state-of-the-art knowledge in key technologies. The series also captures fundamental principles and practices developed during decades of space exploration at JPL, and celebrates the successes achieved. These books will serve to guide a new generation of scientists and engineers.

We would like to thank the Office of the Chief Scientist and Chief Technologist for their encouragement and support. In particular, we would like to acknowledge the support of Thomas A. Prince, former JPL Chief Scientist; Erik K. Antonsson, former JPL Chief Technologist; Daniel J. McCleese, JPL Chief Scientist; and Paul E. Dimotakis, JPL Chief Technologist.

Joseph H. Yuen, Editor-in-Chief
JPL Space Science and Technology Series
Jet Propulsion Laboratory
California Institute of Technology

x

Foreword

I am very pleased to commend the Jet Propulsion Laboratory (JPL) Space Science and Technology Series, and to congratulate and thank the authors for contributing their time to these publications. It is always difficult for busy scientists and engineers, who face the constant pressures of launch dates and deadlines, to find the time to tell others clearly and in detail how they solved important and difficult problems, so I applaud the authors of this series for the time and care they devoted to documenting their contributions to the adventure of space exploration.

JPL has been NASA's primary center for robotic planetary and deep-space exploration since the Laboratory launched the nation's first satellite, Explorer 1, in 1958. In the 50 years since this first success, JPL has sent spacecraft to all the planets except Pluto, studied our own planet in wavelengths from radar to visible, and observed the universe from radio to cosmic ray frequencies. Current plans call for even more exciting missions over the next decades in all these planetary and astronomical studies, and these future missions must be enabled by advanced technology that will be reported in this series. The JPL Deep Space Communications and Navigation book series captured the fundamentals and accomplishments of these two related disciplines, and we hope that this new series will expand the scope of those earlier publications to include other space science, engineering, and technology fields in which JPL has made important contributions.

I look forward to seeing many important achievements captured in these books.

CHARLES ELACHI, Director
Jet Propulsion Laboratory
California Institute of Technology

Preface

The fundamentals of radar polarimetry were built on a long history of research in optics. However, the fact that the radar is an active instrument, allowing one the extra flexibility to change the polarization of the transmitted wave in addition to optimizing the receiving antenna polarization, opened new doors to more powerful analysis of scattering from different types of terrain.

This book describes the application of polarimetric synthetic aperture radar to earth remote sensing based on our research at the Jet Propulsion Laboratory (JPL). Many important contributions to the field of radar polarimetry have been made long before we joined the field. Giants in the field include Kennaugh, Sinclair, Huynen, Boerner and many others. Our contribution is to put their work to practice in the field of synthetic aperture radar (SAR), but we owe thanks to these pioneers for pointing the way.

There is a vast amount of literature available on radar polarimetry. Here we did not try to reproduce or summarize all of these. Instead we concentrated our effort on compiling a subset of the knowledge into a reference that we hope would prove useful to both the newcomer and the expert in radar polarimetry. We provide a concise description of the mathematical fundamentals illustrated with many examples using SAR data. Our treatment of the subject is focused on remote sensing of the earth, and the examples are chosen to illustrate this application.

We start with the basics of synthetic aperture radar to provide the basis for understanding how polarimetric SAR images are formed. We follow this introduction with the fundamentals of radar polarimetry. We next discuss some of the more advanced polarimetric concepts that allow one to infer more information about the terrain being imaged. In order to analyze data quantitatively, however, the signals must be calibrated carefully. We included a

chapter summarizing the basic calibration algorithms. We conclude our discussion with an example of applying polarimetric analysis to scattering from rough surfaces with the aim to infer soil moisture from the radar signals.
Much still remains to be discovered about the best ways to extract all the information out of polarimetric SAR data. We hope that by preparing this work we have helped to accelerate this process by providing the next generation of researchers with some of the tools to make those discoveries.

Jakob van Zyl and Yunjin Kim
Pasadena, California
December 2010

Acknowledgments

We would like to express our appreciation to Joseph H. Yuen for his encouragement of us in generating this book and to Roger Carlson, Judi Dedmon, and John Kovacic for their editing and layout. We also thank all of our colleagues who contributed by providing us with the excellent synthetic aperture radar data. We have been privileged to work with them over the years. Likewise, through the years, we have collaborated with many researchers all over the globe. We thank them for their help in shaping our insight into the analysis of polarimetric SAR data.

December 2010

Authors

Jakob van Zyl received his Hons. B. Engineering (*Cum Laude*) in electronics engineering from the University of Stellenbosch, South Africa in 1979. Dr. Van Zyl received his MS and Ph D in electrical engineering from the California Institute of Technology (Pasadena, California) in 1983 and 1986, respectively. He has been with the Jet Propulsion Laboratory in Pasadena, California since 1986. At JPL he contributed to the design and development of many synthetic aperture radar (SAR) systems, including SIR-C, SRTM, AIRSAR, TOPSAR, and GeoSAR. In 1997 he received the Fred Nathanson Memorial Radar Award for advancement of radar polarimetry, radar interferometry, and synthetic aperture radar from the Aerospace and Electronics Society of the IEEE. In 2010 he received the Distinguished Achievement Award from the Geoscience and Remote Sensing Society of the IEEE for his contributions to polarimatric SAR remote sensing. He is currently the Director for Astronomy and Physics at the Jet Propulsion Laboratory.

Yunjin Kim received his BS in electrical engineering from the California State University, Sacramento, in 1983. Dr. Kim received his MS and PhD in electrical engineering from the University of Pennsylvania (Philadelphia, Pennsylvania) in 1985 and 1987, respectively. From 1987 to 1989, he was with the Department of Electrical Engineering, New Jersey Institute of Technology (Newark), as an assistant professor. Since 1989, Dr. Kim has been with the Jet Propulsion Laboratory (JPL), California Institute of Technology. At JPL, Dr. Kim has contributed to the development of several radar systems such as SIR-C, AIRSAR, GeoSAR, and SRTM. Currently, he is the Nuclear Spectroscopic Telescope Array (NuSTAR) project manager.

Chapter 1
Synthetic Aperture Radars (SAR) Imaging Basics

The word "radar" is an acronym for Radio Detection and Ranging. A radar measures the distance, or range, to an object by transmitting an electromagnetic signal to and receiving an echo reflected from the object. Since electromagnetic waves propagate at the speed of light, one only has to measure the time it takes the radar signal to propagate to the object and back to calculate the range to the object. The total distance traveled by the signal is twice the distance between the radar and the object, since the signal travels from the radar to the object and then back from the object to the radar after reflection. Therefore, once we measured the propagation time (t), we can easily calculate the range (R) as

$$R = \frac{1}{2}ct , \tag{1-1}$$

where c is the speed of light in vacuum. The factor ½ accounts for the fact that the radar signal actually traveled twice the distance measured: first from the radar to the object and then from the object to the radar. If the electric property of the propagation medium is different from that of vacuum, the actual propagation velocity has to be estimated for advanced radar techniques, such as Synthetic Aperture Radars (SAR) interferometry.

Radars provide their own signals to detect the presence of objects. Therefore, radars are known as active, remote-sensing instruments. Because radars provide their own signal, they can operate during day or night. In addition, radar signals typically penetrate clouds and rain, which means that radar images can be acquired not only during day or night, but also under (almost) all weather conditions. For these reasons, radars are often referred to as all-weather

instruments. Imaging, remote-sensing radars, such as SAR, produce high-resolution (from sub meter to a few tens of meters) images of surfaces. The geophysical information can be derived from these high-resolution images by using proper post-processing techniques.

This book focuses on a specific class of implementation of synthetic aperture radar with particular emphasis on the use of polarization to infer the geophysical properties of the scene. As mentioned above, SAR is a way to achieve high-resolution images using radio waves. We shall first describe the basics of radar imaging. This shall be followed by a description of the synthetic aperture principle. Finally, we shall discuss some advanced SAR implementations, such as SAR polarimetry and polarimetric SAR interferometry.

1.1 Basic Principles of Radar Imaging

Imaging radars generate surface images that are at first glance very similar to the more familiar images produced by instruments that operate in the visible or infrared parts of the electromagnetic spectrum. However, the principle behind the image generation is fundamentally different in the two cases. Visible and infrared sensors use a lens or mirror system to project the radiation from the scene on a "two-dimensional array of detectors," which could be an electronic array or, in earlier remote-sensing instruments, a film using chemical processes. The two-dimensionality can also be achieved using scanning systems or by moving a single line array of detectors. This imaging approach — an approach with which we are all familiar from taking photographs with a camera — conserves the relative angular relationships between objects in the scene and their images in the focal plane, as shown in Fig. 1-1. Because of this conservation of angular relationships, the resolution of the images depends on how far away the camera is from the scene it is imaging. The closer the camera, the higher the resolution and the smaller the details that can be recognized in the images. As the camera moves further away from the scene, the resolution degrades and only larger objects can be discerned in the image.

Imaging radars use a quite different mechanism to generate images, with the result that the image characteristics are also quite different from that of visible and infrared images. There are two different mechanisms by which radars can be used to produce images; the two types of radars are broadly classified as real aperture and synthetic aperture radars. We shall discuss the differences between these two types in more detail later in this chapter.

Radar images are typically acquired in strips as the satellite or aircraft carrying the radar system moves along its flight path. These strips are often referred to as swaths or tracks. To separate objects in the cross-track direction and the

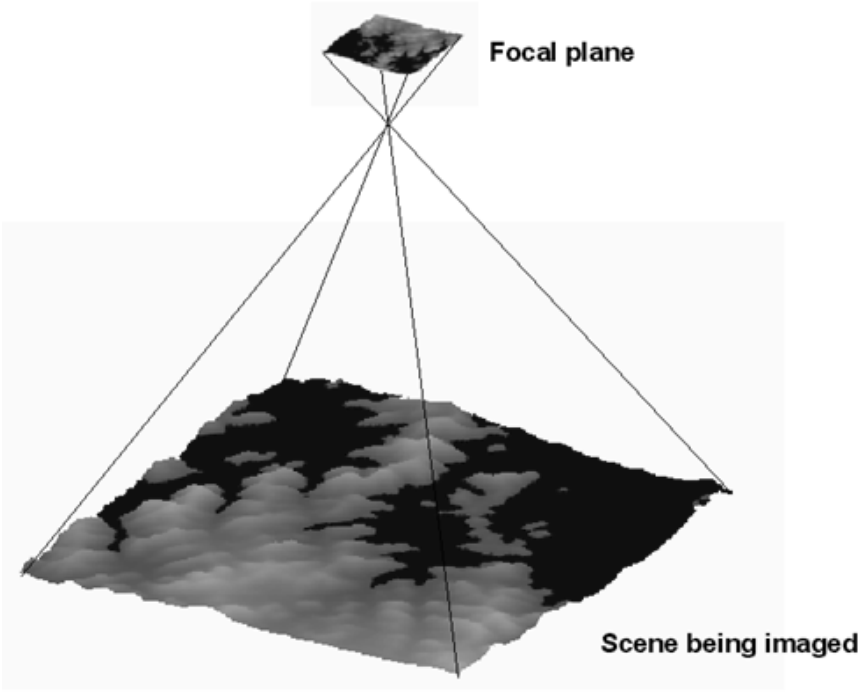


Fig. 1-1. Passive imaging systems conserve the angular relationships between objects in the scene and their images in the focal plane of the instrument.

along-track direction within a radar image, two different methods must be implemented. The cross-track direction, also known as the range direction in radar imaging, is the direction perpendicular to the direction in which the imaging platform is moving. In this direction, radar echoes are separated using the time delay between the echoes that are back-scattered from the different surface elements. This is true for both real aperture and synthetic aperture radar imagers. The along-track direction, also known as the azimuth direction, is the direction parallel to the movement of the imaging platform. The angular size (in the case of the real aperture radar) or the Doppler history (in the case of the synthetic aperture radar) is used to separate surface pixels in the along-track dimension in the radar images. As we will see later, only the azimuth imaging mechanism of real aperture radars is similar to that of regular cameras. Using the time delay and Doppler history results, SAR images have resolutions that are independent of how far away the radar is from the scene it is imaging. This fundamental advantage enables high-resolution, spaceborne SAR without requiring an extremely large antenna.

Another difference between images acquired by cameras operating in the visible and near infrared part of the electromagnetic spectrum and radar images

is the way in which they are acquired. Cameras typically look straight down, or at least have no fundamental limitation that prevents them from taking pictures looking straight down from the spacecraft or aircraft. Not so for imaging radars. To avoid so-called ambiguities, which we will discuss in more detail later, the imaging radar sensor has to use an antenna that illuminates the surface to one side of the flight track. Usually, the antenna has a fan beam that illuminates a highly elongated, elliptically shaped area on the surface, as shown in Fig. 1-2. The illuminated area across track generally defines the image swath.

Within the illumination beam, the radar sensor transmits a very short effective pulse of electromagnetic energy. Echoes from surface points farther away along the cross-track coordinate will be received at proportionally later time (see Fig. 1-2). Thus, by dividing the receive time in increments of equal time bins, the surface can be sub-divided into a series of range bins. The width in the along-track direction of each range bin is equal to the antenna footprint along the track x_a. As the platform moves, the sets of range bins are covered sequentially, thereby allowing strip mapping of the surface line by line. This is comparable to strip mapping with a so-called pushbroom imaging system using a line array in the visible and infrared part of the electromagnetic spectrum. The brightness associated with each image pixel in the radar image is proportional to the echo power contained within the corresponding time bin. As we will see later, the real difference between real aperture radars and synthetic aperture radars lies in the way in which the azimuth resolution is achieved.

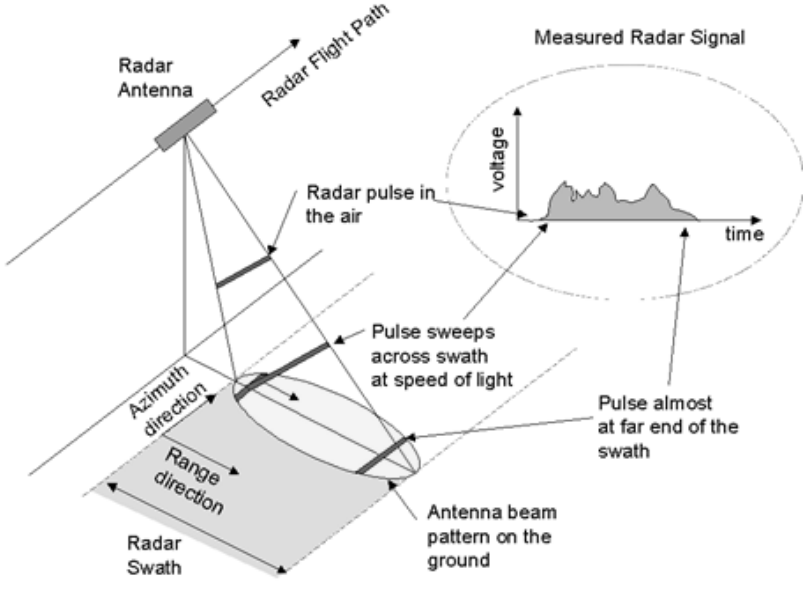


Fig. 1-2. Imaging geometry for a side-looking radar system.

This is also a good time to point out that there are two different meanings for the term range in radar imaging. The first is the so-called slant range and refers to the range along the radar line-of-sight, as shown in Fig. 1-3. Slant ranges are measured along the line connecting the radar and the object being imaged, often called the target or the scatterer. The second use of the term range is for the ground range, which refers to the range along a smooth surface (the ground) to the scatterer. The ground range is measured from the so-called nadir track, which represents the line described by the position directly underneath the radar imaging platform. One has to be careful to take topography into account when re-sampling radar images from slant range to ground range. This will be discussed in more detail in Section 1.6.2.

Before looking at radar resolutions, let us define a few more terms commonly encountered in radar imaging. The look angle is defined as the angle between the vertical direction and the radar beam at the radar platform. The incidence angle is defined as the angle between the vertical direction and the radar wave propagation vector at the surface (as shown in Fig. 1-3). When surface curvature effects are neglected, the look angle is equal to the incidence angle at the surface when the surface is flat. In the case of space-borne systems, surface curvature must be taken into account; this leads to an incidence angle that is always larger than the look angle for flat surfaces. It is quite common in the literature to find authors using the terms look angle and incidence angle interchangeably. That is only correct for low-flying aircraft, and only when there is no topography present in the scene. As we will see next, if topography is present (i.e., if the surface is not flat), the local incidence angle might vary in the radar image from pixel to pixel.

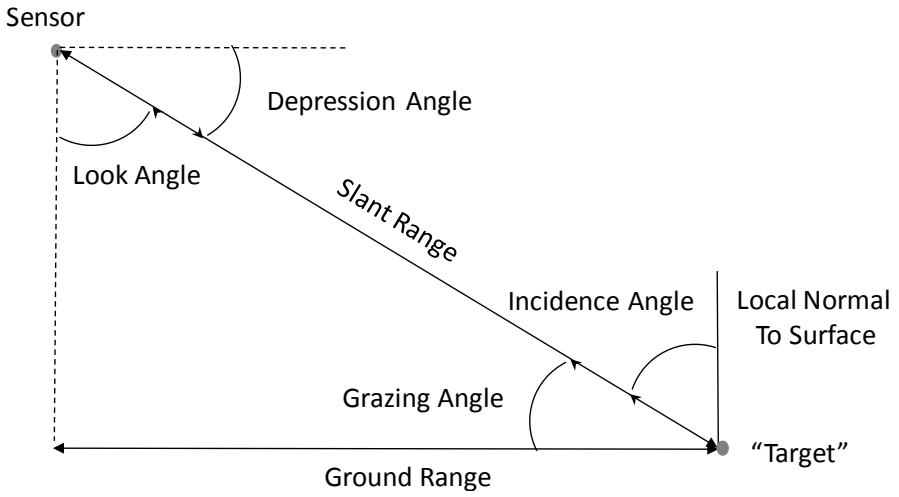


Fig. 1-3. Definition of some common radar imaging terms.

Consider the simple case of a single hill illuminated by a radar system, as is shown in Fig. 1-4. Also shown in the figure are the local normal to the surface for several positions on the hill. Relative to a flat surface, it is clear that for points on the hill facing the radar, the local normal tilts more towards the radar; therefore, the local incidence angle will be smaller than for a point at the same ground range, but on a flat surface.

A term commonly encountered in military literature is depression angle. This is the angle between the radar beam and the horizontal at the radar platform. The depression angle is, therefore, related to the look angle in that one is equal to 90 degrees minus the other. A small look angle is equivalent to a large depression angle, and vice versa. Similarly, one often finds the term grazing angle describing the angle between the horizontal at the surface and the incident wave in the military literature. The grazing angle is, therefore, related to the incidence angle in the same way that the depression angle is related to the look angle. In this text, we shall use look angle and incidence angle to describe the imaging geometry.

1.2 Radar Resolution

The resolution of an image is defined as that separation between the two closest features that can still be resolved in the final image. First, consider two point targets that are separated in the slant range direction by x_r. Because the radar waves propagate at the speed of light, the corresponding echoes will be separated by a time difference Δt equal to:

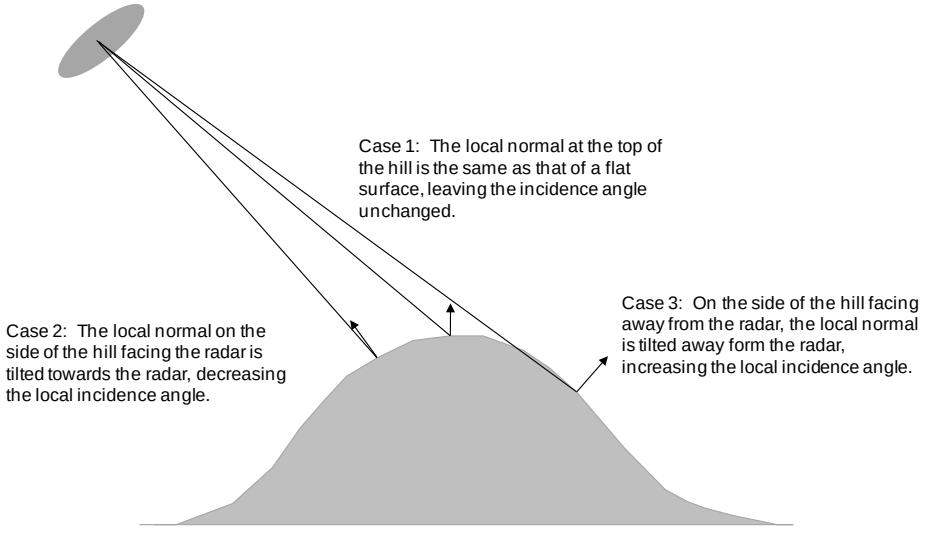


Fig. 1-4. Topographic variations in the image will cause the local incidence angle to be different from that expected for a flat surface with no relief.

$$\Delta t = 2x_r / c \,, \tag{1.2-1}$$

where c is the speed of light and the factor 2 is included to account for the signal round trip propagation, as was previously described. Radar waves are usually not transmitted continuously; instead, a radar usually transmits short bursts of energy known as radar pulses. The two features can be discriminated if the leading edge of the pulse returned from the second object is received later than the trailing edge of the pulse received from the first feature, as shown in Fig. 1-5.

Therefore, the smallest separable time difference in the radar receiver is equal to the effective time length τ of the pulse. Thus, the slant range resolution of a radar is:

$$2x_r / c = \tau \quad \Rightarrow \quad x_r = \frac{c\tau}{2} \tag{1.2-2}$$

Now let us consider the case of two objects separated by a distance x_g on the ground. The corresponding echoes will be separated by a time difference Δt equal to:

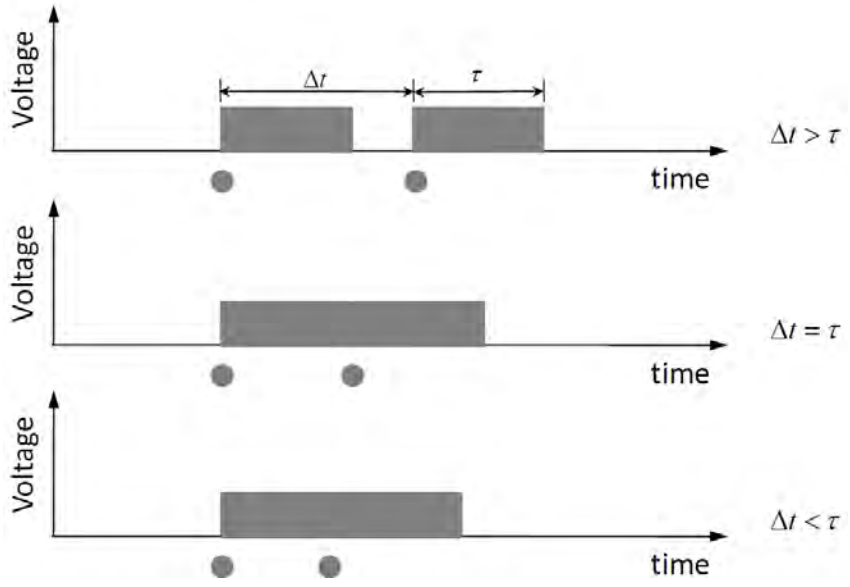


Fig. 1-5. If the radar echoes from two point targets are separated in time by more than or equal to the length of the radar pulse, it is possible to recognize the echoes as those from two different scatterers, as shown in the top two panels. If the time difference between the echoes is less than the radar pulse length, it is not possible to recognize two distinct scatters, as in the case of the bottom panel.

$$\Delta t = 2x_g \sin\theta / c \,. \tag{1.2-3}$$

The angle θ in Eq. (1.2-3) is the local incidence angle. (This should actually be called the incident angle, or angle of incidence. Since incidence angle is used almost universally in the literature, we shall continue to use that term to avoid confusion.). As in the case of the slant range discussed above, the two features can be discriminated if the leading edge of the pulse returned from the second object is received later than the trailing edge of the pulse received from the first feature. Therefore, the ground range resolution of the radar is given by

$$2x_g \sin\theta / c = \tau \quad \Rightarrow \quad x_g = \frac{c\tau}{2\sin\theta} \,. \tag{1.2-4}$$

In other words, the ground range resolution is equal to half the footprint of the radar pulse on the surface. Note that Eq. (1.2-4) implies that the ground range resolution is different for different incidence angles. This also means that the local slopes in the images will affect the ground range resolution. For slopes facing the radar, the ground range resolution will be poorer than that for slopes facing away from the radar.

Occasionally, the effective pulse length is described in terms of the system bandwidth B. As we will show in the next section, to a good approximation,

$$\tau = 1/B \,. \tag{1.2-5}$$

A pulsed radar determines the range by measuring the round trip time by transmitting a pulse signal. In designing the signal pattern for a radar sensor, there is usually a strong requirement to have as much energy as possible in each pulse in order to enhance the signal-to-noise ratio. This can be done by increasing the transmitted peak power or by using a longer pulse. However, the peak power is usually strongly limited by the available power sources, particularly in the case of spaceborne sensors. On the other hand, an increased pulse length leads to a worse range resolution (see Eq. (1.2-4)). This dilemma is usually resolved by using modulated pulses, which have the property of a wide bandwidth even when the pulse is very long. After so-called pulse-compression, a short effective pulse length is generated, increasing the resolution. One way to modulate the pulse is to vary the radar signal frequency linearly while the pulse is being transmitted. This waveform is known as chirp. Although other ways of modulating pulses are occasionally used, the chirp modulation is by far the most common; we shall, therefore, use this scheme to illustrate how signal modulation can be used to enhance radar range resolution.

In a chirp, the signal frequency within the pulse is linearly changed as a function of time. If the frequency is linearly changed from f_0 to $f_0 + \Delta f$, the effective bandwidth would be equal to:

$$B = \left| \left(f_0 + \Delta f \right) - f_0 \right| = \left| \Delta f \right|, \tag{1.2-6}$$

which is independent of the pulse length τ_p. Thus, a pulse with long duration (i.e., high energy) and wide bandwidth (i.e., high range resolution) can be constructed. The instantaneous frequency for such a signal is given by:

$$f(t) = f_0 + \frac{B}{\tau_p} t \quad \text{for} \quad -\tau_p/2 \le t \le \tau_p/2 . \tag{1.2-7}$$

The corresponding signal amplitude is:

$$A(t) \sim \Re\left\{ \exp\left[-i2\pi \int f(t)\, dt \right] \right\} = \cos\left[2\pi \left(f_0 t + \frac{B}{2\tau_p} t^2 \right) \right], \tag{1.2-8}$$

where $\Re(x)$ means the real part of x. Note that the instantaneous frequency is the derivative of the instantaneous phase. A pulse signal such as shown in equation (1.2-8) has a physical pulse length τ_p and a bandwidth B. The product $\tau_p B$ is known as the time bandwidth product of the radar system. In typical radar systems time bandwidth products of several hundred are used.

At first glance, it might seem that using a pulse of the form shown in Eq. (1.2-8) cannot be used to separate targets that are closer than the projected physical length of the pulse (as shown in the previous section). It is, indeed, true that the echoes from two neighboring targets that are separated in the range direction by much less than the physical length of the signal pulse will overlap in time. If the modulated pulse and, therefore, the echoes have a constant frequency, it will not be possible to resolve the two targets. However, if the frequency is modulated as described in Eq. (1.2-7), the echoes from the two targets will have different frequencies at any instant of time and, therefore, can be separated by frequency filtering.

In actual radar systems, a matched filter is used to compress the returns from the different targets. Consider an example where we transmit a signal of the form described in Eq. (1.2-8). The signal received from a single point scatterer at a range R is a scaled replica of the transmitted signal delayed by a time $t = 2R/c$. The output of the matched filter for such a point scatterer is mathematically described as the convolution of the returned signal with a

replica of the transmitted signal. Being careful about the limits of the integration, one finds that for large time-bandwidth products,

$$V_o(t) = \tau_p E_r \exp(i\omega t) \exp(-i4\pi R/\lambda) \frac{\sin(\pi B(t - 2R/c))}{\pi B(t - 2R/c)} \qquad (1.2\text{-}9)$$

where

$$\omega = 2\pi f_0$$

and

E_r = electric field received by radar from a point source.

This compressed pulse has a half power width of approximately $1/B$ and its peak position occurs at time $2R/c$. Therefore, the achievable range resolution using a modulated pulse of the kind described in Eq. (1.2-8) is a function of the chirp bandwidth and not the physical pulse length. In typical spaceborne and airborne SAR systems, physical pulse lengths of several tens of micro-seconds are used, while bandwidths of several tens of megahertz are no longer uncommon for spaceborne systems, and several hundreds of megahertz are common in airborne systems.

So far, we have seen the first major difference between radar imaging and that used in passive imaging systems. The cross-track resolution in the radar case is independent of the distance between the scene and the radar instrument and is a function of the system bandwidth. Before looking at the imaging mechanisms in the along-track direction, we will examine the general expression for the amount of reflected power that the radar receiver would measure. This is described through the so-called radar equation, which we will examine in the next section.

1.3 Radar Equation

One of the key factors that determine the quality of the radar imagery is the corresponding signal-to-noise ratio (SNR), commonly called SNR. This is the equivalent of the brightness of a scene being photographed with a camera versus the sensitivity of the film or detector. Here, we consider the effect of thermal noise on the sensitivity of radar imaging systems. The derivation of the radar equation is graphically shown in Fig. 1-6.

In addition to the target echo, the received signal also contains noise, which results from the fact that all objects at temperatures higher than absolute zero emit radiation across the whole electromagnetic spectrum. The noise

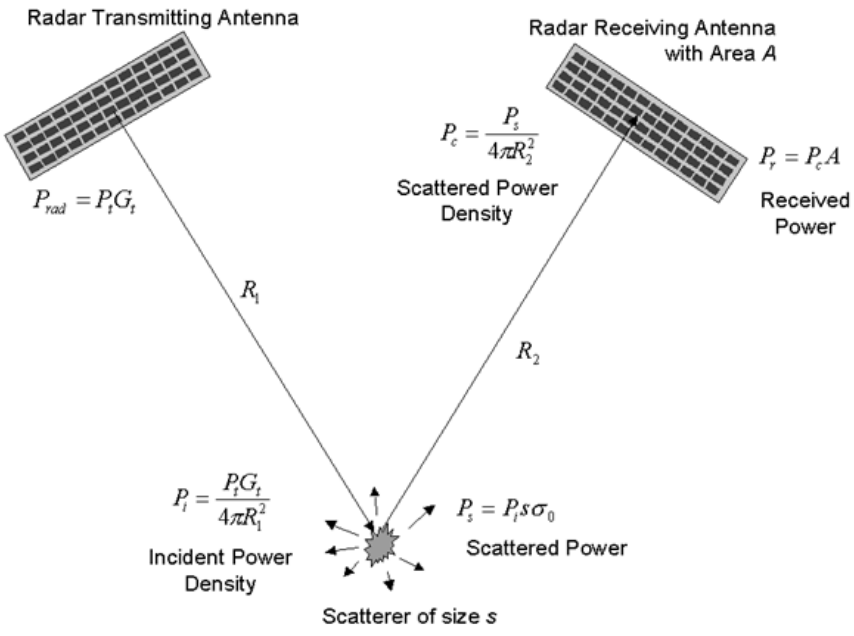


Fig. 1-6. Schematic of the derivation of the radar equation.

component that is within the spectral bandwidth B of the sensor is passed through with the signal. The receiver electronics also generates noise that contaminates the signal. The thermal noise power is given by:

$$P_N = kTB\,, \tag{1.3-1}$$

where k is Boltzmann's constant ($k = 1.6 \times 10^{-23}$ W/K/Hz) and T is the total equivalent noise temperature in kelvin. The resulting SNR is then:

$$SNR = P_r / P_N\,. \tag{1.3-2}$$

Notice that the noise bandwidth is usually larger than the transmit bandwidth, due to the hardware limitation.

One common way of characterizing an imaging radar sensor is to determine the surface backscatter cross section σ_N, which gives an $SNR = 1$. This is called the noise equivalent backscatter cross section. It defines the weakest surface return that can be detected and, therefore, identifies the range of surface units that can be imaged.

1.4 Real Aperture Radar

The real aperture imaging radar sensor also uses an antenna that illuminates the surface to one side of the flight track. As mentioned before, the antenna usually has a fan beam that illuminates a highly elongated elliptical shaped area on the surface (see Fig. 1-2). As shown in Fig. 1-2, the illuminated area across track defines the image swath. For an antenna of width W operating at a wavelength λ, the beam angular width in the range plane is given by:

$$\theta_r \approx \lambda / W \tag{1.4-1}$$

and the resulting surface footprint or swath S is given by

$$S \approx \frac{h\theta_r}{\cos^2\theta} = \frac{\lambda h}{W\cos^2\theta}, \tag{1.4-2}$$

where h is the sensor height above the surface, θ is the angle from the center of the illumination beam to the vertical (the look angle at the center of the swath), and θ_r is assumed to be very small. Note that Eq. (1.4-2) ignores the curvature of the Earth. For spaceborne radars, this effect should not be ignored. If the antenna beam width is large, one needs to use the law of cosines to solve for the swath width.

A real aperture radar relies on the resolution afforded by the antenna beam in the along-track direction for imaging. This means that the resolution of a real aperture radar in the along-track direction is determined by the size of the antenna as well as the range to the scene. Assuming an antenna length of L, the antenna beam width in the along-track direction is

$$\theta_a \approx \frac{\lambda}{L}. \tag{1.4-3}$$

At a distance R from the antenna, this means that the antenna beam width illuminates an area with the along-track dimension equal to

$$x_a \approx R\theta_a \approx \frac{\lambda R}{L} \approx \frac{\lambda h}{L\cos\theta}. \tag{1.4-4}$$

To illustrate, for h = 800 km, λ = 23 cm, L = 12 m and $\theta = 20^\circ$, then x_a = 16 km. Even if λ is as short as 2 cm and h as low as 200 km, x_a will still be equal to about 360 meters, which is considered to be a relatively poor resolution, even for remote sensing. This has led to very limited use of the real-aperture technique for surface imaging, especially from space. A real aperture radar uses the same imaging mechanism as a passive optical system for the

along-track direction. However, because of the small value of λ (about 1 μm), resolutions of a few meters can be achieved from orbital altitudes with an aperture only a few tens of centimeters in size. From aircraft altitudes, however, reasonable azimuth resolutions can be achieved if higher frequencies (typically X-band or higher) are used. For this reason, real aperture radars are not commonly used in spaceborne remote sensing (except in the case of scatterometers and altimeters that do not need high-resolution data).

In terms of the radar equation, the area responsible for reflecting the power back to the radar is given by the physical size of the antenna illumination in the along-track direction and the projection of the pulse on the ground in the cross-track direction. This is shown in Fig. 1-2 for the pulses in the radar swath. The along-track dimension of the antenna pattern is given by Eq. (1.4-4). If the pulse has a length τ_p in time, and the signal is incident on the ground at an angle θ_i the projected length of the pulse on the ground is

$$l_g = \frac{c\tau_p}{2\sin\theta_i}. \tag{1.4-5}$$

Therefore, the radar equation in the case of a real aperture radar becomes

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi)^3 R^4} \frac{\lambda R}{L} \frac{c\tau_p}{2\sin\theta_i} \sigma_0 \ . \tag{1.4-6}$$

The normalized backscattering cross section (σ_0) is defined as

$$\sigma_0 = \lim_{R \to \infty} \frac{4\pi R^2}{A_0} \frac{|E_r|^2}{|E_i|^2} \tag{1.4-7}$$

where A_0 is the illuminated surface area, E_r is the reflected electric field, and E_i is the incident electric field.

This shows that when a real aperture radar images an extended area, the received power decreases as the range to the third power. In terms of the physical antenna sizes, we can rewrite this expression as

$$P_r = \frac{P_t W^2 L c\tau_p \sigma_0}{8\pi\lambda R^3 \sin\theta_i} \ . \tag{1.4-8}$$

This is the radar equation for a so-called distributed target for the real aperture radar case. From Eq. (1.4-7), it is clear that the received power increases as the square of the width of the antenna. However, increasing the antenna width also decreases the swath width. The received power only increases linearly with an increase in antenna length. Increasing the antenna length also improves the along-track resolution of the real aperture radar. For this reason, real aperture radars usually operate with antennas that are the longest that could be practically accommodated.

In summary, a real aperture radar uses the same imaging mechanism as passive imaging systems to achieve along-track resolution. The practically achievable resolutions are usually poorer than what is generally required for remote sensing applications. Real aperture radars are, therefore, not commonly used for remote sensing applications.

1.5 Synthetic Aperture Radar

Synthetic aperture radar refers to a particular implementation of an imaging radar system that utilizes the movement of the radar platform and specialized signal processing to generate high-resolution images. Prior to the discovery of synthetic aperture radar, principle imaging radars operated using the real-aperture principle and were known as side-looking aperture radars (SLAR).

Carl Wiley of the Goodyear Aircraft Corporation is generally credited as being the first person to describe the use of Doppler frequency analysis of signals from a moving coherent radar to improve along-track resolution. He noted that two targets at different along-track positions will be at different angles relative to the aircraft velocity vector, resulting in different Doppler frequencies. (The Doppler effect is the well-known phenomenon that causes a change in the pitch of a car horn as it travels past a stationary observer.) Using this effect, targets can be separated in the along-track direction on the basis of their different Doppler frequencies. This technique was originally known as Doppler beam sharpening, but later became known as synthetic aperture radar (SAR).

The main difference between real and synthetic aperture radars is, therefore, in the way in which the azimuth resolution is achieved. The range resolution and radar equation derived previously for a real aperture radar is still valid here. The along-track imaging mechanism and the resulting along-track resolution is, however, quite different for the real and synthetic aperture radar case.

As the radar moves along the flight path, it transmits pulses of energy and records the reflected signals, as shown in Fig. 1-2. When the radar data are processed, the position of the radar platform is taken into account when adding the signals to integrate the energy for the along-track direction. Consider the geometry shown in Fig. 1-7. As the radar moves along the flight path, the

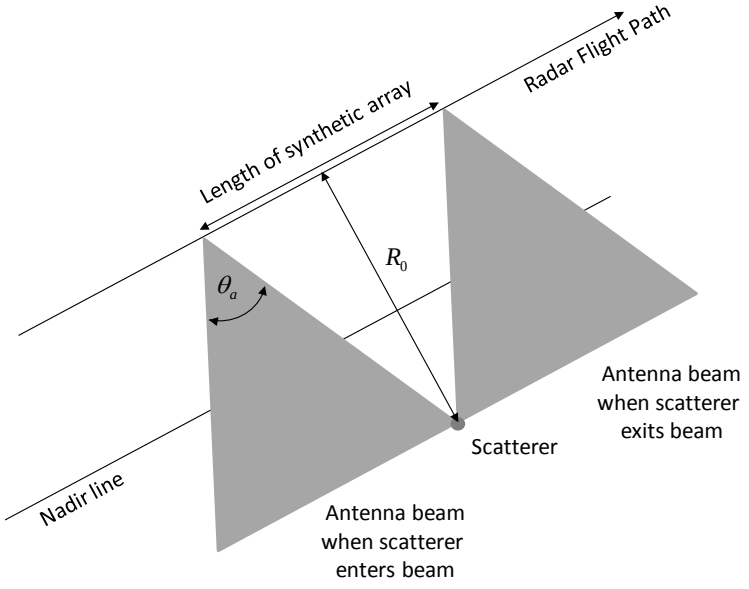


Fig. 1-7. The synthetic aperture radar integrates the signal from the scatter for as long as the scatterer remains in the antenna beam.

distance between the radar and the scatterer changes, with the minimum distance occurring when the scatterer is directly broadside of the radar platform. The phase of the radar signal is given by $-4\pi R(s)/\lambda$. The changing distance between the radar and the scatterer means that after range compression, the phase of the signal will be different for the different positions along the flight path.

This changing distance can be written as

$$R(s) = \sqrt{R_0^2 + v^2 s^2}\,, \tag{1.5-1}$$

where R_0 is the range at closest approach to the scatterer, v is the velocity of the radar platform, and s is the time along the flight path (so-called slow time) with zero time at the time of closest approach. To a good approximation for remote sensing radars, we can assume that $vs << R_0$. (This might not be true for very-high-resolution radars; however, the basic principle remains the same.) In this case, we can approximate the range as a function of slow time as

$$R(s) \approx R_0 + \frac{v^2}{2R_0} s^2 . \tag{1.5-2}$$

The phase of the signal after range compression, as shown in the second exponential in Eq. (1.2-9), then becomes

$$\phi(s) = -\frac{4\pi R(s)}{\lambda} \approx -\frac{4\pi R_0}{\lambda} - \frac{2\pi v^2}{R_0 \lambda} s^2 . \tag{1.5-3}$$

The instantaneous frequency of this signal is

$$f(s) = \frac{1}{2\pi}\frac{\partial \phi(s)}{\partial s} = -\frac{2v^2}{R_0 \lambda} s . \tag{1.5-4}$$

This is the expression of a linear frequency chirp. To find the bandwidth of this signal, we have to find the maximum time that we can use in the signal integration. This maximum "integration time" is given by the amount of time that the scatterer will be in the antenna beam. For an antenna with a physical length L, the half-power horizontal beam width is $\theta_a = \lambda / L$, so that the scatterer at the range of closest approach R_0 is illuminated for a time

$$s_{tot} = \frac{\lambda R_0}{L v} . \tag{1.5-5}$$

Half of this time occurs when the radar is approaching the range of closest approach and half of it is spent traveling away from the range of closest approach. Therefore, the bandwidth of the signal shown in Eq. (1.5-4), which is the Doppler bandwidth of the SAR signal, is

$$B_D = \frac{2v}{L} . \tag{1.5-6}$$

If this signal is filtered using a matched filter, as described earlier under signal modulation, the resulting compressed signal will have a width in time of $1/B_D$. Since the radar platform moves at a speed of v, this leads to an along-track resolution of

$$\Delta_a = \frac{v}{B_D} = \frac{L}{2} . \tag{1.5-7}$$

This result shows that the azimuth (or along-track) surface resolution for a synthetic aperture radar is equal to half the size of the physical antenna and is independent of the distance between the sensor and the surface! At first glance, this result seems most unusual: it shows that a smaller antenna gives better resolution. This can be explained in the following way: the smaller the physical antenna is, the larger its footprint. This allows a longer observation time for

each point on the surface (i.e., a longer array can be synthesized). This longer synthetic array allows a larger Doppler bandwidth and, hence, a finer surface resolution. Similarly, if the range between the sensor and surface increases, the physical footprint increases, leading to a longer observation time and larger Doppler bandwidth, which counter balances the increase in the range.

As mentioned earlier, the imaging radar transmits a series of pulsed electromagnetic waves. Thus, the Doppler history from a scatterer is not measured continuously but sampled on a repetitive basis. To get an accurate record of the Doppler history, the Nyquist sampling criterion requires that sampling occurs at least at twice the highest frequency in the Doppler bandwidth. Thus, the pulse repetition frequency, usually called PRF, must be larger than

$$PRF \geq 2\left(\frac{B_D}{2}\right) = \frac{2v}{L}. \tag{1.5-8}$$

Note that we used half the Doppler bandwidth as the highest frequency in the Doppler signal in Eq. (1.5-8). The reason for this is that the Doppler frequency varies linearly from $-B_D/2$ to $+B_D/2$. Therefore, even though the total bandwidth of the signal is B_D, the highest frequency in the bandwidth is only $B_D/2$.

Eq. (1.5-8) means that at least one sample (i.e., one pulse) should be taken every time the sensor moves by half an antenna length. As an example, for a spaceborne imaging system moving at a speed of 7 km/sec and using an antenna 10 meters in length, the corresponding minimum PRF is 1.4 kHz. As we will see in the next section, the requirement to cover a certain swath size provides and upper bound on the PRF. Interpreted in a different way, the requirement to adequately sample the signal bandwidth limits the size of the swath that could be imaged.

1.6 Radar Image Artifacts and Noise

Radar images could contain a number of anomalies that result from the way imaging radars generate the image. Some of these are similar to what is encountered in optical systems, such as blurring due to defocusing or scene motion, and some (such as range and azimuth ambiguities) are unique to radar systems. This section summarizes the anomalies that are most commonly encountered in radar images.

1.6.1 Range and Azimuth Ambiguities

As mentioned earlier (see Fig. 1-2), a radar images a surface by recording the echoes line by line with successive pulses. The leading edge of each echo corresponds to the near edge of the image scene; the tail end of the echo corresponds to the far edge of the scene. The length of the echo (i.e., swath width of the scene imaged) is determined by the antenna beam width or the size of the data window used in the recording of the signal. The exact timing of the echo reception depends on the range between the sensor and the surface being imaged. If the timing of the pulses or the extent of the echoes is such that the leading edge of one echo overlaps with the tail end of the previous one, then the far edge of the scene is folded over the near edge of the scene. This is called range ambiguity. The temporal extent of the echo is equal to:

$$T_e \approx 2\frac{R}{c}\theta_r \tan\theta = 2\frac{h\lambda}{cW}\frac{\sin\theta}{\cos^2\theta} \tag{1.6-1}$$

To avoid overlapping echoes, this time extent should be shorter than the time separating two pulses (i.e., $1/PRF$). Thus, we must have

$$PRF < \frac{cW}{2h\lambda}\frac{\cos^2\theta}{\sin\theta} \tag{1.6-2}$$

In addition, the sensor parameters, specifically the PRF, should be selected such that the echo is completely within an interpulse period (i.e., no echoes should be received during the time that a pulse is being transmitted). The above equation gives an upper limit for the PRF as mentioned before. The SAR designers have to trade off system parameters to maximize the swath, while at the same time transmitting a high enough PRF to adequately sample the signal Doppler spectrum.

Another kind of ambiguity present in SAR imagery also results from the fact that the target's return in the azimuth direction is sampled at the PRF. This means that the azimuth spectrum of the target return repeats itself in the frequency domain at multiples of the PRF. In general, the azimuth spectrum is not a band limited signal; instead, the spectrum is weighted by the antenna pattern in the azimuth direction. This means that parts of the azimuth spectrum might be aliased and high-frequency data will actually appear in the low-frequency part of the spectrum. In actual images, these azimuth ambiguities appear as ghost images of a target repeated at some distance in the azimuth direction, as shown in Fig. 1-8. To reduce the azimuth ambiguities, the PRF of a SAR has to exceed the lower limit given by Eq. (1.5-8).

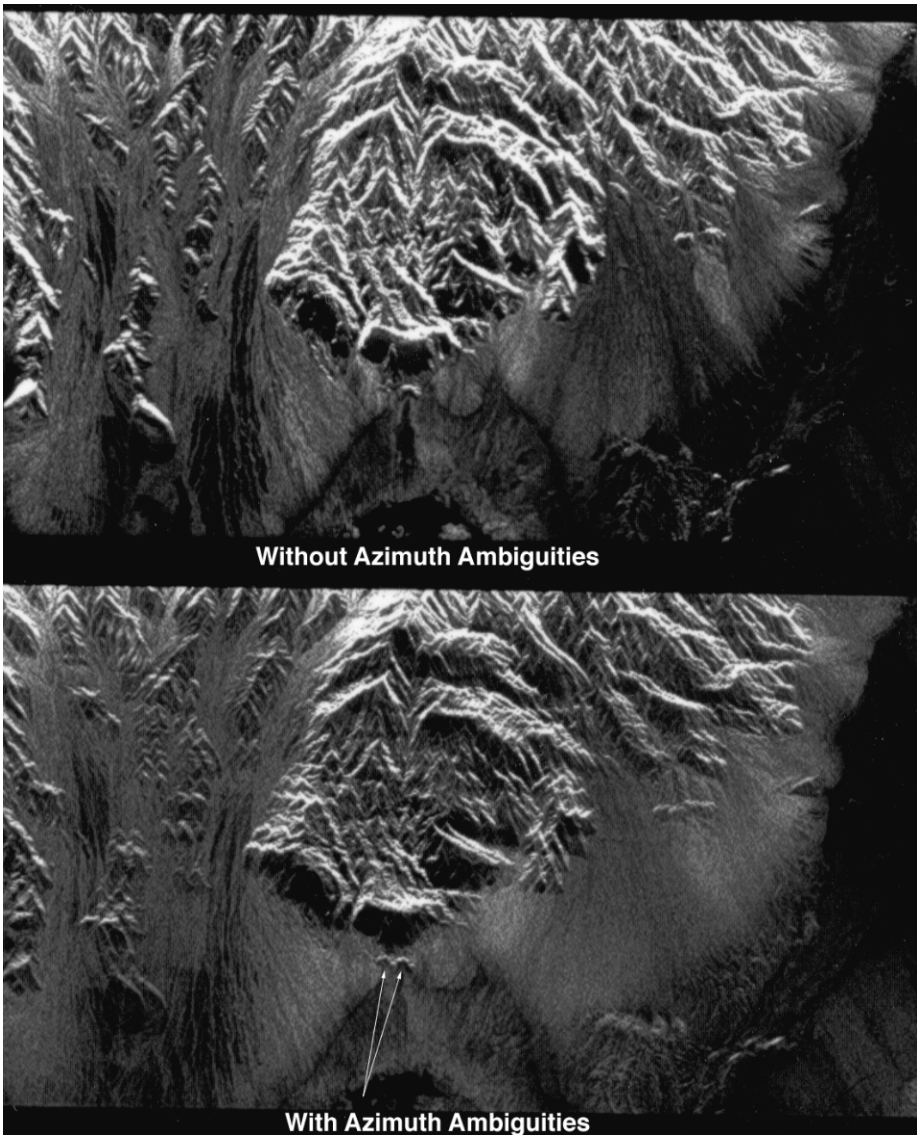


Fig. 1-8. Azimuth ambiguities result when the radar pulse repetition frequency is too low to sample the azimuth spectrum of the data adequately. In this case, the edges of the azimuth spectrum fold over themselves, creating ghost images, as shown in this figure.

To reduce both range and azimuth ambiguities, the PRF must, therefore, satisfy both the conditions expressed by Eq. (1.5-8) and Eq. (1.6-2). Therefore, we must insist that

$$\frac{cW}{2h\lambda}\frac{\cos^2\theta}{\sin\theta} > \frac{2v}{L}, \tag{1.6-3}$$

from which we derive a lower limit for the antenna size as

$$LW > \frac{4vh\lambda}{c}\frac{\sin\theta}{\cos^2\theta}. \tag{1.6-4}$$

A word of caution about the use of Eq. (1.6-4): This expression is derived assuming the SAR processor uses the full Doppler bandwidth in the processing and a swath that covers the full antenna beam width is imaged. This might not always be the case. For many reasons, SAR images are sometimes processed to only a fraction of the achievable resolution, or the swath width might be limited artificially by recording only that section of the returned echo that falls in a so-called data window. When either of these conditions is used, it is not appropriate to limit the antenna size as given by Eq. (1.6-4). In fact, in the case where both the swath is artificially limited and the resolution is decreased, antennas significantly smaller that that given by Eq. (1.6-4) may be used with perfectly good results.

Another type of artifact in radar images results when a very bright surface target is surrounded by a dark area. As the image is being formed, some spill over from the bright target, although weak, could exceed the background and become visible. It should be pointed out that this type of artifact is not unique to radar systems. They are common in optical systems, where they are known as the side lobes of the point spread function. The difference is that in optical systems, the side lobe characteristics are determined by the characteristics of the imaging optics (i.e., the hardware). Conversely, in the case of a SAR, the side lobe characteristics are determined mainly by the characteristics of the processing filters. In the radar case, the side lobes might, therefore, be reduced by suitable weighting of the signal spectra during matched filter compression. The equivalent procedure in optical systems is called apodization of the telescope aperture.

The vast majority of these artifacts and ambiguities can be avoided with proper selection of the sensor and processor parameters. However, the interpreter should be aware of their occurrence, because in some situations they might be difficult, if not impossible, to suppress.

1.6.2 Geometric Effects and Projections

The time delay/Doppler history basis of SAR image generation leads to an image projection different than in the case of optical sensors. Even though at first look radar images seem very similar to optical images, close examination

quickly shows that geometric shapes and patterns are projected in a different fashion by the two sensors. This difference is particularly acute in rugged terrain. If the topography is known, a radar image can be re-projected into a format identical to an optical image, thus allowing image pixel registration. In extremely rugged terrain, however, the nature of the radar image projection leads to distortions that sometimes cannot be corrected without knowledge of the terrain elevations.

In the radar image, two neighboring pixels in the range dimension correspond to two areas in the scene with slightly different range to the sensor. This has the effect of projecting the scene in a cylindrical geometry on the image plane, which leads to distortions (as shown in Fig. 1-9). Areas that slope toward the sensor look shorter in the image, while areas that slope away from the sensor look longer in the image than horizontal areas. This effect is called foreshortening. In the extreme case where the slope is larger than the incidence angle, layover occurs. In this case, a hill would look as if it is projected over the region in front of it. Layover cannot be corrected and can only be avoided by having an incidence angle at the surface larger than any expected surface slopes. When the slope facing away from the radar is steep enough such that the radar waves do not illuminate it, shadowing occurs and the area on that slope is not imaged. Note that in the radar images, shadowing is always away from the sensor flight line and is not dependent on the time of data acquisition or the Sun angle in the sky. As in the case of optical images, shadowing can be beneficial for highlighting surface morphologic patterns. Fig. 1-10 contains some examples of foreshortening and shadowing.

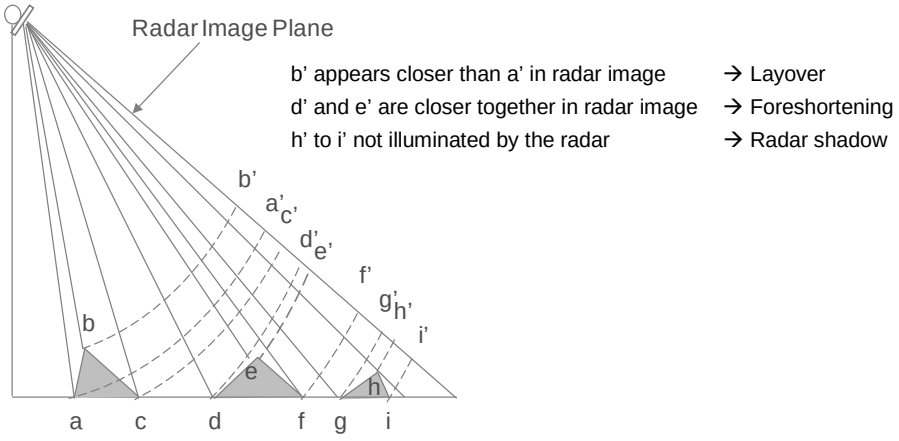


Fig. 1-9. Radar images are cylindrical projections of the scene onto the image plane, leading to characteristic distortions.

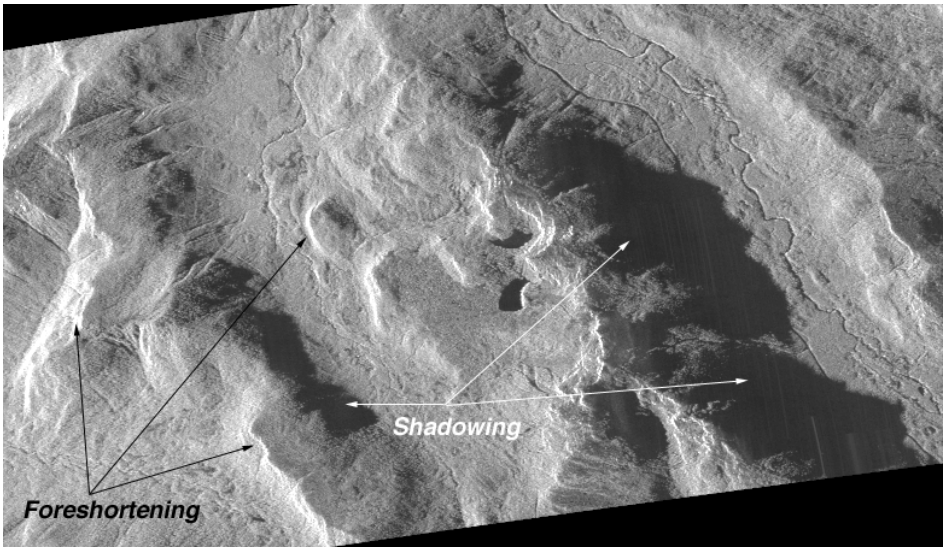


Fig. 1-10. This NASA/JPL AIRSAR image shows examples of foreshortening and shadowing. Note that since the radar provides its own illumination, radar shadowing is a function of the radar look direction relative to the terrain and does not depend on the Sun angle. This image was illuminated from the left. (Image courtesy of Jet Propulsion Laboratory.)

1.6.3 Signal Fading and Speckle

A close examination of a synthetic-aperture radar image shows that the brightness variation is not smooth but, instead, has a granular texture that is called speckle. Even for an imaged scene that has a constant backscatter property, the image will have statistical variations of the brightness on a pixel-by-pixel basis, but with a constant mean over many pixels. This effect is identical to when a scene is observed optically under laser illumination. It is a result of the coherent nature (or very narrow spectral width) of the illuminating signal.

Rigorous mathematical analysis shows that the noise-like radar signal has well defined statistical properties. The measured signal amplitude has a Rayleigh distribution and the signal power has an exponential distribution [1]. To narrow the width of these distributions (i.e., reduce the brightness fluctuations), successive signals or neighboring pixels can be averaged incoherently (i.e., their power values are added). This would lead to a more accurate radiometric measurement (and a more pleasing image) at the expense of degradation in the image resolution.

Another approach to reducing speckle is to combine images acquired at neighboring frequencies. In this case, the exact interference patterns lead to independent signals but with the same statistical properties. Incoherent

averaging would then result in a smoothing effect. In fact, this is why a scene illuminated with white light does not show speckled image behavior.

In most imaging SARs, the smoothing is done by averaging the brightness of neighboring pixels in azimuth, or range, or both. The number of statistically independent pixels averaged is called the number of looks N. We can show that the signal standard deviation S_N is related to the mean signal power $\overline{P}$ by:

$$S_N = \frac{1}{\sqrt{N}} \overline{P} . \tag{1.6-5}$$

The larger the number of looks N, the better the quality of the image from the radiometric point of view. However, this degrades the spatial resolution of the image. Noted that for N larger than about 25, large increase in N leads to only a small decrease in the signal fluctuation. This small improvement in the radiometric resolution should be traded off against the large degradation in the spatial resolution. For example, if one were to average 10 resolution cells in a four-look image, the speckle noise will be reduced to about 0.5 dB. At the same time, however, the image resolution will be reduced by an order of magnitude. Whether this loss in resolution is worth the reduction in speckle noise depends on both the aim of the investigation and the kind of scene imaged.

Fig. 1-11 shows the effect of multi-look averaging. An image acquired by the NASA/JPL AIRSAR system is shown displayed at one, four, sixteen, and thirty-two looks, respectively. This figure clearly illustrates the smoothing effect, as well as the decrease in resolution resulting from the multi-look process. In one early survey of geologists, the results showed that even though the optimum number of looks depended on the scene type and resolution, the majority of the responses preferred 2-look images. However, this survey dealt with images that had rather poor resolution to begin with. One may well find that with today's higher-resolution systems, analysts might be asking for a larger number of looks.

1.7 Summary

In this chapter we introduced some of the terms required to understand SAR imaging as described in the rest of this text. For a more in-depth treatment of SAR imaging and system engineering, the reader is referred to the two texts in the bibliography [2,3]. In the rest of this text, we shall concentrate specifically on polarimetric SAR and its applications.

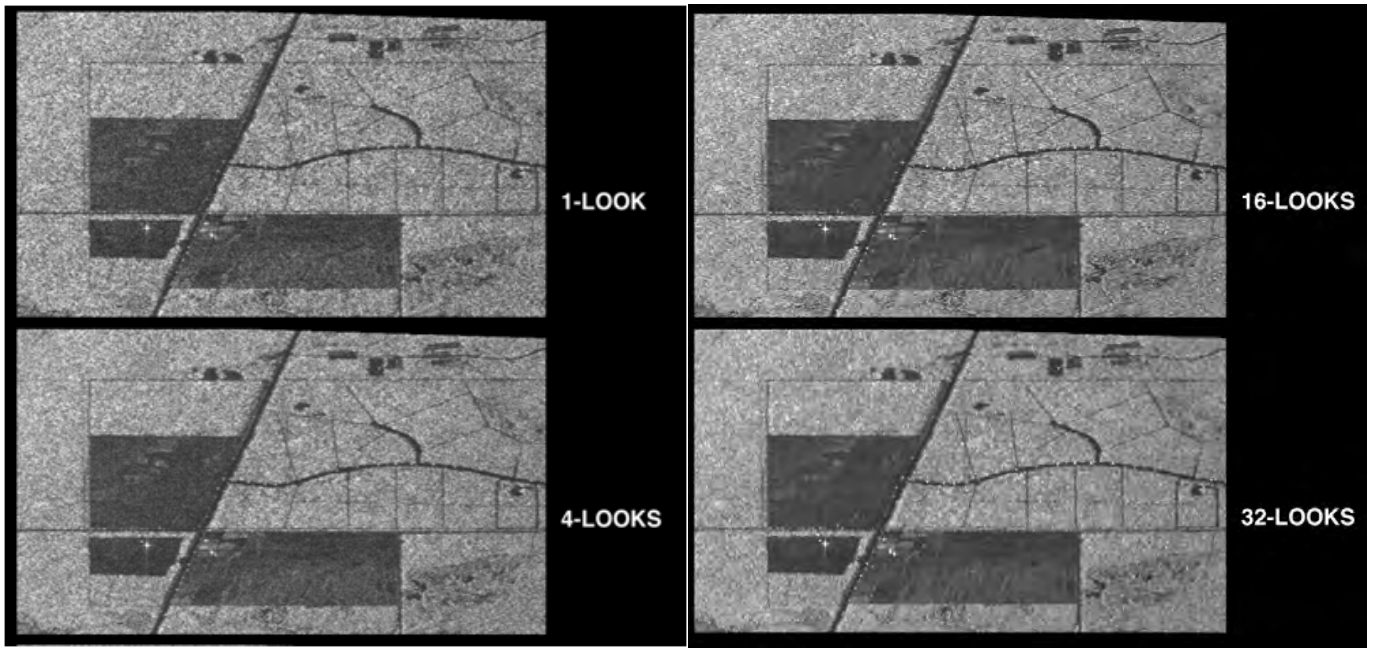


Fig. 1-11. The effects of speckle can be reduced by incoherently averaging pixels in a radar image, a process known as multi-looking. Shown in this image is the same image, processed at 1 look, 4 looks, 16 looks, and 32 looks. Note the reduction in the granular texture as the number of looks increase, while at the same time the resolution of the image decreases. Some features, such as those in the largest dark patch, might be completely masked by the speckle noise. (Image courtesy of the Jet Propulsion Laboratory.)

References

[1] F.T. Ulaby, R. K. Moore and A. K. Fung, *Microwave Remote Sensing Active and Passive*, Vol. II, pp. 583–595, ARTECH House, Norwood, Massachusetts, 1986.

[2] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar System and Signal Processing*, John Wiley & Sons, New York, 1991.

[3] C. Elachi, Spaceborne Radar Remote Sensing: Applications and Techniques, IEEE Press, New York, 1988.

Chapter 2
Basic Principles of SAR Polarimetry

The field of synthetic aperture radar changed dramatically in the early 1980s with the introduction of advance radar techniques, such as polarimetry and interferometry. While both of these techniques had been demonstrated much earlier, radar polarimetry only became an operational research tool with the introduction of the NASA/JPL Airborne Synthetic Aperture Radar (AIRSAR) system in the early 1980s. Radar polarimetry was proven from space with the two Spaceborne Imaging Radar C-band and X-band (SIR-C/X) SAR flights on board the space shuttle Endeavour in April and October 1994. In this chapter, we describe the basic principles of SAR polarimetry and, thereby, provide tools necessary to understand SAR polarimetry applications, such as land classification.

2.1 Polarization of Electromagnetic Waves

In SAR polarimetry, information is transmitted from an object to a sensor by electromagnetic waves. The information could be encoded in the frequency content, intensity, or polarization of the electromagnetic wave. The electromagnetic waves propagate at the velocity of light from the object directly through free space or indirectly by reflection, scattering, and radiation to the sensor. The interaction of electromagnetic waves with natural surfaces and atmospheres is strongly dependent on the frequency of the waves.

An electromagnetic wave consists of a coupled electric and magnetic force field. In free space, these two fields are at right angles to each other and transverse to the direction of propagation. The direction and magnitude of only one of the fields (usually the electric field) is sufficient to completely specify the direction and magnitude of the other field in free space using Maxwell's equations. The polarization of the electromagnetic wave is contained in the

elements of the vector amplitude $\mathbf{A}$ of the electric field. For a transverse electromagnetic wave, this vector is orthogonal to the direction in which the wave is propagating; we can, therefore, completely describe the amplitude of the electric field by writing $\mathbf{A}$ as a two-dimensional complex vector:

$$\mathbf{A} = a_h e^{i\delta_h}\hat{\mathbf{h}} + a_v e^{i\delta_v}\hat{\mathbf{v}}. \tag{2.1-1}$$

Here, we denote the two orthogonal basis vectors as $\hat{\mathbf{h}}$ for horizontal and $\hat{\mathbf{v}}$ for vertical. Horizontal polarization is usually defined as the state where the electric vector is perpendicular to the plane of incidence. Vertical polarization is orthogonal to both horizontal polarization and the direction of propagation and corresponds to the case where the electric vector is in the plane of incidence. Any two orthogonal basis vectors could be used to describe the polarization; in some cases, the right- and left-handed circular basis is used. The amplitudes, a_h and a_v, and the relative phases, δ_h and δ_v, are real numbers. The polarization of the wave can be thought of as the shape that the tip of the electric field would trace over time at a fixed point in space. Taking the real part of Eq. (2.1-1), we find that the polarization figure is the locus of all the points in the h-v plane that have the coordinates $E_h = a_h \cos\delta_h;\quad E_v = a_v \cos\delta_v$. It can easily be shown that the points on the locus satisfy the expression

$$\left(\frac{E_h}{a_h}\right)^2 + \left(\frac{E_v}{a_v}\right)^2 - 2\frac{E_h}{a_h}\frac{E_v}{a_v}\cos\left(\delta_h - \delta_v\right) = \sin^2\left(\delta_h - \delta_v\right). \tag{2.1-2}$$

This is the expression of an ellipse (shown in Fig. 2-1). In the general case, therefore, electromagnetic waves are elliptically polarized. In tracing the ellipse, the tip of the electric field can rotate either clockwise or counter-clockwise; this direction is denoted by the handedness of the polarization. The definition of handedness accepted by the Institute for Electrical and Electronics Engineers (IEEE) is that a wave is said to have right-handed polarization if the tip of the electric field vector rotates clockwise when the wave is viewed receding from the observer. If the tip of the electric field vector rotates counter-clockwise when the wave is viewed in the same way, it has a left-handed polarization. It is worth pointing out that a different definition of handedness is often encountered in optics literature. Specifically, in optics literature, a wave is said to have a right-handed (left-handed) polarization when the wave is viewed approaching the observer and the tip of the electric field vector rotates in the clockwise (counter-clockwise) direction.

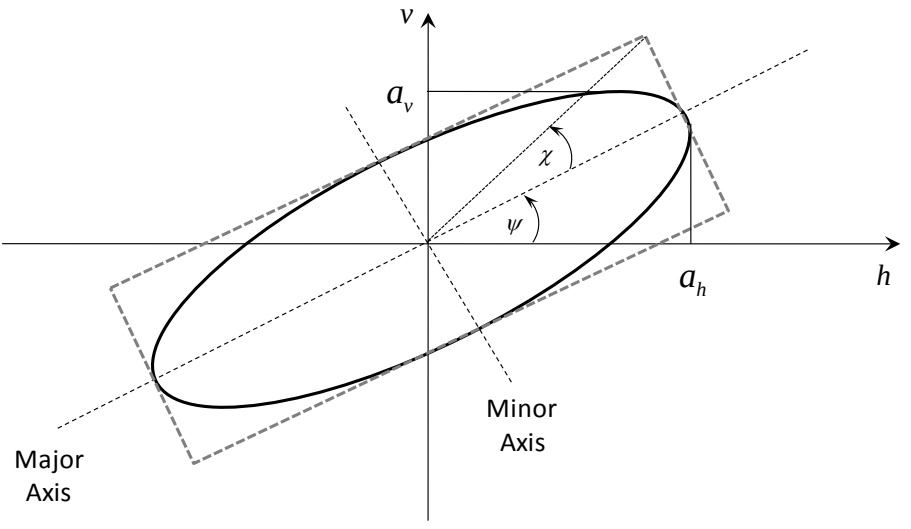


Fig. 2-1. A polarization ellipse.

In the special case where the ellipse collapses to a line, which happens when $\delta_h - \delta_v = n\pi$ with n any integer, the wave is said to be linearly polarized. Another special case is encountered when the two amplitudes are the same $(a_h = a_v)$ and the relative phase difference $\delta_h - \delta_v$ is either $\pi/2$ or $-\pi/2$. In this case, the wave is circularly polarized.

The polarization ellipse (see Fig. 2-1) can also be characterized by two angles known as the ellipse orientation angle (ψ in Fig. 2-1, $0 \le \psi \le \pi$) and the ellipticity angle, shown as χ ($-\pi/4 \le \chi \le \pi/4$) in Fig. 2-1. These angles can be calculated as follows:

$$\tan 2\psi = \frac{2a_h a_v}{a_h^2 - a_v^2}\cos(\delta_h - \delta_v); \quad \sin 2\chi = \frac{2a_h a_v}{a_h^2 + a_v^2}\sin(\delta_h - \delta_v). \qquad (2.1\text{-}3)$$

Note that linear polarizations are characterized by an ellipticity angle $\chi = 0$. Note also that two waves are orthogonally polarized: that is, the scalar product of the two polarization vectors will be zero if the two polarization ellipses have orientation angles that are 90 degrees (deg) different and the handedness of the two waves are opposite.

So far, it was implied that the amplitudes and phases shown in Eq. (2.1-1) and Eq. (2.1-2) are constant in time. This might not always be the case. If these quantities vary with time, the tip of the electric field vector will not trace out a smooth ellipse. Instead, the figure will, in general, be a noisy version of an

ellipse that after some time might resemble an "average" ellipse. In this case, the wave is said to be partially polarized, and it can be considered that part of the energy has a deterministic polarization state. The radiation from some sources, such as the Sun, does not have any clearly defined polarization. The electric field assumes different directions at random as the wave is received. In this case, the wave is called randomly polarized or unpolarized. In the case of some man-made sources, such as lasers and radio/radar transmitters, the wave usually has a well-defined polarized state.

Another way to describe the polarization of a wave that is particularly appropriate for the case of partially polarized waves is through the use of the Stokes parameters of the wave. For a monochromatic wave, these four parameters are defined as

$$\begin{aligned} S_0 &= |a_h|^2 + |a_v|^2 \\ S_1 &= |a_h|^2 - |a_v|^2 \\ S_2 &= 2a_h a_v \cos(\delta_h - \delta_v) \\ S_2 &= 2a_h a_v \sin(\delta_h - \delta_v) \end{aligned} \quad . \tag{2.1-4}$$

Note that for such a fully polarized wave, only three of the Stokes parameters are independent, since $S_0^2 = S_1^2 + S_2^2 + S_3^2$. Using the relations in Eq. (2.1-3) between the ellipse orientation and ellipticity angles and the wave amplitudes and relative phases, it can be shown that the Stokes parameters can also be written as

$$\begin{aligned} S_1 &= S_0 \cos 2\chi \cos 2\psi \\ S_2 &= S_0 \cos 2\chi \sin 2\psi \\ S_3 &= S_0 \sin 2\chi \end{aligned} \quad . \tag{2.1-5}$$

If two ellipse orientations differ by 90 deg and the handedness of the ellipses are opposite (that is, the ellipticity angles are equal but of opposite sign), it follows from Eq. (2.1-5) that the Stokes parameters of two orthogonally polarized waves are the same magnitudes, but opposite in sign.

The relations in Eq. (2.1-5) lead to a simple geometric interpretation of polarization states. The Stokes parameters S_1, S_2 and S_3 can be regarded as the Cartesian coordinates of a point on a sphere, known as the Poincaré sphere, of radius S_0 (see Fig. 2-2). There is, therefore, a unique mapping between the position of a point in the surface of the sphere and a polarization state. Linear polarizations map to points on the equator of the Poincaré sphere, while the circular polarizations map to the poles. Orthogonal polarizations are anti-podal

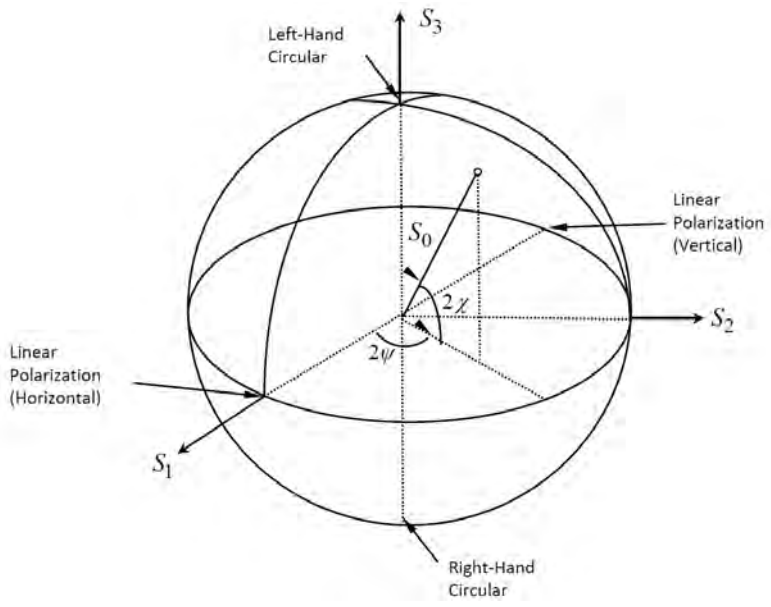


Fig. 2-2. Polarization represented as a point on the Poincaré sphere.

on the Poincaré sphere, which means they lie on opposite sides of the sphere and the line connecting the orthogonal polarizations runs through the center of the sphere. See, for example, the positions of horizontally and vertically polarized linear polarizations or the two circular polarizations in Fig. 2-2.

In the case of partially polarized waves, all four Stokes parameters are required to fully describe the polarization of the wave. In general, the Stokes parameters are related by $S_0^2 \geq S_1^2 + S_2^2 + S_3^2$, with equality holding only for fully polarized waves. In the extreme case of an unpolarized wave, the Stokes parameters are $S_0 > 0$ and $S_1 = S_2 = S_3 = 0$. It is always possible to describe a partially polarized wave by the sum of a fully polarized wave and an unpolarized wave. The magnitude of the polarized wave is given by $\sqrt{S_1^2 + S_2^2 + S_3^2}$; the magnitude of the unpolarized wave is $S_0 - \sqrt{S_1^2 + S_2^2 + S_3^2}$. Finally, it should be pointed out that the Stokes parameters of an unpolarized wave can be written as the sum of two fully polarized waves, as follows:

$$\begin{pmatrix} S_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} S_0 \\ -S_1 \\ -S_2 \\ -S_3 \end{pmatrix}. \tag{2.1-6}$$

These two fully polarized waves have orthogonal polarizations. This important result shows that when an antenna with a particular polarization is used to receive unpolarized radiation, the amount of power received by the antenna will be only that half of the power in the unpolarized wave that aligns with the antenna polarization. The other half of the power will not be absorbed because its polarization is orthogonal to that of the antenna.

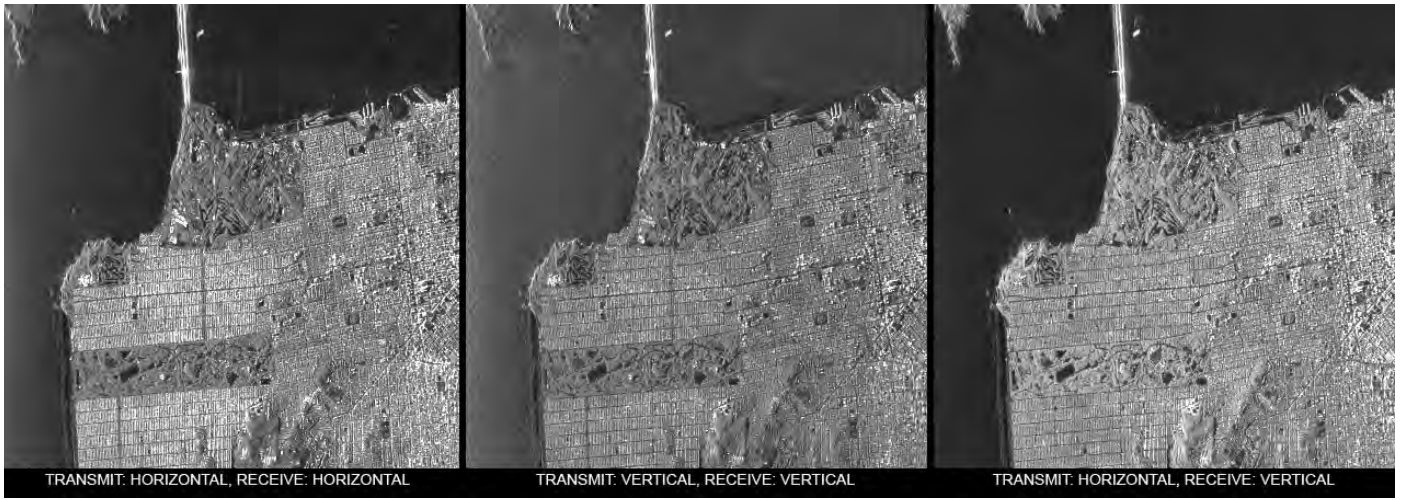
2.2 Mathematical Representations of Scatterers

If a radiated electromagnetic wave is scattered by an object and one observes this wave in the far-field of the scatterer, the scattered wave can, again, be adequately described by a two-dimensional vector. In this abstract way, one can consider the scatterer as a mathematical operator that takes one two-dimensional complex vector (the wave impinging upon the object) and changes that into another two-dimensional vector (the scattered wave). Mathematically, therefore, a scatterer can be characterized by a complex 2 × 2 scattering matrix:

$$\mathbf{E}^{sc} = \begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix}\mathbf{E}^{tr} = [\mathbf{S}]\mathbf{E}^{tr} \tag{2.2-1}$$

where $\mathbf{E}^{tr}$ is the electric field vector that was transmitted by the radar antenna, $[\mathbf{S}]$ is the 2 × 2 complex scattering matrix that describes how the scatterer modified the incident electric field vector, and $\mathbf{E}^{sc}$ is the electric field vector that is incident on the radar receiving antenna. This scattering matrix is also a function of the radar frequency and the viewing geometry. The scatterer can, therefore, be thought of as a polarization transformer, with the transformation given by the scattering matrix. Once the complete scattering matrix is known and calibrated, one can synthesize the radar cross-section for any arbitrary combination of transmit and receive polarizations.

Fig. 2-3 shows a number of such synthesized images for the San Francisco Bay area in California. The data were acquired with the NASA/JPL AIRSAR system.



(a) Horizontal and vertical transmission

Fig. 2-3. This series of L-band images of San Francisco were synthesized from a single polarimetric image acquired by the NASA/JPL AIRSAR system at L-band. The nine images show the co-polarized (transmit and receive polarizations are the same) and the cross-polarized (transmit and receive polarizations are orthogonal) images for the three axes of the Poincaré sphere. They are (a)horizontal and vertical transmission, (b) different linear polarization angles, and (c) different circular polarizations. Note the relative change in brightness between the city of San Francisco, the ocean, and the Golden Gate Park, which is the large rectangle about 1/3 from the bottom of the images. The radar illumination is from the left.

(b) Different linear polarization angles

Fig. 2-3. This series of L-band images of San Francisco were synthesized from a single polarimetric image acquired by the NASA/JPL AIRSAR system at L-band. The nine images show the co-polarized (transmit and receive polarizations are the same) and the cross-polarized (transmit and receive polarizations are orthogonal) images for the three axes of the Poincaré sphere. They are (a)horizontal and vertical transmission, (b) different linear polarization angles, and (c) different circular polarizations. Note the relative change in brightness between the city of San Francisco, the ocean, and the Golden Gate Park, which is the large rectangle about 1/3 from the bottom of the images. The radar illumination is from the left.

(c) Different circular polarizations

Fig. 2-3. This series of L-band images of San Francisco were synthesized from a single polarimetric image acquired by the NASA/JPL AIRSAR system at L-band. The nine images show the co-polarized (transmit and receive polarizations are the same) and the cross-polarized (transmit and receive polarizations are orthogonal) images for the three axes of the Poincaré sphere. They are (a)horizontal and vertical transmission, (b) different linear polarization angles, and (c) different circular polarizations. Note the relative change in brightness between the city of San Francisco, the ocean, and the Golden Gate Park, which is the large rectangle about 1/3 from the bottom of the images. The radar illumination is from the left.

To specify the polarization vectors and electric field vectors, we need to define the radar coordinate system. In this book, we will use the backscatter alignment coordinate system as shown in Fig. 2-4; this is the coordinate system in which radar measurements are performed. (See Ulaby and Elachi, [1], Chapter 2, for a more detailed discussion of coordinate systems.)

The voltage measured by the radar system is proportional to the scalar product of the radar antenna polarization and the incident wave electric field, i.e.:

$$V = c_a \mathbf{p}^{rec} \cdot [\mathbf{S}] \mathbf{p}^{tr} , \tag{2.2-2}$$

where $\mathbf{p}^{tr}$ and $\mathbf{p}^{rec}$ are the normalized polarization vectors describing the transmitting and receiving radar antennas expressed in the backscatter alignment coordinate system, and c_a is a factor that includes the transmitting antenna gain, the receiving antenna effective area and the distance to the scatterer (see the derivation of the radar equation in Chapter 1). For our purposes here, we are interested in the properties of the scatterer, so we shall

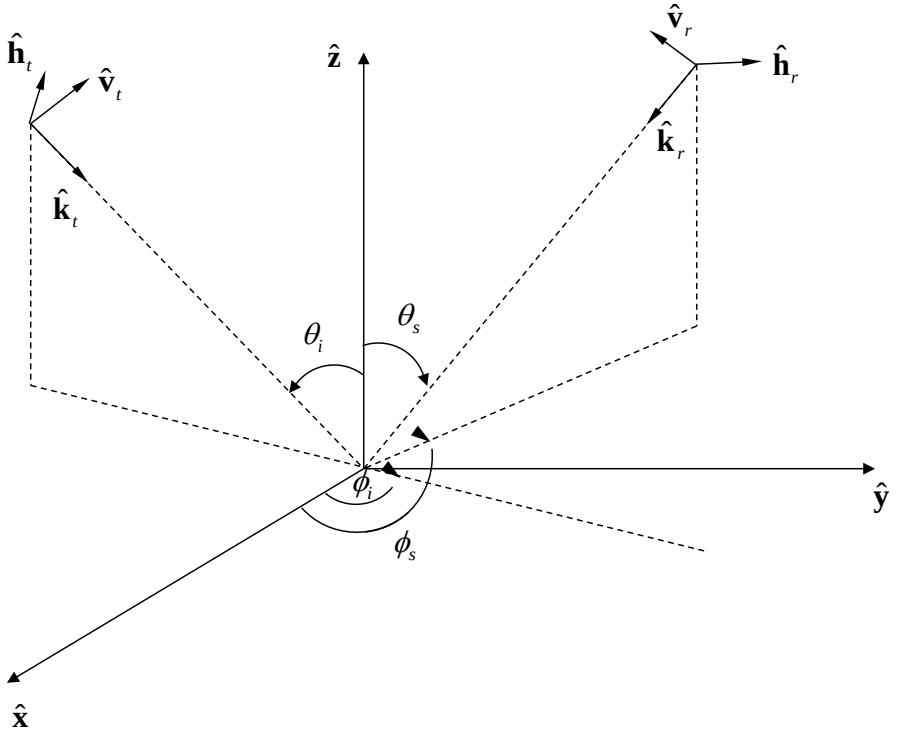


Fig. 2-4. A backscatter alignment coordinate system. Notice that the transmitting and receiving polarizations coincide for the backscattering case where the transmitting and receiving antennas are located at the same place.

ignore c_a in the rest of the discussion. The power received by the radar is the magnitude of the voltage squared (Kennaugh [2]; Kostinski and Boerner [3]; van Zyl et al. [4]; Zebker et al. [5]:

$$P = VV^* = \left| \mathbf{p}^{rec} \cdot [\mathbf{S}] \mathbf{p}^{tr} \right|^2 . \tag{2.2-3}$$

Expanding the expression inside the magnitude sign in Eq. (2.2-3), it can be shown that the received power can also be written in terms of the scatterer covariance matrix, as follows:

$$P = VV^* = \left(\tilde{\mathbf{A}} \mathbf{T} \right) \left(\tilde{\mathbf{T}} \mathbf{A} \right)^* = \tilde{\mathbf{A}} \mathbf{T} \tilde{\mathbf{T}}^* \mathbf{A}^* = \mathbf{A} \cdot [\mathbf{C}] \mathbf{A}^*; \quad [\mathbf{C}] = \mathbf{T} \tilde{\mathbf{T}}^* , \tag{2.2-4}$$

where $\tilde{\mathbf{A}} = \begin{pmatrix} p_h^{rec} p_h^{tr} & p_h^{rec} p_v^{tr} & p_v^{rec} p_h^{tr} & p_v^{rec} p_v^{tr} \end{pmatrix}$ represents the transpose of the antenna polarization vector elements and $\tilde{\mathbf{T}} = \begin{pmatrix} S_{hh} & S_{hv} & S_{vh} & S_{vv} \end{pmatrix}$ represents only the scatterer. The superscript * denotes complex conjugation. The covariance matrix characterization is particularly useful when analyzing multi-look radar images, since the covariance matrix of a multi-look pixel is simply the average covariance matrix of all the individual measurements contained in the multi-look pixel.

Recall that multi-looking is performed by averaging the power from adjacent pixels together to reduce speckle. This averaging process can be written as

$$\langle P \rangle = \frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} P_{ij} = \frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} \mathbf{A} \cdot [C]_{ij} \mathbf{A}^* = \mathbf{A} \cdot \langle [\mathbf{C}] \rangle \mathbf{A}^* , \tag{2.2-5}$$

where the two subscripts denote averaging in the range and azimuth directions, respectively. The angular brackets $\langle \ \rangle$ denote this spatial averaging.

Eq. (2.2-4) shows the covariance matrix to be a 4 × 4 complex Hermetian matrix. In the case of radar backscatter, reciprocity dictates that $S_{hv} = S_{vh}$ and the covariance matrix can, in general, be written as a 3 × 3 complex Hermetian matrix. Also note that it is always possible to calculate the covariance matrix from the scattering matrix. However, the inverse is not true: it is not always possible to calculate an equivalent scattering matrix from the covariance matrix. This follows from the fact that the off-diagonal terms in the covariance matrix involve cross-products of the scattering matrix elements (for example, $S_{hh} S_{hv}^*$). For a single scattering matrix, there is a definite relationship between this term and the two diagonal terms $S_{hh} S_{hh}^*$ and $S_{hv} S_{hv}^*$. However, once the covariance

matrix elements are averaged spatially (such as during multi-looking of an image), this definite relationship no longer holds and we cannot uniquely find an equivalent S_{hh} and S_{hv} that would satisfy all three cross-products $\left\langle S_{hh}S_{hv}^{*}\right\rangle$, $\left\langle S_{hh}S_{hh}^{*}\right\rangle$, and $\left\langle S_{hv}S_{hv}^{*}\right\rangle$.

The power expression shown in Eq. (2.2-3) can also be written in terms of the antenna Stokes vectors. First, consider the following form of the power equation (van Zyl [6]; van Zyl et al. [7]):

$$\begin{aligned}
P &= \left(\mathbf{p}^{rec}\cdot\mathbf{E}^{sc}\right)\left(\mathbf{p}^{rec}\cdot\mathbf{E}^{sc}\right)^{*} \\
&= \left(p_h^{rec}E_h^{sc}+p_v^{rec}E_v^{sc}\right)\left(p_h^{rec}E_h^{sc}+p_v^{rec}E_v^{sc}\right)^{*} \\
&= \left(p_h^{rec}p_h^{rec*}\right)\left(E_h^{sc}E_h^{sc*}\right)+\left(p_v^{rec}p_v^{rec*}\right)\left(E_v^{sc}E_v^{sc*}\right) \\
&\quad +\left(p_h^{rec}p_v^{rec*}\right)\left(E_h^{sc}E_v^{sc*}\right)+\left(p_v^{rec}p_h^{rec*}\right)\left(E_v^{sc}E_h^{sc*}\right) \\
&= \begin{pmatrix} p_h^{rec}p_h^{rec*} \\ p_v^{rec}p_v^{rec*} \\ p_h^{rec}p_v^{rec*} \\ p_v^{rec}p_h^{rec*} \end{pmatrix}\cdot\begin{pmatrix} E_h^{sc}E_h^{sc*} \\ E_v^{sc}E_v^{sc*} \\ E_h^{sc}E_v^{sc*} \\ E_v^{sc}E_h^{sc*} \end{pmatrix} \\
&= \mathbf{g}^{rec}\cdot\mathbf{X}
\end{aligned} \tag{2.2-6}$$

Here, the vector $\mathbf{X}$ in Eq. (2.2-6) is a function of the transmit antenna parameters and the scattering matrix elements. Using the fact that $\mathbf{E}^{sc}=[\mathbf{S}]\mathbf{p}^{tr}$, it can be shown that $\mathbf{X}=[\mathbf{W}]\mathbf{g}^{tr}$, where

$$[\mathbf{W}]=\begin{pmatrix} S_{hh}S_{hh}^{*} & S_{hv}S_{hv}^{*} & S_{hh}S_{hv}^{*} & S_{hv}S_{hh}^{*} \\ S_{vh}S_{vh}^{*} & S_{vv}S_{vv}^{*} & S_{vh}S_{vv}^{*} & S_{vv}S_{vh}^{*} \\ S_{hh}S_{vh}^{*} & S_{hv}S_{vv}^{*} & S_{hh}S_{vv}^{*} & S_{hv}S_{vh}^{*} \\ S_{vh}S_{hh}^{*} & S_{vv}S_{hv}^{*} & S_{vh}S_{hv}^{*} & S_{vv}S_{hh}^{*} \end{pmatrix}. \tag{2.2-7}$$

Therefore, the measured power can be expressed as

$$P=\mathbf{g}^{rec}\cdot[\mathbf{W}]\mathbf{g}^{tr}. \tag{2.2-8}$$

The Stokes vector of a wave can be written as

$$\mathbf{S} = \begin{pmatrix} p_h p_h^* + p_v p_v^* \\ p_h p_h^* - p_v p_v^* \\ p_h p_v^* + p_h^* p_v \\ -i(p_h p_v^* - p_h^* p_v) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -i & i \end{pmatrix} \begin{pmatrix} p_h p_h^* \\ p_v p_v^* \\ p_h p_v^* \\ p_h^* p_v \end{pmatrix} = [\mathbf{R}]\mathbf{g} \quad . \tag{2.2-9}$$

From Eq. (2.2-9), $\mathbf{g} = [\mathbf{R}]^{-1}\mathbf{S}$. Then, after straightforward calculations, it can be shown that

$$P = \mathbf{S}^{rec} \cdot [\mathbf{M}]\mathbf{S}^{tr} . \tag{2.2-10}$$

The matrix $[\mathbf{M}]$ is known as the Stokes scattering operator and is given by

$$[\mathbf{M}] = \left[[\mathbf{R}]^{-1} \right]^T [\mathbf{W}][\mathbf{R}]^{-1}, \tag{2.2-11}$$

where the superscript *T* indicates the transpose of the matrix. Note that, like the covariance matrix, the average power can be written in terms of the average Stokes scattering operator.

2.3 Implementation of a Radar Polarimeter

Polarimetric radars must measure the full scattering matrix to preserve the information regarding the scatterer. From Eq. (2.2-2), it can be seen that setting one transmit vector element equal to zero allows us to measure two components of the scattering matrix at a time. Mathematically, this operation is expressed as

$$\begin{pmatrix} S_{hh} \\ S_{vh} \end{pmatrix} = \begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{tr} ; \begin{pmatrix} S_{hv} \\ S_{vv} \end{pmatrix} = \begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}^{tr} . \tag{2.3-1}$$

Eq. (2.3-1) represents the typical implementation of a radar polarimeter: that is, a radar polarimeter transmits a wave of one polarization and receiving echoes in two orthogonal polarizations simultaneously. This is followed by transmitting a wave with a second polarization and, again, receiving echoes with both polarizations simultaneously (as is shown in Fig. 2-5). In this way, all four elements of the scattering matrix are measured. This implementation means that the transmitter is in slightly different positions when measuring the two columns of the scattering matrix, but this distance between the two positions is typically small compared to a synthetic aperture and, therefore, does not lead to a significant decorrelation of the signals. The more important aspect of this

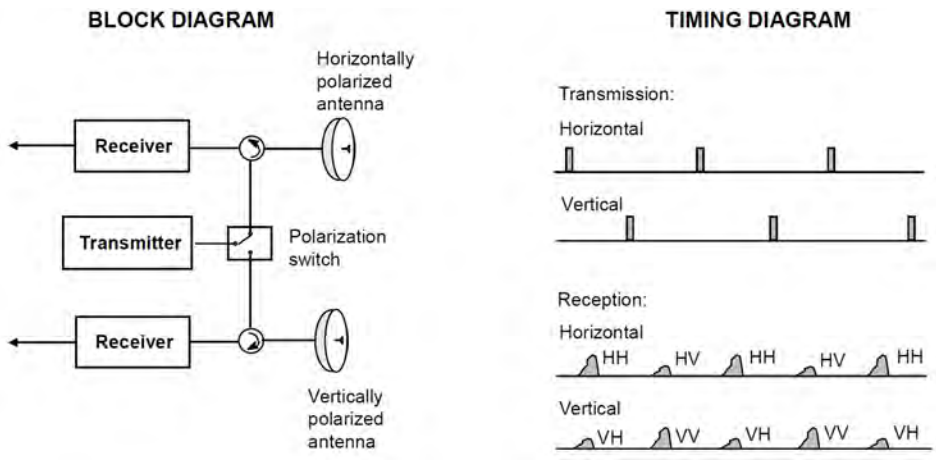


Fig. 2-5. A polarimetric radar is implemented by alternatively transmitting signals out of horizontally and vertically polarized antennas and receiving at both polarizations simultaneously. Two pulses are needed to measure all the elements in the scattering matrix.

implementation is that the pulse repetition frequency (PRF) must be high enough to ensure that each polarimetric channel is sampled adequately. Therefore, each channel must independently satisfy the minimum PRF requirement. Since we are interleaving two measurements, this means that the master PRF for a polarimetric system runs twice the rate of a single-channel SAR. The NASA/JPL AIRSAR system pioneered this implementation for SAR systems. Subsequently, the same implementation was used in the SIR-C part of the SIR-C/X-SAR radars.

A polarimetric SAR implemented in this fashion actually acquires four images: one each for the horizontal-horizontal (HH), horizontal-vertical (HV), vertical-horizontal (VH), and vertical-vertical (VV) combinations. The basic measurement for each pixel in the highest resolution image is, therefore, a complete scattering matrix, or four complex numbers. If the SAR operates in the backscatter mode, reciprocity dictates that $S_{hv} = S_{vh}$ and that there are only three independent images. In practice, the HV and VH measurements are made at different times and through different receivers, so thermal noise in the system will cause these numbers to be different. Once the channels are properly calibrated, any remaining differences are due to thermal noise. Therefore, one could, in fact, average these two channels together coherently to increase the signal-to-noise ratio in the cross-polarized image. After this operation, one is left with three independent complex numbers per pixel.

As discussed in the previous chapter, all SAR images suffer from speckle noise, which is the result of coherent interference from individual scatterers that might be present inside a pixel. To reduce this speckle noise, the power from adjacent

pixels are averaged; this process is known as multi-looking. We have shown in Eq. (2.2-5) that all the polarimetric information can be retained by performing this multi-looking operation by averaging the covariance matrices of adjacent pixels. A similar operation follows from Eq. (2.2-10) for the Stokes scattering operator case, as follows:

$$P_m = \langle P \rangle = \frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} P_{ij} = \frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} \mathbf{S}^{rec} \cdot [\mathbf{M}]_{ij} \mathbf{S}^{tr} = \mathbf{S}^{rec} \cdot \langle [\mathbf{M}] \rangle \mathbf{S}^{tr} . \qquad (2.3\text{-}2)$$

This multi-looking operation can be done once; all subsequent analyses would then be performed on the multi-looked data set. In fact, the polarimetric data from the NASA/JPL AIRSAR system is distributed in a multi-looked format with some special compression formatting to reduce the data volume further. The multi-looked polarimetric data from the SIR-C radar was distributed as cross-products of the scattering matrix, which are the elements of the covariance matrix.

Polarimetric SAR systems place additional restrictions on the pulse repetition frequency (PRF) used to operate the radar. Each transmit polarization channel must satisfy the normal constraints imposed on SAR systems using a single transmit polarization. The result is that polarimetric systems operate with a master PRF that runs twice as fast as that of a single transmit channel SAR. Additionally, the range ambiguities of a polarimetric SAR system are more complicated than those of a single channel SAR. In the HV channel, for example, one would measure ambiguous signals from the next and the previous pulses (in fact, at all odd numbers of ambiguous pulses) at HH. Given that HV is usually much smaller than HH to begin with, we then have to place even more stringent requirements on the overall ambiguity levels to measure HV accurately in the presence of the ambiguities. This severely limits the useful swath width that can be achieved with polarimetric SAR systems from space.

One way to achieve much of the desirable information from a polarimetric measurement with reduced requirements on the PRF and the ambiguity level is to use so-called compact polarimetry (Souyris et al. [8]). Compact polarimetry essentially is a special, dual-polarization mode in which one polarization only is transmitted and two orthogonal polarizations are used to measure the return. In the original mode proposed by Souyris et al. [8], a 45 deg linear polarization signal is transmitted and horizontal and vertical polarizations are used to receive the signal. In this case, the received signal is simply

$$V_h = \frac{1}{\sqrt{2}} \left(S_{hh} + S_{hv} \right); \quad V_v = \frac{1}{\sqrt{2}} \left(S_{vh} + S_{vv} \right). \qquad (2.3\text{-}3)$$

The average covariance matrix is then

$$\left\langle \left[C_p \right] \right\rangle \approx \frac{1}{2} \begin{pmatrix} \left\langle S_{hh} S_{hh}^* \right\rangle + \left\langle S_{hv} S_{hv}^* \right\rangle & \left\langle S_{hh} S_{vv}^* \right\rangle + \left\langle S_{hv} S_{hv}^* \right\rangle \\ \left\langle S_{vv} S_{hh}^* \right\rangle + \left\langle S_{hv} S_{hv}^* \right\rangle & \left\langle S_{hv} S_{hv}^* \right\rangle + \left\langle S_{vv} S_{vv}^* \right\rangle \end{pmatrix}. \quad (2.3\text{-}4)$$

In deriving this covariance matrix, we have assumed that the terrain exhibits reflection symmetry. Compact polarimetry allows us to relax the requirements on the PRF to be the same as that of a conventional SAR system. Compact polarimetry also balances the ambiguity levels better than that of a regular polarimetric system. The drawback, of course, is that we no longer have "pure" measurements of co-polarized and cross-polarized terms. Instead, we have mixtures of co- and cross-polarized terms in all components of the covariance matrix.

Following the original proposed compact polarimetry mode, Raney [9] suggested transmitting circular polarization to be more advantageous in the presence of Faraday rotation. The basic expressions are similar to those derived above, however, and will not be repeated here.

2.4 Polarization Response

Once the scattering matrix, the covariance matrix, or the Stokes matrix is known, one can synthesize the received power for any transmit and receive antenna polarizations using the power equations ((Eq. (2.2-3), Eq. (2.2-4), and Eq. (2.2-10)). This is known as polarization synthesis and is discussed in more detail in Chapter 2 of Ulaby and Elachi [1]. Note that if we allow the polarization of the transmit and receive antennas to be varied independently, the polarization response of the scene would be a four-dimensional space. This is most easily understood by representing each of the two polarizations by the orientation and ellipticity angles of the respective polarization ellipses. The polarization response is, therefore, a function of these four angles. Visualizing such a four-dimensional response is not easy. To simplify the visualization, the so-called polarization response (van Zyl [6]; Agrawal and Boerner [10]; Ulaby and Elachi [1]) was introduced. The polarization response is displayed as a three-dimensional figure, and the transmit and receive polarizations are either the same (the co-polarized response) or they are orthogonal (the cross-polarized response). One can also display the maximum or minimum received power as a function of transmit polarization or the polarized and unpolarized component of the power using this same display. Agrawal and Boerner [10] also used this method to display the relative phase of the received signal as a function of polarization.

We shall introduce the polarization response through the example of a trihedral corner reflector that has been used extensively for the polarimetric radar

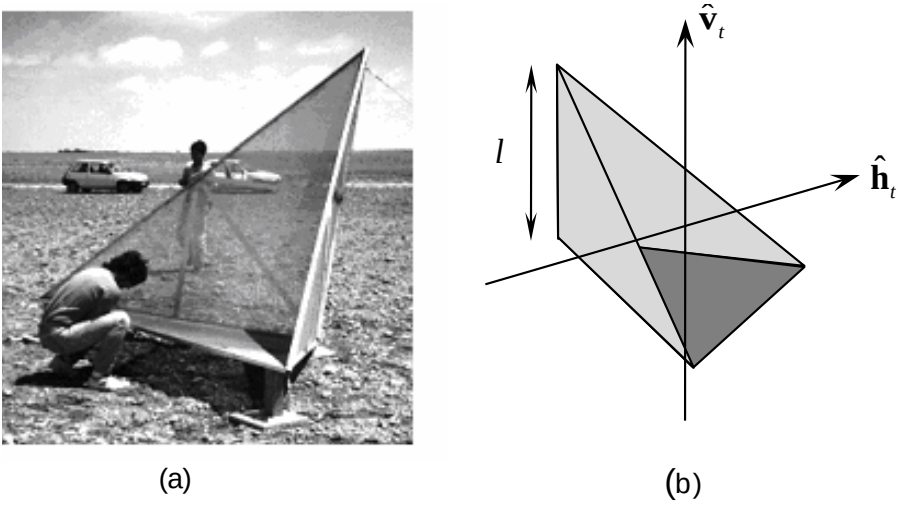


Fig. 2-6. (a) A trihedral corner reflector is being deployed at the calibration site. (b) The trihedral corner reflector geometry.

calibration. A picture of a trihedral corner reflector is shown in Fig. 2-6. The scattering matrix of a trihedral corner reflector is given by

$$[\mathbf{S}] = c_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.4-1}$$

where $c_1 = \dfrac{k_0 l^2}{\sqrt{12\pi}}$. From Eq. (2.4-1), the characteristics of a trihedral corner reflector are:

1. No cross polarization components are generated (HV = VH = 0) for the linear polarization case.
2. Horizontal and vertical backscattering cross sections are identical (HH = VV).
3. Horizontal and vertical co-polarized components are in phase.

These are the desired properties of a calibration target to balance the co-polarized elements (i.e., the diagonal terms) of the scattering matrix. In addition, trihedral corner reflectors provide relatively large radar cross sections with a large, 3-dB beamwidth independent of the radar wavelength and the corner reflector size.

From Eq. (2.4-1), Eq. (2.2-4), and Eq. (2.2-11), we can calculate the covariance matrix and the Stokes scattering operator as

$$[\mathbf{C}] = |c_1|^2 \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \tag{2.4-2}$$

and

$$[\mathbf{M}] = \frac{1}{2}|c_1|^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.4-3}$$

Using Eq. (2.2-10) and Eq. (2.4-3), the received power can be calculated explicitly, as shown in Eq. (2.4-4) for the case of co-polarized and cross-polarized antennas:

$$\begin{aligned} P &= \frac{k_0^2 l^4}{24\pi^2}\left\{1 \pm \cos^2(2\psi)\cos^2(2\chi) \pm \sin^2(2\psi)\cos^2(2\chi) \mp \sin^2(2\chi)\right\} \\ &= \frac{k_0^2 l^4}{24\pi^2}\left\{1 \pm \left[\cos^2(2\chi) - \sin^2(2\chi)\right]\right\} \\ &= \frac{k_0^2 l^4}{24\pi^2}\left\{1 \pm \cos(4\chi)\right\} \end{aligned} \quad , \tag{2.4-4}$$

where the top sign is for the co-polarization case and the bottom sign is for the cross-polarization case. (Recall that orthogonally polarized waves are anti-podal on the Poincaré sphere.) The ellipse orientation and ellipticity angles refer to those of the transmitting antenna. It is immediately apparent that the received power is independent of the ellipticity angle. Taking the derivative with respect to the ellipse orientation angle, it is easily shown that for the co-polarized response, the maximum is found when $\chi = 0$ (linear polarizations) and the minimum occurs at $\chi = 45°$ (circular polarizations). The locations of the maxima and minima are reversed in the case of the cross-polarized response. This behavior is explained by the fact that the reflected waves have the opposite handedness than the transmitted ones when analyzed in the backscatter alignment coordinate system. Therefore, if either circular polarization is transmitted, the reflected wave is polarized orthogonally to the transmitted wave, leading to maximum reception in the cross-polarized and minimum reception in the co-polarized case.

The co- and cross-polarized responses are shown in Fig. 2-7. These three-dimensional displays show the normalized radar cross-section (power) in the vertical direction plotted as a function of the transmitting antenna ellipse orientation and ellipticity angles. For co-polarized responses, the receiving antenna has the same polarization as that of the transmitting antenna. The left-hand figure in Fig. 2-7 shows the positions of some of the more common antenna combinations for this type of display. For the cross-polarized response, the receiving antenna polarization is orthogonal to that of the transmitting antenna. Some of the common combinations are shown on the right in Fig. 2-7. Note that the displays in Fig. 2-7 confirm what we already discussed using Eq. (2.4-4). First, the received power is not a function of the ellipse orientation angle, as seen by the constant value of the radar cross-section in Fig. 2-7 for changing ellipse orientation. The maximum of the co-polarization signature occurs for the linear polarizations, and the minima at the circular polarizations. For the cross-polarization response, the maxima occur at the circular polarizations, while the minima occur at the linear polarizations.

A slightly different display of the polarization response involves displaying the normalized radar cross-section in spherical coordinates using the Poincaré sphere. The polarization is defined by the angles as shown in Fig. 2-2, but the radius is plotted as the normalized received power. This three-dimensional version of the polarization responses for the trihedral corner reflector is shown in Fig. 2-8. The co-polarized response clearly shows the nulls at the circular polarizations; the cross-polarized response shows the corresponding maximum for the circular polarizations.

The second example of polarization responses is that of a dihedral corner reflector, shown in Fig. 2-9. The scattering matrix of a dihedral corner reflector

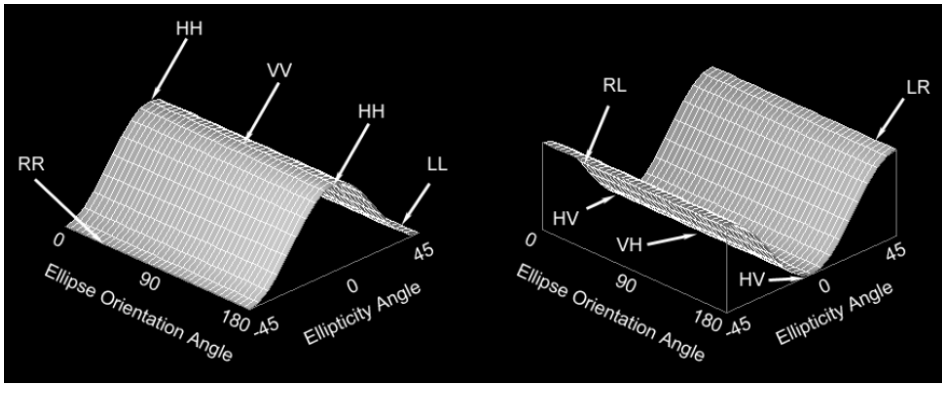


Fig. 2-7. The polarization response of a trihedral corner reflector. Two responses are displayed: (right) identical transmit and receive polarizations (co-polarized) and (left) orthogonal transmit and receive polarizations (cross-polarized).

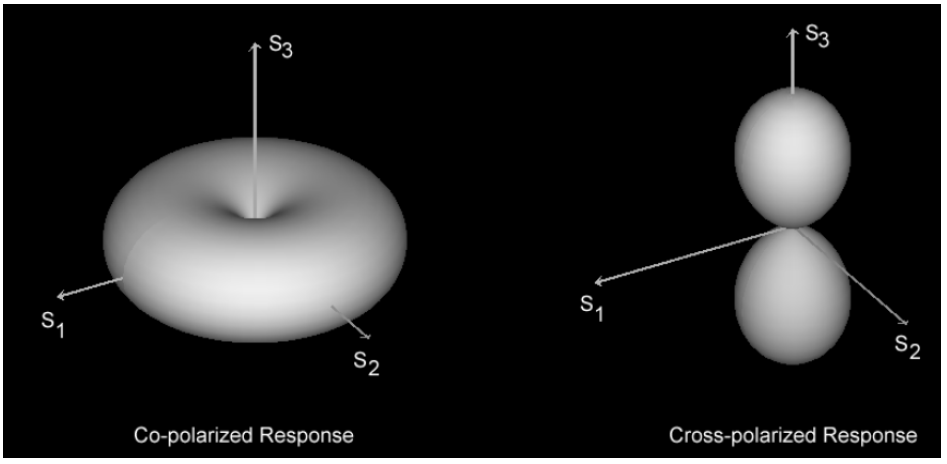


Fig. 2-8. Three-dimensional polarization response of a trihedral corner reflector.

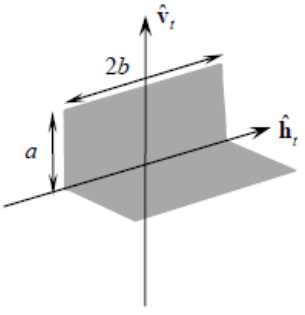


Fig. 2-9. A dihedral corner reflector.

is given by

$$[\mathbf{S}] = c_2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.4-5}$$

where $c_2 = \dfrac{k_0 ab}{\pi}$. From Eq. (2.4-5), we find the following characteristics for a dihedral corner reflector:

1. No cross polarization components are generated (HV = VH = 0) when it is illuminated by a purely horizontal or vertical radar signal.
2. Horizontal and vertical backscattering cross sections are identical (HH = VV).
3. Horizontal and vertical co-polarization components are out of phase.

From Eq. (2.4-5), we can calculate the covariance matrix and the Stokes scattering operator as

$$[\mathbf{C}] = |c_2|^2 \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \tag{2.4-6}$$

and

$$[\mathbf{M}] = \frac{1}{2}|c_2|^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.4-7}$$

The received power from a dihedral corner reflector for the co- and cross-polarized cases can be written as

$$\begin{aligned} P &= \frac{k_0^2 a^2 b^2}{\pi^2}\left\{1 \pm \cos^2(2\psi)\cos^2(2\chi) \mp \sin^2(2\psi)\cos^2(2\chi) \pm \sin^2(2\chi)\right\} \\ &= \frac{k_0^2 a^2 b^2}{\pi^2}\left\{1 \pm \cos^2(2\chi)\cos(4\psi) \pm \sin^2(2\chi)\right\} \end{aligned}. \tag{2.4-8}$$

Taking the derivatives of this expression, one finds that the co-polarized maxima occur at horizontal linear, vertical linear, and the circular polarizations. The co-polarized minima occur at 45 deg and 135 deg linear polarizations. The cross-polarized maxima occur at the same polarizations as the co-polarized minima. The cross-polarized minima occur at the same polarizations as the co-polarized maxima. This information allows us to conclude that the minima observed in these responses occur because the polarization of the received wave is orthogonal to that of the transmitting antenna.

Figure 2-10 shows the co- and cross-polarized responses of the dihedral corner reflector. Note that the minima in one figure correspond to the maxima in the other. While not obvious when comparing Fig. 2-7 and the top row in Fig. 2-10, a comparison of Fig. 2-8 and the bottom row of Fig. 2-10 clearly shows that the polarization responses of the dihedral and trihedral corner reflectors are rotated versions of each other in the Poincaré space. In fact, the responses of the dihedral corner reflector are exactly those of the trihedral rotated 90 deg about the S_1 axis. In that sense, we can consider these responses as “orthogonal” to each other.

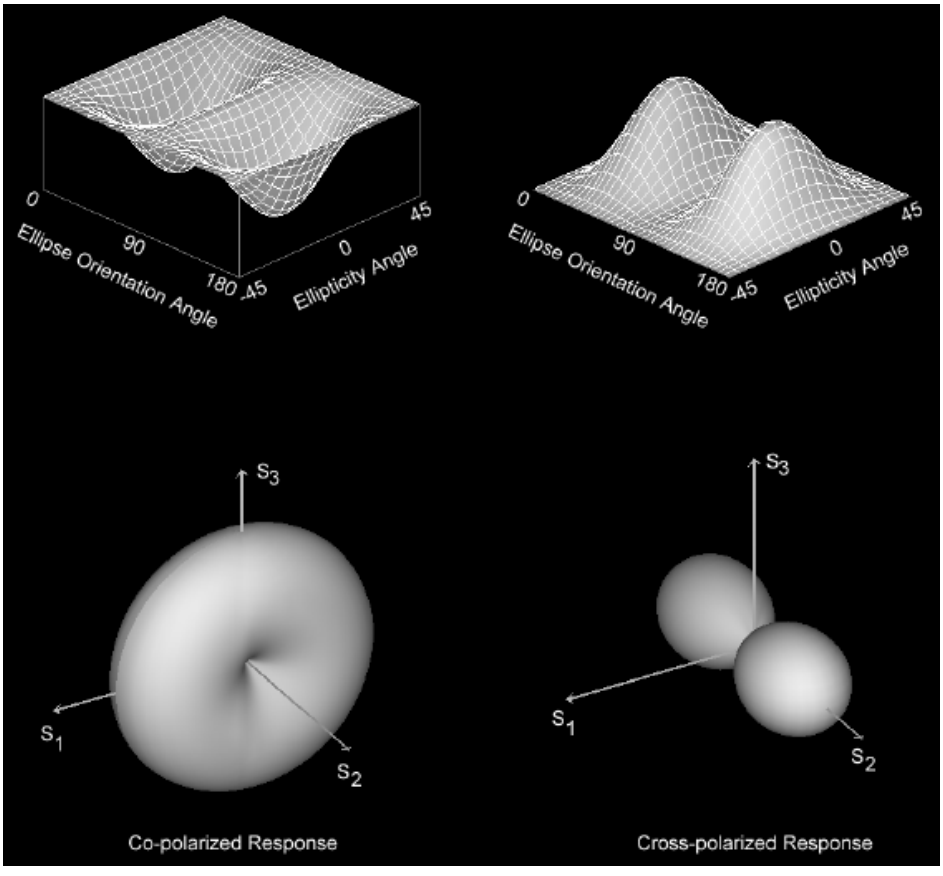


Fig. 2-10. Standard (top row) and three-dimensional (bottom row) polarization responses of a dihedral corner reflector.

Next, consider the case of a dihedral corner reflector, such as the one showed in Fig. 2-9; for this case, however, we rotate the base of the corner reflector 45 deg relative to the horizontal axis. Consequently, the scattering matrix becomes

$$[\mathbf{S}] = c_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{2.4-9}$$

where $c_2 = \dfrac{k_0 ab}{\pi}$, as before. The covariance matrix and the Stokes scattering operator are

$$[\mathbf{C}] = |c_2|^2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.4-10}$$

and

$$[\mathbf{M}] = \frac{1}{2}|c_3|^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.4-11}$$

The received power from a rotated dihedral corner reflector for the co- and cross-polarized cases can be written as

$$\begin{aligned} P &= \frac{k_0^2 a^2 b^2}{\pi^2}\left\{1 \mp \cos^2(2\psi)\cos^2(2\chi) \pm \sin^2(2\psi)\cos^2(2\chi) \pm \sin^2(2\chi)\right\} \\ &= \frac{k_0^2 a^2 b^2}{\pi^2}\left\{1 \mp \cos^2(2\chi)\cos(4\psi) \pm \sin^2(2\chi)\right\} \end{aligned}. \tag{2.4-12}$$

The polarization responses for this rotated dihedral corner reflector are shown in Fig. 2-11. Note that the cross-polarized maxima are now at HV and VH, with corresponding zeros at HH and VV. The three-dimensional signature in Fig. 2-11 shows that the responses of this rotated dihedral are the same shape as those of the trihedral, but rotated by 90 deg about the S_2 axis. They are in some sense "orthogonal" to the responses of both the trihedral and the non-rotated dihedral.

The scattering from vegetation is often modeled using a combination of randomly oriented cylinders. For the short, thin, vertical-conducting cylinder shown in Fig. 2-12, the scattering matrix is given by

$$[\mathbf{S}] = c_3 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.4-13}$$

where $c_3 = \dfrac{k_0^2 l^3}{6\left[\ln(4l/a) - 1\right]}$ and a is the radius of a cylinder.

From Eq. (2.4-13), we can calculate the covariance matrix and the Stokes scattering operator as

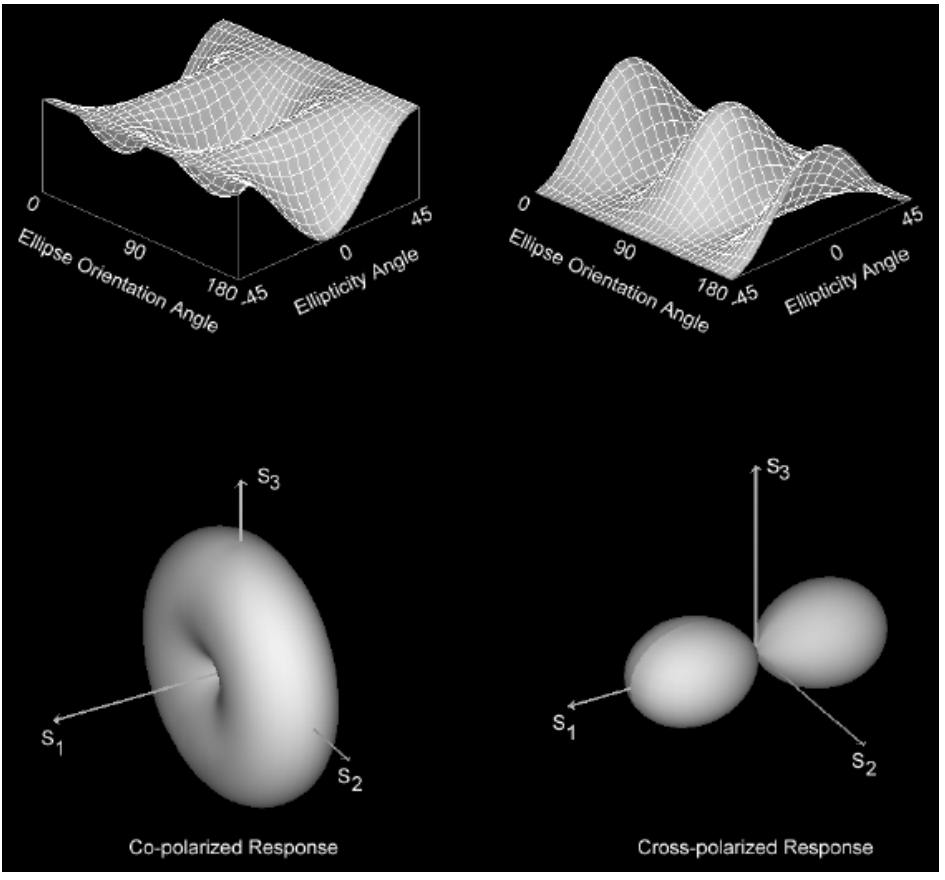


Fig. 2-11. Standard (top row) and three-dimensional (bottom row) polarization responses of a dihedral corner reflector rotated by an angle of 45 deg relative to the horizontal direction.

$$[\mathbf{C}] = |c_3|^2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.4-14}$$

and

$$[\mathbf{M}] = \frac{|c_3|^2}{4} \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.4-15}$$

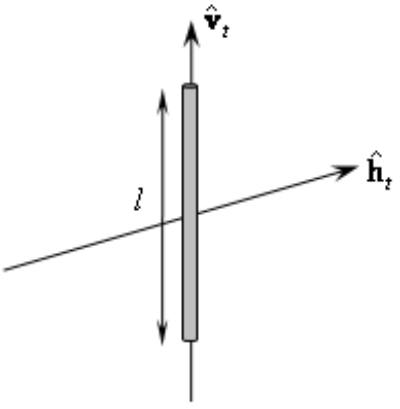


Fig. 2-12. A short, thin vertical conducting cylinder. The length of a cylinder is *l*.

The corresponding polarization responses are shown in Fig. 2-13. Note that both the co-polarized response is zero for HH polarization and the cross-polarized response is zero for VH polarization. This is the first example where both the co- and cross-polarized responses are zero for a given transmit polarization (horizontal polarization in this case). The reason for this is that when we transmit a horizontally polarized wave at this scatterer, there is no coupling to the cylinder and no scattered wave. Contrast this with the case of Fig. 2-11 for a horizontally polarized wave. The co-polarized response is zero, but the cross-polarized response is a not. This means that the scattered wave is non zero, but polarized orthogonally to the transmitted wave, and, hence, the maximum in the cross-polarized response.

Next, let us consider the case of a short, thin cylinder oriented at 45 deg from the vertical direction shown in Fig. 2-14. In this case, the scattering matrix is given by

$$[\mathbf{S}] = \frac{c_3}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{2.4-16}$$

where $c_3 = \dfrac{k_0^2 l^3}{6\left[\ln(4l/a) - 1\right]}$ and a is the radius of a cylinder.

From Eq. (2.4-16), we can calculate the covariance matrix and the Stokes scattering operator as

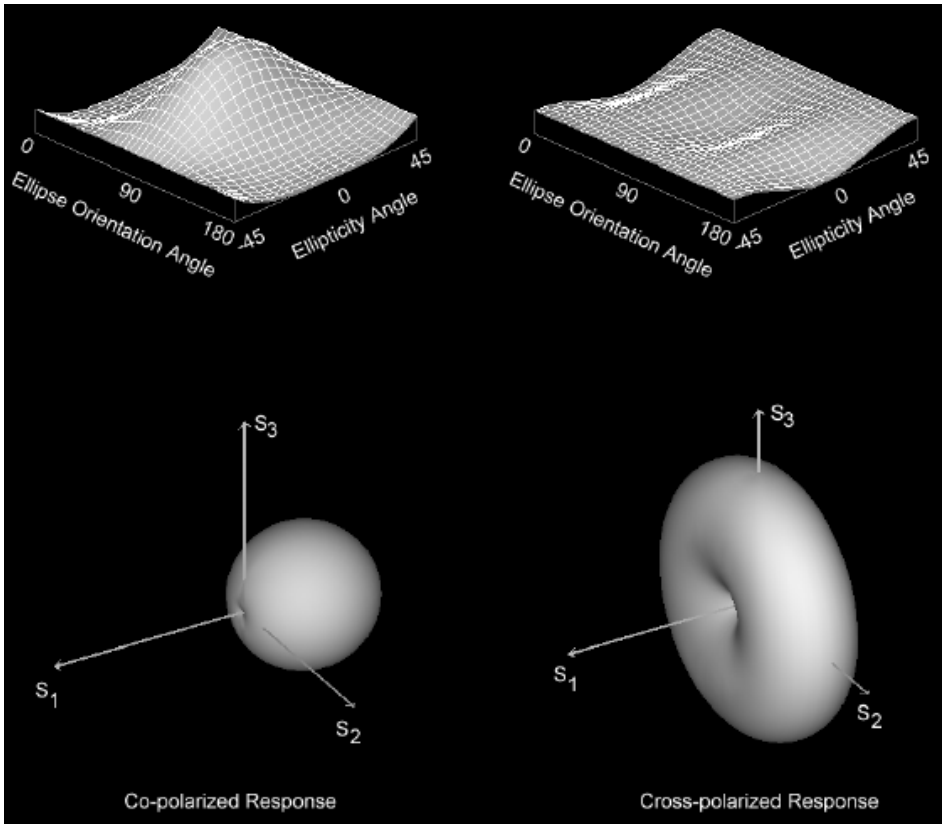


Fig. 2-13. The polarization responses of a short, thin, vertically oriented conducting cylinder.

$$[\mathbf{C}] = \frac{|c_3|^2}{4}\begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \tag{2.4-17}$$

and

$$[\mathbf{M}] = \frac{|c_3|^2}{4}\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.4-18}$$

The corresponding polarization response is shown in Fig. 2-15. Note that the position of the maximum in the co-polarized responses has shifted to be at 45 deg linear polarization. This indicates that the polarization response is sensitive

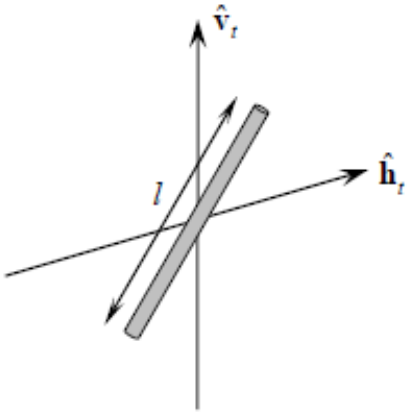


Fig. 2-14. A short, thin conducting cylinder oriented at 45 deg from the vertical direction. The length of a cylinder is *l*.

to the orientation of the short cylinder; we will exploit this fact when analyzing scattering from vegetation. Also, note that the cross-polarized response in Fig. 2-15 has the same shape as the co-polarized response in Fig. 2-10. This illustrates an important point: these polarization responses are not unique. In fact, these responses represent only a subset of the entire polarization space, and it is possible that more than one type of scatterer might have the same response. It would, therefore, be a mistake to consider the polarization responses as a unique way of identifying a specific type of scatterer.

Thus far, we have discussed only cases where the scatterer is characterized by a unique, single scattering matrix. In those cases, we have shown theoretically (we shall discuss this in more detail later) that the co-polarized and cross-polarized responses will each have at least one polarization for which the measured power goes to zero. When we analyze multi-looked data, however, we are working with the average power received from a collection of pixels. The composite polarization response can be thought of as the (properly normalized) sum of the individual polarization responses representing individual pixels. Unless all the individual responses have their null responses at exactly the same polarization, the composite signal will no longer have a null polarization. Instead, there might be some polarization for which the composite radar return is a minimum, but not zero.

To illustrate this further, let us consider the case of randomly oriented cylinders. This model has been used to describe scattering from randomly oriented tree branches that are thin compared to the radar wavelength. We shall start with a thin, conducting cylinder oriented at an angle α with respect to the

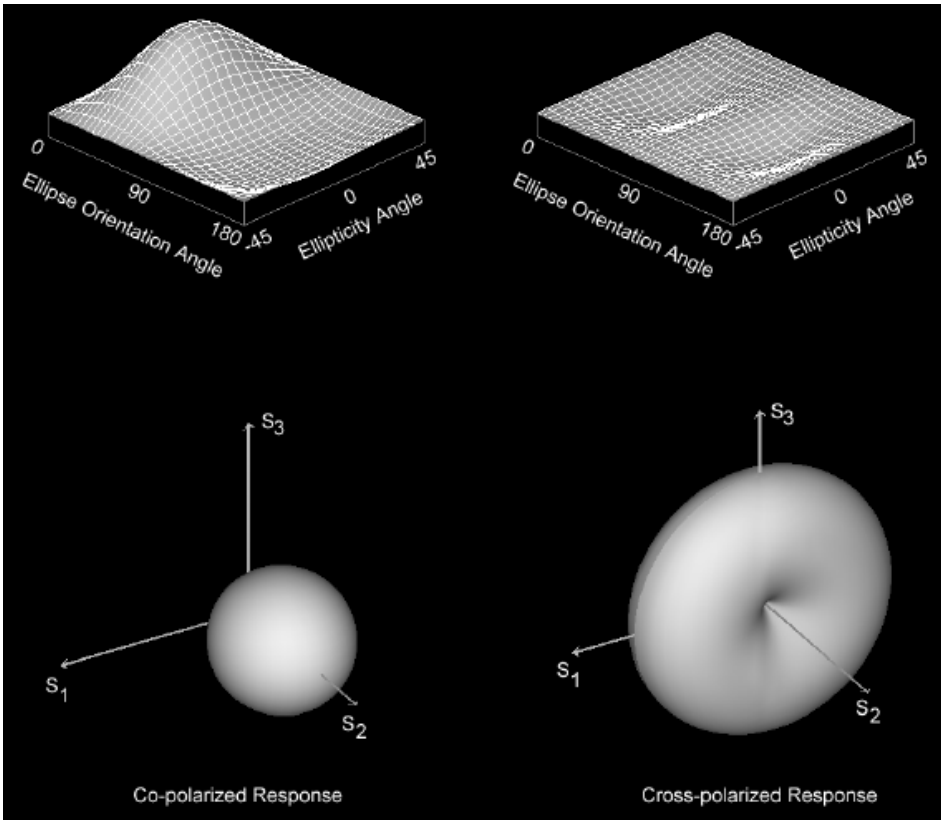


Fig. 2-15. The polarization response of a short, thin conducting cylinder oriented at 45 deg from the vertical direction. Note the difference between the responses in this figure and the responses shown in Fig. 2-13 for the vertically oriented cylinder.

vertical axis. The scattering matrix can be calculated by considering the rotation of the local coordinate axes defining the cylinder orientation, resulting in

$$\begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} = c_3 \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}. \qquad (2.4\text{-}19)$$

Here we assume that the cylinder lies in a plane that is orthogonal to the direction in which the radar waves travel. After straightforward matrix multiplications, the scattering matrix can be written as

$$\begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} = c_3 \begin{bmatrix} \sin^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \cos^2\alpha \end{bmatrix}. \qquad (2.4\text{-}20)$$

To illustrate the affect of randomness on the polarization response, we shall now analyze two different cases. In the first case, we assume that the cylinders are oriented uniformly randomly; i.e., we assume that α to be a random variable with a uniform distribution over [0,2π]. In this case, we find that

$$\begin{aligned}
&\left\langle S_{hh}S_{hh}^{*}\right\rangle = \frac{|c_3|^2}{2\pi}\int_0^{2\pi}\sin^4\alpha\, d\alpha = \frac{3}{8}|c_3|^2 \\
&\left\langle S_{vv}S_{vv}^{*}\right\rangle = \frac{|c_3|^2}{2\pi}\int_0^{2\pi}\cos^4\alpha\, d\alpha = \frac{3}{8}|c_3|^2 \\
&\left\langle S_{hv}S_{hv}^{*}\right\rangle = \left\langle S_{hh}S_{vv}^{*}\right\rangle = \left\langle S_{hv}S_{vh}^{*}\right\rangle = \left\langle S_{vh}S_{vh}^{*}\right\rangle = \frac{|c_3|^2}{2\pi}\int_0^{2\pi}\sin^2\alpha\cos^2\alpha\, d\alpha = \frac{1}{8}|c_3|^2 \\
&\left\langle S_{hh}S_{hv}^{*}\right\rangle = \left\langle S_{hv}S_{vv}^{*}\right\rangle = 0
\end{aligned} \quad . \qquad (2.4\text{-}21)$$

The covariance matrix for this case is

$$\left\langle [\mathbf{C}]\right\rangle = |c_3|^2\begin{pmatrix} 3/8 & 0 & 0 & 1/8 \\ 0 & 1/8 & 1/8 & 0 \\ 0 & 1/8 & 1/8 & 0 \\ 1/8 & 0 & 0 & 3/8 \end{pmatrix}. \qquad (2.4\text{-}22)$$

The Stokes scattering operator is

$$\left\langle [\mathbf{M}]\right\rangle = |c_3|^2\begin{pmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/8 & 0 & 0 \\ 0 & 0 & 1/8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \qquad (2.4\text{-}23)$$

The polarization responses for this cloud of randomly oriented cylinders are shown in Fig. 2-16. Comparing these responses to those of the trihedral corner reflector, it is clear that they are similar except for the addition of a “pedestal” on which the responses sit. This “pedestal height” can be defined as the ratio of the minimum power in the response to the maximum power in the response. Using this definition, the pedestal height of a single scatterer will be zero, while the pedestal height for this example of the randomly oriented cylinders will be a non-zero value between 0 and 1. Therefore, the pedestal height can be used to infer the amount of randomness exhibited by multi-looked pixels.

To illustrate this point further, let us look at a second example involving the thin cylinders. Instead of assuming a uniform distribution in the angle α, we shall assume that the cylinders are mostly vertically oriented in a plane

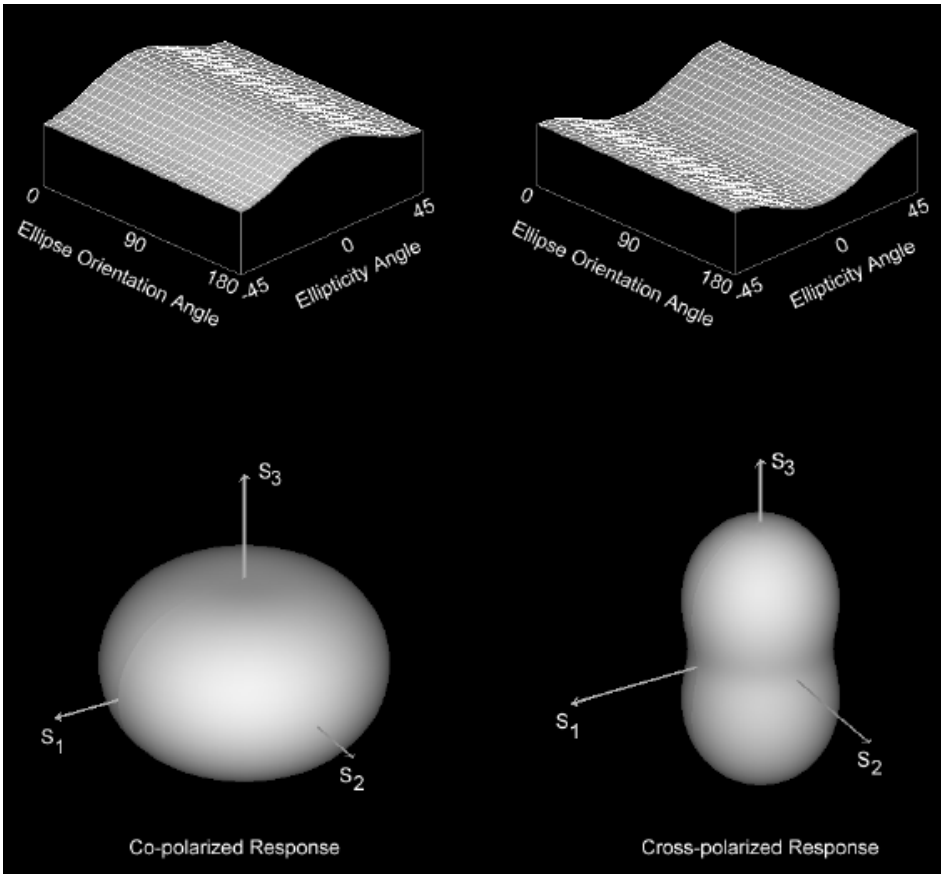


Fig. 2-16. Polarization responses of a cloud of thin conducting cylinders that are oriented uniformly randomly in a plane orthogonal to the radar look direction. Note that these shapes are similar to those shown for the trihedral corner reflector, except that the nulls in the responses have been replaced by non-zero minima.

orthogonal to the direction in which the radar signal is propagating. We shall assume the following probability density function for the cylinders:

$$p(\alpha) = \frac{1}{\pi}\cos^2\alpha; \quad 0 \le \alpha \le 2\pi \,. \tag{2.4-24}$$

This distribution has been used to describe scattering from vegetation, such as grass or wheat, which grows mostly vertically. Performing the averaging, we find

$$\left\langle S_{hh}S_{hh}^{*}\right\rangle = \frac{|c_3|^2}{2\pi}\int_0^{2\pi}\cos^2\alpha\sin^4\alpha d\alpha = \frac{1}{8}|c_3|^2$$

$$\left\langle S_{vv}S_{vv}^{*}\right\rangle = \frac{|c_3|^2}{2\pi}\int_0^{2\pi}\cos^6\alpha d\alpha = \frac{5}{8}|c_3|^2 \quad . \quad (2.4\text{-}25)$$

$$\left\langle S_{hv}S_{hv}^{*}\right\rangle = \left\langle S_{hh}S_{vv}^{*}\right\rangle = \left\langle S_{hv}S_{vh}^{*}\right\rangle = \left\langle S_{vh}S_{vh}^{*}\right\rangle = \frac{|c_3|^2}{2\pi}\int_0^{2\pi}\sin^2\alpha\cos^4\alpha d\alpha = \frac{1}{8}|c_3|^2$$

$$\left\langle S_{hh}S_{hv}^{*}\right\rangle = \left\langle S_{hv}S_{vv}^{*}\right\rangle = 0$$

The covariance matrix for this case is

$$\left\langle [\mathbf{C}]\right\rangle = |c_3|^2\begin{pmatrix} 1/8 & 0 & 0 & 1/8 \\ 0 & 1/8 & 1/8 & 0 \\ 0 & 1/8 & 1/8 & 0 \\ 1/8 & 0 & 0 & 5/8 \end{pmatrix}. \quad (2.4\text{-}26)$$

The Stokes scattering operator is

$$\left\langle [\mathbf{M}]\right\rangle = |c_3|^2\begin{pmatrix} 1/4 & -1/8 & 0 & 0 \\ -1/8 & 1/8 & 0 & 0 \\ 0 & 0 & 1/8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (2.4\text{-}27)$$

The polarization responses for this cloud of randomly oriented cylinders are shown in Fig. 2-17. Compare the co-polarized response in the top left to that of a single, vertically oriented cylinder as shown in Fig. 2-13. These two responses are quite similar, with the exception that the response in Fig. 2-17 has a pedestal added to it. The relative height of this pedestal in the co-polarized response in Fig. 2-17 is 20 percent; however, the pedestal in Fig. 2-16 is 66 percent, clearly indicating the relationship between pedestal height and randomness.

So far, we have assumed that all the matrices are measured without any noise added to the measurements. In reality, of course, this is not the case. Before looking at the effect of system thermal noise on these responses, however, let us first examine the case of noise only. As mentioned in the previous section, scattering matrices are measured using two transmit pulses and two independent receiver channels.

Because measurements in a particular receiver are made at different times, and because the pulse repetition frequency is low enough so that the time between pulses (i.e., measurements) is much longer than the noise coherence time, the noise measurements in the four images corresponding to the scattering matrix

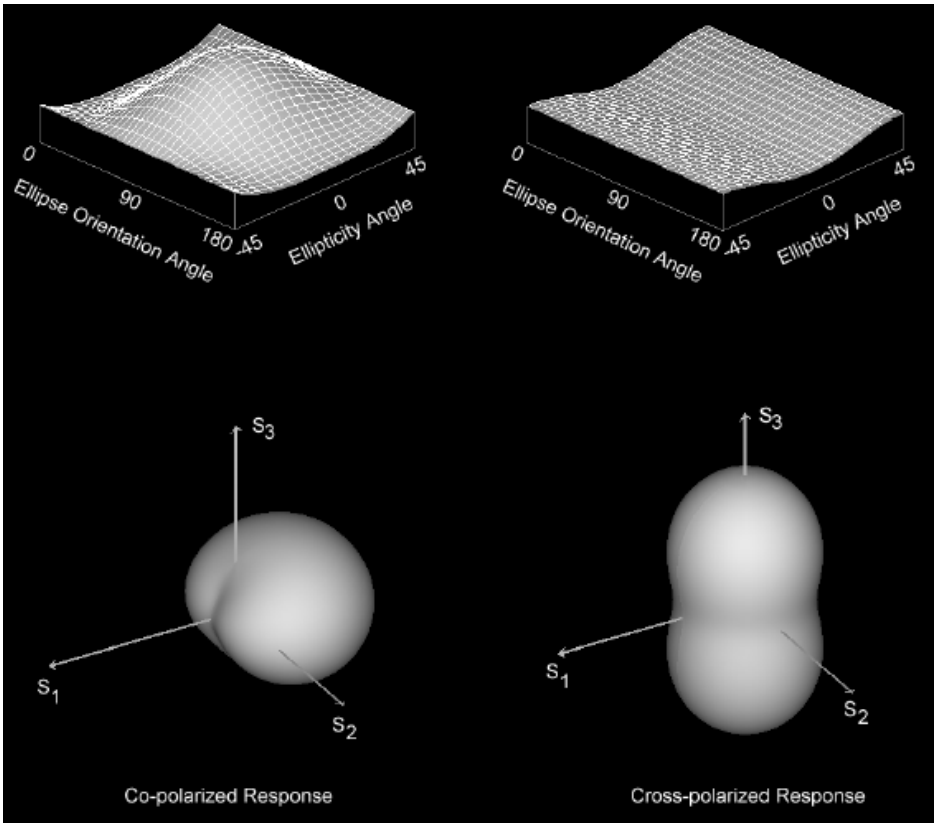


Fig. 2-17. Polarization responses for a cloud of dipoles with orientations according to the statistical distribution shown in Eq. (2.4-24). Note the similarity with the responses of a single, vertically oriented cylinder shown in Fig. 2-13, except for the addition of a pedestal. Also note the relative size of the pedestal in this figure, as compared to that of a uniform random orientation distribution shown in Fig. 2-16.

elements can be considered statistically independent. We shall further assume that the data, including the noise-only data, are well calibrated such that the channels are properly balanced in amplitude. In the absence of a return signal, after averaging, the covariance matrix will be an identity matrix, and the corresponding Stokes scattering operator will be

$$\langle [\mathbf{M}] \rangle = P_n \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.4-28}$$

This means, as expected, that the polarization response of a noise-only measurement is a constant value, independent of the transmit or receive

polarization. These polarization responses are shown in Fig. 2-18. Note that, in this case, the pedestal height is 100 percent.

In the backscatter case, reciprocity dictates that the two cross-polarized measurements should be identical. In practice, they differ because of the additive thermal noise in the receiver channels. As mentioned in the previous section, we can exploit the fact that the two cross-polarized measurements are made through different receivers and at different times, resulting in uncorrelated noise. If we average these two measurements before multi-looking, the noise power is effectively halved. Assuming that the noise power in each channel is P_n, the resulting noise power in the averaged cross-polarized channel will be

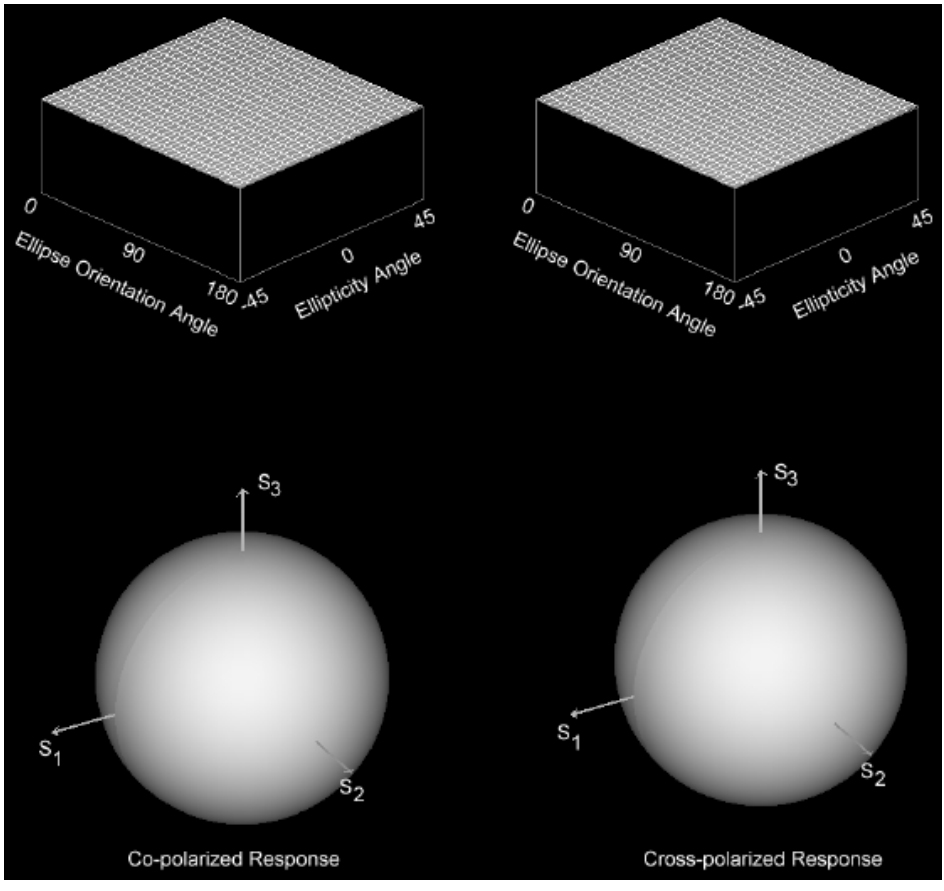


Fig. 2-18. The polarization responses of noise-only measurements when the noise power is identical in the two receiver channels. It is assumed that the cross-polarized channels are not averaged to increase the signal-to-noise ratio.

$$\begin{aligned}\left\langle S_{hv}S_{hv}^{*}\right\rangle &= \left\langle \frac{1}{4}\left(S_{hv}^{m}+S_{vh}^{m}\right)\left(S_{hv}^{m}+S_{vh}^{m}\right)^{*}\right\rangle \\ &= \left\langle \frac{1}{4}\left(S_{hv}^{m}\right)\left(S_{hv}^{m}\right)^{*}\right\rangle + \left\langle \frac{1}{4}\left(S_{hv}^{m}\right)\left(S_{vh}^{m}\right)^{*}\right\rangle + \left\langle \frac{1}{4}\left(S_{vh}^{m}\right)\left(S_{hv}^{m}\right)^{*}\right\rangle + \left\langle \frac{1}{4}\left(S_{vh}^{m}\right)\left(S_{vh}^{m}\right)^{*}\right\rangle \\ &= \frac{1}{4}P_n + 0 + 0 + \frac{1}{4}P_n \\ &= \frac{1}{2}P_n\end{aligned} \quad (2.4\text{-}29)$$

The superscript *m* is added to indicate individual noise measurements. The second and third terms are zero because the noise in the two channels is uncorrelated. In this case, the covariance matrix is

$$\left\langle [\mathbf{C}]\right\rangle_N = P_n \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.4\text{-}30)$$

and the Stokes scattering operator is

$$\left\langle [\mathbf{M}]\right\rangle_N = P_n \begin{pmatrix} 3/4 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}. \quad (2.4\text{-}31)$$

The polarization responses of this case are shown in Fig. 2-19. Note that the pedestal is still 100 percent, but the cross-polarized responses are now less than the co-polarized ones.

To illustrate the effect of system noise on the polarization signatures, we note that the thermal noise and the radar signal are uncorrelated. The measured scattering matrix for each pixel can be written as

$$[\mathbf{S}]_m = [\mathbf{S}] + [\mathbf{N}]. \quad (2.4\text{-}32)$$

The first term on the right is the actual scattering matrix; the second term represents thermal noise in the system. Since the signal and noise are uncorrelated, the multi-looked covariance matrix is

$$\left\langle [\mathbf{C}]\right\rangle_m = \left\langle [\mathbf{C}]\right\rangle + \left\langle [\mathbf{C}]\right\rangle_N \quad (2.4\text{-}33)$$

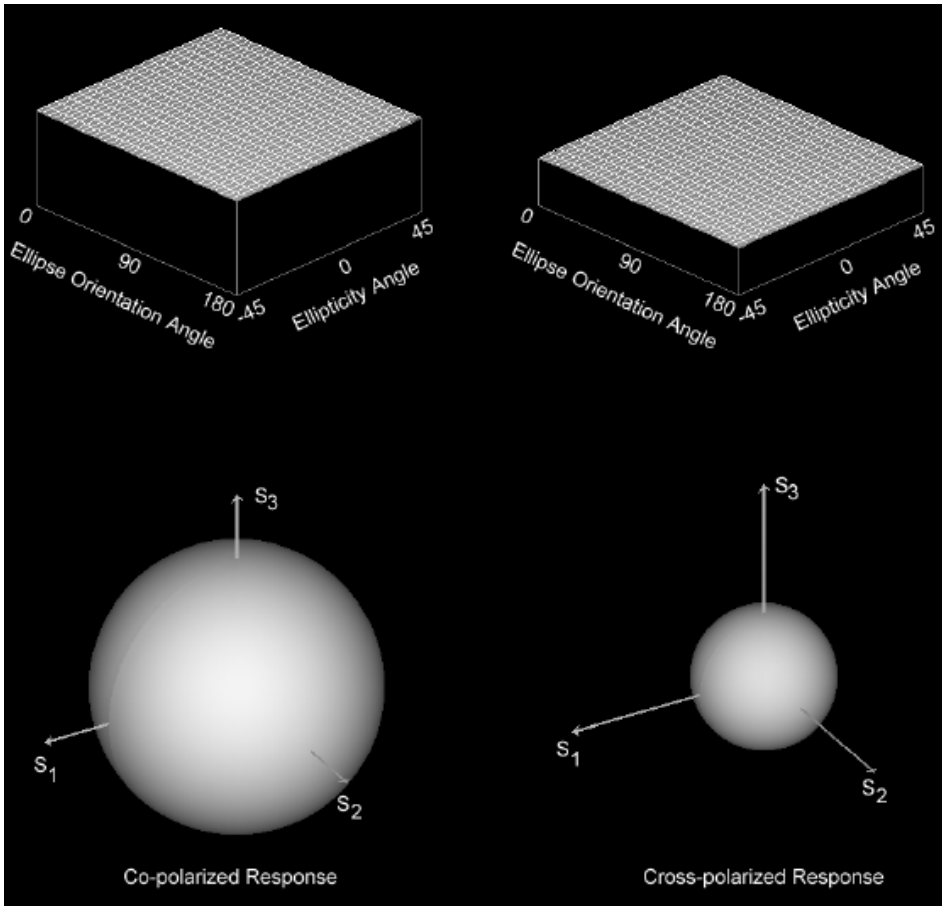


Fig. 2-19. Polarization responses of noise-only measurements when the cross-polarized channels are averaged before multi-looking to increase the signal-to-noise ratio in those channels.

and the multi-looked Stokes scattering operator is

$$\langle [\mathbf{M}] \rangle_m = \langle [\mathbf{M}] \rangle + \langle [\mathbf{M}] \rangle_N . \tag{2.4-34}$$

The noise matrices are the same as those in Eq. (2.4-30) and Eq. (2.4-31). The composite polarization response in the presence of noise will, therefore, be the sum of the actual scatterer response and that of the system noise. Since the system noise response is just a pedestal, the presence of the system noise will simply add a pedestal to the scatterer response. This is illustrated for the case of a vertically oriented thin cylinder with different amounts of thermal noise present in Fig. 2-20.

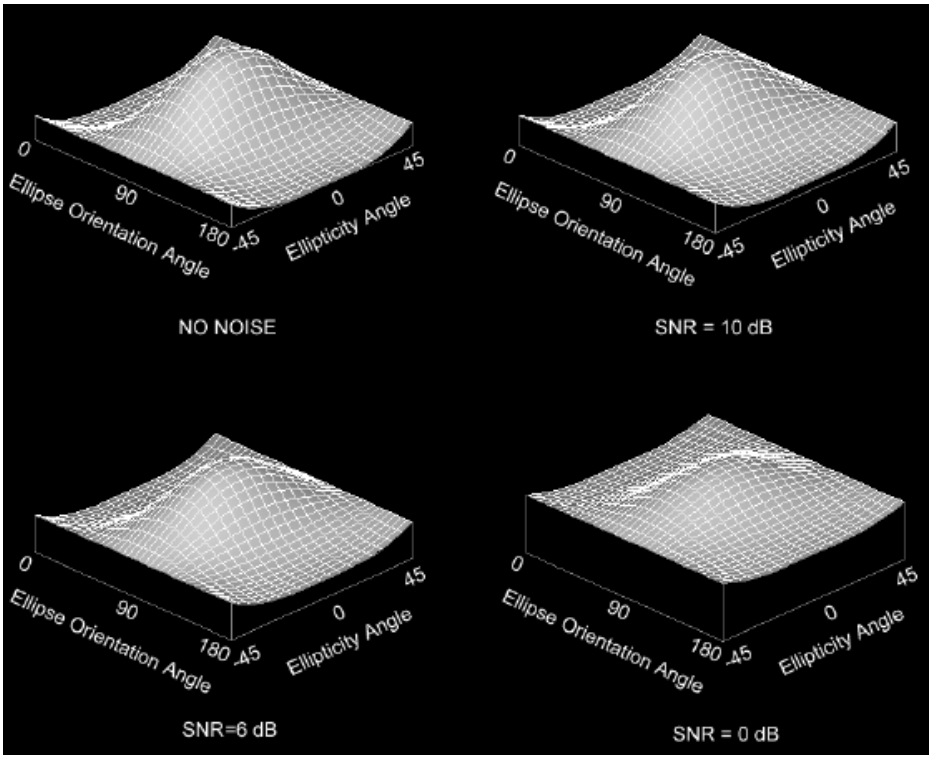


Fig. 2-20. Co-polarized responses of a thin conducting cylinder with various signal-to-noise ratios. Note the similarity of the SNR = 6dB response to that of the cosine squared random orientation shown in Fig. 2-17.

Note the similarity of the polarization response in the bottom left of Fig. 2-20 to that of the cloud of thin cylinders with a cosine squared distribution around the vertical direction. These figures illustrate an important point: the presence of a pedestal implies randomness, but the randomness can be due to many different causes. Care should, therefore, be exercised not to assume that the presence of a pedestal implies random orientation of scatterers without first verifying that there is an adequate signal-to-noise ratio (SNR) for this interpretation.

2.5 Optimum Polarizations

In the previous section, we introduced the concept of polarization responses as a way of visualizing the response of a scatterer, or a collection of scatterers after multi-looking. We shall now look at how to determine which polarization combination would optimize (i.e., either maximize or minimize) the received power from such a scatterer or collection of scatterers. We shall use the Stokes scattering operator representation of the scatterer for two reasons. First, it can be used to represent both a single scatterer or the ensemble average, or multi-

looked data. Second, it provides a simple, intuitive way of understanding how to optimize the received power by changing either the transmit or receive polarization. In the rest of this section, the matrix $[\mathbf{M}]$ will be used to mean either a single scatterer or a multi-looked average.

2.5.1 General (Bistatic) Case

This problem was first analyzed by Ioannidids and Hammers [11] in the context of maximizing radar signal strength in the presence of unwanted clutter. From Eq. (2.2-10) and Eq. (2.3-2), the received power can be written as

$$P = \mathbf{S}^{rec} \cdot [\mathbf{M}] \mathbf{S}^{tr}, \tag{2.5-1}$$

where $\mathbf{S}^{rec}$ and $\mathbf{S}^{tr}$ are the Stokes vectors representing the fully polarized waves of the receiving and transmitting antennas, respectively. Because these vectors represent fully polarized waves, each of these vectors has to satisfy Eq. (2.1-5). Therefore,

$$\mathbf{S} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \mathbf{S} = 0 . \tag{2.5-2}$$

We can also write the Stokes scattering operator in the following form:

$$[\mathbf{M}] = \begin{bmatrix} m & \tilde{\mathbf{u}} \\ \mathbf{v} & \mathbf{Q} \end{bmatrix}, \tag{2.5-3}$$

with $\mathbf{u}$ and $\mathbf{v}$ three-dimensional real vectors and $\mathbf{Q}$ a 3 × 3 real matrix. The ~ sign indicates the transpose of the vector or matrix. Note that in the backscatter case, $\mathbf{u} = \mathbf{v}$ and $\mathbf{Q}$ is a symmetrical matrix. Now let us write the Stokes vectors of the antennas as

$$\mathbf{S}^{rec} = \begin{pmatrix} 1 \\ \mathbf{y} \end{pmatrix}; \quad \mathbf{S}^{tr} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix}; \quad \mathbf{x} \cdot \mathbf{x} = 1; \quad \mathbf{y} \cdot \mathbf{y} = 1 . \tag{2.5-4}$$

Note that these vectors satisfy Eq. (2.5-2). Using Eq. (2.5-4) and Eq. (2.5-3) in Eq. (2.5-1), we find that

$$P = m + \mathbf{u} \cdot \mathbf{x} + \mathbf{v} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{Q}\mathbf{x} . \tag{2.5-5}$$

This is the quantity we need to optimize, but subject to the constraints in Eq. (2.5-4). In the special case of co- and cross-polarized responses, either

$\mathbf{y} = \mathbf{x}$ or $\mathbf{y} = -\mathbf{x}$. To find the optima of *P* subject to the constraints in Eq. (2.5-4), we use the Lagrange multiplier method and construct the function

$$F = m + \mathbf{u}\cdot\mathbf{x} + \mathbf{v}\cdot\mathbf{y} + \mathbf{y}\cdot\mathbf{Q}\mathbf{x} + \frac{\lambda_x}{2}\left(1 - \mathbf{x}\cdot\mathbf{x}\right) + \frac{\lambda_y}{2}\left(1 - \mathbf{y}\cdot\mathbf{y}\right). \tag{2.5-6}$$

Let us first look at the receiving antenna polarization. The optimum polarizations are found by taking the derivatives of *F:*

$$\begin{aligned} \nabla_x F = 0 \;&\Rightarrow\; \lambda_x \mathbf{x} = \mathbf{u} + \tilde{\mathbf{Q}}\mathbf{y} \\ \nabla_y F = 0 \;&\Rightarrow\; \lambda_y \mathbf{y} = \mathbf{v} + \mathbf{Q}\mathbf{x} \end{aligned}. \tag{2.5-7}$$

The value of the Lagrange multiplier λ_y is found from the constraint in Eq. (2.5-4) that the magnitude of the vectors $\mathbf{y}$ and $\mathbf{x}$ must be 1. Using this constraint for the receiving antenna and the first equation in Eq. (2.5-7), we find that the optimum receive antenna polarization is given by

$$\mathbf{y} = \pm\frac{\mathbf{v} + \mathbf{Q}\mathbf{x}}{\sqrt{\left(\mathbf{v} + \mathbf{Q}\mathbf{x}\right)\cdot\left(\mathbf{v} + \mathbf{Q}\mathbf{x}\right)}}. \tag{2.5-8}$$

Let us examine this equation in more detail. The vector on the right is simply the polarized part of the vector that is the product of the Stokes scattering operator and the polarized part of the Stokes vector representing the transmitting antenna. Therefore, Eq. (2.5-8) simply states that the optimum receiving polarization is that Stokes vector that is either aligned with this vector or anti-podal to this vector on the Poincaré sphere. Further recall from our earlier discussion that orthogonal polarizations are anti-podal on the Poincaré sphere. Therefore, Eq. (2.5-8) simply states that the optimum receive polarizations for a given transmit polarization are those that either match the polarization of the incoming wave or are orthogonal to the polarization of the incoming wave.

From Eq. (2.5-7), we can now derive the following two expressions for the optimum transmit and receive antenna polarizations:

$$\begin{aligned} \left(\tilde{\mathbf{Q}}\mathbf{Q} - \lambda_x\lambda_y\mathbf{I}\right)\mathbf{x} &= -\left(\lambda_y\mathbf{u} + \tilde{\mathbf{Q}}\mathbf{v}\right) \\ \left(\mathbf{Q}\tilde{\mathbf{Q}} - \lambda_x\lambda_y\mathbf{I}\right)\mathbf{y} &= -\left(\lambda_y\mathbf{v} + \mathbf{Q}\mathbf{u}\right) \end{aligned}. \tag{2.5-9}$$

We can solve for the values of λ_x and λ_y using the constraint in Eq. (2.5-4) that the magnitude of the vectors $\mathbf{y}$ and $\mathbf{x}$ must be 1. This gives the following two expressions

$$\begin{aligned}\left|\left(\tilde{\mathbf{Q}}\mathbf{Q}-\lambda_x\lambda_y\mathbf{I}\right)^{-1}\left(\lambda_y\mathbf{u}+\tilde{\mathbf{Q}}\mathbf{v}\right)\right|&=1\\ \left|\left(\mathbf{Q}\tilde{\mathbf{Q}}-\lambda_x\lambda_y\mathbf{I}\right)^{-1}\left(\lambda_y\mathbf{v}+\mathbf{Q}\mathbf{u}\right)\right|&=1\end{aligned}. \tag{2.5-10}$$

Once the values for λ_x and λ_y are known, we can use Eq. (2.5-9) to find the optimum polarization vectors.

These non-linear equations are not easy to solve explicitly. In practice, it is easier (and faster computationally) to use Eq. (2.5-8) in Eq. (2.5-5) to write the received power in the following form

$$P=m+\mathbf{u}\cdot\mathbf{x}\pm\sqrt{\left(\mathbf{v}+\mathbf{Q}\mathbf{x}\right)\cdot\left(\mathbf{v}+\mathbf{Q}\mathbf{x}\right)} \tag{2.5-11}$$

and to solve for the maxima or minima of Eq. (2.5-11) numerically, because the polarization response is a slowly varying function, as can be seen from the figures in the previous section. Once the optimum transmit polarizations have been identified, the optimum receive polarizations can be calculated using Eq. (2.5-8).

2.5.2 Backscatter (Monostatic) Case

So far, we have made no assumptions on the form of the Stokes scattering operator. In the special case of a monostatic radar, reciprocity dictates that the scattering matrix be symmetrical. In that case, the Stokes scattering operator is also symmetrical; i.e.,

$$\mathbf{u}=\mathbf{v};\quad \mathbf{Q}=\tilde{\mathbf{Q}}\,. \tag{2.5-12}$$

For this special case, the two expressions in Eq. (2.5-9) are identical, implying that the optimum receive polarization is related to the optimum transmit polarization by

$$\mathbf{y}=\pm\mathbf{x}\,. \tag{2.5-13}$$

In other words, the optimum values of the received power are found using co-polarized or cross-polarized antennas. We can then rewrite the power expression in Eq. (2.5-5) as follows

$$\begin{aligned}P_c&=m+2\mathbf{u}\cdot\mathbf{x}+\mathbf{x}\cdot\mathbf{Q}\mathbf{x}\\ P_x&=m-\mathbf{x}\cdot\mathbf{Q}\mathbf{x}\end{aligned}, \tag{2.5-14}$$

where the subscripts *c* and *x* refer to the co-polarized power and the cross-polarized power, respectively. We shall analyze the two cases separately. First,

consider the cross-polarized case. Solving for the optimum polarization vector subject to the constraint that the magnitude of this vector should be 1, gives the following expression

$$\mathbf{Q}\mathbf{x} = \lambda\mathbf{x}\,. \tag{2.5-15}$$

In other words, the optimum polarizations for the cross-polarized case are the normalized eigenvectors of the matrix $\mathbf{Q}$ and the optimum cross-polarized power values are

$$P_X = m - \lambda\,, \tag{2.5-16}$$

where λ represents the eigenvalues of the matrix $\mathbf{Q}$. Since $\mathbf{Q}$ is a symmetrical real matrix, all the eigenvectors are orthogonal to each other. This means that the optimum cross-polarized vectors are orthogonal when plotted on the Poincaré sphere. Note that the Stokes vectors, not the electric field vectors, are orthogonal.

The case of the co-polarized power is a bit more complicated. Taking the derivatives of the first equation in Eq. (2.5-14) subject to the constraint that the magnitude of the polarization vector should be 1 and setting the result equal to zero gives the following result

$$\left(\mathbf{Q} - \upsilon\mathbf{I}\right)\mathbf{x} = -\mathbf{u}\,. \tag{2.5-17}$$

As long as υ is not an eigenvalue of $\mathbf{Q}$, this equation has a unique solution:

$$\mathbf{x} = -\left(\mathbf{Q} - \upsilon\mathbf{I}\right)^{-1}\mathbf{u}\,. \tag{2.5-18}$$

In the special case where $\mathbf{u}$ is the null vector, the solutions to Eq. (2.5-17) are the normalized eigenvectors of $\mathbf{Q}$. In general, however, the optimum vectors $\mathbf{x}$ will not be eigenvectors of $\mathbf{Q}$. The values of υ are found from the constraint that the magnitude of $\mathbf{x}$ must be 1. The most intuitive solution is to write both $\mathbf{x}$ and $\mathbf{u}$ in the basis formed by the normalized eigenvectors of the matrix $\mathbf{Q}$. Since $\mathbf{Q}$ is a symmetrical matrix, these normalized eigenvectors are orthogonal and can be used as a basis to express these vectors. If we write $\mathbf{u}$ as

$$\mathbf{u} = b_1\hat{\mathbf{e}}_1 + b_2\hat{\mathbf{e}}_2 + b_3\hat{\mathbf{e}}_3\,, \tag{2.5-19}$$

with $\hat{\mathbf{e}}_i\,; i = 1,2,3$ representing the normalized eigenvectors, we can solve for $\mathbf{x}$ from Eq. (2.5-17). The result is

$$\mathbf{x} = \frac{b_1}{(\upsilon - \lambda_1)}\hat{\mathbf{e}}_1 + \frac{b_2}{(\upsilon - \lambda_2)}\hat{\mathbf{e}}_2 + \frac{b_3}{(\upsilon - \lambda_3)}\hat{\mathbf{e}}_3, \tag{2.5-20}$$

where $\lambda_i; i = 1,2,3$ are the eigenvalues of $\mathbf{Q}$. The magnitude of this vector is

$$|\mathbf{x}| = \sqrt{\frac{b_1^2}{(\upsilon - \lambda_1)^2} + \frac{b_2^2}{(\upsilon - \lambda_2)^2} + \frac{b_3^2}{(\upsilon - \lambda_3)^2}}\,. \tag{2.5-21}$$

The values of υ are found by requiring that this magnitude be equal to 1. Solving for υ from this equation results in a sixth order polynomial in υ (van Zyl et al. [4]). The real valued roots of this polynomial are the optimum values of υ. The polynomial is

$$\upsilon^6 + d_1\upsilon^5 + d_2\upsilon^4 + d_3\upsilon^3 + d_4\upsilon^2 + d_5\upsilon + d_6 = 0, \tag{2.5-22}$$

with

$$\begin{aligned}
d_1 &= -2(\lambda_1 + \lambda_2 + \lambda_3) = -2m \\
d_2 &= (\lambda_1 + \lambda_2 + \lambda_3)^2 + 2(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) - b_1 - b_2 - b_3 \\
d_3 &= -2\left[(\lambda_1 + \lambda_2 + \lambda_3)(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) + \lambda_1\lambda_2\lambda_3\right] \\
&\quad + 2b_1(\lambda_2 + \lambda_3) + 2b_2(\lambda_1 + \lambda_3) + 2b_3(\lambda_1 + \lambda_2) \\
d_4 &= (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)^2 + 2\lambda_1\lambda_2\lambda_3(\lambda_1 + \lambda_2 + \lambda_3) \\
&\quad - b_1(\lambda_2^2 + \lambda_3^2 + 4\lambda_2\lambda_3) - b_2(\lambda_1^2 + \lambda_3^2 + 4\lambda_1\lambda_3) - b_3(\lambda_1^2 + \lambda_2^2 + 4\lambda_1\lambda_2) \\
d_5 &= -2\lambda_1\lambda_2\lambda_3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3) + 2b_1\lambda_2\lambda_3(\lambda_2 + \lambda_3) \\
&\quad + 2b_2\lambda_1\lambda_3(\lambda_1 + \lambda_3) + 2b_3\lambda_1\lambda_2(\lambda_1 + \lambda_2) \\
d_6 &= \lambda_1^2\lambda_2^2\lambda_3^2 - b_1\lambda_2^2\lambda_3^2 - b_2\lambda_1^2\lambda_3^2 - b_3\lambda_1^2\lambda_2^2
\end{aligned} \tag{2.5-23}$$

So how do we know that the polynomial in Eq. (2.5-22) has any real-valued roots? We can show that, in general, it will have at least two real-valued roots. This follows from the behavior of $|\mathbf{x}|$ as a function of υ, which is easiest seen from Eq. (2.5-21). As the value of υ approaches both $\pm\infty$, it is clear that $|\mathbf{x}| \to 0$. On the other hand, as the value of υ approaches any of the real-valued eigenvalues of the symmetrical matrix $\mathbf{Q}$, $|\mathbf{x}| \to \infty$. Therefore, there must be at least one value of υ between $-\infty$ and the smallest eigenvalue of $\mathbf{Q}$ for which $|\mathbf{x}| = 1$. Likewise, there must be at least one value of υ between the largest eigenvalue of $\mathbf{Q}$ and $\pm\infty$ and for which $|\mathbf{x}| = 1$. This proves that there will always be at least two real-valued roots to the polynomial in Eq. (2.5-22).

The one root will be less than the smallest eigenvalue of $\mathbf{Q}$, and the other will be larger than the largest eigenvalue of $\mathbf{Q}$.

2.5.3 Special Case: Single Scatterer in Backscatter (Monostatic) Case

Now let us consider the special case of the return from a single scatterer measured in the backscatter configuration. In this case, the scattering matrix $\mathbf{S}$ is symmetrical and there is a definite relationship between the elements of the matrix $\mathbf{Q}$. In fact, straightforward but tedious algebra shows that in this case, $\mathbf{Q}$ has three eigenvalues

$$\begin{aligned}\lambda_1 &= m \\ \lambda_2 &= \sqrt{\left(m-|\mathbf{u}|\right)\left(m+|\mathbf{u}|\right)} \geq 0 \\ \lambda_3 &= -\sqrt{\left(m-|\mathbf{u}|\right)\left(m+|\mathbf{u}|\right)} \leq 0\end{aligned} \quad . \tag{2.5-24}$$

The normalized eigenvector corresponding to the first eigenvalue is $\pm\mathbf{u}/|\mathbf{u}|$. Using this eigenvalue in Eq. (2.5-14) shows that the vector $\pm\mathbf{u}/|\mathbf{u}|$ corresponds to a cross-polarized null. It also is a solution to Eq. (2.5-17) with

$$\upsilon = m \pm |\mathbf{u}| . \tag{2.5-25}$$

Using these in the top equation in Eq. (2.5-14) gives the two co-polarized powers of

$$P_c = 2\left(m \pm |\mathbf{u}|\right). \tag{2.5-26}$$

The plus sign corresponds to the maximum co-polarized power. Therefore, the polarization vector that maximizes the co-polarized power also results in a cross-polarized null. Also note from Eq. (2.5-26) that

$$P_{cs} = 2\left(m - |\mathbf{u}|\right) \geq 0 \Rightarrow m \geq |\mathbf{u}| . \tag{2.5-27}$$

As we shall see shortly, this co-polarized power represents the saddle point in the co-polarized response.

To find the vectors corresponding to the co-polarized minima, we define a normalized vector

$$\mathbf{x} = a\hat{\mathbf{e}}_1 + b\hat{\mathbf{e}}_3 , \tag{2.5-28}$$

where $\hat{\mathbf{e}}_1 = -\mathbf{u}/|\mathbf{u}|$ and $\hat{\mathbf{e}}_3$ are normalized eigenvectors of the matrix $\mathbf{Q}$ and a and b are real numbers. Since $\mathbf{Q}$ is a symmetrical matrix, these eigenvectors are orthogonal. Since $\mathbf{x}$ must have unit amplitude, it follows that we must require that

$$a^2 + b^2 = 1. \tag{2.5-29}$$

The vector $\mathbf{x}$ must also be a solution to Eq. (2.5-17). If we use Eq. (2.5-28) in Eq. (2.5-17), we arrive at the following two equations:

$$\begin{aligned} \upsilon &= \lambda \\ a &= \frac{|\mathbf{u}|}{(m-\lambda)}. \end{aligned} \tag{2.5-30}$$

Here, λ is the eigenvalue corresponding to $\hat{\mathbf{e}}_3$. Combining the bottom line of Eq. (2.5-30) with Eq. (2.5-29) allows us to derive the value of b as

$$b = \pm\sqrt{1 - \frac{|\mathbf{u}|^2}{(m-\lambda)^2}}. \tag{2.5-31}$$

Using $\mathbf{x}$ with these values of a and b in the expression for the co-polarized power shows that the received power will be zero if λ corresponds to either λ_2 or λ_3 as given in Eq. (2.5-24). To decide which of the eigenvectors should be used in Eq. (2.5-28), we note that we can use Eq. (2.5-17) in the top line of Eq. (2.5-14) to write the co-polarized power as

$$P_{cn} = m + \mathbf{u}\cdot\mathbf{x} + \lambda = 0 \Rightarrow \lambda = -m - \mathbf{u}\cdot\mathbf{x}. \tag{2.5-32}$$

However,

$$-m - \mathbf{u}\cdot\mathbf{x} \le -m + |\mathbf{u}| = -\frac{1}{2}P_{cs} \le 0 \Rightarrow \lambda = \lambda_3. \tag{2.5-33}$$

Therefore, we have two solutions that would lead to zero co-polarized power. These two are

$$\mathbf{x}_{n1} = -a\frac{\mathbf{u}}{|\mathbf{u}|} + \sqrt{1-a^2}\,\hat{\mathbf{e}}_3; \quad \mathbf{x}_{n1} = -a\frac{\mathbf{u}}{|\mathbf{u}|} - \sqrt{1-a^2}\,\hat{\mathbf{e}}_3, \tag{2.5-34}$$

where

$$a = \frac{|\mathbf{u}|}{m + \sqrt{\left(m - |\mathbf{u}|\right)\left(m + |\mathbf{u}|\right)}}. \tag{2.5-35}$$

Equation (2.5-34) shows that the two co-polarized nulls and the co-polarized maximum and saddle point all lie in the same plane. Furthermore, the two co-polarized nulls are equal distances from the co-polarized saddle. This is easily proven by taking the dot products

$$-\frac{\mathbf{u}}{|\mathbf{u}|} \cdot \mathbf{x}_{n1} = -\frac{\mathbf{u}}{|\mathbf{u}|} \cdot \mathbf{x}_{n2} = a = \cos\alpha\,. \tag{2.5-36}$$

This is the famous polarization fork originally proven by Huynen in his landmark thesis (Huynen [12]).

Note that while the cross-polarized null polarization correspond to the co-polarized maximum polarization, the same is not true for the co-polarized null and the cross-polarized maximum. The co-polarized null polarizations are given in Eq. (2.5-34) and the cross-polarized maximum polarization vector is $\pm\hat{\mathbf{e}}_3$. The value of the maximum cross-polarized power is

$$P_{xm} = m + \sqrt{\left(m - |\mathbf{u}|\right)\left(m + |\mathbf{u}|\right)}\,. \tag{2.5-37}$$

Figure 2-21 shows these optimum polarizations on the Poincaré sphere.

2.5.4 Special Case: Multiple Scatterers with Reflection Symmetry

Here we shall consider the special case of the return from a group of scatterers with a special orientation distribution measured in the backscatter configuration. The orientation distribution is such that it is symmetrical about the vertical direction. In this case, it can be shown that

$$\left\langle S_{hv} S_{hh}^{*} \right\rangle = \left\langle S_{hv}^{*} S_{vv} \right\rangle = 0\,. \tag{2.5-38}$$

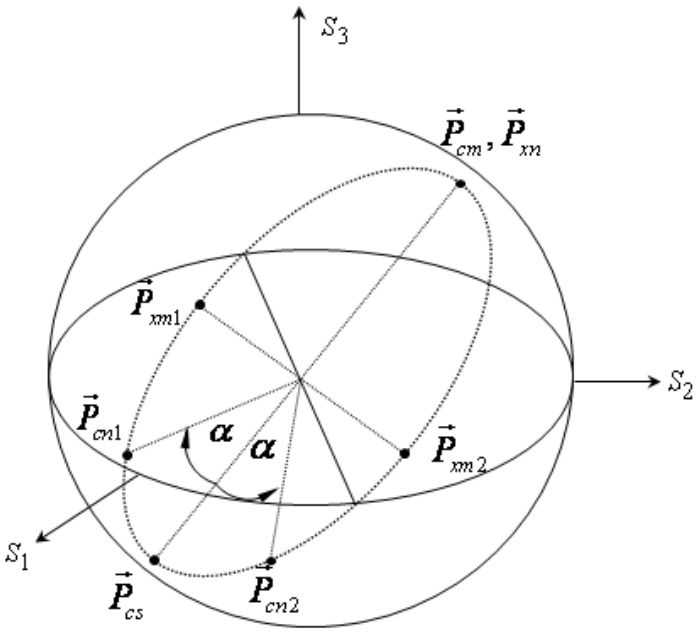


Fig. 2-21. The optimum polarizations plotted on the Poincaré sphere. The vectors corresponding to the co-polarized maximum, saddle point, and minima form the Huynen polarization fork. All these vectors lie in the same plane.

This leads to a special form of the Stokes scattering operator;

$$m = \frac{1}{4}\left\{\left\langle|S_{hh}|^2\right\rangle + \left\langle|S_{vv}|^2\right\rangle + 2\left\langle|S_{hv}|^2\right\rangle\right\}$$

$$\mathbf{u} = \frac{1}{4}\left\{\left\langle|S_{hh}|^2\right\rangle - \left\langle|S_{vv}|^2\right\rangle\right\}\begin{bmatrix}1\\0\\0\end{bmatrix}$$

$$Q = \frac{1}{4}\begin{bmatrix}\left\langle|S_{hh}|^2\right\rangle + \left\langle|S_{vv}|^2\right\rangle - 2\left\langle|S_{hv}|^2\right\rangle & 0 & 0\\ 0 & 2\left\langle|S_{hv}|^2\right\rangle + 2\,\mathrm{Re}\left\langle S_{hh}^{*}S_{vv}\right\rangle & 2\,\mathrm{Im}\left\langle S_{hh}^{*}S_{vv}\right\rangle\\ 0 & 2\,\mathrm{Im}\left\langle S_{hh}^{*}S_{vv}\right\rangle & 2\left\langle|S_{hv}|^2\right\rangle - 2\,\mathrm{Re}\left\langle S_{hh}^{*}S_{vv}\right\rangle\end{bmatrix}$$

,

where $\langle x\rangle$ means the ensemble average of x. In this case, **Q** has three eigenvalues

$$\lambda_1 = Q_{11} = m - \left\langle |S_{hv}|^2 \right\rangle$$
$$\lambda_2 = \frac{1}{2}\left[\left\langle |S_{hv}|^2 \right\rangle + \left| \left\langle S_{hh}^* S_{vv} \right\rangle \right| \right]. \tag{2.5-39}$$
$$\lambda_3 = \frac{1}{2}\left[\left\langle |S_{hv}|^2 \right\rangle - \left| \left\langle S_{hh}^* S_{vv} \right\rangle \right| \right]$$

The normalized eigenvector corresponding to the first eigenvalue is $\pm\mathbf{u}/|\mathbf{u}|$. Using this eigenvalue in Eq. (2.5-14) shows that the vector $\mathbf{u}/|\mathbf{u}|$ corresponds to a cross-polarized power value equal to $\left\langle |S_{hv}|^2 \right\rangle$. It also is a solution to Eq. (2.5-17) with

$$\upsilon = m - \left\langle |S_{hv}|^2 \right\rangle \pm |\mathbf{u}|. \tag{2.5-40}$$

Using these in the top equation in Eq. (2.5-14) gives the two co-polarized powers of

$$P_c = \left\langle |S_{hh}|^2 \right\rangle; \left\langle |S_{vv}|^2 \right\rangle. \tag{2.5-41}$$

To find the vectors corresponding to the co-polarized minima, we follow the same procedure as in the previous section and define a normalized vector

$$\mathbf{x} = a\hat{\mathbf{e}}_1 + b\hat{\mathbf{e}}_3, \tag{2.5-42}$$

where $\hat{\mathbf{e}}_1 = -\mathbf{u}/|\mathbf{u}|$ and $\hat{\mathbf{e}}_3$ are normalized eigenvectors of the matrix $\mathbf{Q}$ and a and b are real numbers. Since $\mathbf{Q}$ is a symmetrical matrix, these eigenvectors are orthogonal. In fact, the two eigenvectors corresponding to the second and third eigenvalues lie entirely in the plane containing the S_2 and S_3 axes of the Poincaré sphere. If $\mathrm{Im}\left\langle S_{hh}^* S_{vv} \right\rangle = 0$, these eigenvectors coincide with the S_2 and S_3 axes. Since $\mathbf{x}$ must have unit amplitude, it follows that we have to require that

$$a^2 + b^2 = 1. \tag{2.5-43}$$

The vector $\mathbf{x}$ must also be a solution to Eq. (2.5-17). If we use Eq. (2.5-28) in Eq. (2.5-17), we arrive at the following two equations:

$$\upsilon = \lambda$$
$$a = \frac{|\mathbf{u}|}{\left(m - \left\langle |S_{hv}|^2 \right\rangle - \lambda\right)}. \tag{2.5-44}$$

Here, λ is the eigenvalue corresponding to $\hat{\mathbf{e}}_3$. Combining the bottom line of Eq. (2.5-30) with Eq. (2.5-29) allows us to derive the value of b as

$$b = \pm\sqrt{1 - \frac{|\mathbf{u}|^2}{\left(m - \left\langle |S_{hv}|^2 \right\rangle - \lambda\right)^2}}. \tag{2.5-45}$$

Using $\mathbf{x}$ with these values of a and b in the expression for the co-polarized power shows that the received power will be

$$P_c = m + \lambda - \frac{|\mathbf{u}|^2}{\left(m - \left\langle |S_{hv}|^2 \right\rangle - \lambda\right)}. \tag{2.5-46}$$

The three values of the cross-polarized power are

$$P_x = \left\langle |S_{hv}|^2 \right\rangle; m - \frac{1}{2}\left\{\left\langle |S_{hv}|^2 \right\rangle \pm \left|\left\langle S_{hh}^* S_{vv} \right\rangle\right|\right\}. \tag{2.5-47}$$

Two of these values lie in the plane containing the S_2 and S_3 axes of the Poincaré sphere and one lies on the S_1 axis.

2.5.5 A Numerical Example

As an example of optimizing the power using independent transmit and receive antenna polarizations, let us consider the case of the randomly oriented cylinders from the previous section with a cosine squared distribution around the vertical direction. From Eq. (2.4-27), the normalized Stokes scattering operator have the following values

$$
\begin{aligned}
&m = 1 \\
&\tilde{\mathbf{u}} = \begin{pmatrix} -1/2 & 0 & 0 \end{pmatrix} \\
&\mathbf{v} = \mathbf{u} \\
&\mathbf{Q} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
\qquad (2.5\text{-}48)
$$

In this case, the three eigenvalues of the matrix $\mathbf{Q}$ are 0, 1/2, and 1/2. The maximum cross-polarized return occurs for the polarization corresponding to the smallest eigenvalue; i.e., for circular polarizations. The minimum cross-polarized return occurs for the eigenvector corresponding to the largest eigenvalue, which, in this case, is any polarization that lies in the S_1-S_2 plane of the Poincaré sphere; i.e., any linear polarization.

The co-polarized solutions are simply the HH and VV returns, with the maximum occurring for VV and the minimum for HH. Note that for the case of $\lambda = 0$, we find from Eq. (2.5-45) and Eq. (2.5-48) that $b = 0$, which means that the minimum co-polarized power also coincides with the saddle point.

In this section, we concentrated on those polarization combinations that optimize (either maximizes or minimizes) the received power from a particular scatter or ensemble of scatterers. However, maximizing or minimizing the power from a single pixel, or even an area in an image, is of limited utility. This is best illustrated using the polarimetric L-band image of San Francisco acquired with the NASA/JPL AIRSAR system that we discussed earlier. Figure 2-22 shows the total power image (the trace of the covariance matrix, or the M_{11} element of the Stokes scattering operator) of this image with three co-polarized signatures of the ocean (top), an urban area (middle), and an area in Golden Gate Park (bottom). The maximum for the ocean occurs near VV. The other two signatures have their maxima near HH polarization. These two images are shown in Fig. 2-23. The HH image shows that the urban areas and park are bright relative to the ocean. The VV image shows the ocean to be bright relative to the urban areas. In image analysis, it is more common to require the contrast between two different areas to be maximized so as to be able to clearly identify similar areas. One of the earliest papers on this topic is the one by Ioannidis and Hammers [11] that reported on the optimum polarization to discriminate a target in the presence of clutter. We shall look at this problem in the next section.

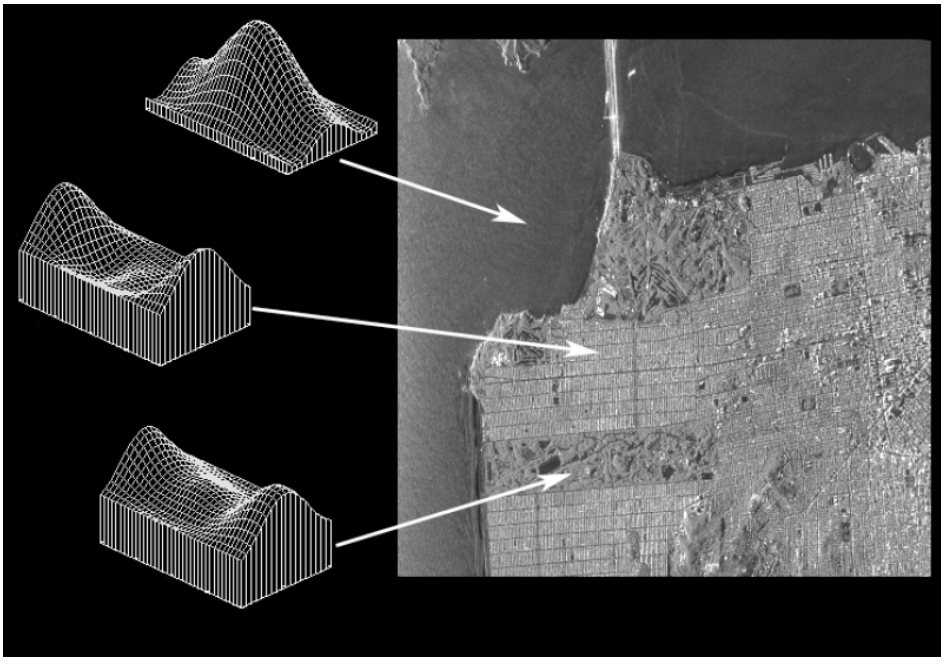


Fig. 2-22. An L-band total power image of San Francisco acquired with the NASA/JPL AIRSAR system. The three co-polarized signatures for the ocean, urban area, and Golden Gate Park are shown on the left.

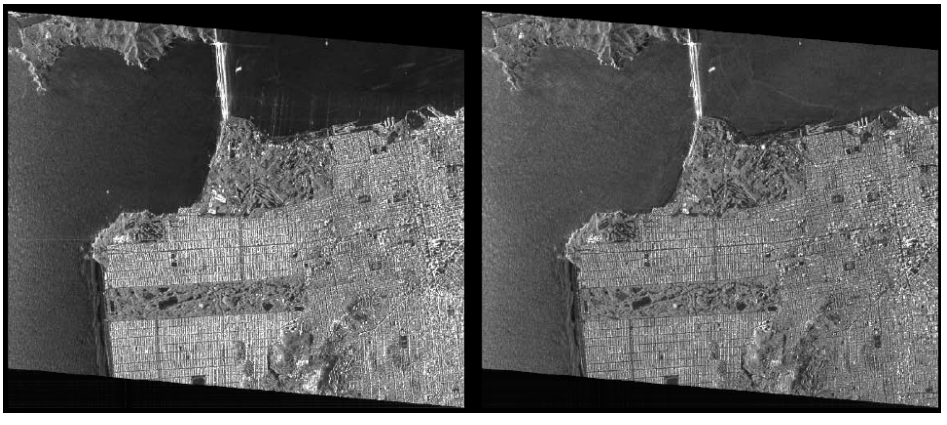


Fig. 2-23. HH polarized image (left) and VV polarized image (right) of San Francisco. The image on the left maximizes the power from the urban and park areas. The image on the right maximizes the power from the ocean. Note, however, that even though the power from the ocean is maximized in the image on the right, the contrast between the ocean and the urban areas is larger on the left.

2.6 Contrast Enhancement

Suppose we want to maximize or minimize the contrast between two areas in an image. We shall assume that each area is represented by an average Stokes scattering operator. The contrast can be written as

$$C_E = \frac{P_1}{P_2} = \frac{\mathbf{S}^{rec} \cdot \mathbf{M}_1 \mathbf{S}^{tr}}{\mathbf{S}^{rec} \cdot \mathbf{M}_2 \mathbf{S}^{tr}} = \frac{m_1 + \mathbf{u}_1 \cdot \mathbf{x} + \mathbf{v}_1 \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{Q}_1 \mathbf{x}}{m_2 + \mathbf{u}_2 \cdot \mathbf{x} + \mathbf{v}_2 \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{Q}_2 \mathbf{x}}, \tag{2.6-1}$$

where the subscripts 1 and 2 refer to the two areas. We now have to optimize this quantity subject to the constraints that the Stokes vectors have unit amplitude as shown in Eq. (2.5-4). To find the optimum values of the contrast function, we again use the Lagrange multiplier method as before and calculate the gradient with respect to $\mathbf{x}$ and $\mathbf{y}$. The optimum values are those for which the gradient is zero. This leads to the following two coupled equations:

$$\begin{aligned} \mathbf{y} &= \pm \frac{\mathbf{v}_1 + \mathbf{Q}_1 \mathbf{x} - C_E \left(\mathbf{v}_2 + \mathbf{Q}_2 \mathbf{x} \right)}{\left| \mathbf{v}_1 + \mathbf{Q}_1 \mathbf{x} - C_E \left(\mathbf{v}_2 + \mathbf{Q}_2 \mathbf{x} \right) \right|} \\ \mathbf{x} &= \pm \frac{\mathbf{u}_1 + \tilde{\mathbf{Q}}_1 \mathbf{y} - C_E \left(\mathbf{u}_2 + \tilde{\mathbf{Q}}_2 \mathbf{y} \right)}{\left| \mathbf{u}_1 + \tilde{\mathbf{Q}}_1 \mathbf{y} - C_E \left(\mathbf{u}_2 + \tilde{\mathbf{Q}}_2 \mathbf{y} \right) \right|} \end{aligned}. \tag{2.6-2}$$

The solution to these equations is a set of two non-linear equations in either $\mathbf{x}$ or $\mathbf{y}$ that is quite complex to solve. To provide more insight into the solution, a simpler approach is to consider the problem in two steps. First, we fix the transmit polarization and then derive the receive polarization that would optimize the contrast for that particular transmit polarization $\mathbf{x}$. For this, we introduce the shorthand notations:

$$\begin{aligned} S_{01} &= m_1 + \mathbf{u}_1 \cdot \mathbf{x} \\ \mathbf{s}_1 &= \mathbf{v}_1 + \mathbf{Q}_1 \mathbf{x} \\ S_{02} &= m_2 + \mathbf{u}_2 \cdot \mathbf{x} \\ \mathbf{s}_2 &= \mathbf{v}_2 + \mathbf{Q}_2 \mathbf{x} \end{aligned}. \tag{2.6-3}$$

It is then clear from the top line in Eq. (2.6-2) that we can write

$$\mathbf{y} = \frac{\mathbf{s}_1 - C_E \mathbf{s}_2}{\sqrt{\left(\mathbf{s}_1 - C_E \mathbf{s}_2 \right) \cdot \left(\mathbf{s}_1 - C_E \mathbf{s}_2 \right)}}. \tag{2.6-4}$$

Using this expression in the original expression for the contrast gives

$$C_E = \left\{ S_{01} + \frac{\mathbf{s}_1 \cdot \mathbf{s}_1 - C_E \mathbf{s}_1 \cdot \mathbf{s}_2}{\sqrt{(\mathbf{s}_1 - C_E \mathbf{s}_2) \cdot (\mathbf{s}_1 - C_E \mathbf{s}_2)}} \right\} \Bigg/ \left\{ S_{02} + \frac{\mathbf{s}_2 \cdot \mathbf{s}_1 - C_E \mathbf{s}_2 \cdot \mathbf{s}_2}{\sqrt{(\mathbf{s}_1 - C_E \mathbf{s}_2) \cdot (\mathbf{s}_1 - C_E \mathbf{s}_2)}} \right\}. \quad (2.6\text{-}5)$$

The only unknown in this expression is the contrast ratio C_E. Rearranging terms leads to the following quadratic expression for this ratio

$$C_E^2 \left(S_{02}^2 - \mathbf{s}_2 \cdot \mathbf{s}_2 \right) - 2C_E \left(S_{01} S_{02} - \mathbf{s}_1 \cdot \mathbf{s}_2 \right) + \left(S_{01}^2 - \mathbf{s}_1 \cdot \mathbf{s}_1 \right) = 0. \quad (2.6\text{-}6)$$

The optimum contrast ratios are then

$$C_{\max}(\mathbf{x}) = \frac{S_{01} S_{02} - \mathbf{s}_1 \cdot \mathbf{s}_2}{S_{02}^2 - \mathbf{s}_2 \cdot \mathbf{s}_2} + \sqrt{\left(\frac{S_{01} S_{02} - \mathbf{s}_1 \cdot \mathbf{s}_2}{S_{02}^2 - \mathbf{s}_2 \cdot \mathbf{s}_2} \right)^2 - \frac{S_{01}^2 - \mathbf{s}_1 \cdot \mathbf{s}_1}{S_{02}^2 - \mathbf{s}_2 \cdot \mathbf{s}_2}} \quad (2.6\text{-}7)$$

and

$$C_{\min}(\mathbf{x}) = \frac{S_{01} S_{02} - \mathbf{s}_1 \cdot \mathbf{s}_2}{S_{02}^2 - \mathbf{s}_2 \cdot \mathbf{s}_2} - \sqrt{\left(\frac{S_{01} S_{02} - \mathbf{s}_1 \cdot \mathbf{s}_2}{S_{02}^2 - \mathbf{s}_2 \cdot \mathbf{s}_2} \right)^2 - \frac{S_{01}^2 - \mathbf{s}_1 \cdot \mathbf{s}_1}{S_{02}^2 - \mathbf{s}_2 \cdot \mathbf{s}_2}}. \quad (2.6\text{-}8)$$

These expressions are still quite complex to solve analytically for the $\mathbf{x}$ that optimizes either the maximum or minimum contrast ratio. A simpler approach, proposed by Dubois and van Zyl (1989) [13], is to solve for the maximum of Eq. (2.6-8) over all $\mathbf{x}$ numerically. Once this maximum ratio is found, the optimum receive polarization follows from the top equation in Eq. (2.6-2) to be

$$\mathbf{y} = \frac{\mathbf{s}_1 - C_{opt} \mathbf{s}_2}{\left| \mathbf{s}_1 - C_{opt} \mathbf{s}_2 \right|}. \quad (2.6\text{-}9)$$

Note that these expressions turn out to be this complicated because we are insisting that the resulting Stokes vectors represent fully polarized waves; i.e., those of actual antennas. This, of course would be necessary if one wanted to implement a real radar system that operated with the optimum polarization combinations. However, for pure signal processing applications after data collection, one might want to relax this limitation. This approach was first reported by Swartz et al. [14] using the covariance approach for the backscatter case.

To illustrate the difference, let us assume that the two vectors $\mathbf{S}^{rec}$ and $\mathbf{S}^{tr}$ in Eq. (2.6-1) are simply 4-dimensional vectors. In that case, we solve for the optimum contrast by introducing (following Swartz et al. [14]) an artificial, but

general, constraint that the denominator of Eq. (2.6-1) should be equal to 1. The solution in this case then becomes

$$\begin{aligned}\mathbf{M}_1\mathbf{S}^{tr} &= \lambda_{tr}\mathbf{M}_2\mathbf{S}^{tr} \\ \tilde{\mathbf{M}}_1\mathbf{S}^{rec} &= \lambda_{rec}\tilde{\mathbf{M}}_2\mathbf{S}^{rec}\end{aligned}, \tag{2.6-10}$$

where λ_{tr} and λ_{rec} are the Lagrange multipliers. If the inverse of $\mathbf{M}_2$ exists, these expressions can be written as

$$\begin{aligned}\mathbf{M}_2^{-1}\mathbf{M}_1\mathbf{S}^{tr} &= \lambda_{tr}\mathbf{S}^{tr} \\ \tilde{\mathbf{M}}_2^{-1}\tilde{\mathbf{M}}_1\mathbf{S}^{rec} &= \lambda_{rec}\mathbf{S}^{rec}\end{aligned}, \tag{2.6-11}$$

Therefore, if we relax the requirement that the two vectors $\mathbf{S}^{rec}$ and $\mathbf{S}^{tr}$ represent fully polarized Stokes vectors, the solutions simply are the eigenvectors of $\mathbf{M}_2^{-1}\mathbf{M}_1$ and $\tilde{\mathbf{M}}_2^{-1}\tilde{\mathbf{M}}_1$. Note that, in the backscatter case, the Stokes scattering operators are symmetrical matrices and these two expressions become identical. This is also true for the solution to Eq. (2.6-2). If the inverse of $\mathbf{M}_2$ does not exist, we can rewrite Eq. (2.6-10) in a form that has the solution similar to Eq. (2.6-11), but with the roles of $\mathbf{M}_1$ and $\mathbf{M}_2$ interchanged.

2.6.1 Numerical Example

Consider the case of maximizing the contrast between the return from a dihedral corner reflector and that of a randomly oriented cloud of thin cylinders. In this case, the normalized Stokes scattering operators are

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad \mathbf{M}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.6-12}$$

First we look at the case where we insist that the Stokes vectors represent fully polarized waves. Fig. 2-24 shows graphically the maximum and minimum contrast ratios for each transmit polarization. In each case, we performed the calculations that follow. For each of the transmit polarizations shown, we calculated the contrast ratio for all possible receive polarizations. We then found the maximum and minimum values and plotted those two values for that particular transmit polarization in these two figures. From Fig. 2-24 we see that the maximum contrast occurs for linear transmit polarizations that have

orientations of either 45 deg or 135 deg. The corresponding maximum contrast ratio is 4. Using this information, Eq. (2.6-9) gives the optimum receive polarization as either the co-polarized companion to the transmit polarization (in which case the contrast ratio is 0 [minimum]) or the cross-polarized companion (in which case the contrast ratio is 4, or a maximum).

This result is not unexpected. We know (see Fig. 2-10 and Fig. 2-16) that a dihedral corner reflector has its maximum cross-polarized return at 45 deg or 135 deg linear polarizations. At the same time, the linear polarizations are where the randomly oriented cylinders have their minimum in their cross-polarized return. This combination argues that the maximum contrast should then be at the 45 deg or 135 deg linear cross-polarized pair.

Next, let us consider the case where we do not insist that the solutions only be Stokes vectors representing fully polarized waves. In this case, the inverse of $\mathbf{M}_2$ does not exist. However, we can rewrite Eq. (2.6-10) by multiplying both sides by the inverse of $\mathbf{M}_1$ to give

$$\mathbf{M}_1^{-1}\mathbf{M}_2\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\mathbf{S} = \lambda\mathbf{S}\,. \tag{2.6-13}$$

The eigenvalues are 2, 1, 0, and -2, and the corresponding four eigenvectors are the four principal dimensions of the vector $\mathbf{S}$. The Notice that this type of solution allows not only vectors that are not Stokes vectors, but also allows negative contrast ratios. The vector $\tilde{\mathbf{S}} = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}$ corresponds to the largest eigenvalue, but the largest contrast is actually achieved with the eigenvector $\tilde{\mathbf{S}} = \begin{pmatrix} 0 & 0 & 0 & \pm 1 \end{pmatrix}$ because, in this case, the denominator of the contrast is

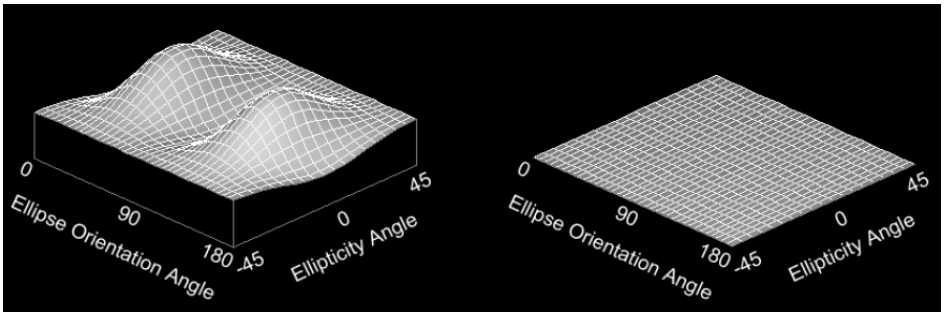


Fig. 2.24. The maximum (left) and minimum (right) contrast ratios calculated as a function of the transmit polarization orientation and ellipticity angles. The maximum contrast is achieved using a linear transmit polarization with an orientation of either 45 deg or 135 deg.

zero, theoretically allowing the contrast to approach infinity. Of course, in general, if there is any polarization combination that would zero out the power from the denominator, one could argue that the contrast ratio would be maximized. The problem with this approach is that the exact polarization combination at which the zero occurs might vary from one pixel in an image to the next, causing the result to be quite noisy.

2.6.2 Image Example

In the late 1980s, two different groups investigated the optimum polarizations to maximize the contrast between man-made structures (such as buildings) and natural terrain. Swartz et al. [14] used data from an area in the city of San Francisco to represent man-made structures and data from an area in Golden Gate Park for natural terrain. They found that the optimum polarizations were a transmitting antenna with parameters $(\psi_t, \chi_t) = (-41.3°, -6.44°)$ and a receiving antenna with parameters $(\psi_r, \chi_r) = (60.3°, 3.51°)$. Notice that these antennas are very nearly orthogonally polarized and close to the combination we derived in the previous section.

Dubois and van Zyl (1989) [13] made a more systematic study of the same problem, choosing many different pairs of areas and repeating the experiment at different incidence angles. They concluded that the optimum polarization varies little with the change in incidence angles and that a 45 deg cross-polarized pair provides the optimum contrast. They defined the polarization enhancement as the contrast calculated with normalized Stokes scattering operators and found the enhancement factor to be about 3 dB independent of the incidence angle and of the size of the image area used to characterize either class of scatterer. Furthermore, they found the same results if they used bare surfaces as the natural terrain.

Figure 2-25 illustrates this using the data from the NASA/JPL AIRSAR system that are shown in Fig. 2-23. Notice the excellent contrast between the urban areas and the vegetated Golden Gate Park, as well as the contrast between the urban areas and the ocean. Comparing to the image on the left in Fig. 2-23, we note some improvement in contrast between the urban area and the park. In fact, Swartz et al. [14] found that among the canonical polarization combinations, HH provided the best contrast, but was still about 2 dB less than what can be achieved with the optimum polarization combination.

Figure 2-26 shows the "image" using the non-physical polarization that theoretically zeros the power from a cloud of randomly oriented thin cylinders, as discussed in the previous section. This image, which is the M_{44} element of the Stokes scattering operator, shows the areas that are dominated by double

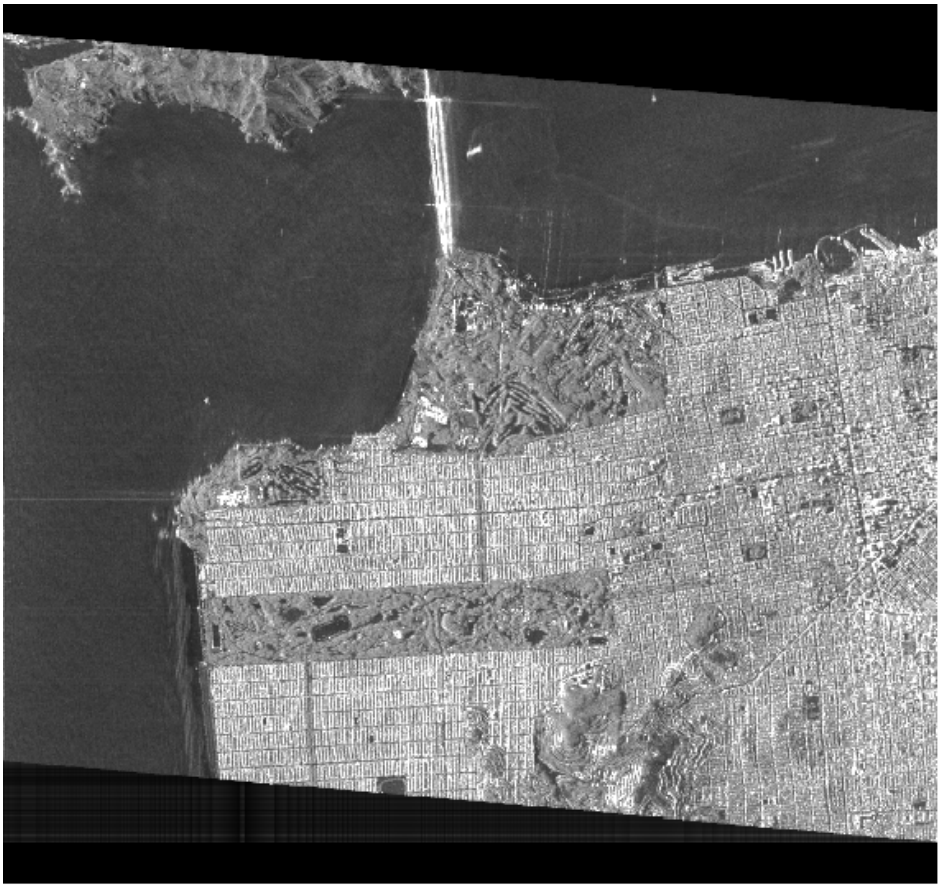


Fig. 2-25. Image of San Francisco, California, synthesized using a 45-deg linearly polarized antenna to transmit the signals and a 135-deg linearly polarized antenna to receive. Notice the increased contrast between the urban areas and the vegetated Golden Gate Park, as well as the contrast between the urban areas and the ocean.

reflections to be very bright, areas that are vegetated medium gray, and areas that have near specular scattering as very dark. To understand why this is the case, recall that for specular reflection the scattering matrix is the identity matrix. In that case, the M_{44} element of the normalized Stokes scattering operator is -1. The same element is zero for a randomly oriented cloud of thin cylinders and +1 for double reflections from a dihedral corner reflector.

Note, however, that this image is quite noisy, making it less useful than, for example, Fig. 2-25, for image-photo-interpretation purposes. It does, however, provide a simple check on whether a bright point in an image is so because it represents specular scattering or because it represents a double reflection resembling a dihedral corner reflector.

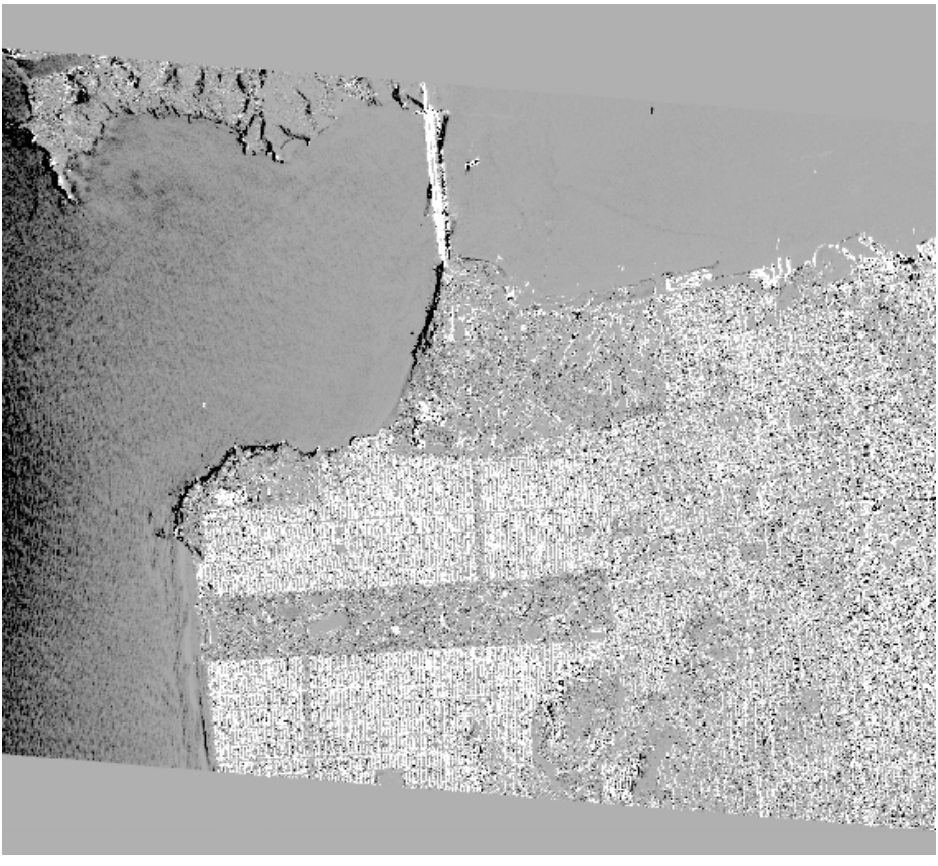


Fig. 2-26. "Image" of the M_{44} element of the Stokes scattering operator of San Francisco, California. Bright areas represent double reflections, while dark areas represent near-specular scattering. The image is quite noisy, reducing it utility for image-photo interpretation.

2.7 Summary

In this chapter, we introduced the basic principles of radar polarimetry. We introduced the important concepts of the various mathematical representations of scatterers and briefly reviewed the implementation of a radar polarimeter. We also discussed the concept of the polarization responses, and provided many examples illustrating this concept of visualizing the information contained in the polarization properties of a scatterer. Next, we looked at how to optimize the power received from a particular scatterer or group of scatterers by varying the polarization. We concluded with the case of using polarization to optimize the contrast between scatterers.

The discussion and examples presented in this chapter show the utility of polarization to extract more information regarding the properties of scatterers. In the rest of this book, we shall build upon these concepts and discuss more sophisticated analysis techniques to extract detailed geophysical information from polarimetric radar data.

References

[1] F. T. Ulaby and C. Elachi (Eds), *Radar Polarimetry for Geoscience Applications*, Artech House, Inc., Norwood, Massachusetts, 1990.

[2] E. M. Kennaugh, *Effects of the type of polarization on echo characteristics*, Report 389-9, Antenna Laboratory, Ohio State University, Columbus, Ohio, 1951.

[3] A. B. Kostinski and W. M. Boerner, "On Foundations of Radar Polarimetry," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, pp. 1395-1404, 1986.

[4] J. J. van Zyl, H. A. Zebker, and C. Elachi, "Imaging Radar Polarization Signatures: Theory and Observation," *Radio Science*, vol. 22, pp. 529-543, 1987.

[5] H. A. Zebker, J. J. van Zyl, and D. N. Held, "Imaging Radar Polarimetry from Wave Synthesis," *Journal of Geophysical Research*, vol. 92, pp. 683-701, 1987.

[6] J. J. van Zyl, "On the Importance of Polarization in Radar Scattering Problems," Ph.D. Thesis, Caltech Antenna Laboratory Report No. 120, California Institute of Technology, Pasadena, California, 1985.

[7] J. J. van Zyl, C. H. Papas, and C. Elachi, "On the Optimum Polarizations of Incoherently Reflected Waves," *IEEE Transactions on Antennas and Propagation*, vol. AP-35, pp. 818-825, 1987.

[8] J. C. Souyris, P. Imbo, R. Fjortoft, S. Mingot, and J. S. Lee, "Compact Polarimetry Based on Symmetry Properties of Geophysical Media: the $\pi/4$ Mode," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-43, pp. 634-646, 2005.

[9] R. K. Raney, "Hybrid-Polarity SAR Architecture," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-45, pp. 3397-3404, 2007.

[10] A. P. Agrawal and W. M. Boerner, "Redevelopment of Kennaugh's Target Characteristic Polarization State Theory Using the Polarization Transformation Ratio for the Coherent Case," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-27, pp. 2-14, 1988.

[11] G. A. Ioannidids and D. E. Hammers, "Optimum Antenna Polarizations for Target Discrimination in Clutter," *IEEE Transaction on Antennas and Propagation*, vol. AP-27, pp. 357-363, 1979.

[12] J. R. Huynen, *Phenomenological Theory of Radar Targets*, Ph.D. Dissertation, Drukkerij Bronder-Offset, N.V., Rotterdam, The Netherlands, 1970.

[13] P. C. Dubois and J. J. van Zyl, "Polarization Filtering of SAR data," *IGARSS '89 Digest*, vol. 3, pp. 1816-1819, 1989.

[14] A. A. Swartz, H. A. Yueh, J. A. Kong, L. M. Novak, and R. T. Shin, "Optimal polarizations for achieving maximum contrast in radar images," *Journal of Geophysical Research*, vol. 93, no. B12, pp. 15252-15260, 1988.

Chapter 3
Advanced Polarimetric Concepts

Since the introduction of polarimetric radar data in the 1980s, many different analysis techniques have been investigated. Many of these are application specific. In this chapter we shall discuss the theoretical background for many of these techniques and compare the information that is provided by the different approaches. As we shall see, many of these techniques provide very similar information, with the result that the choice of analysis technique becomes more one of personal preference.

3.1 Vector-Matrix Duality of Scatterer Representation

In the previous Chapter we demonstrated that the received power can be written in terms of the scatterer covariance matrix, as follows:

$$P = \mathbf{A} \cdot [\mathbf{C}] \mathbf{A}^*; \quad [\mathbf{C}] = \mathbf{T}\tilde{\mathbf{T}}^* \tag{3.1-1}$$

The superscript * denotes complex conjugation and ~ denotes the transpose operation. The vector $\mathbf{T}$ contains the same information as the original scattering matrix.

If we restrict ourselves to the backscatter direction where $S_{hv} = S_{vh}$, the usual forms for the antenna and scatterer vectors are

$$\tilde{\mathbf{A}} = \left(p_h^{rec} p_h^{tr} \quad 1/\sqrt{2} \left\{ p_h^{rec} p_v^{tr} + p_v^{rec} p_h^{tr} \right\} \quad p_v^{rec} p_v^{tr} \right)$$

and

$$\tilde{\mathbf{T}} = \left(S_{hh} \quad \sqrt{2} S_{hv} \quad S_{vv} \right).$$

The covariance matrix as defined in this expression is a positive semi-definite Hermetian matrix. This means that all the eigenvalues are real and that the eigenvectors are orthogonal. To prove that the eigenvalues have to be non-negative, recall that we can diagonalize the covariance matrix using a unitary transformation, as follows

$$[\mathbf{\Lambda}] = [\mathbf{U}]^{-1}[\mathbf{C}][\mathbf{U}]. \tag{3.1-2}$$

In this equation, $[\mathbf{\Lambda}]$ is a 3 × 3 diagonal matrix containing the non-negative real eigenvalues of the covariance matrix and $[\mathbf{U}]$ is a 3 × 3 complex matrix with columns equal to the normalized eigenvectors of the covariance matrix. Because the covariance matrix is Hermitian, however, the matrix $[\mathbf{U}]$ also satisfies

$$[\mathbf{U}]^{\dagger}[\mathbf{U}] = [\mathbf{I}] \quad \Rightarrow \quad [\mathbf{U}]^{\dagger} = [\mathbf{U}]^{-1}, \tag{3.1-3}$$

where the † sign denotes the adjoint (complex conjugate transpose) of the matrix. The normalized eigenvectors (or their complex conjugates) form an orthonormal basis; we can, therefore, write any antenna vector as a linear combination of these vectors. Specifically,

$$\mathbf{A} = b_1\hat{\mathbf{e}}_1^* + b_2\hat{\mathbf{e}}_2^* + b_3\hat{\mathbf{e}}_3^* = [\mathbf{U}]^* \mathbf{B}. \tag{3.1-4}$$

Using Eq. (3.1-4) in Eq. (3.1-1), we find the following expression for the power

$$P = [\mathbf{U}]^* \mathbf{B} \cdot [\mathbf{C}][\mathbf{U}]\mathbf{B}^* = \mathbf{B} \cdot [\mathbf{U}]^{-1}[\mathbf{C}][\mathbf{U}]\mathbf{B}^* = |b_1|^2 \lambda_1 + |b_2|^2 \lambda_2 + |b_3|^2 \lambda_3 \geq 0. \tag{3.1-5}$$

The received power must be non-negative for all antenna vectors, which means that all the eigenvalues must be non-negative.

In the case where the covariance matrix represents a single scatterer (that is, it was calculated from a scattering matrix, as shown in Eq. (3.1-1)), it is easy to show that the eigenvalues are

$$\lambda_1 = S_{hh}S_{hh}^* + S_{vv}S_{vv}^* + 2S_{hv}S_{hv}^*, \lambda_2 = \lambda_3 = 0. \tag{3.1-6}$$

In this case, two of the three eigenvalues are zero. This, in fact, is the test of whether one could calculate an equivalent scattering matrix from any given covariance matrix.

The covariance matrix characterization is particularly useful when analyzing multi-look radar images, since the covariance matrix of a multi-look pixel is simply the average covariance matrix of all the individual measurements

contained in the multi-look pixel. Recall that multi-looking is performed by averaging the power from adjacent pixels together in order to reduce speckle. This averaging process can be written as

$$\langle P \rangle = \frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} P_{ij} = \frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} \mathbf{A} \cdot [C]_{ij} \mathbf{A}^* = \mathbf{A} \cdot \langle [\mathbf{C}] \rangle \mathbf{A}^*, \tag{3.1-7}$$

where the two subscripts denote averaging in the range and azimuth directions, respectively. The angular brackets $\langle\ \rangle$ denote this spatial averaging. In general, this average covariance matrix will have more than one non-zero eigenvalue. All eigenvalues must, however, still be non-negative.

Cloude (1992) [1] was the first to use the orthonormality of the eigenvectors of the covariance matrix (in the context of radar polarimetry) to propose the decomposition of the covariance matrix in terms of its eigenvalues and eigenvectors. Specifically:

$$\langle [C] \rangle = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^{\dagger} . \tag{3.1-8}$$

The decomposition proposed by Cloude [1], as shown in Eq. (3.1-8), is unique. That is, since the eigenvectors of the covariance matrix are orthogonal, they form a natural basis in which to express the scattering. In some sense, this breaks the covariance matrix into orthogonal components, as one would normally do for a vector. However, as a vector can be expressed in many different coordinate systems, so too can a covariance matrix. For example, we could also choose to write the average covariance matrix in the following form

$$\langle [C] \rangle = \sum_{i=1}^{3} w_i \hat{\mathbf{p}}_i \hat{\mathbf{p}}_i^{\dagger} , \tag{3.1-9}$$

where the vectors $\hat{\mathbf{p}}_i$ form an orthonormal set. An example of such a set is the vectors representing the Pauli spin matrices:

$$\hat{\mathbf{p}}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}; \quad \hat{\mathbf{p}}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}; \quad \hat{\mathbf{p}}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{3.1-10}$$

Note that any set of orthonormal vectors can be used when decomposing the averaged covariance matrix. Just as in the vector case, the choice of coordinate system depends on the application. In the case of the covariance matrix, this

choice is often dictated by the fact that we are trying to interpret the total scattering in terms of known models or scattering mechanisms. For example, the three Pauli vectors represent scattering from a metallic trihedral corner reflector, a metallic dihedral corner reflector, and a metallic dihedral corner reflector rotated by 45 degrees (deg) about the line of sight, respectively, as discussed in Chapter 2.

The Cloude eigenvector decomposition is a special case of the general decomposition. While it is mathematically unique, its interpretation is not necessarily straightforward. The reason for this is that there is no guarantee that the eigenvectors will represent any known physical scattering mechanism directly. An additional complication comes from the fact that if this decomposition is done on every pixel in a multi-looked image, the eigenvectors that form the coordinate system for this decomposition might be different from pixel to pixel. This means that the coordinate system generally varies from pixel-to-pixel unless all the covariance matrices have identical eigenvectors. Therefore, the value of any eigenvalue might vary from pixel to pixel, and it is not easy to tell if the variation is due to the strength of the scattering or to the fact that the eigenvectors are different. One common way to reduce this problem is to express the scattering vector in the Pauli basis and then calculate the equivalent covariance matrix and perform the decomposition in this basis. This does not, however, actually overcome the fundamental issue of the coordinate system varying from pixel to pixel; instead, this approach facilitates interpreting the eigenvectors in terms of Pauli vectors.

Many authors have proposed so-called target decomposition schemes in which a covariance matrix is decomposed into separate matrices on the basis of simple models. In those cases, care must be taken to ensure that the individual matrices that are used in the decomposition all satisfy the condition that their individual eigenvalues must be non-negative. We shall discuss this further in a later section.

Whether one thinks of a scatterer in terms of a covariance matrix or a set of scattering vectors is a matter of personal choice. In the rest of this chapter we shall explore various interpretations of the average scattering based on different applications. Before doing so, we shall discuss a number of polarimetric parameters that are often encountered in the literature.

3.2 Eigenvalue and Eigenvector-Based Polarimetric Parameters

Cloude (1992) [1] and, later, Cloude and Pottier (1995) [2] introduced a number of polarimetric parameters that are derived from the eigenvalues and eigenvectors that are commonly used today. In this section, we shall discuss

these parameters, as well as others that are commonly encountered in the literature.

3.2.1 Parameters Used to Describe Randomness in Scattering

One such parameter, intended to measure target randomness, is the entropy, which defined as

$$H_T = -\sum_{i=1}^{3} P_i \log_3 P_i; \quad P_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \lambda_3}. \tag{3.2-1}$$

As pointed out by Cloude [1], the target entropy is a measure of target disorder, with $H_T = 1$ for random targets with three equal eigenvalues and $H_T = 0$ for simple (single, non-random) targets.

Recall from the discussion of polarization responses in Chapter 2 that the amount of variation (that is, randomness) in the scattering properties manifests itself in the form of a "pedestal" in the polarization response. While we normally refer to the pedestal height in the context of the co-polarized response, Durden et al. (1990) [3] showed that measuring the pedestal height is equivalent to measuring the ratio of the minimum eigenvalue to the maximum eigenvalue; that is

$$Pedestal\ Height = \frac{\min(\lambda_1, \lambda_2, \lambda_3)}{\max(\lambda_1, \lambda_2, \lambda_3)}. \tag{3.2-2}$$

In reality, this ratio over-estimates the total variation in the observed radar cross-section as a function of polarization. The reason for this is that in order for this minimum and maximum to be realized, both eigenvectors corresponding to those eigenvalues must be valid antenna vectors (see (Eq. 3.1-5)). This is not necessarily the case, possibly resulting in a smaller ratio. Nevertheless, this definition of the pedestal height is a useful measure of the randomness of the scattering process.

Using the model of randomly oriented thin cylinders (we will discuss this in more detail later), Kim and van Zyl (2001) [4] introduced the so-called radar thin vegetation index (RVI):

$$RVI = \frac{4\min\left(\lambda_1, \lambda_2, \lambda_3\right)}{\lambda_1 + \lambda_2 + \lambda_3} = \frac{8\sigma_{hv}}{\sigma_{hh} + \sigma_{vv} + 2\sigma_{hv}}. \tag{3.2-3}$$

(The equation of the right is true for media with reflection symmetry.) This parameter, which is also a measure of the randomness in the scattering, generally varies between 0 and 1. This ratio decreases as the cylinders become

thick compared the radar wavelength. In the limiting case where the cylinders are very thick compared to the radar wavelength, this ratio approaches zero. The factor 4 in Eq. (3.2-3) is arbitrary: it was chosen so that the RVI for a cloud of randomly oriented thin cylinders would be equal to 1.

To illustrate the similarity between these parameters with real image data, we shall first consider the image of San Francisco we discussed in Chapter 2. From the signatures shown in Fig. 2-22, we expect little randomness for the ocean scattering, and significant randomness for the vegetation scattering in the Golden Gate Park area. The urban area should show intermediate randomness. Figure 3-1 shows the comparison of the three measures of randomness introduced above. What is immediately obvious is that these three images convey the same basic information. The only real difference is the scaling from the bare ocean surface to the vegetated areas. The entropy image shows a more compressed scale with less variation in color than the other two. The pedestal height image based on the ratio of the eigenvalues shows the largest dynamic range. Overall, however, there is little reason to prefer one display over the other.

Note how both the entropy and the thin vegetation index show an increase in randomness in the ocean from left to right in the image. Recall that the radar illumination is from the left and that the angle of incidence increases from left to right. The observed increase in randomness in the ocean is due to the decrease in signal-to-noise ratio as the angle of incidence increases. Also, note that not all urban areas show the same amount of randomness. This can be

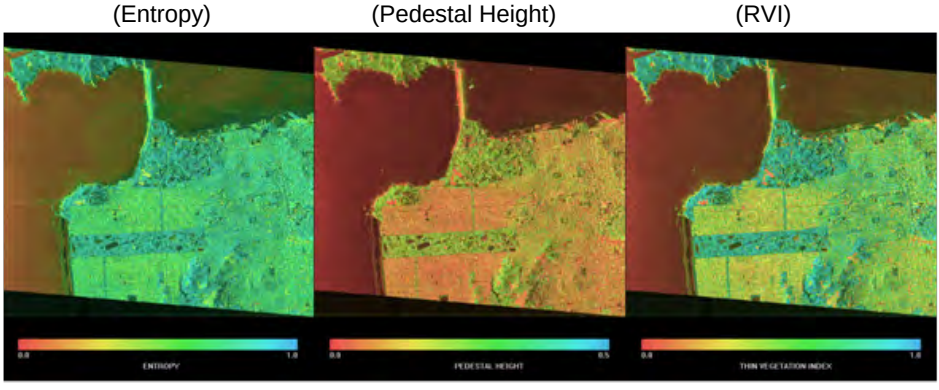


Fig. 3-1. L-band randomness images of San Francisco acquired with the NASA/JPL AIRSAR system. The Golden Gate Bridge is visible in the top center of the image linking the Presidio of San Francisco to the Golden Gate National Recreational Area north of the entrance to San Francisco Bay. Golden Gate Park is the rectangular feature in the lower half of the image in the left portion of the city. The image on the left shows the entropy scaled from 0 to 1. The middle image shows the pedestal height scaled from 0 to 0.5. The image on the right shows the RVI scaled from 0 to 1.

explained by the orientation of the buildings relative to the radar look direction. If a building is oriented such that it presents a large face to the radar, in general, the dominant scattering mechanism is a double reflection off the street onto the face of the building and back to the radar. This is the case above and below Golden Gate Park in the image. If the building is turned slightly, however, such that the front face of the building no longer is orthogonal to the direction in which the radar waves propagate, this double reflection signal no longer travels back to the radar. The result is that other direct reflections from the street and, possibly, the roofs of the buildings begin to dominate, and the signals appear more random. This effect is visible in the middle right portion of the image.

As a second example, we consider an image of a portion of the Black Forest in Germany acquired with the NASA/JPL AIRSAR system during the summer of 1991. The L-band image is shown for reference in Fig. 3-2. The bright feature in the left portion of the image is the town of Vilingen. The brighter right-hand portion of the image is a mixed forest consisting of spruce (*Picea abies*), pine (*Pinus sylvestris*), and fir (*Abies alba*) trees. The dry weight biomass ranges up to 50 kilograms per square meter (kg/m^2). The darker areas in the upper portion

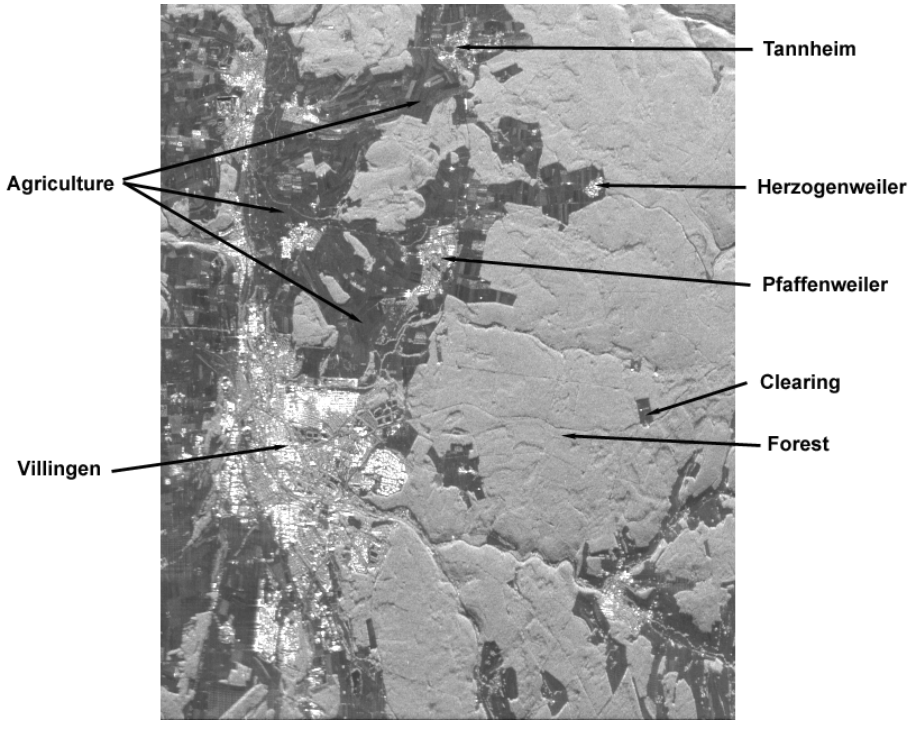


Fig. 3-2. L-band total power image of a portion of the Black Forest in Germany acquired with the NASA/JPL AIRSAR system in the summer of 1991.

and to the left in the image are mostly agricultural fields with varying amounts of biomass, depending on the crop type and maturity of the plants.

Figure 3-3 shows the three parameters calculated from the L-band data. In all cases, the forested areas show significant randomness, and the urban areas show little randomness. The agricultural areas show variations consistent with the amount of vegetation present on a field scale.

Figure 3-4 compares the radar vegetation index for the three frequencies (C-band, L-band, and P-band) that the AIRSAR system uses to acquire images. The C-band images shows much more detail in the agricultural areas because the shorter wavelength is more sensitive to the smaller biomass in these fields. The P-band image, on the other hand, shows a large variation in the forested area. This is due to the increased penetration through the canopy at the longer wavelength with a resulting increase in double reflections from the ground to the trunks of the trees and back to the radar. The variation in RVI is due to the effect that the underlying topography has on the resulting mixture of scattering mechanisms, as discussed by van Zyl (1993) [5].

3.2.2 Alpha Angle

Cloude and Pottier (1996) [6] proposed the following description for the eigenvectors of the covariance matrix:

$$\tilde{\mathbf{e}} = \begin{pmatrix} \cos\alpha & \sin\alpha\cos\beta e^{i\delta} & \sin\alpha\sin\beta e^{i\gamma} \end{pmatrix} \tag{3.2-4}$$

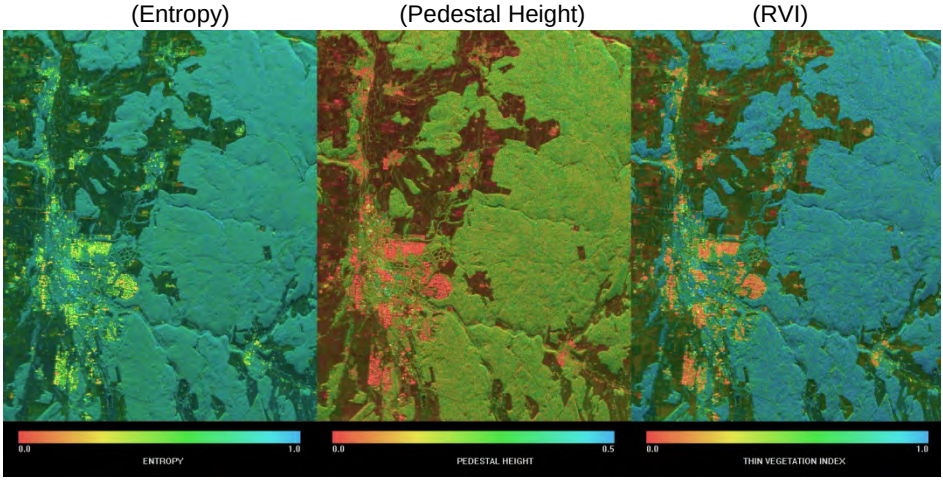


Fig. 3-3. L-band randomness images of the area shown in Fig. 3-2. Forested areas show the highest amount of randomness.

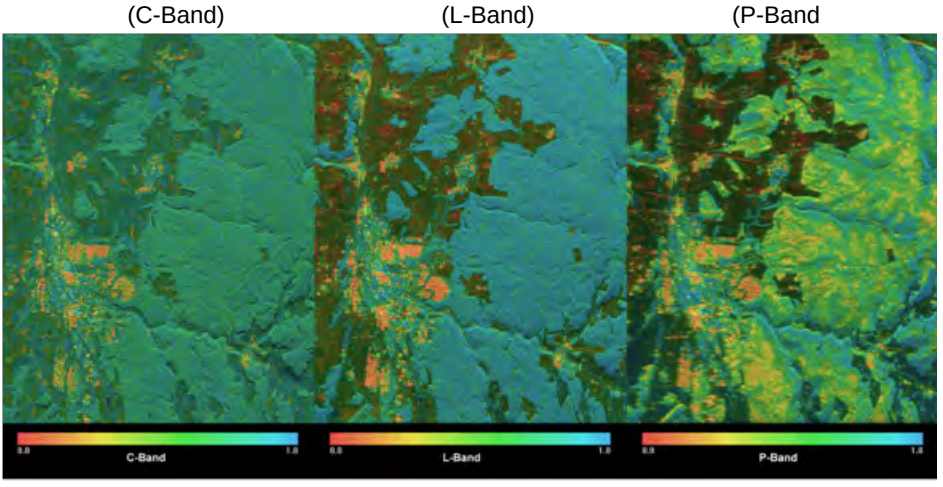


Fig. 3-4. Radar vegetation index images of the area shown in Fig. 3-2 for different frequencies. See the text for a discussion of the images.

in a basis formed by the Pauli matrices. The average angles are then calculated using

$$\bar{\alpha} = \sum_{i=1}^{3} \alpha_i P_i \,, \tag{3.2-5}$$

where P_i is defined in Eq. (3.2-1). The α angle in particular has received significant attention, and, together with the entropy, has been proposed as a way to perform an unsupervised classification of polarimetric SAR images. To investigate the meaning of these angles a bit further, consider that the normalized eigenvector (expressed in Pauli basis) can be written as follows

$$\tilde{\mathbf{e}} = \frac{e^{i\varphi}}{\sqrt{2}\left(q_{hh} + q_{vv} + 2q_{hv}\right)}\left(\left(S_{hh} + S_{vv}\right) \quad \left(S_{hh} - S_{vv}\right)e^{-i\delta} \quad 2S_{hv}e^{-i\gamma}\right), \tag{3.2-6}$$

where $q_{xy} = S_{xy}S_{xy}^{*}$ and the phase angle φ represents the phase of $S_{hh} + S_{vv}$. A comparison of Eq. (3.2-6) and Eq. (3.2-4) now shows that the angle δ is the relative phase between $S_{hh} - S_{vv}$ and $S_{hh} + S_{vv}$. Similarly, the angle γ is the relative phase between S_{hv} and $S_{hh} + S_{vv}$. Next, notice that

$$\tan\beta = \frac{2\left|S_{hv}\right|}{\left|S_{hh} - S_{vv}\right|} = \frac{2\left|S_{hv}/S_{vv}\right|}{\left|1 - S_{hh}/S_{vv}\right|} = \frac{2\left|R_x\right|}{\left|1 - R_{co}\right|}. \tag{3.2-7}$$

The angle β is, therefore, a function of the ratio of the cross-polarized term to the co-polarized VV term in the scattering matrix, as well as the ratio of the two co-polarized terms. The angle α can similarly be written as

$$\tan\alpha = \frac{\sqrt{\left|1-R_{co}\right|^2 + 4\left|R_x\right|^2}}{\left|1+R_{co}\right|}. \tag{3.2-8}$$

At this point, it is useful to compare these parameters for a few canonical cases. These are summarized in Table 3-1.

The results in Table 3-1 show that the angle α varies from zero deg for trihedral scattering to $\pi/2$ for dihedral scattering. Dipole scattering represents an intermediate case where $\alpha = \pi/4$. Note that this value of α for dipole scattering does not depend on the physical orientation of the dipole. The angle β is near zero for all cases where the cross-polarized return is small compared to the co-polarized returns. For the case of a single dipole, the β angle is related to the physical orientation of the dipole.

Also note that when the cross-polarized term is small compared to the co-polarized ones, the angle α is basically dependent on the co-polarized ratio. For example, in the case of a bare slightly rough surface, this ratio is a function

Table 3-1. Comparison of α and β angles for canonical scatterers.

Scatterer	Matrix Elements	R_{co}	R_x	α	β
Vertical Dipole	$S_{vv} \neq 0; S_{hh} = S_{hv} = 0,$	0	0	$\pi/4$	0
Horizontal Dipole	$S_{hh} \neq 0; S_{vv} = S_{hv} = 0,$	∞	0	$\pi/4$	0
Trihedral	$S_{hh} = S_{vv} = 1, S_{hv} = 0$	1	0	0	0
Dihedral	$S_{hh} = -S_{vv} = 1, S_{hv} = 0$	-1	0	$\pi/2$	0
Dipole oriented at angle ψ w.r.t. horizontal direction	$S_{hh} = \cos^2\psi$ $S_{vv} = \sin^2\psi$ $S_{hv} = \sin\psi\cos\psi$	$\tan^2\psi$	$\tan\psi$	$\pi/4$	2ψ
Slightly Rough Surface	$S_{vv} \geq S_{hh}; R_x \ll 1$	$0 \leq R_{co} \leq 1$	$R_x \ll 1$	$\approx \tan^{-1}\left(\frac{\left\|1-R_{co}\right\|}{\left\|1+R_{co}\right\|}\right)$	≈ 0
Dielectric Double Bounce	$S_{vv} \leq S_{hh}; R_x \ll 1$	$0 \geq R_{co} \geq -1$	$R_x \ll 1$	$\approx \tan^{-1}\left(\frac{\left\|1-R_{co}\right\|}{\left\|1+R_{co}\right\|}\right)$	≈ 0

of the surface dielectric constant. Therefore, for such a surface, α would be a function of the dielectric constant. For wet surfaces at high incidence angles, dipole scattering. Figure 3-5 shows the alpha angle for different dielectric constants and different incidence angles. Note that as the surface dielectric constant increases, the alpha angle increases as discussed above. Also, because of the change in the co-polarized ratio, as the angle of incidence increases, so does the alpha angle.

In the case of double reflections from dielectric surfaces, the co-polarized ratio will approach infinity at the Brewster angle. For that case, the angle α will also approach $\pi/4$. Figure 3-6 illustrates this point further. The figure illustrates an example where we calculated the alpha angle for a dielectric dihedral reflection assuming that both surfaces have the same dielectric constant. When the dielectric constant becomes very large, the alpha angle approaches $\pi/2$, which is the expected value for a metallic dihedral. For low dielectric constants, however, the alpha angle is closer to 45 deg. Note that for some dielectric constants and angle-of-incidence ranges the alpha angle can actually be less than $\pi/4$. This will mostly happen for low dielectric constant values at larger angles of incidence. These examples illustrate that care must be exercised when interpreting the values of α.

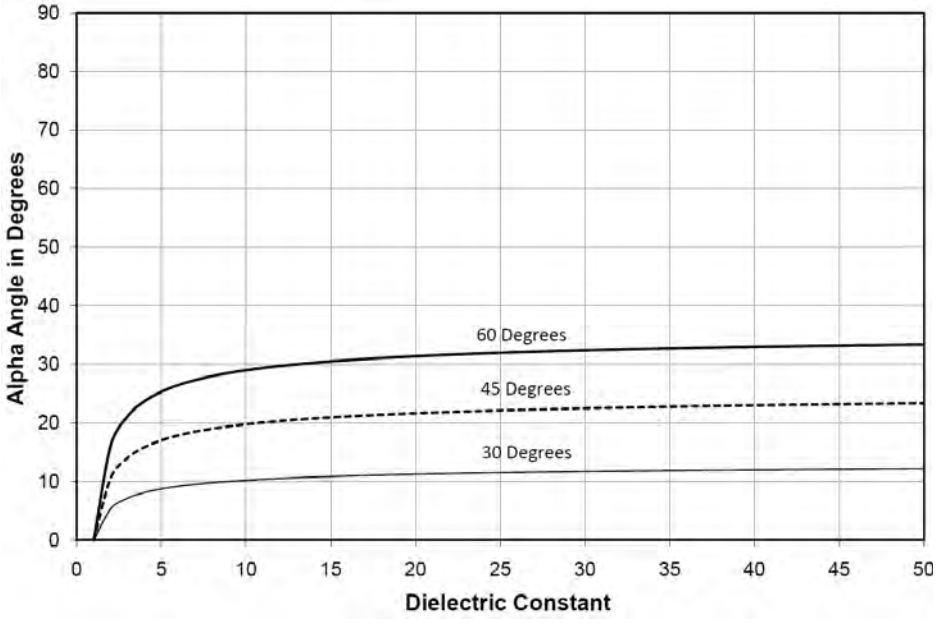


Fig. 3-5. Alpha angle as a function of dielectric constant for a slightly rough dielectric surface for three different incidence angles. Note that as the incidence angle and the dielectric constant increases, the alpha angle also increases. At very large dielectric constants and incidence angles, the alpha angle will approach 45 deg.

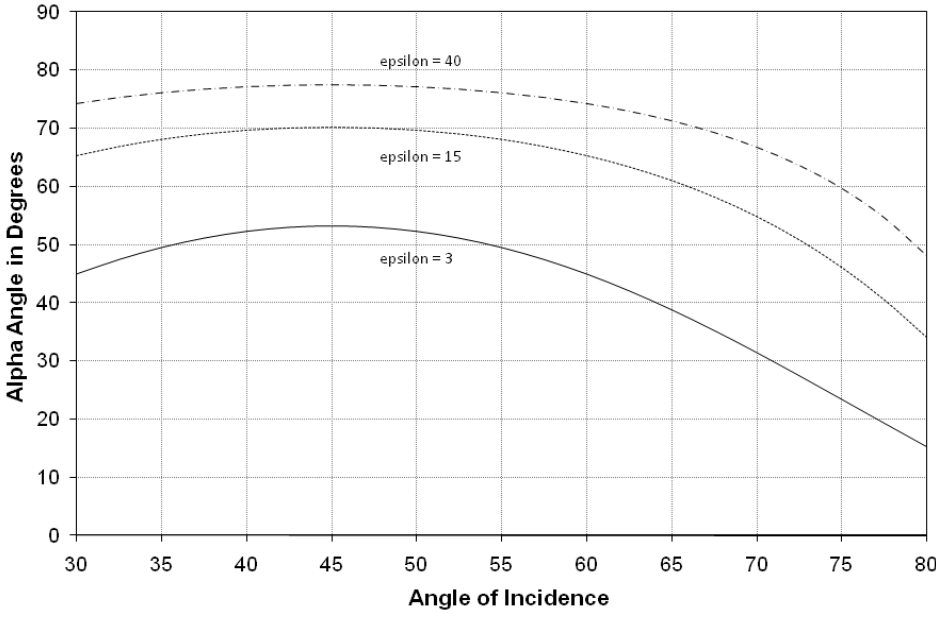


Fig. 3-6. Alpha angle as a function of angle of incidence for a double reflection from two dielectric surfaces with the same dielectric constant (epsilon). At angles of incidence equal to the Brewster angle of the dielectric surfaces, the alpha angle will be 45 deg.

To illustrate the meaning of the alpha angle in image data, Fig. 3.7 shows the L-band alpha angle image for San Francisco. The alpha angle for the ocean is mostly less than 45 deg, consistent with the expectation for a slightly rough surface. Note the increase in the alpha angle with increasing angle of incidence (the angle of incidence increases from left to right across the image) in the ocean, consistent with the predictions shown in Fig. 3-5. The vegetated areas all show alpha angles near 45 deg, consistent with dipole scattering. The urban areas consistently show alpha angles larger than 45 deg, consistent with the expectations for a double-reflection signal from a non-metallic surface.

Figure 3-8 shows the alpha angles at different frequencies for the Black Forest image. As in the case of San Francisco, the urban areas consistently show alpha angles larger than 45 deg. We also see the effect of frequency very clearly in the agricultural areas. At C-band, the alpha angle is mostly near 45 deg, while at L- and P-band, the values are closer to zero, indicating bare (or nearly bare) surfaces. The forested areas at L- and P-band show alpha angles either near 45 deg (especially at L-Band) or larger than 45 deg where there is an appreciable amount of double reflection signal. The C-band image, interestingly, shows the alpha angle in the forested areas to be less than 45 deg and generally less than the values observed in the agricultural areas where there is vegetation. The

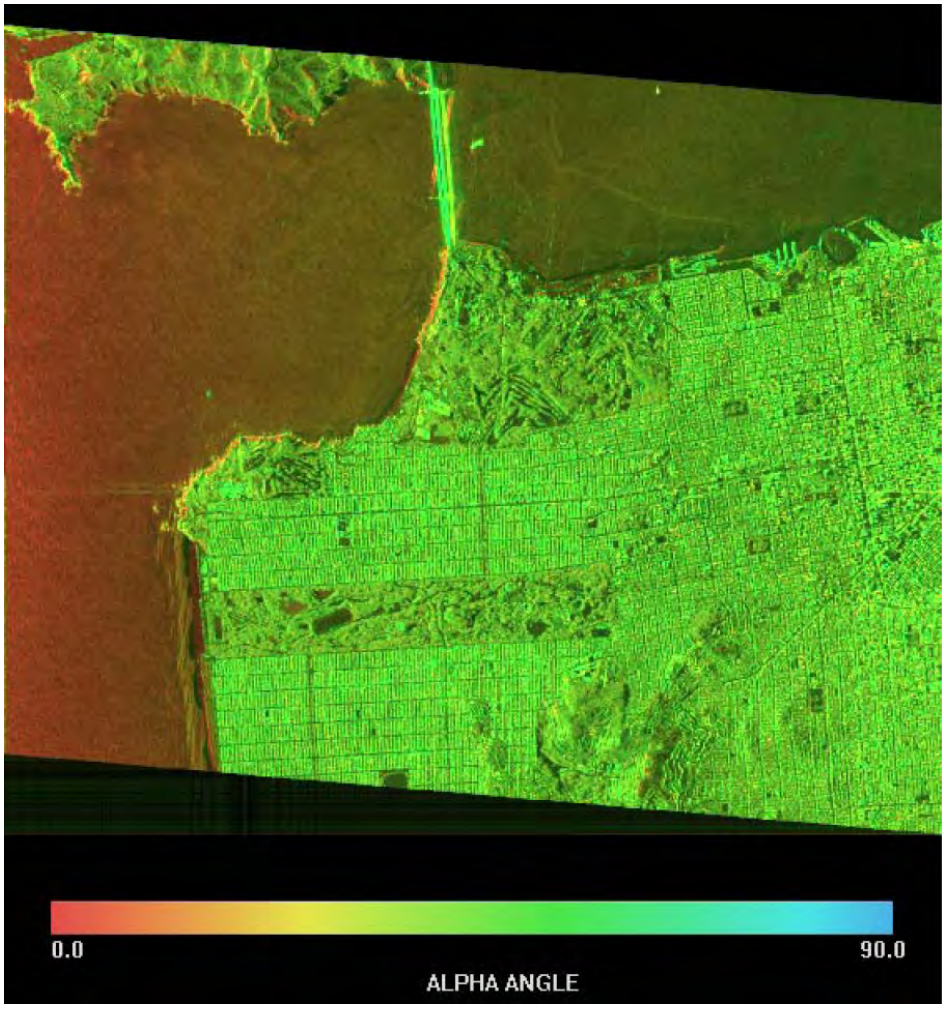


Fig. 3-7. Alpha angle for the San Francisco image. See the text for discussion.

explanation for this lies in the size of the branches relative to the radar wavelength. When we used the term “dipole” before, we could have substituted “cylinder that is thin compared to the radar wavelength.” Let us consider such a cylinder that is oriented vertically. In the thin cylinder limit, we will observe the scattering shown in Table 3-1. As the cylinder becomes thicker compared to the radar wavelength, we observe an increase in the HH term relative to the VV term. In fact, in the thick cylinder limit, the HH term will approach the VV term. As the cylinder radius increase relative to the radar wavelength, therefore, the co-polarized ratio will increase from zero to 1 in the limit of a thick cylinder. Equation (3.2-8) predicts that this will cause the alpha angle to decrease from the 45-deg range as the cylinders increase in thickness relative to

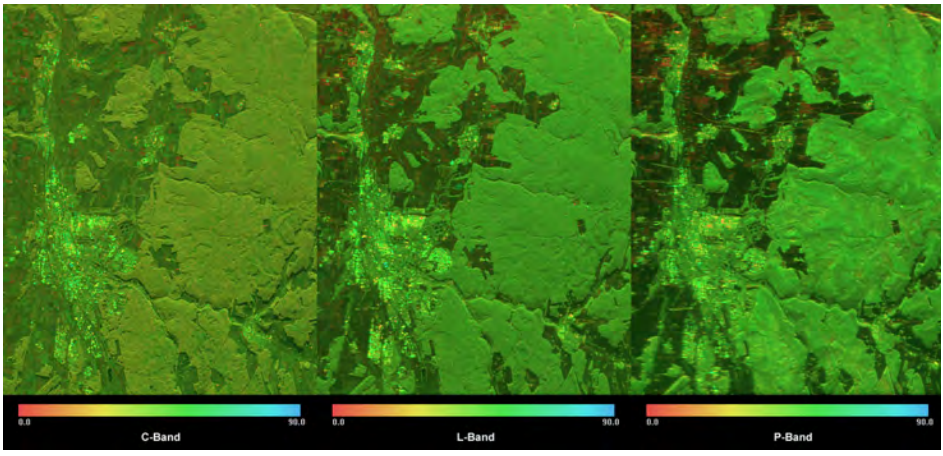


Fig. 3-8. Multi-frequency alpha angle for the Black Forest image. See the text for discussion.

the radar wavelength. In these images, the cylinders are of fixed size, but the radar wavelength changes. At C-band, the radius of a given cylinder will be larger relative to the radar wavelength than at L-band. We would, therefore, expect to see a lower alpha angle at C-band than at L-band.

3.3 Decomposition of Polarimetric Scattering

Earlier in this chapter we discussed the duality between representing a scatterer by its scattering matrix or by an equivalent vector containing the elements of the scattering matrix. In this section, we shall explore this concept in greater detail.

First, let us consider the case of a single scatterer that can be represented by its scattering matrix. We shall further restrict our discussion to the backscatter case where reciprocity is assumed to hold. If we assume that the scattering matrix was measured in the linear basis, we can write

$$\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix} = S_{hh}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + S_{hv}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + S_{vv}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.3-1}$$

or, in the equivalent vector form,

$$\begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vv} \end{pmatrix} = S_{hh}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + S_{hv}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + S_{vv}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.3-2}$$

In writing the scattering matrix in vector form, we used an orthogonal basis to express the elements of the scattering vector. It should be immediately obvious

that there are an infinite number of such orthogonal bases that one could choose to represent the scattering vector. Any combination of the form

$$\mathbf{S} = a_1\mathbf{p_1} + a_2\mathbf{p_2} + a_3\mathbf{p_3} \tag{3.3-3}$$

would be permissible as long as

$$\mathbf{p}_i \cdot \mathbf{p}_j = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}. \tag{3.3-4}$$

The logical question is then: Is any basis (other than the one used to make the measurement to begin with) that should be considered for such decomposition? The answer depends on the specific application, but the basis derived from the Pauli spin matrices has a nice intuitive interpretation. This basis is shown in Eq. (3.1-10) in vector form. In matrix form, the basis is

$$\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix} = \frac{a}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{b}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.3-5}$$

This decomposition was also used by Krogager (1993) [7] in his thesis. The first two terms involve only the co-polarized elements of the scattering matrix and can be interpreted as scattering by an odd number of reflections from a metallic structure and an even number of reflections from a metallic structure, respectively. The first matrix represents, therefore, scattering from a flat plate, a sphere, or a metallic trihedral corner reflector. The second matrix represent scattering from a metallic dihedral corner reflector.

The third matrix can be interpreted in different ways. Since it only involves the cross-polarized component of the scattering matrix, it is usually interpreted as indicating the amount of random scattering. While this interpretation certainly has some merit in the practical sense that scattering from vegetated areas usually shows a large amount of cross-polarized return, it is not strictly correct from a theoretical point of view. The third matrix, as was pointed out in Chapter 2, is also the scattering matrix of a dihedral corner reflector rotated by 45 deg about the line of sight. The resulting scattered energy is fully polarized, but the polarization vector has been rotated. Admittedly, this is a special case. In most practical applications, a large cross-polarized component is also typically associated with significant depolarization of the scattered energy.

To illustrate the usefulness of this approach for interpreting scattering, we display a color image of San Francisco in Fig. 3-9 in which we assigned the blue color as the magnitude of *a* in Eq. (3.3-5), the red color as the magnitude of *b*, and the green color as the magnitude of *c*. This image shows that the urban area shows a large fraction of the scattering in the red color, which corresponds

to the dihedral component. On the other hand, the ocean shows much more blue on the left, consistent with the single scattering mechanism. The vegetated areas, on the other hand, show significant cross-polarized return. Note the interesting change in color in the ocean from the left to the right in the image, where there is significantly more red visible in the ocean. The explanation for this lies in the fact that for a dielectric surface like the ocean, the co-polarization ratio is a function of both the dielectric constant and the angle of incidence. Figure 3-10 shows the expected co-polarization ratio for a surface with a dielectric constant of 81. The figure also shows the co-polarized ratio to be less than 1, especially at the larger angles of incidence. The Pauli basis, however, forces the HH and VV terms to be equal. Therefore, if the co-polarized ratio is less than 1, a dihedral component is needed to explain the difference between the HH and VV terms.

Fig. 3-9. Color overlay of San Francisco displaying $|S_{hh}+S_{vv}|$ in blue, $|S_{hh}-S_{vv}|$ in red, and $|S_{hv}|$ in green. These three colors are the magnitudes of the scattering matrix elements when they are expressed in the Pauli basis.

The smaller the co-polarization ratio, the stronger the dihedral component required to explain the difference. This is shown as the dashed curve in Fig. 3-10.

It is important to appreciate that the first-order small perturbation model only includes single scattering terms. The fact that the ocean scattering is interpreted to have a significant amount of double reflections occurs only because of the basis under which we have chosen to interpret the scattering. This basis forces the HH and VV components to be the same for the "single scattering" term, leading to this interpretation. This is a fundamental issue with many of the target decomposition schemes proposed in the literature. All of these schemes try to interpret scattering based on an assumption of an underlying basis. If the scattering fits the basis, the interpretation is obviously appropriate. If the basis is not consistent with the actual scattering, however, the interpretation should be modified to take this fact into account. Nevertheless, the Pauli basis provides a good general purpose framework for interpreting polarimetric radar images (as the San Francisco example shows).

Figure 3-11 shows the three-frequency Pauli decompositions for the Black Forest image. These images show a consistent interpretation of the scattering from what we discussed before. The scattering from the randomly oriented

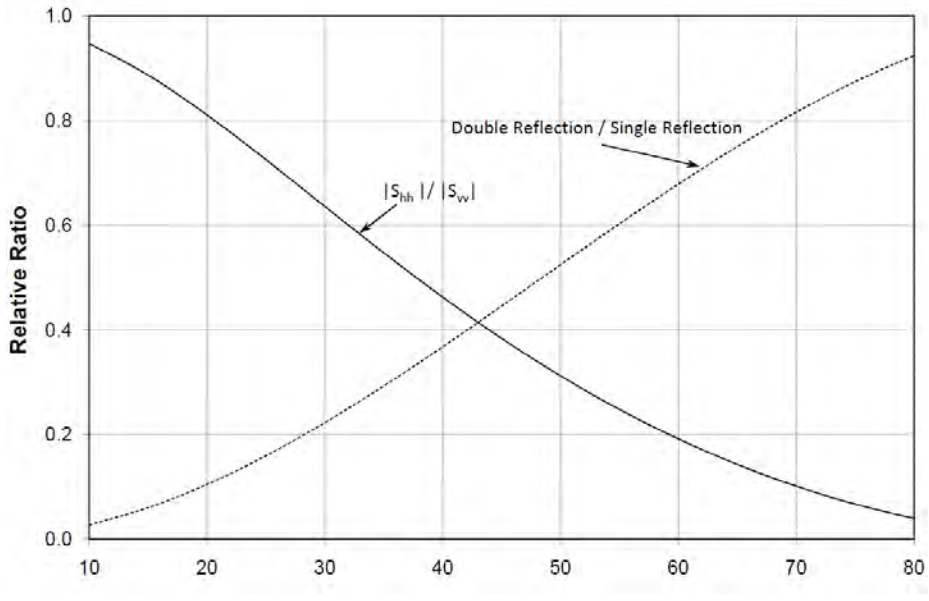


Fig. 3-10. Expected co-polarization ratio and the ratio of double reflection scattering to single reflection scattering as a function of the angle of incidence for a surface with a dielectric constant of 81. The values were calculated using the first-order small perturbation model and assuming the Pauli basis to calculate double and single reflection components.

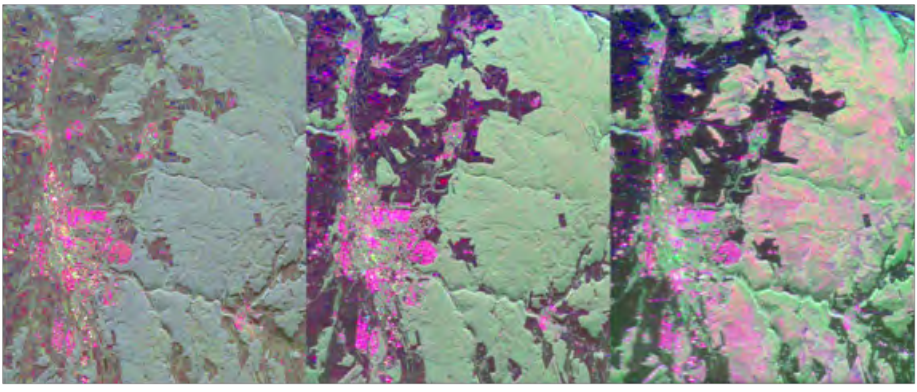


Fig. 3-11. Pauli basis color overlays for the Black Forest image. The color scheme is the same as that in Fig. 3-9. The C-band image is on the left, the L-Band in the middle, and the P-Band on the right.

vegetation is relatively strong in the cross-polarized return. The urban areas, on the other hand, are dominated by dihedral-type reflections at all wavelengths. The increased amount of penetration at P-band shows, relatively speaking, a stronger double reflection signal in much of the forest. The C-band signals interact more with the shorter agricultural crops than the longer wavelengths, resulting in increased cross-polarized returns in those areas.

The discussion so far was concerned with the decomposition of the scattering matrix, or its associated vector form, into orthogonal components. As pointed out before, there is potentially an infinite set of bases we can use for this decomposition. A more important question is the following: What about the case where we have an average covariance matrix? What is the most appropriate way to decompose this observed scattering into simpler parts? We shall discuss this in detail in the next section.

3.3.1 Scattering Decomposition in the Incoherent Case Using Orthonormal Bases

As mentioned in the beginning of this chapter, after multi-looking to reduce speckle, we can write the average covariance matrix as

$$\langle [\mathbf{C}] \rangle = \frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} [\mathbf{C}]_{ij} \ , \tag{3.3-6}$$

where the two sums indicate averaging in the range and azimuth directions, respectively. We shall restrict our discussion to the backscatter case, where the individual covariance matrices are defined as

$$[\mathbf{C}] = \begin{pmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{pmatrix} \begin{pmatrix} S_{hh} & \sqrt{2}S_{hv} & S_{vv} \end{pmatrix}^* = \begin{pmatrix} S_{hh}S_{hh}^* & \sqrt{2}S_{hh}S_{hv}^* & S_{hh}S_{vv}^* \\ \sqrt{2}S_{hv}S_{hh}^* & 2S_{hv}S_{hv}^* & \sqrt{2}S_{hv}S_{vv}^* \\ S_{vv}S_{hh}^* & \sqrt{2}S_{vv}S_{hv}^* & S_{vv}S_{vv}^* \end{pmatrix}. \quad (3.3\text{-}7)$$

The covariance matrix is Hermitian. The matrix, therefore, contains, at most, three independent complex numbers and three real numbers, for a total of nine real numbers. The scattering matrix, on the other hand, contains, at most, three complex numbers. If we remove an absolute phase number from one of the elements of the scattering matrix, we would be left with one real and two complex numbers (a total of five real numbers). There must, therefore, be at least four relations between the elements of the covariance matrix of a single scatterer. These are

$$\begin{aligned} C_{00}C_{11} - C_{01}C_{10} &= 2S_{hh}S_{hh}^*S_{hv}S_{hv}^* - \sqrt{2}S_{hh}S_{hv}^*\sqrt{2}S_{hv}S_{hh}^* = 0 \\ C_{00}C_{22} - C_{02}C_{20} &= S_{hh}S_{hh}^*S_{vv}S_{vv}^* - S_{hh}S_{vv}^*S_{vv}S_{hh}^* = 0 \\ C_{11}C_{22} - C_{12}C_{21} &= 2S_{hv}S_{hv}^*S_{vv}S_{vv}^* - \sqrt{2}S_{hv}S_{vv}^*\sqrt{2}S_{vv}S_{hv}^* = 0 \\ C_{00}C_{11}C_{22} - C_{01}C_{02}^*C_{12} &= 2S_{hh}S_{hh}^*S_{hv}S_{hv}^*S_{vv}S_{vv}^* - \sqrt{2}S_{hh}S_{hv}^*S_{vv}S_{hh}^*\sqrt{2}S_{hv}S_{vv}^* = 0 \end{aligned} \quad . \quad (3.3\text{-}8)$$

Once we perform the averaging process shown in Eq. (3.3-6) during the multi-looking process, these relations will no longer hold. Instead, the equal signs should be replaced with greater than or equal to signs. In fact, that is simply a statement of the Cauchy-Schwarz inequality applied to complex numbers. Therefore, in general,

$$\begin{aligned} C_{00}C_{11} - C_{01}C_{10} &\geq 0 \\ C_{00}C_{22} - C_{02}C_{20} &\geq 0 \\ C_{11}C_{22} - C_{12}C_{21} &\geq 0 \\ C_{00}C_{11}C_{22} - C_{01}C_{02}^*C_{12} &\geq 0 \end{aligned} \quad . \quad (3.3\text{-}9)$$

Unless equality holds in each of the four cases, we cannot find an equivalent scattering matrix to fully represent the scattering described by the covariance matrix as shown in Eq. (3.3-7). A reasonable question then is whether we can find a set of scattering matrices that, when transformed to their covariance matrices, could be added to fully describe the observed covariance matrix. To answer this question, consider the decomposition proposed by Cloude (1992) [1]:

$$\langle [\mathbf{C}] \rangle = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^{\dagger} . \quad (3.3\text{-}10)$$

This decomposition is unique. That is, since the eigenvectors of the covariance matrix are orthogonal, they form a natural basis in which to express the scattering. There is one potential problem with this decomposition. Since the decomposition is performed for each covariance matrix (that is, on a pixel-by-pixel basis in an image), the resulting basis for the decomposition changes from pixel to pixel. This could make it more difficult to compare the meaning of a given eigenvalue in different areas without also looking at the associated eigenvectors that form the basis. To illustrate what we mean, consider the special case of the covariance matrix of terrain with reflection symmetry. In that case, the covariance matrix has the special form (Borgeaud et al., 1985) [8]:

$$\left\langle [\mathbf{C}] \right\rangle = \begin{pmatrix} \xi & 0 & \rho \\ 0 & \eta & 0 \\ \rho^* & 0 & \zeta \end{pmatrix}, \tag{3.3-11}$$

where

$$\begin{aligned} \xi &= \left\langle S_{hh} S_{hh}^* \right\rangle \\ \rho &= \left\langle S_{hh} S_{vv}^* \right\rangle \\ \eta &= 2\left\langle S_{hv} S_{hv}^* \right\rangle \\ \zeta &= \left\langle S_{vv} S_{vv}^* \right\rangle \end{aligned}. \tag{3.3-12}$$

The parameters ξ, η, ζ and ρ all depend on the size, shape, and electrical properties of the scatterers, as well as their statistical distribution. It is easily shown that the eigenvalues of $\left\langle [\mathbf{C}] \right\rangle$ are

$$\begin{aligned} \lambda_1 &= \frac{1}{2}\left\{ \xi + \zeta + \sqrt{(\xi - \zeta)^2 + 4|\rho|^2} \right\} \\ \lambda_2 &= \frac{1}{2}\left\{ \xi + \zeta - \sqrt{(\xi - \zeta)^2 + 4|\rho|^2} \right\} \\ \lambda_3 &= \eta \end{aligned}. \tag{3.3-13}$$

All these are real numbers, as expected for a Hermitian matrix. It follows from inspection that the first and third eigenvalues are positive. To show that the second eigenvalue is also positive, note that, from the second relationship in Eq. (3.3-10),

$$\xi\zeta \geq |\rho|^2 \quad \Rightarrow \quad (\xi-\zeta)^2 + 4|\rho|^2 \leq (\xi+\zeta)^2 . \tag{3.3-14}$$

Hence, the second eigenvalue is also positive.

The corresponding three eigenvectors are

$$\mathbf{k}_1 = \sqrt{\frac{\left[\zeta-\xi+\sqrt{\Delta}\right]^2}{\left[\zeta-\xi+\sqrt{\Delta}\right]^2 + 4|\rho|^2}} \begin{pmatrix} 2\rho \Big/ \left[\zeta-\xi+\sqrt{\Delta}\right] \\ 0 \\ 1 \end{pmatrix}$$

$$\mathbf{k}_2 = \sqrt{\frac{\left[\zeta-\xi-\sqrt{\Delta}\right]^2}{\left[\zeta-\xi-\sqrt{\Delta}\right]^2 + 4|\rho|^2}} \begin{pmatrix} 2\rho \Big/ \left[\zeta-\xi-\sqrt{\Delta}\right] \\ 0 \\ 1 \end{pmatrix} . \tag{3.3-15}$$

$$\mathbf{k}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

In these expressions, we used the shorthand notation

$$\Delta = (\zeta-\xi)^2 + 4|\rho|^2 . \tag{3.3-16}$$

We note that Δ is always positive. Also note that we can write the ratio of the first elements of the first two eigenvectors as

$$\frac{k_{11}}{k_{21}} = -\sqrt{\frac{\left\{\left[\zeta-\xi+\sqrt{\Delta}\right]^2\right\}\left\{\left[\zeta-\xi-\sqrt{\Delta}\right]^2 + 4|\rho|^2\right\}}{\left\{\left[\zeta-\xi-\sqrt{\Delta}\right]^2\right\}\left\{\left[\zeta-\xi+\sqrt{\Delta}\right]^2 + 4|\rho|^2\right\}}} \frac{\left[\zeta-\xi-\sqrt{\Delta}\right]^2}{4|\rho|^2}, \tag{3.3-17}$$

which is always negative. This means that the first two eigenvectors represent scattering matrices that can be interpreted in terms of odd and even numbers of reflections. Without looking explicitly at the eigenvectors, however, we will not know which eigenvalue to associate with which scattering mechanism.

Figure 3-12 illustrates this with an example. On the left, we display the image of San Francisco previously discussed with the three eigenvalues as defined in Eq. (3.3-13) colored blue for the first eigenvalue, red for the second, and green for the third. Note that the third eigenvalue and eigenvector are identical to the third Pauli element encountered before. The ocean is dominated by a blue color, indicating that the scattering is dominated by whatever scattering mechanism

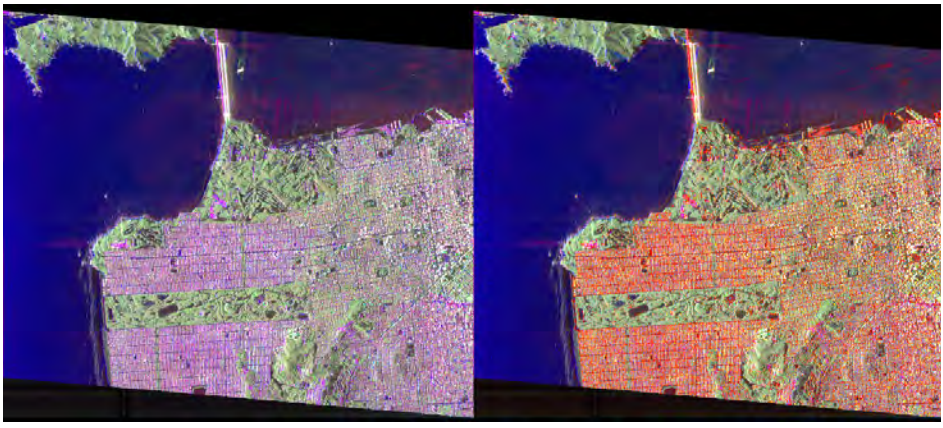


Fig. 3-12. Two color overlays for the image of San Francisco. In the image on the left, colors are assigned to the three eigenvalues without further examining the eigenvectors. The first eigenvalue is colored blue, the second red, and the third green. In the image on the right, the co-polarized phase of the first eigenvector is used to determine the color of the first eigenvalue. If the phase is more consistent with odd numbers of reflections, the eigenvalue is colored blue. Otherwise it is colored red. The color of the second eigenvector is determined using the fact that the two co-polarized phases of the first two eigenvectors are 180 deg different.

the first eigenvector represents. The urban areas, however, are also dominated by a blue color, indicating the same. Vegetated areas consistently show a green color, indicating that depolarization is, relatively speaking, high. On the right panel of Fig. 3-12, we show the same image, but this time we examined the eigenvectors for each pixel. If the co-polarized phase is in the range [-90°;90°], we assign a blue color to the eigenvalue corresponding to that eigenvector. If the co-polarized phase is in the range [90°;270°], we assign a red color. We already showed that the phase difference between the first two eigenvectors is 180 deg; once we identified the color for the first eigenvector, therefore, the color for the second is automatically determined. This image on the right is quite different from the one on the left. In particular, the urban area in the image on the right is now dominated by red, indicating scattering more consistent with double reflections.

Comparing the results on the right in Fig. 3-12 with those in Fig. 3-9, we note many similarities between the Pauli basis display and the eigenvalue display. The main difference is that the eigenvalue display colors appear more pure. Also, note that the gradient from blue to red in the ocean is not as pronounced in the image on the right in Fig. 3-12. The reason is that while the Pauli basis insists that the co-polarized terms have the same amplitude, the eigenvectors do not (see Eq. (3.3-15)).

A natural question is then: Is there a preferred basis in which to do the decomposition? In general, we could also choose to write the average covariance matrix in the following form

$$\begin{aligned} \langle [\mathbf{C}] \rangle &= \sum_{i=1}^{3} w_i \mathbf{p}_i \mathbf{p}_i^{\dagger} \\ \mathbf{p}_i \mathbf{p}_j^{\dagger} &= \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \end{aligned} . \tag{3.3-18}$$

The Pauli and the eigenvalue bases are two special cases of this decomposition. The answer to this question is not obvious. It depends largely on the goal of the analysis. As our discussion shows, the decomposition is not the hard part; interpreting the results is. This desire to be able to interpret the results of such decomposition is what led many researchers to propose decompositions based on specific models, rather than orthogonal bases as discussed so far. We shall look at this in more detail in the next section.

3.3.2 Model Based Scattering Decomposition in the Incoherent Case

The basic idea behind model-based decompositions is to hypothesize that the measured covariance matrix can be modeled as the combination of a number of individual matrices representing scattering as predicted by models. In this section we shall examine a number of different model-based decomposition schemes in more detail.

3.3.2.1 Freeman-Durden Three-Component Scattering Decomposition**.** For vegetated terrain, we could hypothesize that the dominant scattering mechanisms might be direct scattering from randomly oriented branches, plus double reflections from the ground/trunk combination, plus direct (although attenuated) scattering from the underlying ground surface. This is the basic idea behind the three-component scattering decomposition proposed by Freeman and Durden (1998) [9]. This decomposition can be written as follows

$$\langle [\mathbf{C}] \rangle = f_s \left[\mathbf{C}_{ground} \right] + f_d \left[\mathbf{C}_{trunk-ground} \right] + f_v \left[\mathbf{C}_{branches} \right], \tag{3.3-19}$$

with

$$\left[\mathbf{C}_{ground}\right]=\begin{pmatrix}|\beta|^2 & 0 & \beta\\ 0 & 0 & 0\\ \beta^* & 0 & 1\end{pmatrix}$$
$$\left[\mathbf{C}_{trunk-ground}\right]=\begin{pmatrix}|\alpha|^2 & 0 & \alpha\\ 0 & 0 & 0\\ \alpha^* & 0 & 1\end{pmatrix} \quad . \tag{3.3-20}$$
$$\left\langle\left[\mathbf{C}_{branches}\right]\right\rangle=\begin{pmatrix}1 & 0 & 1/3\\ 0 & 2/3 & 0\\ 1/3 & 0 & 1\end{pmatrix}$$

The matrix representing branch scattering assumes that the branches are thin compared to the radar wavelength and that the branches are uniformly randomly oriented. We previously discussed this case in Chapter 2. From Eq. (3.3-19) one can then derive the following four equations:

$$\begin{aligned}\left\langle|S_{hh}|^2\right\rangle &= f_s|\beta|^2+f_d|\alpha|^2+f_v\\ \left\langle|S_{vv}|^2\right\rangle &= f_s+f_d+f_v\\ \left\langle|S_{hv}|^2\right\rangle &= f_v/3\\ \left\langle\left|S_{hh}S_{vv}^*\right|\right\rangle &= f_s\beta+f_d\alpha+f_v/3\end{aligned} \quad . \tag{3.3-21}$$

Freeman and Durden (1998) [9] point out that there are four equations and five unknowns. They then make the following crucial suggestion: since neither the ground reflection nor the double reflection terms add to the predicted cross-polarized return, they can use the measured cross-polarized return to solve for the parameter f_v. They continue to suggest that the volume contribution can then be subtracted from the measured matrix before solving for the other terms. In other words, we can write Eq. (3.3-19) as

$$\left\langle[\mathbf{C}]\right\rangle-\begin{pmatrix}3\left\langle|S_{hv}|^2\right\rangle & 0 & \left\langle|S_{hv}|^2\right\rangle\\ 0 & 2\left\langle|S_{hv}|^2\right\rangle & 0\\ \left\langle|S_{hv}|^2\right\rangle & 0 & 3\left\langle|S_{hv}|^2\right\rangle\end{pmatrix}=f_s\left[\mathbf{C}_{ground}\right]+f_d\left[\mathbf{C}_{trunk-ground}\right]. \tag{3.3-22}$$

Once the subtraction has been done, there are three remaining equations in four unknowns, as follows:

$$\begin{aligned} \left\langle |S_{hh}|^2 \right\rangle' &= f_s |\beta|^2 + f_d |\alpha|^2 \\ \left\langle |S_{vv}|^2 \right\rangle' &= f_s + f_d \\ \left\langle |S_{hh} S_{vv}^*| \right\rangle' &= f_s \beta + f_d \alpha \end{aligned} \quad . \tag{3.3-23}$$

The primes on the left serve to remind us that these are the quantities after the volume-scattering contribution have been subtracted. The phase of the remaining co-polarized component is then used to fix either α or β. The argument is that if the residual co-polarized phase is closer to zero than to π, surface scattering dominates and we should solve for β explicitly. Therefore, we set $\alpha = -1$ (indicating a double reflection) and solve for the remaining parameters. On the other hand, if the residual co-polarized phase is closer to π than to zero, double reflection scattering dominates and we should solve for α explicitly. In this case, we set $\beta = 1$ (indicating a single reflection) and solve for the remaining parameters.

Figure 3-13 shows a color overlay image of the three contributions calculated using the Freeman and Durden (1998) [9] model for the image of the Black Forest at L-band. The image on the left shows the relative strength of the three scattering mechanisms in the color code indicated. Overall, the image clearly shows volume scattering to dominate in the vegetated areas, double reflections to dominate in the urban areas, and some of the agricultural areas to show surface scattering. At this qualitative level, the results appear consistent with our expectations. The image is also similar to the Pauli basis image shown in the middle of Fig. 3-11.

However, a deeper examination shows a significant flaw in this decomposition. The image on the right shows the results of an analysis of the eigenvalues of the matrix on the left in Eq. (3.3-22) after we subtracted the vegetation contribution from the original data. If any of the eigenvalues are negative, we blanked out the pixel in the image on the right in Fig. 3-13. This surprising result shows that the majority of the pixels in the vegetated area end up with negative eigenvalues after we subtracted the scattering from the vegetation, as suggested by Freeman and Durden (1998) [9]. However, this is exactly where we expect the scattering model for scattering from the vegetation to be most applicable! We shall show later that this surprising result is the consequence of assigning all the cross-polarized return to the vegetation scattering.

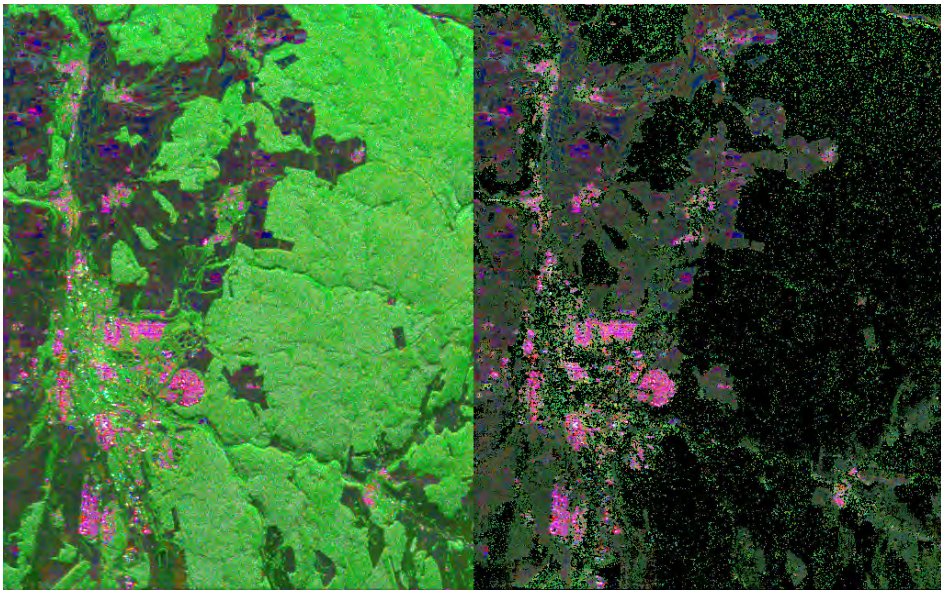


Fig. 3-13. Two color overlays displaying the results of the Freeman and Durden decomposition applied to the L-band image of the Black Forest. The image on the left displays the surface contribution in blue, the double reflection contribution in red, and the vegetation contribution in green. The image on the right shows only those pixels with non-negative eigenvalues once the vegetation contribution has been subtracted. See the text for more discussion.

3.3.2.2 Four-Component Model Proposed by Yamaguchi et al. The three-component model described above assumes that the terrain has reflection symmetry by ignoring the terms in the covariance matrix that involves products of co-polarized and cross-polarized terms. While this assumption seems to be valid for most types of terrain in the sense that these components of the covariance matrix are much smaller than the others, one cannot always assume that reflection symmetry will hold. Yamaguchi et al. (2005) [10] recognized this fact and proposed an extension of the three-component model to include a term that would account for the non-zero products of co-polarized and cross-polarized terms. Their decomposition is written as

$$\langle[\mathbf{C}]\rangle = f_s\left[\mathbf{C}_{ground}\right] + f_d\left[\mathbf{C}_{trunk-ground}\right] + f_v\left[\mathbf{C}_{branches}\right] + f_c\left[\mathbf{C}_{helix}\right], \quad (3.3\text{-}24)$$

with the matrix representing the helix scattering taking one of the following two forms:

$$\left[\mathbf{C}_{r-helix}\right] = \frac{1}{4}\begin{pmatrix} 1 & j\sqrt{2} & -1 \\ -j\sqrt{2} & 2 & j\sqrt{2} \\ -1 & -j\sqrt{2} & 1 \end{pmatrix}; \quad \left[\mathbf{C}_{l-helix}\right] = \frac{1}{4}\begin{pmatrix} 1 & -j\sqrt{2} & -1 \\ j\sqrt{2} & 2 & -j\sqrt{2} \\ -1 & j\sqrt{2} & 1 \end{pmatrix}. \quad (3.3\text{-}25)$$

Note that these two matrices both predict the cross products $\left\langle S_{hh}S_{hv}^{*}\right\rangle$ and $\left\langle S_{hv}S_{vv}^{*}\right\rangle$ to be purely imaginary numbers. This might not be the case in observed data. To get around this, Yamaguchi et al. (2005) [10] recommend using only the imaginary portion of the observed cross-products in the decomposition. Furthermore, the model matrices predict that $\left\langle S_{hh}S_{hv}^{*}\right\rangle = \left\langle S_{hv}S_{vv}^{*}\right\rangle$. Again, this might not be exactly what is observed. Therefore, they recommend using

$$\frac{f_c}{4} = \frac{1}{2}\left|\mathrm{Im}\left\{\left\langle S_{hh}S_{hv}^{*}\right\rangle + \left\langle S_{hv}S_{vv}^{*}\right\rangle\right\}\right| \quad . \text{(3.3-26)}$$

Finally, to decide which matrix in Eq. (3.3-25) to use, they propose

$$\text{if} \quad \begin{cases} \mathrm{Im}\left\{\left\langle S_{hh}S_{hv}^{*}\right\rangle + \left\langle S_{hv}S_{vv}^{*}\right\rangle\right\} > 0 \quad \Rightarrow \quad \text{use } \left[\mathbf{C}_{r-helix}\right] \\ \mathrm{Im}\left\{\left\langle S_{hh}S_{hv}^{*}\right\rangle + \left\langle S_{hv}S_{vv}^{*}\right\rangle\right\} < 0 \quad \Rightarrow \quad \text{use } \left[\mathbf{C}_{l-helix}\right] \end{cases} \quad . \qquad (3.3\text{-}27)$$

Assuming a uniformly oriented canopy, and writing out the covariance matrices in Eq. (3.3-24), we find the following five equations with six unknowns:

$$\begin{aligned} \left\langle \left|S_{hh}\right|^2 \right\rangle &= f_s \left|\beta\right|^2 + f_d \left|\alpha\right|^2 + \frac{3}{8} f_v + \frac{1}{4} f_c \\ \left\langle \left|S_{hv}\right|^2 \right\rangle &= \frac{1}{8} f_v + \frac{1}{4} f_c \\ \left\langle \left|S_{vv}\right|^2 \right\rangle &= f_s + f_d + \frac{3}{8} f_v + \frac{1}{4} f_c \\ \left\langle S_{hh}S_{vv}^{*}\right\rangle &= f_s\beta + f_d\alpha + \frac{1}{8} f_v - \frac{1}{4} f \\ \frac{1}{2}\left|\mathrm{Im}\left\{\left\langle S_{hh}S_{hv}^{*}\right\rangle + \left\langle S_{hv}S_{vv}^{*}\right\rangle\right\}\right| &= \frac{f_c}{4} \end{aligned} \quad . \qquad (3.3\text{-}28)$$

The unknown quantities are $\alpha, \beta, f_s, f_d, f_v$ and f_c. To solve for these unknowns, note that the last expression allows us to find f_c as follows

$$f_c = 2\left|\mathrm{Im}\left\{\left\langle S_{hh}S_{hv}^{*}\right\rangle + \left\langle S_{hv}S_{vv}^{*}\right\rangle\right\}\right| . \qquad (3.3\text{-}29)$$

We can then use the second equation in Eq. (3.3-28) to find f_v, as follows:

$$f_v = 8\left(\left\langle |S_{hv}|^2 \right\rangle - \frac{f_c}{4}\right). \tag{3.3-30}$$

Once these two unknowns are determined, the contributions of the helix scattering and the volume scattering can be subtracted from the observed covariance matrix. Solving for the remaining unknowns then follows the same procedure as outlined by Freeman and Durden (1998) [9] as discussed in the previous section. Specifically,

$$f_s\left[\mathbf{C}_{ground}\right] + f_d\left[\mathbf{C}_{trunk-ground}\right] = \langle[\mathbf{C}]\rangle - \frac{f_v}{8}\begin{pmatrix}3 & 0 & 1\\ 0 & 2 & 0\\ 1 & 0 & 3\end{pmatrix} - \frac{f_c}{4}\begin{pmatrix}1 & \pm j\sqrt{2} & -1\\ \mp j\sqrt{2} & 2 & \pm j\sqrt{2}\\ -1 & \mp j\sqrt{2} & 1\end{pmatrix}. \tag{3.3-31}$$

Yamaguchi et al. (2005) [10] went one step further than the Freeman and Durden (1998) [9] decomposition in that they recognized that not all vegetated terrains are well represented by a uniformly oriented canopy. In some cases, the orientations are preferentially vertical, while in others it might be preferentially horizontal. For these cases, Yamaguchi et al. [10] propose to use cosine-squared distributions around either vertical or horizontal directions, which leads to the following covariance matrices

$$\left[\mathbf{C}_{branches\_h}\right] = \frac{1}{15}\begin{pmatrix}8 & 0 & 2\\ 0 & 4 & 0\\ 2 & 0 & 3\end{pmatrix}; \quad \left[\mathbf{C}_{branches\_v}\right] = \frac{1}{15}\begin{pmatrix}3 & 0 & 2\\ 0 & 4 & 0\\ 2 & 0 & 8\end{pmatrix}. \tag{3.3-32}$$

In their decomposition algorithm, Yamaguchi et al. [10] use the ratio of the VV to HH cross-sections to decide which canopy model to use. If the ratio of VV to HH power is less than –2 dB, they use the cosine-squared distribution around the horizontal direction. If the ratio is between - 2 dB and + 2 dB, they use the uniform distribution. When the ratio is larger than +2 dB, they use the cosine-squared distribution around the vertical direction. In the case where we use a distribution that is preferentially horizontal, the expressions in Eq. (3.3-28) need to be modified as follows:

$$\begin{aligned}
\left\langle |S_{hh}|^2 \right\rangle &= f_s |\beta|^2 + f_d |\alpha|^2 + \frac{8}{15} f_v + \frac{1}{4} f_c \\
\left\langle |S_{hv}|^2 \right\rangle &= \frac{2}{15} f_v + \frac{1}{4} f_c \\
\left\langle |S_{vv}|^2 \right\rangle &= f_s + f_d + \frac{3}{8} f_v + \frac{1}{4} f_c \\
\left\langle S_{hh} S_{vv}^* \right\rangle &= f_s \beta + f_d \alpha + \frac{3}{15} f_v - \frac{1}{4} f \\
\frac{1}{2} \left| \mathrm{Im} \left\{ \left\langle S_{hh} S_{hv}^* \right\rangle + \left\langle S_{hv} S_{vv}^* \right\rangle \right\} \right| &= \frac{f_c}{4}
\end{aligned} \quad . \qquad (3.3\text{-}33)$$

The helix component is still determined by Eq. (3.3-29), but we now have to modify how we determine the volume component, as follows:

$$f_v = \frac{15}{2} \left(\left\langle |S_{hv}|^2 \right\rangle - \frac{f_c}{4} \right). \qquad (3.3\text{-}34)$$

The surface and double reflection components are then determined using the Freeman and Durden (1998) [9] approach after the following subtraction

$$f_s \left[\mathbf{C}_{ground} \right] + f_d \left[\mathbf{C}_{trunk-ground} \right] = \langle [\mathbf{C}] \rangle - \frac{f_v}{15} \begin{pmatrix} 8 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 3 \end{pmatrix} - \frac{f_c}{4} \begin{pmatrix} 1 & \pm j\sqrt{2} & -1 \\ \mp j\sqrt{2} & 2 & \pm j\sqrt{2} \\ -1 & \mp j\sqrt{2} & 1 \end{pmatrix}. \qquad (3.3\text{-}35)$$

In the case where the ratio of VV to HH is larger than +2 dB, we assume a preferentially vertical orientation, the helix component is determined as before, and the volume component is determined by Eq. (3.3-34). The surface and double reflection components are determined after the following subtraction

$$f_s \left[\mathbf{C}_{ground} \right] + f_d \left[\mathbf{C}_{trunk-ground} \right] = \langle [\mathbf{C}] \rangle - \frac{f_v}{15} \begin{pmatrix} 3 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 8 \end{pmatrix} - \frac{f_c}{4} \begin{pmatrix} 1 & \pm j\sqrt{2} & -1 \\ \mp j\sqrt{2} & 2 & \pm j\sqrt{2} \\ -1 & \mp j\sqrt{2} & 1 \end{pmatrix}. \qquad (3.3\text{-}36)$$

The Yamaguchi et al. (2005) [10] algorithm can be summarized as follows:

1) Estimate the helix scattering component using Eq. (3.3-29).
2) Depending on the ratio of VV to HH, use Eq. (3.3-30) (-2 dB < VV/HH < 2 dB) or Eq. (3.3-34) (VV/HH < - 2 dB or VV/HH > + 2 dB) to estimate the volume component.

3) Subtract the helix and volume components from the observation using Eq. (3.3-31) when - 2 dB < VV/HH < 2 dB, Eq. (3.3-35) when VV/HH < –2 dB or Eq. (3.3-35) when VV/HH > +2 dB.

4) Use Eq. (3.3-23) and the process described by Freeman and Durden (1998) [9] to estimate the strength of the surface and double reflection scattering terms.

Figure 3-14 shows the results of applying the Yamaguchi et al. (2005) [10] decomposition to the L-band image of the Black Forest. The image on the left shows the relative strength of the volume, double bounce, and surface scattering with the same color scheme that we used for the Freeman and Durden (1998) [9] decomposition in Fig. 3-13. Also shown in the figure on the right are the pixels with negative eigenvalues after the helix and volume components have been subtracted. Comparing Fig. 3-14 and Fig. 3-13, we note that the qualitative results of the decompositions are very similar. This is not surprising, since the helix components are typically much smaller than the others. The major difference between the two decomposition methods lies in the number of pixels with negative eigenvalues. The Yamaguchi et al. (2005) [10] decomposition results show significantly fewer pixels with negative eigenvalues. It should be pointed out that Yamaguchi et al. (2005) [10]

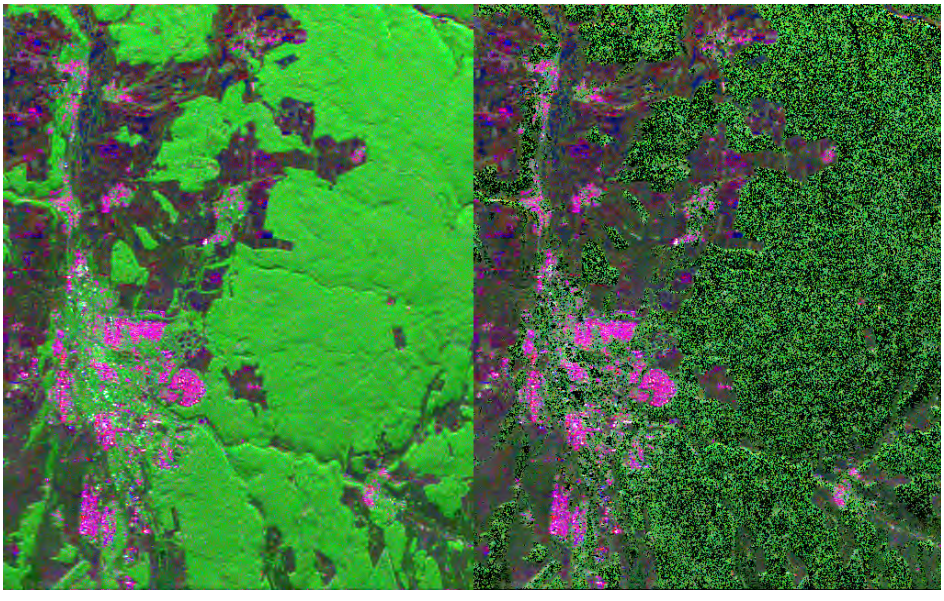


Fig. 3-14. Two color overlays displaying the results of the Yamaguchi et al. [10] decomposition applied to the L-band image of the Black Forest. The image on the left displays the surface contribution in blue, the double reflection contribution in red, and the vegetation contribution in green. The image on the right shows only those pixels with non-negative eigenvalues once the vegetation contribution has been subtracted. See the text for more discussion.

recognized that areas where the HV returns exceed approximately half that of HH or VV could lead to negative powers in the decomposition. Their solution was to apply a slightly different algorithm to pixels that show this high relative value of HV. In the next section we shall describe a systematic way of determining the level of the volume scattering while still ensuring that no negative powers will result.

Figure 3-15 shows the relative strength of the helix term in the Yamaguchi et al. [10] decomposition. Notice that the helix terms are stronger in the vegetated areas, but overall still relatively small, rarely exceeding 10 percent of the scattering. The image in the right shows the strength of the helix term when it is not normalized by the total power. It is interesting to note that the areas indicated by the arrows show modulations that appear to be consistent with the along-track slopes caused by the local topography. This image was analyzed by van Zyl (1993) [5] who showed that the scattering mechanisms at P-band are strongly modulated by the topographic slopes in the range direction. The fact that the strength of the helix term might be influenced by the along-track slopes should not be a surprise because it is well known that along-track tilts lead to

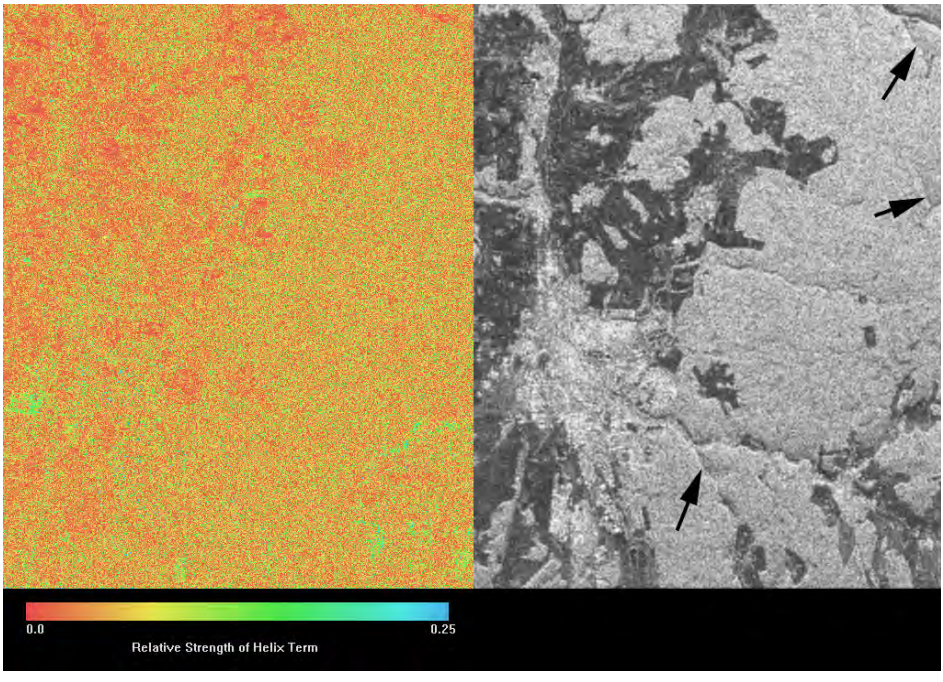


Fig. 3-15. The image on the left displays the relative strength of the helix term in the Yamaguchi et al. [10] decomposition as a fraction of the total power. While the helix term is slightly stronger in the vegetated areas, it is relatively small compared to the overall scattering. The image on the right shows the strength of the helix term when it is not normalized. The arrows indicate areas where the strength of the helix term is modulated by the topographic slopes in the along-track direction. Images were calculated at L-band.

non-zero correlations between co- and cross-polarized components of the scattering matrix (Shuler et al., 1996 [11]; Lee et al., 2000 [12]), even for terrains that otherwise would exhibit reflection symmetry.

To better understand how azimuth slopes affect the co- and cross-polarized correlation, consider the analysis shown in Appendix A for a tilted surface. We show that the scattering matrix of the tilted surface $\left[\mathbf{S}(\theta)\right]$ can be written as a transformation of the scattering matrix of the surface without tilts $\left[\mathbf{S}_l(\theta_l)\right]$, as follows

$$\mathbf{S}(\theta) = \left[\tilde{\mathbf{T}}_R\right]\mathbf{S}_l(\theta_l)\left[\mathbf{T}_R\right]. \tag{3.3-37}$$

The angles refer to the angle of incidence in either the global coordinate system (θ) or the local coordinate system (θ_l) for the surface. For details, please see the appendix. The transformation matrix takes the form of a coordinate rotation

$$\left[\mathbf{T}_R\right] = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}; \quad \tan\varphi = \frac{h_y}{u}; \quad u = \sin\theta - h_x\cos\theta, \tag{3.3-38}$$

where h_x and h_y are the surface slopes in the range (cross-track) and azimuth (along-track) directions, respectively. Note that these expressions are completely general; no assumptions are made about the form of the scattering matrix. In the special case where the range slopes are small, the rotation is directly proportional to the azimuth slope. Performing the transformation, we find that

$$\begin{aligned} \left\langle S_{hh}(\varphi)S_{hv}^*(\varphi)\right\rangle + \left\langle S_{hv}(\varphi)S_{vv}^*(\varphi)\right\rangle = {} & \frac{1}{2}\sin 2\varphi\left(\left\langle S_{hh}S_{hh}^*\right\rangle - \left\langle S_{vv}S_{vv}^*\right\rangle\right) \\ & + \frac{1}{2}(1+\cos 2\varphi)\left(\left\langle S_{hh}S_{hv}^*\right\rangle - \left\langle S_{hv}S_{vv}^*\right\rangle\right) \\ & - \frac{1}{2}(1-\cos 2\varphi)\left(\left\langle S_{hv}S_{hh}^*\right\rangle - \left\langle S_{vv}S_{hv}^*\right\rangle\right) \end{aligned}. \tag{3.3-39}$$

If the terrain exhibits reflection symmetry, the second and third terms are zero and this quantity becomes a real number. In that case, the estimate of the helix component as given by Eq. (3.3-29) will be zero. If the terrain does not exhibit reflection symmetry or if we only average over a relatively small number of pixels (recall that reflection symmetry only says that in the average the co-cross product will be zero) the estimate of the helix component will be

$$\begin{aligned} f_c &= 2\left|\mathrm{Im}\left\{\left\langle S_{hh}(\varphi)S_{hv}^*(\varphi)\right\rangle + \left\langle S_{hv}(\varphi)S_{vv}^*(\varphi)\right\rangle\right\}\right| \\ &= \left|\mathrm{Im}\left\{(1+\cos 2\varphi)\left(\left\langle S_{hh}S_{hv}^*\right\rangle - \left\langle S_{hv}S_{vv}^*\right\rangle\right) - (1-\cos 2\varphi)\left(\left\langle S_{hv}S_{hh}^*\right\rangle - \left\langle S_{vv}S_{hv}^*\right\rangle\right)\right\}\right| \end{aligned}. \tag{3.3-40}$$

This quantity is clearly modulated by the along-track slopes, so the results in the image on the right in Fig. 3-15 are not surprising.

3.3.2.3 The Non-Negative Eigenvalue Decomposition (NNED). The results in the previous two sections show a significant flaw in these decompositions; some negative powers might result after subtraction of the volume components from the observation. This is clearly a non-physical result. If the hypothesis is that the observed radar cross-section is the linear sum of radar cross-sections representing different types of scattering mechanisms, a crucial requirement is that each scattering mechanism must represent a physically realizable scatterer or collection of scatterers. This must also mean that the radar cross-section representing each scattering mechanism must be zero or positive for all polarization combinations. As shown in Eq. (3.1-5), this implies that all eigenvalues of the matrices representing the individual scattering processes must be non-negative. A decomposition method that takes this property into account was proposed by van Zyl et al. (2008) [13] and van Zyl et al. (2010) [14], which we shall call the non-negative eigenvalue decomposition (NNED).

To introduce the NNED technique, let us take a more generic decomposition where we want to express the scattering as follows

$$\left\langle [\mathbf{C}] \right\rangle = a[\mathbf{C}_{model}] + [\mathbf{C}_{remainder}]. \tag{3.3-41}$$

Here, the first term on the right represents the covariance matrix predicted by some model, such as randomly oriented branches. Recognizing that the form of this covariance matrix might be different from the measured matrix, we add the second term, which will contain whatever is in the measured matrix that is not consistent with the model matrix. The question now is what value of a to use in Eq. (3.3-41). To answer this question, we need to recognize that all matrices in Eq. (3.3-41) must represent physically realizable covariance matrices. That is, if we look at each matrix by itself, it should satisfy all the restrictions that we expect for a measured covariance matrix. In particular, we need to insist that the eigenvalues for each matrix are real and greater than or equal to zero. Let us rewrite Eq. (3.3-41) in the form

$$[\mathbf{C}_{remainder}] = \left\langle [\mathbf{C}] \right\rangle - a[\mathbf{C}_{model}]. \tag{3.3-42}$$

The matrix on the left must have eigenvalues that are real and are larger than or equal to zero. This requirement allows us to derive a limit on the values of a. The largest value of a that still ensures that all three eigenvalues of the matrix on the left would be greater than or equal to zero is the maximum value of a that we could use in Eq. (3.3-41).

To derive the general expressions limiting the values of *a*, we start with the average covariance matrix for terrain with reflection symmetry as given in Eq. (3.3-11) and write the model covariance matrix as

$$\left[\mathbf{C}_{model}\right]=\begin{pmatrix}\xi_a & 0 & \rho_a\\ 0 & \eta_a & 0\\ \rho_a^* & 0 & \zeta_a\end{pmatrix}. \tag{3.3-43}$$

Then Eq. (3.3-42) becomes

$$\left[\mathbf{C}_{remainder}\right]=\begin{pmatrix}\xi & 0 & \rho\\ 0 & \eta & 0\\ \rho^* & 0 & \zeta\end{pmatrix}-a\begin{pmatrix}\xi_a & 0 & \rho_a\\ 0 & \eta_a & 0\\ \rho_a^* & 0 & \zeta_a\end{pmatrix}. \tag{3.3-44}$$

The eigenvalues for this matrix are the roots of the following equation

$$(\eta-a\eta_a-\lambda)\left\{\lambda^2-(\xi+\zeta-a\xi_a-a\zeta_a)\lambda+(\xi-a\xi_a)(\zeta-a\zeta_a)-|\rho-a\rho_a|^2\right\}=0, \tag{3.3-45}$$

which are

$$\begin{aligned}\lambda_1&=\frac{1}{2}\left\{\xi+\zeta-a\xi_a-a\zeta_a+\sqrt{(\xi+\zeta-a\xi_a-a\zeta_a)^2-4(\xi-a\xi_a)(\zeta-a\zeta_a)+4|\rho-a\rho_a|^2}\right\}\\ \lambda_2&=\frac{1}{2}\left\{\xi+\zeta-a\xi_a-a\zeta_a-\sqrt{(\xi+\zeta-a\xi_a-a\zeta_a)^2-4(\xi-a\xi_a)(\zeta-a\zeta_a)+4|\rho-a\rho_a|^2}\right\}\\ \lambda_3&=\eta-a\eta_a\end{aligned}. \tag{3.3-46}$$

Since $\lambda_1\ge\lambda_2$, the maximum value of *a* is found when either $\lambda_2=0$ or when $\lambda_3=0$. To find the value of *a* that would make $\lambda_2=0$, we need to solve the equation

$$\begin{aligned}&(\xi-a\xi_a)(\zeta-a\zeta_a)=|\rho-a\rho_a|^2\\ &\quad\Rightarrow\ \left(\xi_a\zeta_a-|\rho_a|^2\right)a^2-\left\{(\xi\zeta_a+\zeta\xi_a)-\rho\rho_a^*-\rho^*\rho_a\right\}a+\xi\zeta-|\rho|^2=0\end{aligned}. \tag{3.3-47}$$

This quadratic has two roots that are both positive. The smaller of the two is

$$\begin{aligned}a=\frac{1}{2\left(\xi_a\zeta_a-|\rho_a|^2\right)}\Big\{&(\xi\zeta_a+\zeta\xi_a)-\rho\rho_a^*-\rho^*\rho_a\\ &-\sqrt{\left[(\xi\zeta_a+\zeta\xi_a)-\rho\rho_a^*-\rho^*\rho_a\right]^2-4\left(\xi_a\zeta_a-|\rho_a|^2\right)\left(\xi\zeta-|\rho|^2\right)}\Big\}\end{aligned}. \tag{3.3-48}$$

The case of $\lambda_3=0$ is straightforward. The resulting maximum *a* is, therefore,

$$a_{\max} = \min\left\{\begin{array}{l}\eta/\eta_a \\ \dfrac{1}{2\left(\xi_a\zeta_a - |\rho_a|^2\right)}\left\{\left(\xi\zeta_a + \zeta\xi_a\right) - \rho\rho_a^* - \rho^*\rho_a\right. \\ \qquad \left. -\sqrt{\left[\left(\xi\zeta_a + \zeta\xi_a\right) - \rho\rho_a^* - \rho^*\rho_a\right]^2 - 4\left(\xi_a\zeta_a - |\rho_a|^2\right)\left(\xi\zeta - |\rho|^2\right)}\right\}\end{array}\right. . \quad (3.3\text{-}49)$$

It should be pointed out that in the general form for the model matrix we assumed in Eq. (3.3-43), we have explicitly assumed that the model matrix represents an average scattering process for terrain with reflection symmetry. It also includes cases of a single scatterer with no cross-polarized component, such as a pure dihedral or first-order scattering from a slightly rough surface. It does not, however, include scattering from terrain that does not exhibit reflection symmetry or single scatterers with cross-polarized components. For those cases, the expressions become significantly more complicated, with the eigenvalues being the roots to a cubic polynomial. However, most observed scattering seems to approximate reflections symmetry well, as evidenced by the relatively small values of the helix component in the Yamaguchi et al. (2005) [10] decomposition discussed in the previous section; therefore, we do not believe the current analysis to be too severely limited by these assumptions. In any case, the fundamental approach remains the same. We always insist that no eigenvalue be negative.

We note that if the model represents a single scatterer with no cross-polarized return, $\xi_a\zeta_a = |\rho_a|^2$ and we can simplify the expressions in Eq. (3.3-49) to

$$a_{\max} = \min\left\{\begin{array}{l}\eta/\eta_a \\ \dfrac{\xi\zeta - |\rho|^2}{\left(\xi\zeta_a + \zeta\xi_a\right) - \rho\rho_a^* - \rho^*\rho_a}\end{array}\right. . \quad (3.3\text{-}50)$$

Let's illustrate this process with an example. For the model, we shall assume a uniformly random oriented cloud of thin cylinders. The expression for the average covariance matrix for this model was derived in Chapter 2 and for the backscatter case is given by

$$\left\langle\left[\mathbf{C}_{cylinders}\right]\right\rangle = \begin{pmatrix} 3/8 & 0 & 1/8 \\ 0 & 2/8 & 0 \\ 1/8 & 0 & 3/8 \end{pmatrix}. \quad (3.3\text{-}51)$$

Putting these together in Eq. (3.3-49), we find the resulting maximum *a* is, therefore,

$$a_{\max} = \min\left\{4\eta,\ \frac{1}{2}\left\{3(\xi+\zeta) - \rho - \rho^* - \sqrt{\left[3(\xi+\zeta) - \rho - \rho^*\right]^2 - 32\left(\xi\zeta - |\rho|^2\right)}\right\}\right\}. \quad (3.3\text{-}52)$$

To illustrate this process with a numerical example, consider the case of a covariance matrix extracted from the Black Forest image at L-band. The normalized matrix is

$$\left\langle [\mathbf{C}] \right\rangle = \begin{pmatrix} 0.472 & 0.008 + i0.010 & 0.056 - i0.029 \\ 0.008 - i0.010 & 0.235 & 0.003 - i0.002 \\ 0.056 + i0.029 & 0.003 + i0.002 & 0.293 \end{pmatrix}. \quad (3.3\text{-}53)$$

Note that this matrix is not the same as the one we assumed in Eq. (3.3-11), but the terms assumed to be zero in Eq. (3.3-11) are indeed much smaller than the others. We shall, therefore, ignore those terms and effectively set them equal to zero, consistent with what Freeman and Durden (1998) [9] assumed. Using Eq. (3.3-52), we find that

$$a_{\max} = \min\left(0.940, 0.752\right) = 0.752. \quad (3.3\text{-}54)$$

Simply setting the strength of the returns from the randomly oriented cylinders based on the cross-polarized return will force us to use the value 0.940 for *a*. Clearly, this is much larger than the allowable value of 0.752. Using this maximum value of *a*, we find that the decomposition would be

$$\begin{pmatrix} 0.472 & 0.008 + i0.010 & 0.056 - i0.029 \\ 0.008 - i0.010 & 0.235 & 0.003 - i0.002 \\ 0.056 + i0.029 & 0.003 + i0.002 & 0.293 \end{pmatrix} \approx \begin{pmatrix} 0.281 & 0 & 0.094 \\ 0 & 0.188 & 0 \\ 0.094 & 0 & 0.281 \end{pmatrix} + \begin{pmatrix} 0.191 & 0 & -0.038 - i0.029 \\ 0 & 0.047 & 0 \\ -0.038 + i0.029 & 0 & 0.012 \end{pmatrix}. \quad (3.3\text{-}55)$$

Figure 3-16 shows a comparison of the fraction of the scattering assigned to the randomly oriented canopy by using the approach suggested by Freeman and Durden (1998) [9] and that using the maximum value of *a* as we have described above. The fraction is expressed as the total power of the scattering from the randomly oriented cloud of cylinders as compared to the total power in the measured covariance matrix. In the case of the Freeman and Durden (1998) [9] decomposition, this ratio is simply the thin vegetation index, as defined in Eq. (3.2-3). We note from this figure that the fraction of scattering that should be assigned to the vegetation is generally quite a bit lower than if we simply assume that all the cross-polarized scattering comes from the randomly oriented branches. In fact, on the average, the maximum amount of scattering from the

branches is between 10 and 20 percent lower if we insist that the eigenvalues remain positive. Nevertheless, the random scattering clearly still dominates in the vegetated areas; this correction would not change the qualitative interpretation of the scattering for this scene significantly.

It is important to remember that the image on the right in Fig. 3-16 represents the maximum amount of scattering we can assign to the randomly oriented cylinders. We are, of course, free to assign less than that, and we would not violate our requirement that the eigenvalues of the remainder not be negative. This result points to one of the difficulties with model-based decompositions: How are we to decide how much of the scattering to assign to the randomly oriented cylinders? Unfortunately, there is no simple answer to this question. A reasonable suggestion is to use the amount of power in the left-over matrix (that is, the one labeled "remainder" in Eq. (3.3-41)) as a guide. One could make the argument that the best value of a to use would be that value of a that results in the minimum amount of power in the remainder matrix. We can then simply compare the total power in the remainder matrix to that in the original measurement and select that value of a that minimizes this ratio. For the single-

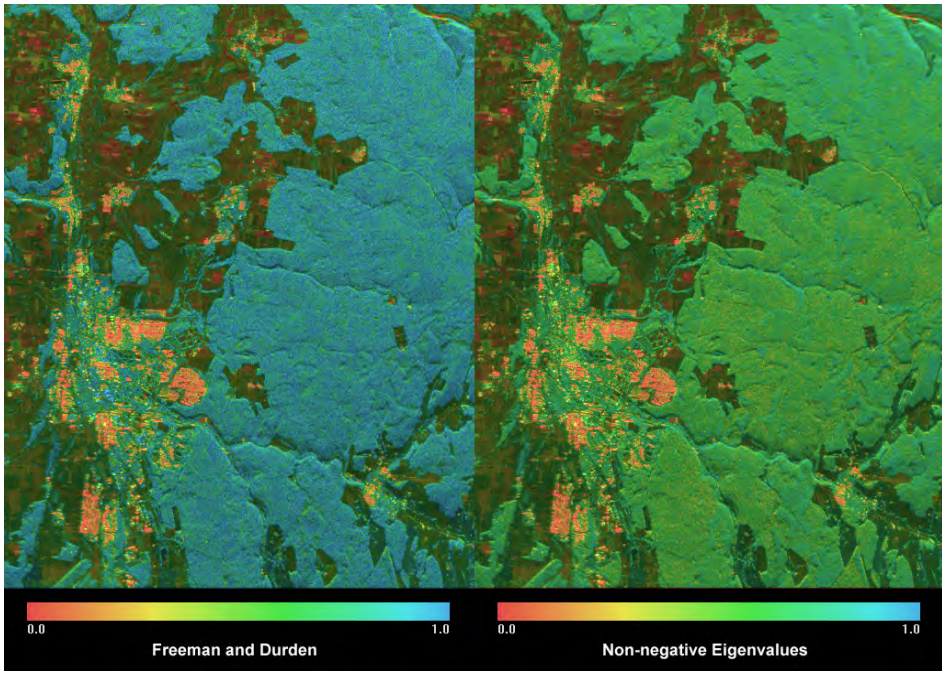


Fig. 3-16. These images display the fraction of the observed total power that can be attributed to scattering from a uniformly randomly oriented layer of cylinders. The image on the left follows the calculation proposed by Freeman and Durden (1998) [9]. The image on the right is calculated by requiring that the remaining eigenvalues not be negative. On the average, the image on the right shows about 10-20 percent lower values for the forested areas than the one on the left.

model case discussed so far, this simply leads to picking the maximum value of *a*.

The question now is whether there is a way to pick the "best" solution for a generic model-based decomposition for vegetation scattering. We shall assume, like Freeman and Durden (1998) [9], that the scattering from vegetated terrain consists of three major components: scattering from the canopy, double reflection scattering, and single reflections from the underlying ground surface. We then propose a hybrid approach that combines model-based decomposition and eigenvalue decomposition. In the case of the Freeman and Durden (1998) [9] decomposition, they must artificially decide whether to explicitly solve for the double-reflection or the single-reflection parameters by examining the phase difference between HH and VV. As we showed before (van Zyl, 1992 [15]), the eigenvalue decomposition itself can be interpreted in terms of single (or odd numbers of) reflections, double (or even numbers of) reflections, and a diffuse part. Our hybrid approach then is as follows. We first subtract a covariance matrix representing a model for canopy scattering from the observed covariance matrix. We pick that multiplicative parameter that minimizes the power in the remainder matrix. We then perform an eigenvalue decomposition on the remainder matrix. Mathematically this translates to

$$\langle [\mathbf{C}] \rangle = f_c \left[\mathbf{C}_{canopy} \right] + \lambda_d \left[\mathbf{C}_{double} \right] + \lambda_s \left[\mathbf{C}_{single} \right] + \lambda_r \left[\mathbf{C}_{remainder} \right]. \quad (3.3\text{-}56)$$

Here, λ_r is the "diffuse" part of the covariance matrix that is left once we subtracted the canopy contribution to the observed scattering. It is clear that if we want to minimize the power in the last matrix on the right, we need to maximize the value of f_c. This is done using the expressions in Eq. (3.3-49), where the subscript *a* in that expression is substituted for the canopy model covariance matrix.

This decomposition needs only one assumption: that of the model to use to represent the canopy scattering. We no longer have to artificially make assumptions about the ratio of the HH to VV scattering strength for the double-reflection or single-reflection terms. These are determined by the eigenvectors of the matrix that is left over once the canopy contribution is subtracted.

Figure 3-17 shows a comparison of the Freeman and Durden (1998) [9] decomposition and NNED using uniformly randomly oriented thin cylinders as the model for the canopy scattering like Freeman and Durden. In our display of the NNED, we use the green color to show the canopy strength, blue to show the single scattering term, and red to show the double reflection term. We do not include the last term on the right in Eq. (3.3-56) in our display. At first look, these images are qualitatively very similar. A closer examination, however, shows that the details in the forested area are quite different. For example, there

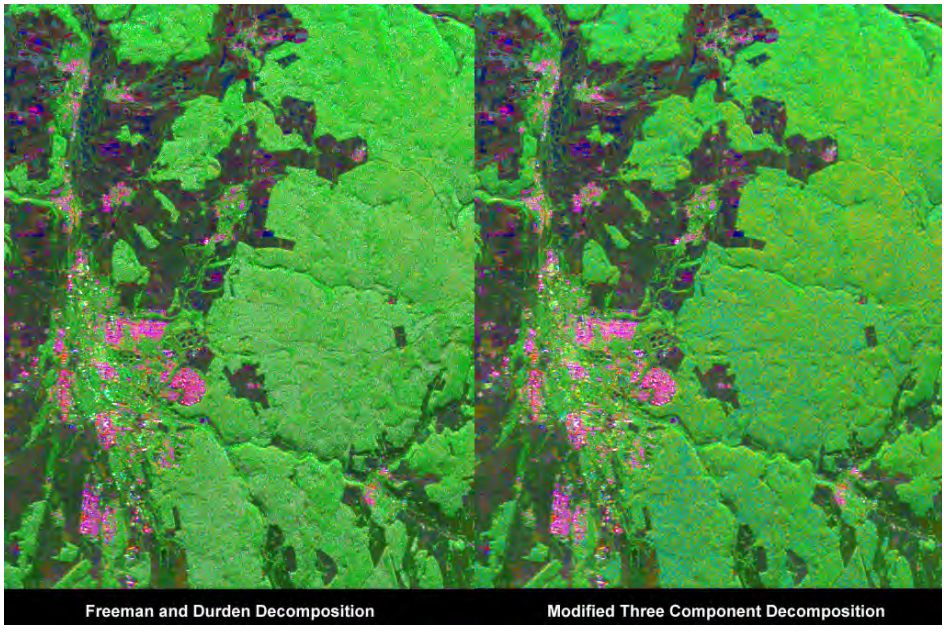


Fig. 3-17. These L-band images show a comparison between a three-component decomposition as proposed by Freeman and Durden (left image) [9] and a modified version that ensures non-negative powers after the vegetation scattering has been subtracted (right image). Note the increased double reflections in the forested areas in the image on the right.

are significantly more double reflections in much of the forest in the modified three-component analysis (NNED) than in the original Freeman and Durden decomposition. This is to be expected, since we have now assigned less of the scattered power to the randomly oriented cylinders. Therefore, while the results are qualitatively the same, these images are quite different. This could be quite significant if one were to use these decomposed signals in further analysis to infer geophysical parameters.

Figure 3-18 shows a similar result calculated using the C-band images. The differences in the forest are now even more evident between the standard Freeman-Durden (1998) [9] decomposition and the modified version that ensures that all eigenvalues remain non-negative. In this case, there is significantly more scattering assigned to the single reflection model, especially in the older, higher biomass parts of the forest. This should not, however, be interpreted that there is more direct scattering observed from the ground surface under the forest. In fact, this increase is due to the fact that the thin cylinder model is less appropriate at C-band because most of the branches are no longer much thinner than the radar wavelength. When this happens, the random cylinder model matrix becomes more like scattering from a flat plate or a sphere. In the limiting case where the branches become much thicker than the

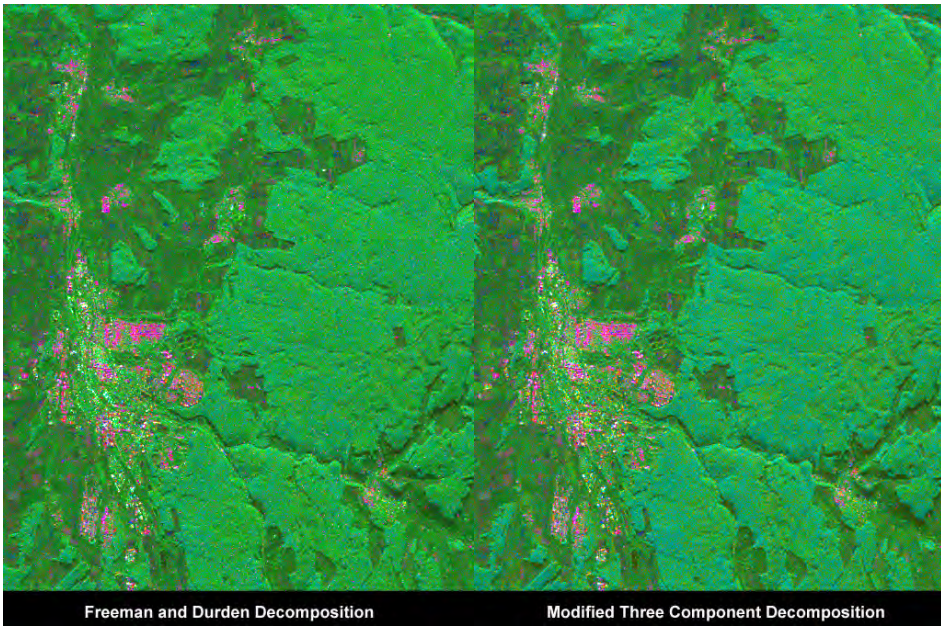


Fig. 3-18. Comparison between a three-component decomposition as proposed by Freeman and Durden (left image) and the modified version that ensures non-negative powers after the vegetation scattering has been subtracted (right image) at C-band. Note the increased single reflections in the forested areas in the image on the right. See the text for a discussion.

radar wavelength, the scattering from the randomly oriented thick branches would be indistinguishable from the scattering from a sphere or a flat plate.

Figure 3-19 shows the comparison at P-band. The modified decomposition shows significantly less random scattering and increased double reflections when compared to the Freeman-Durden (1998) [9] decomposition. This shows again how assigning all the cross-polarized return to the canopy scattering tends to overestimate the contribution from the canopy scattering. Also note that the scattering near the top of the image (the portion where the angle of incidence is quite steep) has a more bluish tint in the image on the right, implying more single scattering from the underlying ground surface. This is consistent with the expectation that at steeper angles of incidence the attenuation loss through the canopy will be lower, increasing the strength of the reflection from the underlying ground surface.

In all the discussion thus far we have implicitly assumed that the covariance matrix terms that involve cross-products of co-polarized returns and cross-polarized returns are negligibly small. The usual argument in favor of this assumption is that natural terrain tends to exhibit reflection symmetry over large scales. Yamaguchi et al. (2005) [10] point out that this is not always the

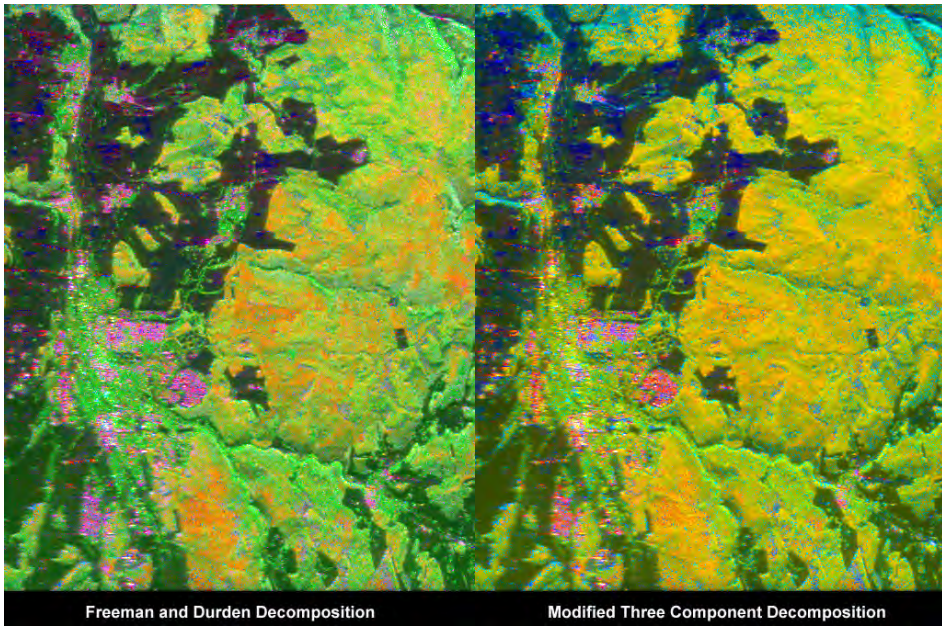


Fig. 3-19. Comparison between a three-component decomposition as proposed by Freeman and Durden (left image) and the modified version that ensures non-negative powers after the vegetation scattering has been subtracted (right image) at P-band. Note the increased double reflections (red color) in the forested areas in the image on the right.

case. In fact, they show examples of urban areas where these terms are not negligible. Their solution is to add a fourth model to the decomposition that includes scattering by helices to account for the observed correlation between the co-polarized and cross-polarized returns. They then proceed to first remove this helix component that is calculated based on the observed correlation. Note that our proposed hybrid decomposition as shown in Eq. (3.3-56) explicitly includes any potential observed cross-correlation between the co-polarized and cross-polarized components. The matrices that are calculated using the eigenvalue decomposition are formed from the eigenvectors of the covariance matrix that results from the subtraction of the assumed canopy model from the observed covariance matrix. Any correlation between the cross-polarized and co-polarized returns will result in these eigenvectors not being "pure" single, double, or diffuse scattering covariance matrices. The advantage of the NNED approach is that we do not have to assume a specific scattering mechanism that accounts for these observed correlations.

3.3.2.4 Adaptive Model-Based Decomposition. In the previous section, we laid out a method for performing a model-based decomposition that ensures that all the individual matrices have non-negative eigenvalues. We shall now show

how this same idea can be used to determine which canopy model is the “best” fit to the observations.

As we pointed out before, the idea is to minimize the amount of power in the matrix that is left once we subtract the canopy scattering. This provides a simple way to compare different canopy models. We simply calculate the amount of power that would be in the remainder matrix once we subtract each canopy model contribution. The model with the smallest amount of power left over is then assumed to be the best fit to the observation. In fact, Yamaguchi et al. (2005) [10] pointed out that the uniformly randomly oriented cloud of thin cylinders might not always be the appropriate model for canopy scattering. As part of their four-component scattering decomposition, they included a test to apply different canopy models.

As an illustration of the idea, let us first look at an example where we will use a limited number of different canopy models to find the best fit for the L-band data of the Black Forest image we have been analyzing so far. In doing the test, we assume three different models: a uniformly randomly oriented cloud of thin cylinders, a cloud of thin cylinders oriented around the horizontal direction with a cosine squared distribution, and a cloud of thin cylinders oriented around the vertical direction with a cosine squared distribution. For each pixel in the image, we then calculate the relative amount of power in the remainder matrix if we subtracted the maximum contribution assuming each of these models. Figure 3-20 shows the result for the three cases at L-band. At the global scale of the image, it appears that the best fit model at L-band might be the cosine

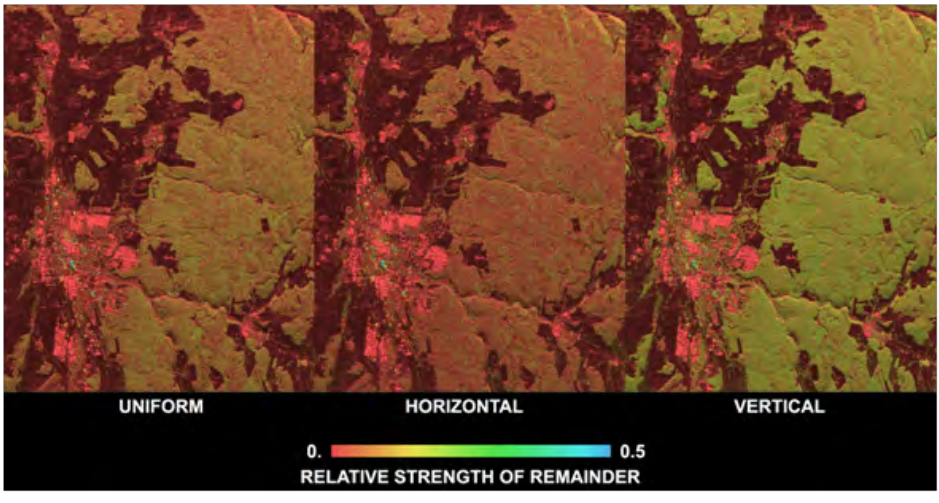


Fig. 3-20. Comparison of different canopy models at L-band. The images show the relative amount of power in the remainder matrix after the canopy contribution has been subtracted from the observation. The results show that the cosine squared distribution around the horizontal orientation provides the best fit over the majority of the image.

squared distribution around the horizontal orientation. Most of the image is, in fact, coniferous trees with branches that are closer to horizontal orientation than uniformly randomly oriented; this result is, therefore, not unexpected.

It should be pointed out that this result cannot be taken as conclusive proof that the canopy contains branches that have orientations that are closer to horizontal. In drawing such conclusions, one has to remember that this result shows the best fit of a canopy model to the observation by first subtracting the canopy contribution. In some sense, this means we assume that the scattering is dominated by the canopy contribution. If, for example, we had the case of a pure double reflection (for which HH is larger than VV) with a small contribution from a uniformly oriented canopy, the resulting covariance matrix would have HH larger than VV. In that case, we might find that the best fit canopy model is a cosine squared around the horizontal orientation because that model predicts that HH will be larger than VV.

We can now extend this concept by adaptively performing the decomposition on each pixel to find the best canopy model for that pixel. The process is the same as we described above, except it is performed for each pixel separately. For each pixel, we now compare the remainders for the set of models and pick the model that leaves the smallest relative amount of power in the remainder after we subtract the canopy contribution. We shall first introduce this concept with just the three models described above (the same three used by Yamagichi et al. (2005) [10]) and later generalize the concept. Figure 3-21 shows the results for the Black Forest image at C-band, L-band, and P-band. At C-band,

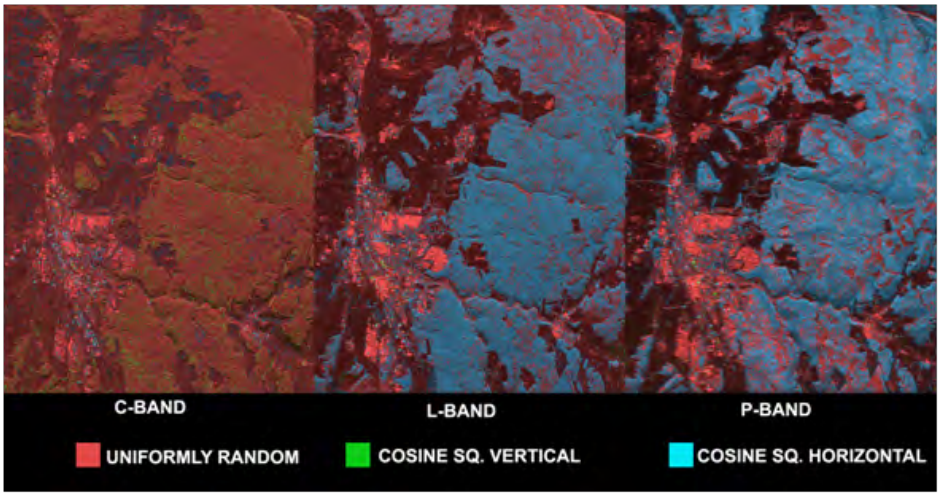


Fig. 3-21. Adaptive non-negative eigenvalue decompositions at different frequencies with three canopy scattering models. At the longer wavelengths, the scattering is best modeled by a cosine squared distribution around the horizontal orientation. At C-band, a uniformly random orientation provides the best fit. See the text for more discussion.

the uniformly randomly oriented model is the best fit, consistent with scattering from the needles in the coniferous trees. At L-band and P-band, we see that most of the pixels are better represented by a model with a cosine-squared distribution around the horizontal orientation. At L-band this result is likely because of the primary branch orientation. At P-band, however, this is probably more indicative that double-bounce scattering from the ground surface and tree trunks likely dominates.

Figure 3-22 shows the L-band comparison of the adaptive decomposition using three models compared to the Yamaguchi et al. (2005) [10] choice of model based on their simplified selection criteria involving the ratio of HH to VV power. We have blanked out pixels for which either the single reflections or the double reflections are stronger than the canopy component since, for those pixels, our original assumption that the canopy scattering dominates clearly does not hold. First, we notice a difference in the number of pixels that are blanked out between the adaptive decomposition and the Yamaguchi et al. decomposition. This is primarily related to the amount of power assigned to the canopy component and the fact that for our adaptive decomposition we insist that all eigenvalues be non-negative. This will typically assign less power to the canopy component. Apart from this difference, however, it is obvious that the two adaptive techniques give very similar results.

Building on these initial results, Arii et al. (2010) [16] and Arii (2009) [17] extended the adaptive decomposition technique to include a generalized canopy

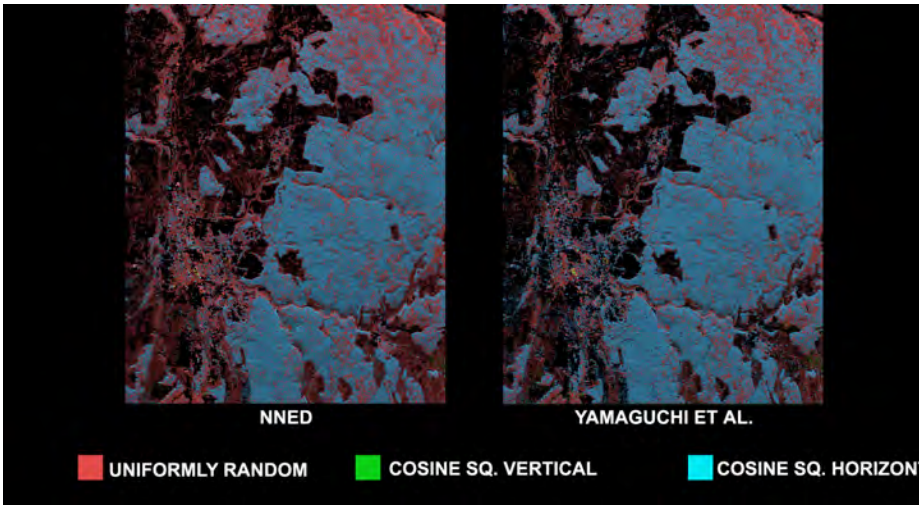


Fig. 3-22. This figure shows a comparison of the optimum model indicated by the color schemes for the adaptive NNED and Yamaguchi et al. models using the L-band image. We have blanked out those pixels for which the canopy scattering is not the strongest component. Note the great similarity between the NNED and Yamaguchi results in terms of which models best represents the canopy scattering.

component that selects the best fit to the canopy scattering over different amounts of randomness and different mean orientations of the canopy. Their characterization of the canopy scattering starts with the scattering matrix of a vertically oriented thin cylinder, which apart from a multiplicative constant that depends on the cylinder length and dielectric constant, can be written as

$$\left[S_{cyl}\right] \propto \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.3-57}$$

They then allow this cylinder to be rotated by an angle θ with respect to the vertical direction about the line of sight, with positive values of θ indicating a clockwise rotation. The resulting covariance matrix is found to be

$$\left[C_{cyl}(\theta)\right] = \frac{1}{4}\begin{pmatrix} (1-\cos 2\theta)^2 & \sqrt{2}\sin 2\theta(1-\cos 2\theta) & \sin^2 2\theta \\ \sqrt{2}\sin 2\theta(1-\cos 2\theta) & 2\sin^2 2\theta & \sqrt{2}\sin 2\theta(1+\cos 2\theta) \\ \sin^2 2\theta & \sqrt{2}\sin 2\theta(1+\cos 2\theta) & (1+\cos 2\theta)^2 \end{pmatrix}. \tag{3.3-58}$$

They write this as the sum of three matrices:

$$\begin{aligned} \left[C_{cyl}(\theta)\right] &= \frac{1}{8}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix} + \frac{2}{8}\begin{pmatrix} -2\cos 2\theta & \sqrt{2}\sin 2\theta & 0 \\ \sqrt{2}\sin 2\theta & 0 & \sqrt{2}\sin 2\theta(1+\cos 2\theta) \\ 0 & \sqrt{2}\sin 2\theta & 2\cos 2\theta \end{pmatrix} \\ &\quad + \frac{1}{8}\begin{pmatrix} \cos 4\theta & -\sqrt{2}\sin 4\theta & -\cos 4\theta \\ -\sqrt{2}\sin 4\theta & -2\cos 4\theta & \sqrt{2}\sin 4\theta \\ -\cos 4\theta & \sqrt{2}\sin 4\theta & \cos 4\theta \end{pmatrix} \\ &= \left[C_\alpha\right] + 2\left[C_\beta(2\theta)\right] + \left[C_\gamma(4\theta)\right] \end{aligned} . \tag{3.3-59}$$

The next step is then to calculate the average covariance matrix for a given probability distribution of the cylinder orientation angles. The resulting average covariance matrix is

$$\left\langle \left[C_{vol}\right] \right\rangle = \int_0^{2\pi} \left[C_{cyl}(\theta)\right] p(\theta) d\theta, \tag{3.3-60}$$

where $p(\theta)$ describes the probability density function of the cylinder orientation angles. In the discussion above, we have used three cases: uniform distribution for which

$$p_{uniform} = \frac{1}{2\pi} \tag{3.3-61}$$

and cosine-squared distributions around the vertical and horizontal directions are as follows

$$p_{\cos\_sq} = \frac{\cos^2(\theta-\theta_0)}{\int_0^{2\pi}\cos^2(\theta-\theta_0)\,d\theta}. \tag{3.3-62}$$

For the distribution peaking around vertical, $\theta_0 = 0$; while for the horizontal case, $\theta_0 = \pi/2$. The uniform distribution represents the most random orientation. In the other extreme, the delta function around a fixed angle

$$p_{delta} = \frac{1}{2}\left\{\delta(\theta-\theta_0)+\delta(\theta-\theta_0-\pi)\right\} \tag{3.3-63}$$

represents the least amount of randomness. Arii (2009) [17] and Arii et al. (2010) [18] suggested that all these cases can be approximated by a generalized *n*-th power cosine squared distribution, as follows:

$$p_v(\theta,\theta_0,n) = \frac{\left\{\cos^2(\theta-\theta_0)\right\}^n}{\int_0^{2\pi}\left\{\cos^2(\theta-\theta_0)\right\}^n d\theta}. \tag{3.3-64}$$

When n = 0, this distribution becomes exactly the same as the uniform distribution; specifically, it approximates the delta function distribution with infinitely large n.

The average covariance matrix for the generalized probability density function can be written as

$$\left\langle\left[C_{vol}(\theta_0,n)\right]\right\rangle = \frac{1}{A_n}\int_0^{2\pi}\left[C_{cyl}(\theta)\right]\cos^{2n}(\theta-\theta_0)\,d\theta, \tag{3.3-65}$$

with

$$A_n = \int_0^{2\pi}\cos^{2n}(\theta-\theta_0)\,d\theta. \tag{3.3-66}$$

To derive expressions for these, we note that we can write

$$\cos^{2n} x = \frac{1}{2^{2n}}\binom{2n}{n} + \frac{1}{2^{2n-1}}\sum_{k=0}^{n-1}\binom{2n}{k}\cos 2(n-k)x\,. \tag{3.3-67}$$

Using the series expansion in Eq. (3.3-67), we find

$$A_n = \frac{\pi}{2^{2n-1}}\binom{2n}{n}. \tag{3.3-68}$$

Also, note that

$$\begin{aligned}\int_0^{2\pi}\left[C_{cyl}(\theta)\right]\cos^{2n}(\theta-\theta_0)\,d\theta &= \frac{1}{2^{2n}}\binom{2n}{n}\int_0^{2\pi}\left[C_{cyl}(\theta)\right]d\theta \\ &+ \frac{1}{2^{2n-1}}\sum_{k=0}^{n-1}\binom{2n}{k}\int_0^{2\pi}\left[C_{cyl}(\theta)\right]\cos\left\{2(n-k)(\theta-\theta_0)\right\}d\theta\end{aligned}\,. \tag{3.3-69}$$

It is easily shown that

$$\frac{1}{2^{2n}}\binom{2n}{n}\int_0^{2\pi}\left[C_{cyl}(\theta)\right]d\theta = \frac{\pi}{2^{2n-1}}\binom{2n}{n}\left[C_\alpha\right]. \tag{3.3-70}$$

Next, we note that

$$\begin{aligned}&\int_0^{2\pi}\left[C_{cyl}(\theta)\right]\cos\left\{2(n-k)(\theta-\theta_0)\right\}d\theta \\ &\quad= \cos\left\{2(n-k)\theta_0\right\}\int_0^{2\pi}\left[C_{cyl}(\theta)\right]\cos\left\{2(n-k)\theta\right\}d\theta \\ &\quad+ \sin\left\{2(n-k)\theta_0\right\}\int_0^{2\pi}\left[C_{cyl}(\theta)\right]\sin\left\{2(n-k)\theta\right\}d\theta\end{aligned}\,. \tag{3.3-71}$$

We easily find that

$$\int_0^{2\pi}\left[C_\alpha\right]\cos\left\{2(n-k)\theta\right\}d\theta = \int_0^{2\pi}\left[C_\alpha\right]\sin\left\{2(n-k)\theta\right\}d\theta = 0\,. \tag{3.3-72}$$

After considerable algebra, we find

$$\int_0^{2\pi}\left[C_\beta(2\theta)\right]\cos\left\{2(n-k)(\theta-\theta_0)\right\}d\theta=\begin{cases}2\pi\left[C_\beta(2\theta_0)\right] & for \quad k=n-1\\ 0 \quad otherwise\end{cases}$$
$$\int_0^{2\pi}\left[C_\gamma(4\theta)\right]\cos\left\{2(n-k)(\theta-\theta_0)\right\}d\theta=\begin{cases}\pi\left[C_\gamma(4\theta_0)\right] & for \quad k=n-2\\ 0 \quad otherwise\end{cases} \quad . \qquad (3.3\text{-}73)$$

Using Eq. (3.3-73), Eq. (3.3-72), Eq. (3.3-70), and Eq. (3.3-68) in Eq. (3.3-65), we find

$$\left\langle\left[C_{vol}(\theta_0,n)\right]\right\rangle=\left[C_\alpha\right]+\frac{2n}{n+1}\left[C_\beta(2\theta_0)\right]+\frac{n(n-1)}{(n+1)(n+2)}\left[C_\gamma(4\theta_0)\right]. \qquad (3.3\text{-}74)$$

Note that we have derived this expression assuming that n is an integer. We have also verified this expression numerically for values of n that are not integers (Arii et al., 2010 [18]). For the three cases we have used so far, we find

$$\begin{aligned}\left\langle\left[C_{vol}^{uniform}\right]\right\rangle&=\left[C_\alpha\right]\\ \left\langle\left[C_{vol}^{\cos\_sq}(\theta_0)\right]\right\rangle&=\left[C_\alpha\right]+\left[C_\beta\right]\\ \left\langle\left[C_{vol}^{delta}(\theta_0)\right]\right\rangle&=\left[C_\alpha\right]+2\left[C_\beta\right]+\left[C_\gamma\right]\end{aligned} \quad . \qquad (3.3\text{-}75)$$

The eigenvalues of the covariance matrix in Eq. (3.3-74) only depend on the value of the power of the cosine squared function and not on the mean orientation angle. The eigenvalues are:

$$\begin{aligned}\lambda_1&=\frac{1}{4(n+1)(n+2)}\left\{2n^2+4n+3+\sqrt{4n^2(n+2)^2+(2n+1)^2}\right\}\\ \lambda_2&=\frac{1}{4(n+1)(n+2)}\left\{2n^2+4n+3-\sqrt{4n^2(n+2)^2+(2n+1)^2}\right\}.\\ \lambda_3&=\frac{1}{2}\frac{2n+1}{(n+1)(n+2)}\end{aligned} \qquad (3.3\text{-}76)$$

Figure 3-23 shows the three eigenvalues as a function of the power of the cosine-squared function. Note that for all values of n, $\lambda_2 \le \lambda_3 \le \lambda_1$. This figure shows that there is little practical difference for distributions for values of $n > 20$. To illustrate this, we show the entropy and the RVI of the resulting average covariance matrix as a function of n in Fig. 3-24. The results confirm that especially the RVI varies very little for values of $n > 20$.

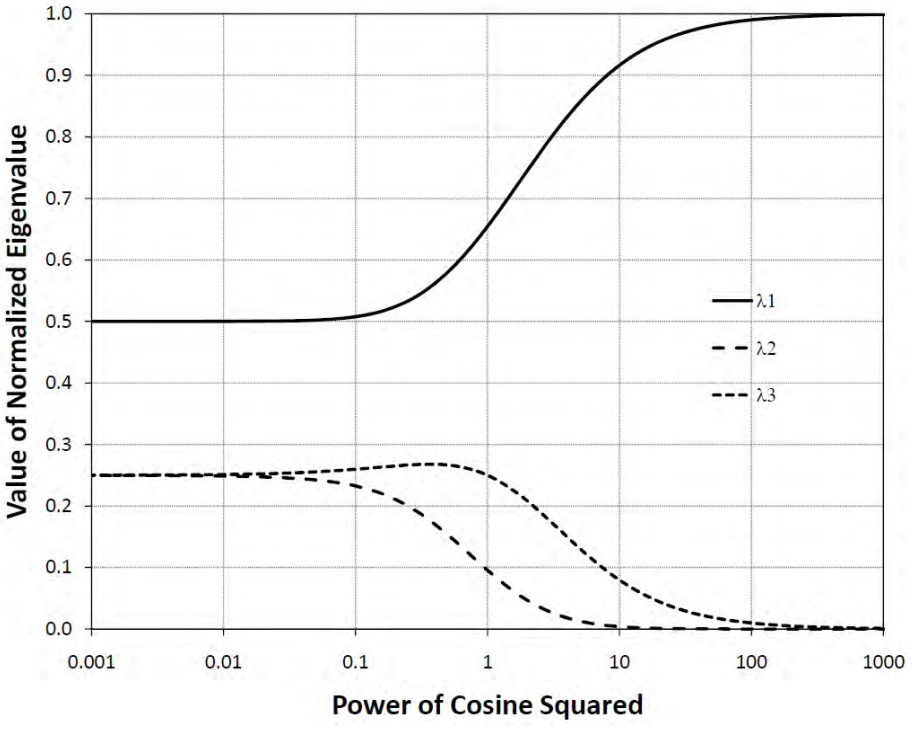


Fig. 3-23. This figure shows the relative strength of the eigenvalues of the covariance matrix representing the generalized volume scattering model. For all values of n, we find that $\lambda_2 \le \lambda_3 \le \lambda_1$..

With this theoretical framework in hand, we can now describe the proposed adaptive model-based decomposition. In fact, it differs from the previously described NNED only in the first step of the decomposition. Instead of choosing a specific model for the canopy scattering, the adaptive decomposition calculates the best fit over all powers *n* and all mean orientation angles θ_0. We decide which pair of parameters represent the best fit by selecting that pair that results in the least amount of power in the "remainder" matrix:

$$\left\langle [C'_{remainder}] \right\rangle = \left\langle [C] \right\rangle - f_v \left\{ [C_\alpha] + \frac{2n}{n+1}[C_\beta] + \frac{n(n-1)}{(n+1)(n+2)}[C_\gamma] \right\}. \qquad (3.3\text{-}77)$$

Note that the canopy distribution function as used here does not assume reflection symmetry. In fact, reflection symmetry only results for special mean orientation angles. As such, Eq. (3.3-77) describes a general decomposition; we do not need to resort to adding special scatterers to generate non-zero correlations between the cross- and co-polarized terms in the covariance matrix.

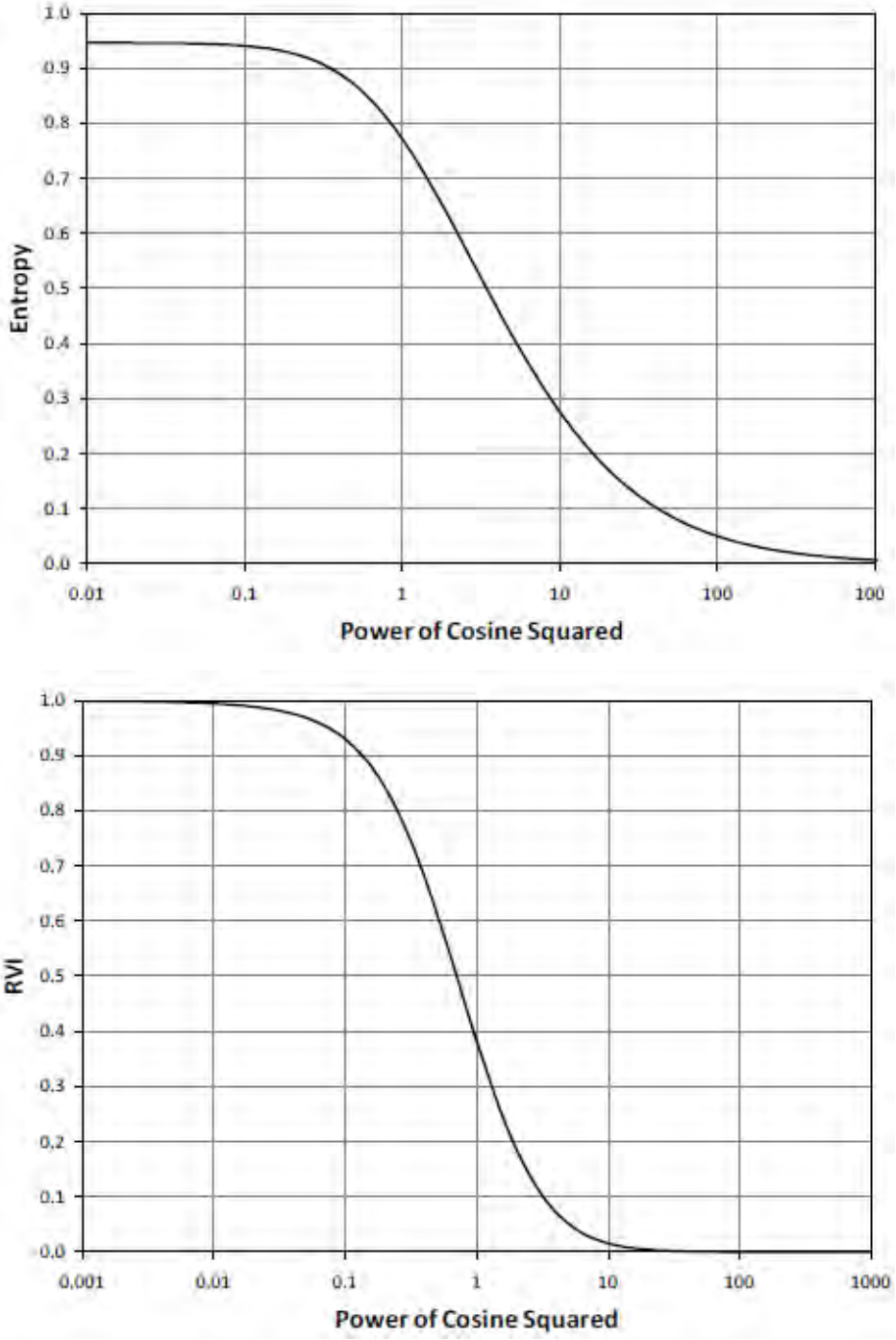


Fig. 3-24. Two charts showing the relative strength of the eigenvalues of the covariance matrix representing the generalized volume-scattering model. The top chart is a plot of entropy, and the bottom is a plot of RVI.

We first illustrate our results with a numerical example. For our comparison, we extracted the average covariance matrix in the forest portion of the Black Forest image at all three frequencies. The resulting matrices are

$$\langle C\rangle_C = \begin{pmatrix} 0.36 & -0.07+i0.0 & -0.18-i0.03 \\ -0.07\text{-}i0.0 & 0.20 & -0.08+i0.0 \\ -0.18\text{+}i0.03 & -0.08-i0.0 & 0.44 \end{pmatrix}$$
$$\langle C\rangle_L = \begin{pmatrix} 0.52 & -0.09-i0.03 & -0.09+i0.08 \\ -0.09\text{+}i0.03 & 0.22 & -0.06+i0.01 \\ -0.09-i0.08 & -0.06-i0.01 & 0.26 \end{pmatrix}. \qquad (3.3\text{-}78)$$
$$\langle C\rangle_P = \begin{pmatrix} 0.67 & -0.07+i0.0 & -0.03+i0.13 \\ -0.07-i0.0 & 0.13 & -0.04+i0.01 \\ -0.03-i0.13 & -0.04-i0.01 & 0.20 \end{pmatrix}$$

We shall illustrate the use of this model by assuming the canopy scattering dominates and that we can consider the canopy as a collection of randomly oriented dipoles. This is equivalent to assuming that the scattering is from a collection of randomly oriented cylinders, but that the cylinders are much thinner than the radar wavelength.

In fitting our observations to the model, we use the methodology described above. The results are:

$$\begin{aligned} C-Band: &\quad n=0.92 \quad \theta_0=143.4^\circ \\ L-Band: &\quad n=1.66 \quad \theta_0=107.7^\circ . \\ P-Band: &\quad n=3.47 \quad \theta_0=99.1^\circ \end{aligned} \qquad (3.3\text{-}79)$$

These results show a decrease in randomness with an increase in wavelength. Also, the longer wavelengths sense a mean orientation angle that is close to horizontal (recall the angles are measured with respect to the orientation angle of a vertically oriented dipole). Since the longer wavelengths interact preferentially with the larger branches, which for the type of trees present in this data set are oriented closer to the horizontal direction, the results appear reasonable. The C-band results indicate more randomness, consistent with interactions with more vertically oriented needles that are also more randomly oriented than the larger branches.

Next we show the results of applying the adaptive decomposition to the Black Forest image. First we show the randomness results in Fig. 3-25. A significant difference is observed in the forested area as a function of wavelength,

consistent with the numerical example as reported in van Zyl et al. (2010) [14]. Randomness values close to that of the uniform distribution are found over the C-band image; conversely, the P-band image shows values with less randomness than cosine-squared distribution. From a physical point of view, this is reasonable considering the orientation distributions of needles, branches, and trunks. The shorter wavelength mainly interacts with the needles, which have higher variance than trunks and branches. Also, the P-band can penetrate needles and branches so that trunks having much lower variance become dominant. Some of the areas of increased randomness visible at P-band have previously been shown to be areas where the trunk scattering is reduced due to topographic effects (van Zyl, 1993 [5]). The L-band result lies in between these two cases; this corresponds mainly to the branch distribution.

Next, we show the mean orientation angle for all three frequencies in Fig. 3-26. Pixels with horizontal orientation are widely distributed in the forested area of the L- and P-band images. Before continuing with the interpretation, it is necessary to look at the decomposition results of all three scattering mechanisms, as in Fig. 3-27. Here, we display each scattering mechanism separately, and the value for each pixel is normalized by the total power. Since the scattering in the forested area at the L-band is mostly contributed by the volume component, the inferred horizontal orientation in the L-band image in Fig. 3-26 should be indicative of scattering by branches. However, the horizontal orientation in the P-band image is not related to the physical orientation of scatterers in the volume layer. It is well-known that the double-bounce scattering raises the HH contribution due to the Fresnel reflections at

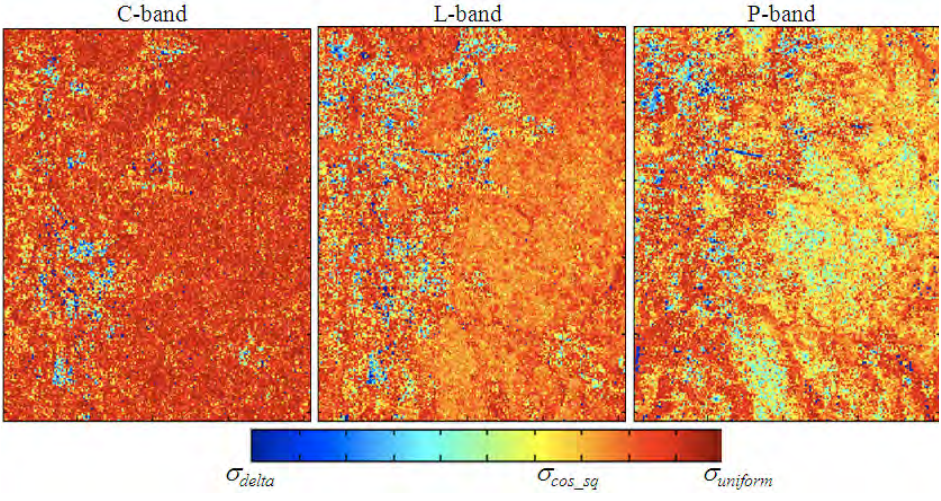


Fig. 3-25. Randomness maps derived from the C-band (left), L-band (center), and P-band (right) AIRSAR images in the Black Forest. As in the numerical example, the randomness decreases for the longer wavelengths (Arii et al., 2009 [16]).

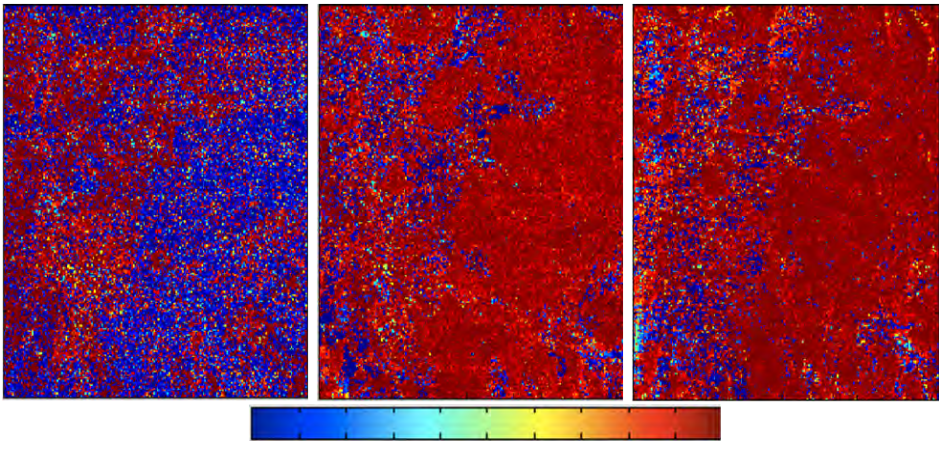


Fig. 3-26. Mean orientation angle maps derived from the C-band (left), L-band (center), and P-band (right) AIRSAR images in the Black Forest (Arii et al., 2009 [16]).

the trunk and ground surface. Since the algorithm subtracts the volume component first and then infers the other scattering mechanisms, strong double-bounce contribution misleads us to interpret a horizontally oriented volume component. One must pay particular attention to interpret the mean orientation angle map when double-bounce scattering contribution is dominant. The C-band mean orientation angle map indicates scattering from more vertically oriented scatterers. Given that the shorter C-band wavelength mostly interacts with the needles and mostly with those needles near the top of the canopy (scattering lower down in the canopy is expected to be attenuated), this result is reasonable.

A clear contrast exists along the river and road in the P-band in Fig. 3-27. The scattering from the forested area is basically contributed by both volume and double-bounce terms. However, the volume component is significant along the river and road, whereas the double-bounce is considerably smaller. This result is due to the local topography, as pointed out by van Zyl [5], where the double-bounce contribution is drastically reduced because the Fresnel reflections no longer dominate when the forest floor is tilted.

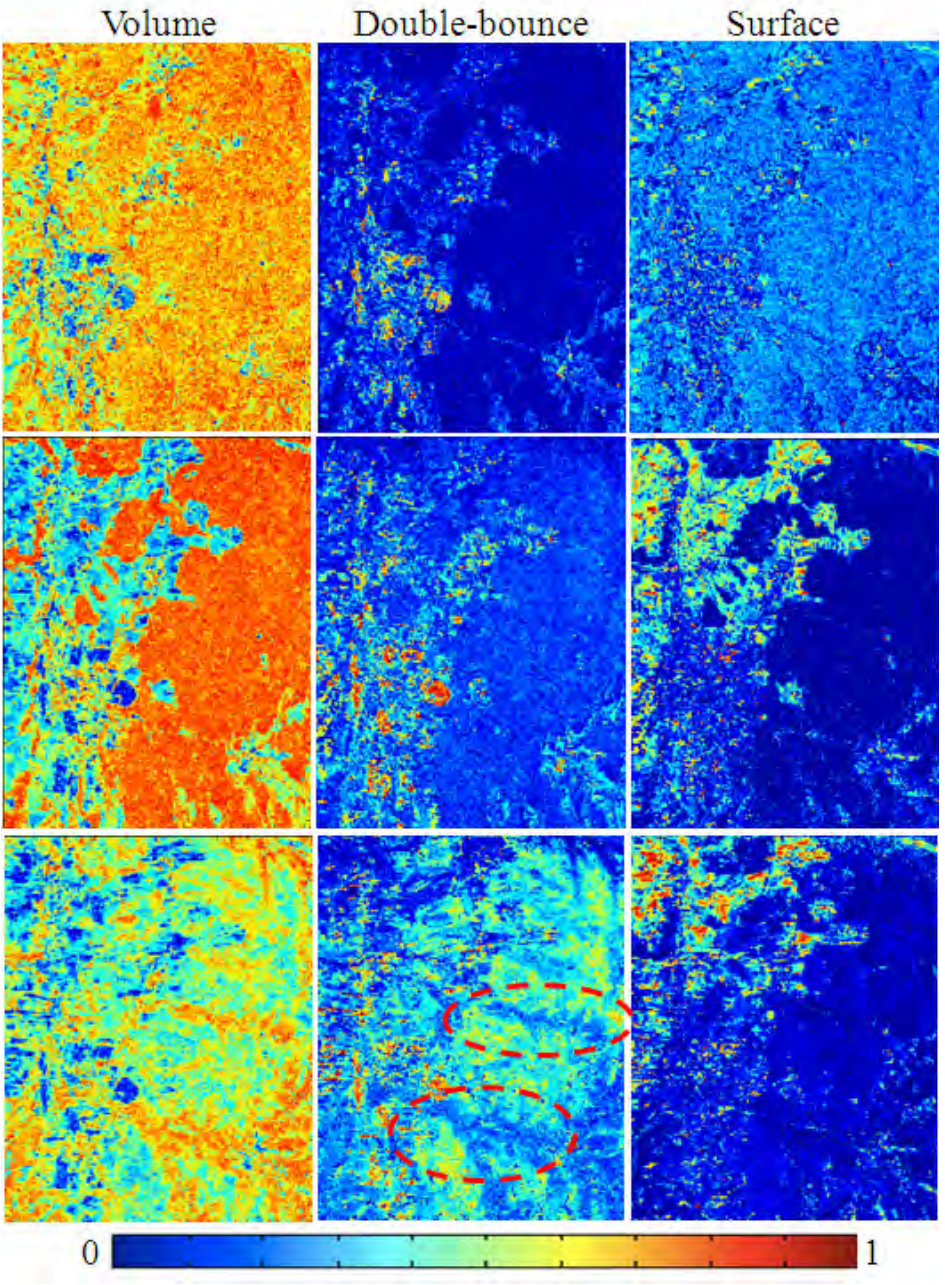


Fig. 3-27. Volume (left), double-bounce (center), and surface (right) components of the adaptive algorithm for the C-band (top), L-band (middle), and P-band (bottom) Black Forest images. Each pixel is normalized by total power. Dotted lines in red indicate river (upper) and road (lower), respectively (Arii et al., 2009 [16]).

3.4 Image Classification

Classification of images involves using a set of rules to decide whether different pixels in an image have similar characteristics. These rules in effect divide the total data space into subsets separated by so-called decision boundaries. All pixels that fall within a volume surrounded by such decision boundaries are then labeled as belonging to a single class.

Two major approaches are used in classifying images; supervised and unsupervised classifications. In the case of supervised classification, a user will specify so-called feature vectors to be used in the comparison process. These vectors can be thought of as defining the centroids of the decision volumes that are separated by the decision boundaries. These feature vectors can be extracted from the image to be classified, or could come from a library of radar signatures either measured in the laboratory or in the field. In the case of unsupervised classification, the computer is allowed to find the feature vectors without help from an image analyst. In the simplest form, known as the *K-means algorithm*, *K* feature vectors are typically selected at random from the data space.

A different way to look at classification approaches is whether they are based purely on the data available in the image, or whether the analyst is adding information based on physics. The former is often referred to as a statistical approach, or more correctly, a data driven approach. The analyst relies purely on the data and the statistics associated with the data to segment the image into classes. In the physics-based approach, the analyst compares the data in the image to known characteristics based on physics. For example, one might be interested in knowing which parts of a scene are covered with vegetation. Based on the physics of vegetation scattering models, one might conclude that such areas would exhibit relatively large entropy or RVI values, and use those parameter with some threshold value to decide whether a pixel should be labeled as vegetated or not.

3.4.1 Supervised Classification

The most popular supervised classification technique is the Bayes classifier. The basic principle of the Bayes classifier can be described using Eqs. (3.4-1) and (3.4-2). If a land class is denoted by ϖ_i where i = 1, 2, …., N classes, a pixel A can be classified as

$$A \in \varpi_i \quad \text{if} \quad P(\varpi_i \mid A) > P(\varpi_j \mid A)\,. \tag{3.4-1}$$

However, it is much easier to evaluate $P(A \mid \varpi_i)$ than $P(\varpi_i \mid A)$ using ground-truth data. Therefore, we transform Eq. (3.4-1) using the Bayes theorem as

$$A \in \varpi_i \quad \text{if} \quad P(A|\varpi_i)P(\varpi_i) > P(A|\varpi_j)P(\varpi_j). \tag{3.4-2}$$

It is convenient to transform Eq. (3.4-2) using a distance measure $d_i(A)$ as

$$A \in \varpi_i \quad \text{if} \quad d_i(A) < d_j(A). \tag{3.4-3}$$

where, typically

$$d_i(A) = -\ln P(A|\varpi_i)P(\varpi_i). \tag{3.4-4}$$

Using Eqs. (3.4-3) and (3.4-4), pixel A is classified to be ϖ_i if the distance measure of ϖ_i (= $-\ln P(A|\varpi_i)P(\varpi_i)$) is the shortest among all possible classes. As the simplest example, if the conditional probability density function (PDF) derived from the ground truth data is given by the normal distribution as

$$P(A|\varpi_i) = \frac{1}{\sqrt{2\pi}s_i}\exp\left[-\left(A-\mu_i\right)^2 / 2s_i^2\right]. \tag{3.4-5}$$

where μ_i and s_i are the mean value and the standard deviation of the conditional PDF of the class ϖ_i. The distance measure can be calculated as

$$d_i(A) = \left[\left(A-\mu_i\right)/\sqrt{2}s_i\right]^2 + \ln(s_i) - \ln P(\varpi_i). \tag{3.4-6}$$

The most challenging step of evaluating Eq. (3.4-6) is to estimate $P(\varpi_i)$ for all classes before we apply a classification method to polarimetric SAR data. One commonly used assumption is that an equal probability is assigned to all classes. Then, Eq. (3.4-6) can be simplified as

$$d_i(A) = \left[\left(A-\mu_i\right)/\sqrt{2}s_i\right]^2 + \ln(s_i). \tag{3.4-7}$$

From Eq. (3.4-7), the distance measure becomes smaller if the distance (normalized by the standard deviation of PDF) between the measured pixel value and the PDF mean value is smaller.

For the fully polarimetric SAR case, Lim et al. (1989) [19] calculated the distance measure as

$$d_i(\mathbf{X}) = \mathbf{X}^+ \left[\mathbf{C}\right]_i^{-1} \mathbf{X} + \ln\left|\left[\mathbf{C}\right]_i\right| - \ln P(\varpi_i). \tag{3.4-8}$$

where the complex vector $\mathbf{X}$ is given by

$$\mathbf{X} = \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vv} \end{bmatrix}; \; [\mathbf{C}]_i = \left\langle \mathbf{X}_i \mathbf{X}_i^{+} \right\rangle. \tag{3.4-9}$$

Here $[\mathbf{C}]_i$ is the covariance matrix of the assumed class feature for the *i*th class. The conditional PDF is given by

$$P(\mathbf{X} \mid \varpi_i) = \frac{1}{\pi^3 \left|[\mathbf{C}]_i\right|} \exp\left(-\mathbf{X}^{+} \cdot [\mathbf{C}]_i^{-1} \cdot \mathbf{X}\right). \tag{3.4-10}$$

The complex Gaussian distribution given in Eq. (3.4-10) is a good approximation for single-look polarimetric data. Lee et al. [20] showed that for multi-look polarimetric SAR data, the complex Wishart distribution is a better approximation. Using this distribution, Lee *et al.* derived the following distance measure

$$d_i(\mathbf{Z}) = n\left[\ln\left|[\mathbf{C}]_i\right| + Tr\left(\mathbf{C}_i^{-1}\mathbf{Z}\right) - \ln P(\varpi_i)\right]. \tag{3.4-11}$$

Here $\mathbf{Z}$ represents the covariance matrix of the pixel being classified, and *n* is the number of looks in the image.

To implement Eq. (3.4-8) or (3.4-11), one must evaluate all $P(\varpi_i)$ or assume equal *a priori* probabilities for all the classes. In the absence of any additional information equal *a priori* probabilities are often assumed. Figure 3-28 shows an example of a supervised Bayesian classification of the Black Forest image we analyzed earlier in the chapter. Here we selected three training sets, one in the agricultural area (displayed in blue), one in the urban area (displayed in red), and one in the forested area (displayed in green). The image on the left in Fig. 3-28 shows the result using the full covariance matrix, including the absolute radar cross-section assuming equal *a-priori* probabilities. The results show that most of the agricultural area is in fact identified as similar to that training area, and similarly for the other two classes. A closer look at the top right-hand corner of the image on the left shows a diagonal line of red pixels, meaning these pixels were classified as being similar to the urban training area. The scattering in this area is not similar to double reflects, however. These pixels were placed in this class mainly because of their large absolute cross-sections.

The image on the right in Fig. 3-28 repeats the calculation, but this time using the normalized covariance matrices. In other words, all brightness information was discarded, and the classification is done purely on the relative strength of

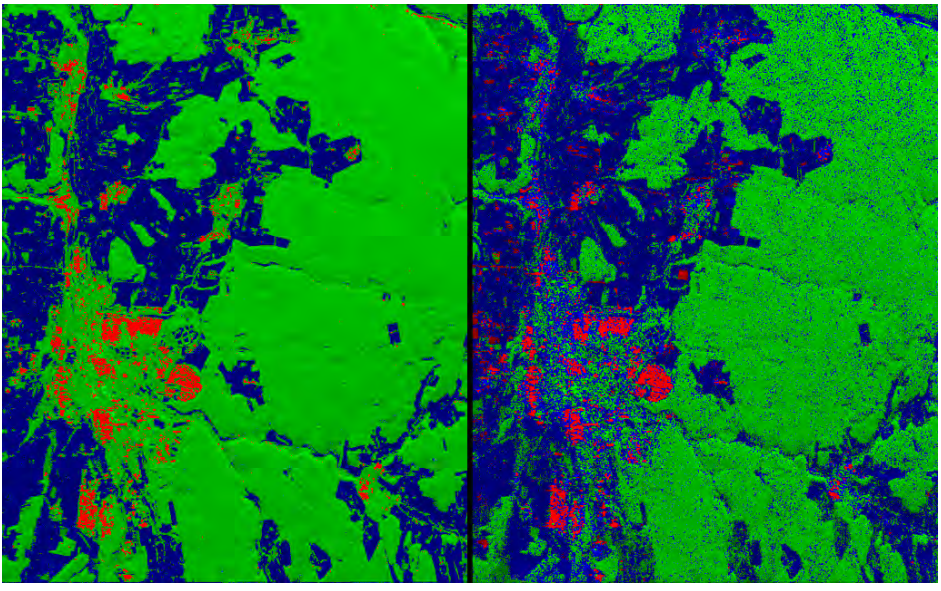


Fig. 3-28. Bayesian classifications of the L-band Black Forest image using training areas in the agricultural area (blue), urban area (red) and the forest (green). The image on the left uses the absolute brightness, while the one on the right uses normalized covariance matrices. See the text for a discussion.

the polarimetric information. Overall, the general results between the two images are similar. However, closer examination of the image on the right shows that several areas, particularly in the agricultural areas, are classified differently. Since the polarimetric information is more directly related to the scattering mechanisms, we expect the image on the right to show more details about the scattering mechanisms. For example, the diagonal line of red pixels in the upper right is now correctly identified as similar to the forest training area. Note, however, that the classification on the right appears noisier than the left.

In calculating the result shown in Fig. 3-28, we assumed equal *a-priori* probabilities for each of the classes. In the absence of additional information, this is a reasonable assumption. Van Zyl and Burnette (1992) [21] proposed an iterative method to evaluate $P(\varpi_i)$. First, they assume equal probabilities for all $P(\varpi_i)$. Then, the results from the first classification (equal probability assumption) are used to estimate all $P(\varpi_i)$. This is done by calculating the number of pixels in a pre-determined box that was placed in each class, divided by the total number of pixels in the box. The successive iteration improves the accuracy of the classification. Rignot and Chellappa (1992) [22] proposed a maximum *a posteriori* (MAP) estimate to segment polarimetric SAR data. The segmentation results using MAP showed a 10–20 percent improvement when

they are compared with the results using the maximum likelihood (ML) method under the equal probability assumption.

Figure 3-29 shows the result of applying the iterative scheme proposed by van Zyl and Burnette to estimate the probabilities for each class. On the left is the original Bayes classification with equal *a-priori* probabilities. On the right is the result after a single iteration using a 5 × 5 box to calculate the probability of finding a specific class in a given pixel. The result is a dramatic decrease in the apparent noise in the classification. Van Zyl and Burnette found that in general very few iterations are needed for the classification to converge [21].

In order to improve the classification accuracy, SAR images are often filtered to suppress speckle noise. Lee at el. (1999) [23] proposed a technique to preserve polarimetric properties without degrading the image quality. To avoid the crosstalk between polarimetric channels, each element of the covariance matrix was filtered independently. The filtering was performed by averaging the covariance matrix of neighboring pixels without deficiency of smearing the edges. To preserve the edge sharpness, the filtering was adaptively applied to a homogeneous area from selected neighboring pixels using edge-aligned windows [23]. Using this filtering, Lee at el. (1999) [23] reported that the classification accuracy was approximately doubled for five classes out of the

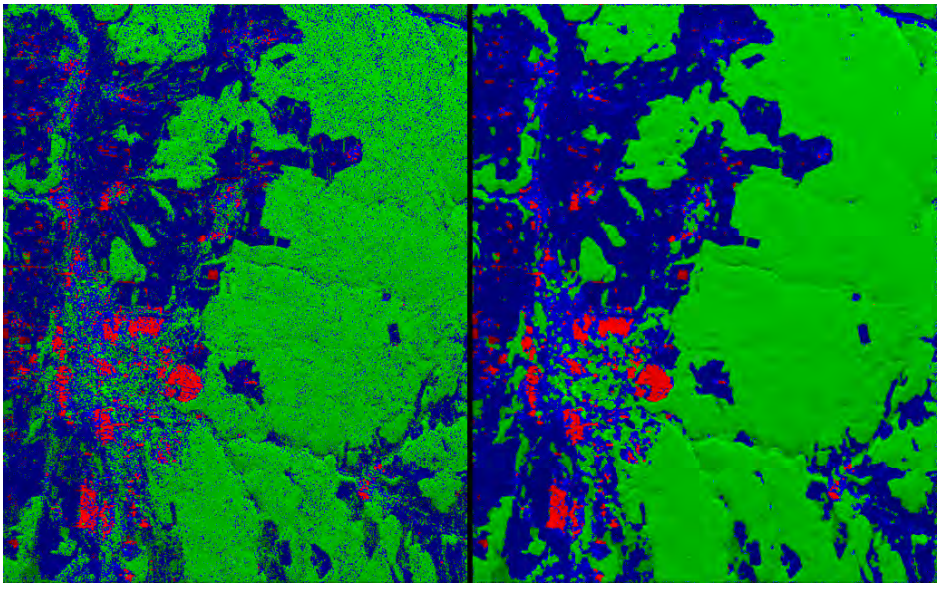


Fig. 3-29. The image on the left was classified using equal *a-priori* probabilities for all classes, and normalized covariance matrices. The image on the right illustrates the result of using adaptive *a-priori* probabilities as proposed by van Zyl and Burnette to perform the classification. The image on the left is used with a 5 × 5 box centered on each pixel to calculate the probability of finding a specific class in that pixel. The image on the right is the result of a single iteration.

total of seven classes when the results were compared with the ones using the data without any filtering.

3.4.2 Physics-Based Unsupervised Classification

Supervised classifications allow the analyst to select feature vectors from the image itself and then use them on the Bayesian algorithm to identify which pixels in the image are the closest to the selected features. One difficulty with this approach is that the results are dependent on how well the person picks the class training areas. If the selected areas are not very homogeneous, the results may not be very satisfying. A different type of approach relies on physics to segment a polarimetric image into different classes of scattering. This approach is completely unsupervised, and does not rely on a human to pick areas to use as the class centroids.

One of the earliest polarimetric unsupervised classification schemes was suggested by van Zyl [24]. Pixels in an image are compared to three scattering mechanisms, odd numbers of reflections, even numbers of reflections, and diffuse scattering based on the expected scattering parameters predicted by simple models. The results showed that bare surfaces are usually similar to the odd number of reflections model, while urban areas typically scatter consistent with an even-number-of-reflections model. Vegetated areas mostly exhibit diffuse scattering, particularly at longer wavelengths. This interpretation is confirmed by the results shown in Fig. 3-30.

This idea was taken further by Cloude and Pottier [25] using the entropy and average alpha angle discussed earlier to divide the data space into nine different regions. They showed that of these nine regions, eight are commonly found in polarimetric radar data. Based on this division of the data space, one can then segment the image based on which region in the data space each pixel falls into. As in the case of van Zyl, this algorithm requires no user input, and the segmentation is done automatically. Results show that the two classes with the largest entropy are similar to the diffuse scattering class defined by van Zyl [24]. The alpha-entropy classification further segments van Zyl's odd numbers of reflections and even number of reflections classes into several classes based on the amount of randomness in the scattering as measured by the entropy. See Fig. 3-30 for the results of this algorithm when applied to the L-band Black Forest image.

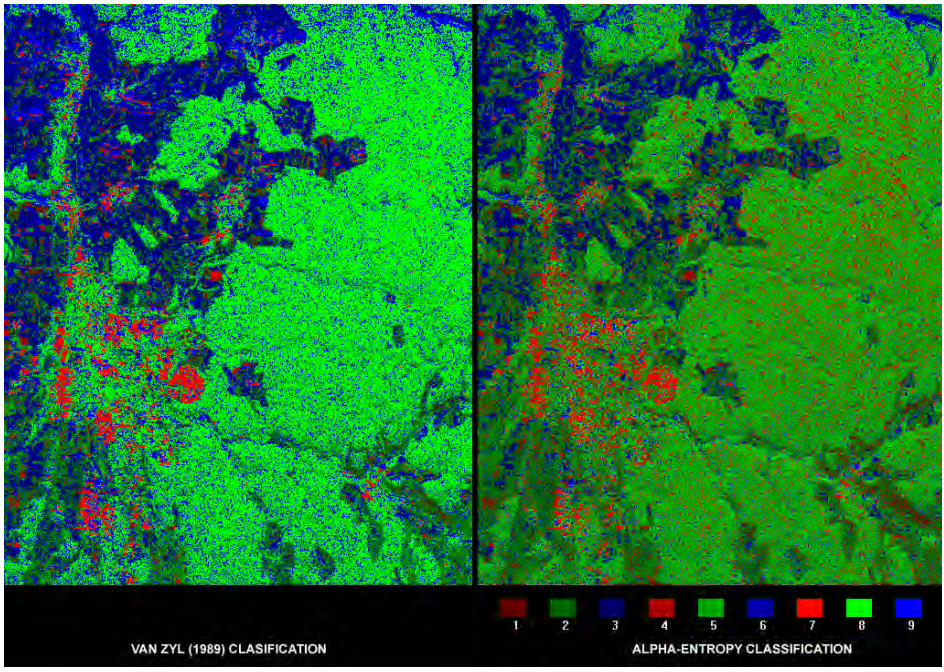


Fig. 3-30. Unsupervised classifications of the L-band image of the Black Forest using the algorithm by van Zyl [24] (left) and the alpha-entropy algorithm (right). The alpha-entropy image further divides the original three classes using the entropy information.

3.4.3 Combined Unsupervised and Bayes Classification Algorithms

As mentioned above, it is difficult to find reliable training data sets to estimate the statistical parameters required for the maximum likelihood classification. To overcome this difficulty, some authors have proposed using an unsupervised classification technique to produce initial training data sets. Lee at el. (1999) [26] proposed an unsupervised classification technique that combines a polarimetric decomposition method (unsupervised) and the complex Wishart classifier [20] (supervised). The "alpha-entropy" algorithm [25, 27] based on the polarimetric decomposition technique was used to provide the initial training data set for the eight zones defined by the alpha angle and the entropy. This training data set was used as an input for the Wishart classifier. The classification result was improved by iteration of this process. When the iteration process satisfies some pre-determined termination criteria, the classification result is finalized.

More recently, Lee at el. (2004) [27] proposed a different approach to combine both unsupervised and supervised classification techniques. In this approach, they applied the model-based decomposition algorithm by Freeman and Durden [9] to the polarimetric SAR data to provide the classification result for three

categories: surface, double bounce, and volume scattering mechanisms. This division was accomplished based on the dominant power associate with each scattering mechanism. Lee et al. (2004) [27] divided each category into 30 clusters based on the backscattering power of their dominant scattering mechanism. The initial clusters were merged based on the between-cluster Wishart distance [27] to form the classes. The Wishart classification method was applied to the polarimetric data iteratively until this process converged. These combined classification approaches usually provide more accurate classification results than unsupervised approach can without requiring the ground truth data.

Figure 3-31 illustrates the idea behind a combined unsupervised and Bayesian classification algorithm. In this example, we used three model-derived covariance matrices to perform the initial unsupervised classification. The three

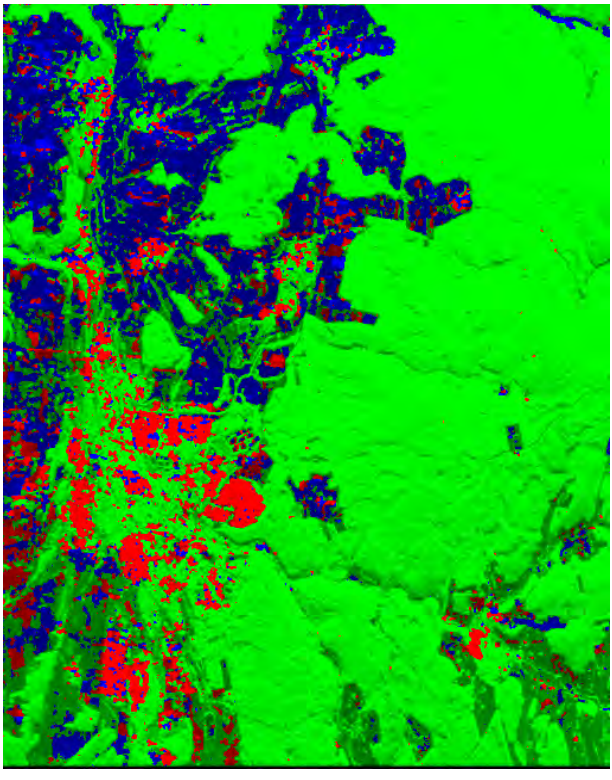


Fig. 3-31. Results using three model-derived covariance matrices to perform an unsupervised classification of the L-band Black Forest data. Pixels displayed in red are similar to a double reflection expected from two dielectric surfaces, those in blue are similar to a bare rough surface, and those in green are similar to a cloud of uniformly randomly oriented thin dielectric cylinders.

models are that of a rough bare surface, a dihedral corner reflector with dielectric surfaces, and a uniformly randomly oriented cloud of thin cylinders. These three model covariance matrices are used in an unsupervised classification using normalized covariance matrices. Once the initial classification is performed, the average covariance matrix for each class is calculated and used in the next step. This process is iterated until the classification converges in the sense that less than .001 percent of the pixels still change class.

We note that there is great similarity between this unsupervised classification and the supervised classification using normalized covariance matrices shown on the right in Fig. 3-28. The main difference is that more areas in the agricultural areas are classified as being more similar to the dielectric dihedral reflection.

One drawback of all classification algorithms is that they force a pixel to belong to one class only. The decision is based on which class most closely resembles the observed scatter. Unfortunately, scattering is seldom purely one mechanism. As we shall see in Chapter 5, scattering from vegetated areas often is a mixture of many different scattering mechanisms. Forcing a pixel to belong to one class ignores this reality. We therefore believe that classification results should be interpreted with care. Our personal preference is to use techniques that preserve the relative fractions of different scattering mechanisms, such as the Pauli matrix decomposition or one of the model-based decompositions.

3.5 Polarimetric SAR Interferometry

SAR interferometry refers to a class of techniques where additional information is extracted from SAR images that are acquired from different vantage points, or at different times. Various implementations allow three types of information to be extracted: 1) topography, 2) surface velocity, and 3) surface deformation. If two SAR images are acquired from slightly different viewing geometries, information about the topography of the surface can be inferred. Conversely, if images are taken at slightly different times, a map of surface velocities can be produced. Finally, if sets of interferometric images are combined, subtle changes in the scene can be measured with extremely high accuracy. These techniques are summarized in more detail in several references [28, 29, 30].

SAR interferometry was first demonstrated by Graham (1974) [31], who demonstrated a pattern of nulls or interference fringes by vectorally adding the signals received from two SAR antennas; one physically situated above the other. Later, Zebker and Goldstein (1986) [32] demonstrated that these interference fringes can be formed after SAR processing of the individual

images if both the amplitude and the phase of the radar images are preserved during the processing.

The basic principles of interferometry can be explained using the geometry shown in Fig. 3-32. Two radar return signals ($E(A_1)$ and $E(A_2)$) are recorded by the two interferometric antennas (A_1 and A_2 shown in Fig. 3-32) separated by the baseline B. A 2x2 Hermitian matrix can be formed using $E(A_1)$ and $E(A_2)$ as

$$[I_n] = \left\langle \begin{bmatrix} E(A_1) \\ E(A_2) \end{bmatrix} \begin{bmatrix} E(A_1)^* & E(A_2)^* \end{bmatrix} \right\rangle = \begin{pmatrix} \left\langle E(A_1)E(A_1)^* \right\rangle & \left\langle E(A_1)E(A_2)^* \right\rangle \\ \left\langle E(A_2)E(A_1)^* \right\rangle & \left\langle E(A_2)E(A_2)^* \right\rangle \end{pmatrix}. \tag{3.5-1}$$

The diagonal components represent two conventional SAR images collected by the interferometric SAR system. The interferometric SAR information can be extracted from the off diagonal component $\left\langle E(A_1)E(A_2)^* \right\rangle$. The phase of this term (known as interferogram) is defined as

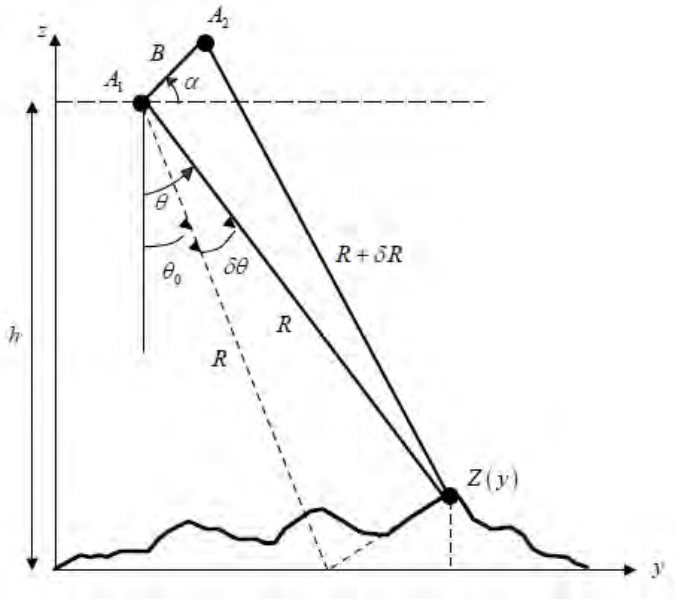


Fig. 3-32. Basic interferometric radar geometry. The path length difference between the signals measured at each of the two antennas is a function of the elevation of the scatterer. (From Elachi and van Zyl (2006) [28])

$$\phi_I = \arg\left\{\left\langle E(A_1)E(A_2)^*\right\rangle\right\}. \tag{3.5-2}$$

Radar interferometry can be implemented in two different ways. In so-called single-pass interferometers, the two antennas are physically separated, but on the same platform. In the case of airborne SAR systems, the two interferometric antennas might be mounted in different places on the fuselage, or under the wings of the aircraft. In the case of the Shuttle Radar Topography Mission (SRTM) that was flown on the Space Shuttle Endeavour in 2000, a 60 meter (m) long boom was used to separate the two antennas. Single-pass interferometers acquire the images from both ends of the baseline at the same time.

An alternative way to implement radar interferometry is to use images from different orbits in the case of spaceborne systems, or in the case of airborne systems, from different flight lines. This implementation is known as repeat-track interferometry. In this case, images at each end of the baseline are acquired at different times, with time separations varying from minutes in the case of airborne systems to days to even years in the case of spaceborne systems. A slight variation of the repeat-track implementation is to use two satellites that follow each other in slightly different orbits. This is knows as tandem interferometry. The time difference for acquiring images is typically seconds to minutes, depending on how far apart the two satellites are positioned.

The scattering center of a pixel can be located using the law of cosines on the triangle formed by the two antennas and the scattering center of a pixel as

$$(R+\delta R)^2 = R^2 + B^2 - 2BR\cos\left(\pi/2 - \theta + \alpha\right). \tag{3.5-3}$$

The variables in Eq. (3.5-3) are shown in Fig. 3-32. If we assume that $R >> B$, (a very good assumption for most interferometers) one finds that

$$\delta R \approx -B\sin(\theta - \alpha). \tag{3.5-4}$$

The radar system does not measure the path length difference explicitly, however. Instead, what is measured, is an interferometric phase difference that is related to the path length difference through

$$\phi_T = \phi_I + 2\pi N = \frac{a2\pi}{\lambda}\delta R = -\frac{a2\pi}{\lambda}B\sin(\theta - \alpha). \tag{3.5-5}$$

where $a = 1$ for the case where signals are transmitted out of one antenna and received through both at the same time, and $a = 2$ for the case where the signal

is alternately transmitted and received through one of the two antennas only. In order to reconstruct the topography, the integer N must be determined using a phase unwrapping technique [33, 34].

From Fig. 3-32, it also follows that the elevation of the point being imaged is given by

$$z(y) = h - R\cos\theta = h - R\cos\left[\alpha - \sin^{-1}\left(\frac{\lambda\phi_T}{a2\pi B}\right)\right]. \tag{3.5-6}$$

with h denoting the height of the reference antenna above the reference plane with respect to which elevations are quoted. As shown in Eq. (3.5-6), the interferometric phase (ϕ_T) provides the information on the location of the scattering center. Since the return signals from two interferometric antennas are collected in a different geometry, there is decorrelation between two signals. This decorrelation can provide the additional information on scattering objects to be imaged.

The interferometric coherence is defined as

$$\gamma_{INT} = \frac{\left|\left\langle E(A_1)E(A_2)^*\right\rangle\right|}{\sqrt{\left\langle E(A_1)E(A_1)^*\right\rangle\left\langle E(A_2)E(A_2)^*\right\rangle}} = \gamma_{SNR}\gamma_T\gamma_B. \tag{3.5-7}$$

The interferometric decorrelation is composed of three contributions: 1) additive noise, 2) slightly different imaging geometry of two antennas, and 3) temporal changes of scattering objects. As shown in Eq. (3.5-7), γ_{SNR} includes decorrelation due to additive noise. If both interferometric channels have the same SNR, γ_{SNR} can be written as

$$\gamma_{SNR} = \frac{1}{1+\frac{1}{SNR}}. \tag{3.5-8}$$

The temporal correlation (γ_T) represents the scattering geometry change in time [35]. If the interferometric signals are collected at the same time, γ_T =1. The baseline decorrelation (1-γ_B) is due to the speckle difference due to the fact that two interferometric signals are collected in a slight different imaging geometry. This term (γ_B) includes the information on scattering objects. If the scattering object is a point target, there is no baseline decorrelation. The baseline decorrelation is a key parameter to understand the scattering

characteristic of each pixel. The baseline correlation (γ_B) can be estimated from the measured correlation coefficient (γ) after removing the effect of γ_{SNR} and γ_T as shown in (3.5-7). The expression for the baseline correlation to estimate the scattering characteristic of each pixel can be found in [29].

Electromagnetic wave propagation is by nature a vector phenomenon. Therefore, in order to capture the complete information about the scattering process, interferometric measurements should really be made in the full polarimetric implementation of a radar system. In this case, there are really three different measurements being made at the same time. First, there are the two polarimetric radar measurements at each end of the baseline, represented below by the two covariance matrices $[\mathbf{C}_{11}]$ and $[\mathbf{C}_{22}]$. Since the baseline is generally short compared to the distance to the scene, these two measurements can be expected to be nearly identical, except for the very small change in the angle of incidence from one end of the baseline to the other. The third measurement, of course, is the full vector interferogram as opposed to the scalar implementation described earlier.

The vector interferogram, which is the complex cross-correlation of the signal from one end of the baseline with that from the other end of the baseline, can be described as

$$V_1 V_2^* = \tilde{\mathbf{A}}_1 \mathbf{T}_1 \tilde{\mathbf{T}}_2^* \mathbf{A}_2^* = \mathbf{A}_1 \cdot [\mathbf{C}_{12}] \mathbf{A}_2^* . \tag{3.5-9}$$

The antenna ($\mathbf{A}$) and scatterer ($\mathbf{T}$) vectors are given in Eq. (3.1-1). The complex correlation of the two signals after averaging is

$$\mu = \frac{\langle V_1 V_2^* \rangle}{\sqrt{\langle V_1 V_1^* \rangle \langle V_2 V_2^* \rangle}} = \frac{\mathbf{A}_1 \cdot \langle [\mathbf{C}_{12}] \rangle \mathbf{A}_2^*}{\sqrt{\left(\mathbf{A}_1 \cdot \langle [\mathbf{C}_{11}] \rangle \mathbf{A}_1^*\right)\left(\mathbf{A}_2 \cdot \langle [\mathbf{C}_{22}] \rangle \mathbf{A}_2^*\right)}} . \tag{3.5-10}$$

Notice that the interferometric phase is a function of the antenna polarization vectors as shown in Eq. (3.5-9).

$$\varphi_{\text{int}} = \arg\left(\mathbf{A}_1 \cdot \langle [\mathbf{C}_{12}] \rangle \mathbf{A}_2^*\right) . \tag{3.5-11}$$

Notice that the interferometric phase will be used to locate the scattering center of different antenna polarization combinations. In addition, the polarimetric correlation coefficient can be used to estimate the scattering characteristics of each pixel. Using this formulation, Cloude and Papathanassiou (1998) [36] showed, using repeat-track SIR-C interferometric data, that polarization diversity can be used successfully to optimize the correlation between images.

They also showed significant differences in the measured elevation in forested areas when using polarization optimization. At present, polarimetric interferometry is a very active research area [37, 38, 39, 40, 41, 42]. Unfortunately, progress is hampered severely by lack of availability of well-calibrated data, as only a hand-full of radar systems have been upgraded to full polarimetric interferometry capability.

To illustrate the information content of polarimetric interferometry, we used data acquired in the repeat-track interferometry mode using the SIR-C system when it flew on the Space Shuttle in October 1994. The data we use were acquired over the Mahantango Watershed near Harrisburg, Pennsylvania. An L-band total power image of part of the area is shown in Fig. 3-33. The Mahantango watershed is part of the Valley and Ridge Physiographic Province of eastern Pennsylvania. The area is characterized by forested ridge tops, while

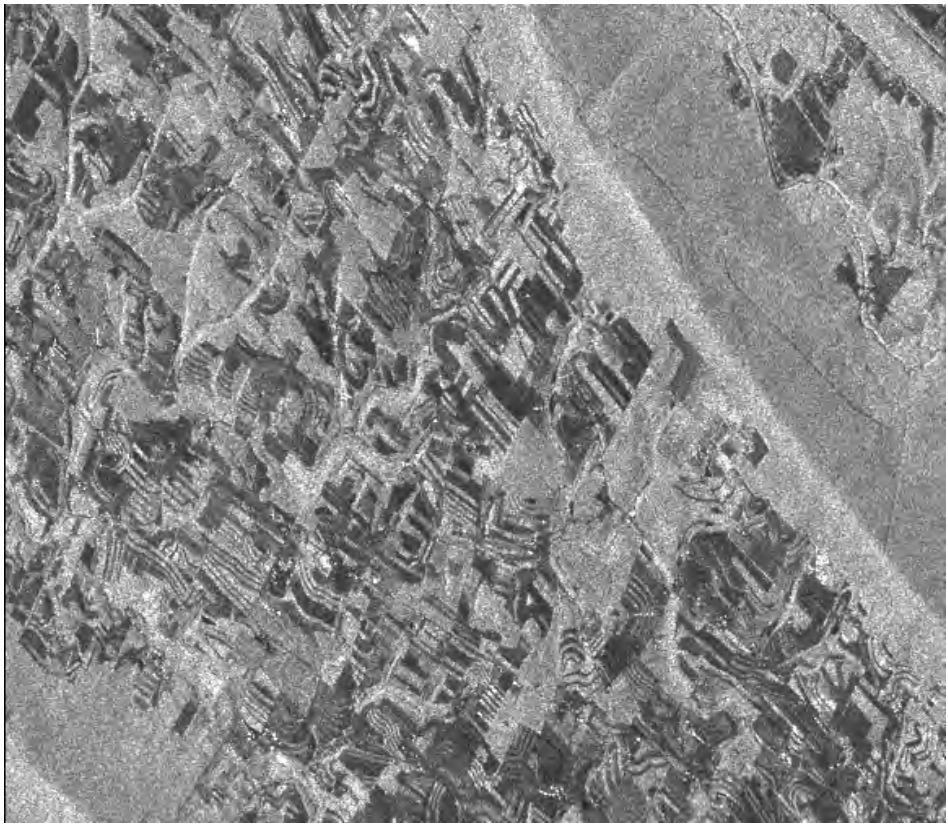


Fig. 3-33. L-band total power image of a portion of the Mahantango watershed in Pennsylvania. The darker areas are agricultural fields, and the brighter grey tones are forested ridge tops.

the valley areas are typically used for agriculture. Approximately 55 percent of the area is forested, and about 45 percent of the area is used for cropland.

To illustrate the additional information contained in the polarimetric data, we choose to display differential interferograms using the image on the left in Fig. 3-34. The differential interferograms are constructed as follows. First we construct the HH interferogram as the reference. We then construct interferograms that are effectively the phase of each of the covariance matrix elements in Eq. (3.5-9). We then subtract the HH interferogram phase from each of these to form nine differential interferograms. These are shown in Fig. 3-34.

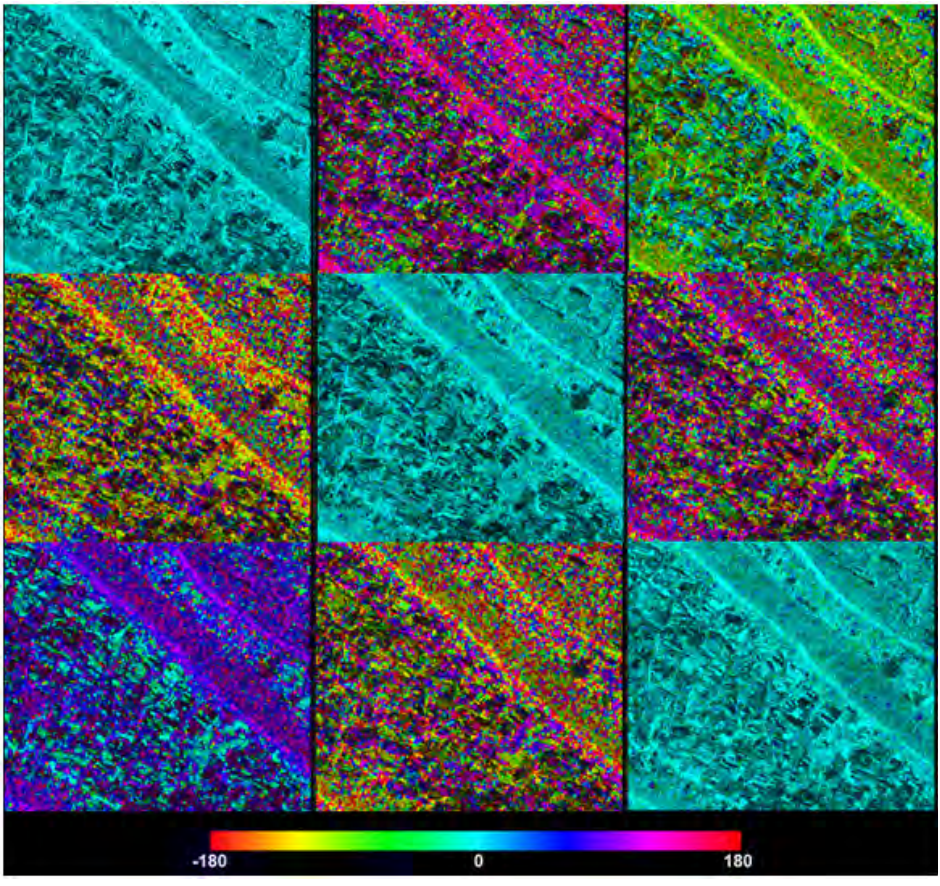


Fig. 3-34. Differential interferograms constructed using the polarimetric information acquired over the Mahantango watershed. The top row represent HH polarization on pass one, and HH on pass two (left), HV on pass two (middle) and VV on pass two (right). The second row is the same, except for HV on pass one, and the third row represents VV on pass one. See the text for how these were generated. Images across the upper left lower right diagonal are the complex conjugates of each other, hence the different colors.

The most striking is the image in the upper right in Fig. 3-34, which displays the differential phase of using HH polarization at one antenna and VV polarization at the other. The image shows that the agricultural areas in general have almost zero phase (which means they have the same phase in the VV and the HH interferogram). The forested areas, however, shows a significant phase angle, in this case near 90 deg.

3.6 Summary

In this chapter we introduced several more advanced polarimetric concepts. The vector-matrix duality allows us to understand decomposition algorithms in more detail. We also discussed several parameters based on the eigenvalues of the covariance matrix. These parameters provide useful information concerning the amount of randomness in the observed scatter, as well as potentially what the actual scattering mechanisms might be.

Finally, we examined different approaches for interpreting the scattering mechanisms in an image. Using orthogonal bases, we can derive unique decompositions. These decompositions, while unique, are not straightforward to interpret. In particular, the decomposition based on the eigenvectors of the covariance matrix suffers from the fact that the basis in which the decomposition is done varies from pixel to pixel in the image. This further complicates the interpretation. At the other end of the spectrum are model-based decompositions. These are not unique, and picking the appropriate solution is not obvious, but their interpretation is more straightforward. We showed that some of the popular model-based decompositions suffer from serious limitations imposed by the assumptions of how the observations are to be decomposed; these assumptions lead to results that are non-physical in the sense that negative powers could be generated. We showed that this limitation can easily be removed using a simple check based on the eigenvalues of the covariance matrix. A simple hybrid approach can be implemented that corrects these limitations. This approach is then easily extended to show a simple way to find the best model to fit the observed canopy scattering.

It is important to remember that all these decomposition techniques are simply tools to make interpretation of the observed scattering easier. The strength of the pure eigenvalue and eigenvector approaches is that the answers are unique, and no assumptions are required to perform the decomposition. The interpretation of the results, however, requires an interpretation of the basis vectors in terms of scattering mechanisms, which might not be unique. The model-based decompositions provide an easy interpretation. But this easy interpretation assumes that the models are indeed applicable to the observations, which might not be the case. Furthermore, the results are not

unique in the sense that many different model combinations could be used. In the final analysis, which tool is used depends on personal preference.

References and Further Reading

[1] S. R. Cloude, "Uniqueness of target decomposition theorems in radar polarimetry," *Direct and Inverse Methods in Radar Polarimetry,* Part 1, *NATO-ARW,* W.M. Boerner, L. A. Cram, W. A. Holm, D. E. Stein, W. Wiebeck, W. Keydel, D. Giuli, D. T. Gjessing, F. A. Molinet, and H. Brand, eds., Kluwer Academic Publishers, Norwell, Massachusetts, pp. 267–296, 1992.

[2] S. R. Cloude and E. Pottier, "The Concept of Polarization Entropy in Optical Scattering," *Optical Engineering,* vol. 34, no. 6, ISSN 0091-3286, pp. 1599–1610, 1995.

[3] S. L. Durden, J. J. van Zyl and H. A. Zebker, "The Unpolarized Component in Polarimetric Radar Observations of Forested Areas," *IEEE Transactions on Geoscience and Remote Sensing,* vol. GE-28, no. 2, pp 268–271, 1990.

[4] Y. Kim and J. van Zyl, "Comparison of Forest Estimation Techniques Using SAR Data," *Proceedings of IGARSS2001 Conference,* Sydney, Australia, IEEE (CDROM), 2001.

[5] J. J. van Zyl, "The Effect of Topography on Radar Scattering from Vegetated Areas," *IEEE Transactions on Geoscience and Remote Sensing,* vol. GE-31, no. 1, pp. 153–160, 1993.

[6] S. R. Cloude and E. Pottier, "A Review of Target Decomposition Theorems in Radar Polarimetry," *IEEE Transactions on Geoscience and Remote Sensing,* vol. GE-34, no. 2, pp. 498–518, 1996.

[7] E. Krogager, *Aspects of Polarimetric Radar Imaging*, Ph.D. thesis, Technical University of Denmark, Copenhagen, 1993.

[8] M. Borgeaud, R. T. Shin, and J. A. Kong, "Theoretical models for polarimetric radar clutter," *Journal of Electromagnetic Waves and Applications,* vol. 1, pp. 73–91, 1987.

[9] A. Freeman and S. Durden, "A Three Component Scattering Model for Polarimetric SAR Data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. GE-36, no. 3, pp. 963–973, 1998.

[10] Y. Yamaguchi, T. Moriyama, M., Ishido, and H. Yamada, "Four-Component Scattering Model for Polarimetric SAR Image Decomposition," *IEEE Transactions on Geoscience and Remote Sensing,* vol. GE-43, no. 8, pp. 1699–1706, 2005.

[11] D. L. Schuler, J-S. Lee, and G. De Grandi, "Measurement of Topography Using Polarimetric SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-34, no. 5, pp. 1266–1277, 1996.

[12] J-S. Lee, D. L. Schuler, and T. L. Ainsworth, "Polarimetric SAR Data Compensation for Terrain Azimuth Slope Variation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-38, no. 5, pp. 2153–2163, 2000.

[13] J. J. van Zyl, Y. Kim, and M. Arii, "Requirements for Model-Based Polarimetric Decompositions," *Proceedings of IGARSS2008 Conf.*, Vol. 5, IEEE International, Boston, Massachusetts, pp. 417–420, 2008.

[14] J. J. van Zyl, M. Arii, and Y. Kim, "Model-Based Decomposition of Polarimetric SAR Covariance Matrices Constrained for Non-Negative Eigenvalues," *IEEE Transactions on Geoscience and Remote Sensing* (in press), 2010.

[15] J. J. van Zyl, "Application of Cloude's Target Decomposition Theorem to Polarimetric Imaging Radar Data," *SPIE, Radar Polarimetry*, vol. 1748, pp. 184–212, 1992.

[16] M. Arii, J. J. van Zyl, and Y. Kim, "Adaptive Model-Based Decomposition of Polarimetric SAR Covariance Matrices," *IEEE Transactions on Geoscience and Remote Sensing* (in press), 2010.

[17] M. Arii, *Soil Moisture Retrieval under Vegetation Using Polarimetric Radar*, Ph.D. dissertation, California Institute of Technology, Pasadena, California, pp. 68–101, 2009.

[18] M. Arii, J. J. van Zyl, and Y. Kim, "A General Characterization for Polarimetric Scattering from Vegetation Canopies," *IEEE Transactions on Geoscience and Remote Sensing* (in press), 2010.

[19] H. H. Lim, A. A. Swartz, H. A. Yueh, J. A. Kong, and R. T. Shin, "Classification of Earth Terrain Using Polarimetric Synthetic Aperture Radar Images," *Journal of Geophysical Research*, vol. 94, no. B6, pp. 7049–7057, 1989.

[20] J. S. Lee, M. R. Grunes, and R. Kwok, "Classification of Multi-Look Polarimetric SAR Imagery Based on Complex Wishart Distribution," *International Journal of Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, 1994.

[21] J. J. van Zyl and C. F. Burnette, "Bayesian Classification of Polarimetric SAR Images Using Adaptive *A Priori* Probabilities," *International Journal of Remote Sensing*, vol. 13, no. 5, pp. 835–840, 1992.

[22] E. Rignot and R. Chellappa, "Segmentation of Polarimetric Synthetic Aperture Radar Data," *IEEE Transactions on Image Processing*, vol. 1, no. 3, pp. 281–300, 1992.

[23] J. S. Lee, M. R. Grunes, and G. De Grandi, "Polarimetric SAR Speckle Filtering and Its Implication for Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2363–2373, 1999.

[24] J. J. van Zyl, "Unsupervised Classification of Scattering Behavior Using Radar Polarimetry Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27, pp. 36–45, 1989.

[25] S. Cloude and E. Pottier, "Entropy Based Classification Scheme for Land Applications," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 68–78, 1997.

[26] J. S. Lee, M. R. Grunes, T. L. Ainsworth, L. J. Du, D. L. Schuler, and S. R. Cloude, "Unsupervised Classification Using Polarimetric Decomposition and the Complex Wishart Classifier," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2249–2258, 1999.

[27] J. S. Lee, M. R. Grunes, E. Pottier, L. Ferro-Famil, "Unsupervised Terrain Classification Preserving Polarimetric Scattering Characteristics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 4, pp. 722–731, 2004.

[28] C. Elachi and J. van Zyl, *Introduction to the Physics and Techniques of Remote Sensing 2nd Edition*, Wiley-Interscience, Hoboken, New Jersey, pp. 268–283, 2006.

[29] E. Rodriguez, J. M. Martin, "Theory and Design of Interferometric Synthetic Aperture Radars," *IEE Proceedings-F*, vol. 139, No.2, pp. 147–159, 1992.

[30] P. A. Rosen, S. Hensley, I. R. Joughin, F. K. Li, S. N. Madsen, E. Rodríguez, and R. M. Goldstein, "Synthetic Aperture Radar Interferometry," *Proceedings of the IEEE*, vol. 88, pp. 333–382, 2000.

[31] L. C. Graham, "Synthetic Interferometer Radar for Topographic Mapping," *Proceedings of the IEEE*, vol. 62, pp. 763–768, 1974.

[32] H. Zebker and R. Goldstein, "Topographic Mapping from Interferometric SAR Observations," *Journal of Geophysical Research*, vol. 91, pp. 4993–4999, 1986.

[33] R. M. Goldstein, H. A. Zebker, and C. Werner, "Satellite Radar Interferometry: Two-Dimensional Phase Unwrapping." *Radio Sci.*, vol. 23, pp. 713–720, 1988.

[34] D. C. Ghiglia and L. A. Romero, "Direct Phase Estimation from Phase Differences Using Fast Elliptic Partial Differential Equation Solvers," *Optics Letters*, vol. 15, pp. 1107–1109, 1989.

[35] H. A. Zebker and J. Villasenor, "Decorrelation in Interferometric Radar Echoes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 950–959, 1992.

[36] S. R. Cloude and K. P. Papathanassiou, "Polarimetric SAR Interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, pp. 1551–1565, 1998.

[37] R. N. Treuhaft and P. R. Siqueira, "Vertical Structure of Vegetated Land Surfaces from Interferometric and Polarimetric Radar," *Radio Science*, vol. 35, pp. 141–177, 2000.

[38] A. Reigber and A. Moreira, "First Demonstration of Airborne SAR Tomography Using Multibaseline L-band Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 2142–2152, 2000.

[39] L. Sagues, J. M. Lopez-Sanchez, J. Fortuny, X. Fabregas, A. Broquetas, and A. Sieber, "Indoor Experiments on Polarimetric SAR Interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 671–684, 2000.

[40] S. Sauer, L. Ferro-Famil, A. Reigber, and E. Pottier, "Polarimetric Dual-Baseline InSAR Building Height Estimation at L-band," *IEEE Geoscience Remote Sensing Letters*, vol. 6, pp. 408-412, 2009.

[41] I. Hajnsek, F. Kugler, S. Lee, and K. P. Papathanassiou, "Tropical-Forest-Parameter Estimation by Means of Pol-InSAR: The INDREX-II Campaign," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 481–493, 2009.

[42] M. Neumann, L. Ferro-Famil, and A. Reigber, "Estimation of Forest Structure, Ground, and Canopy Layer Characteristics from Multibaseline Polarimetric Interferometric SAR Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 1086–1104, 2010.

Chapter 4
Polarimetric SAR Calibration

A polarimetric radar system measures the scattered power from the terrain being imaged. Before these measurements are useful for further quantitative analysis and comparison to models, the measured power values must be converted to normalized radar cross-sections. Polarimetric radar measurements require additional corrections: not only the amplitudes, but the relative phases between channels must also be calibrated. In this Chapter, we shall examine the steps required to calibrate polarimetric radar measurements in detail.

4.1 Polarimetric Radar System Model

To better understand the corrections we need to apply to a radar image during the calibration process, let us first discuss the implementation of such a radar system in the context of potential error sources. Figure 4-1 shows the block diagram of a typical polarimetric radar system.

In reality, the radar system is not perfect. For example, rather than transmitting the signals through an antenna with a single, pure polarization, there might be some leakage into the orthogonal polarization channel. To take this into account, we can write the actual transmitted electric field as

$$\mathbf{E}^t_{actual} = K_t(\gamma)\begin{pmatrix} 1 & \delta^t_1(\gamma) \\ \delta^t_2(\gamma) & f_t(\gamma) \end{pmatrix}\mathbf{E}^t_{ideal}\,. \tag{4.1-1}$$

In this equation, γ is the radar look angle that defines through which part of the antenna radiation pattern the signal is transmitted. During synthetic aperture radar (SAR) data calibration, a flat (airborne case) or spherical (spaceborne case) Earth is usually assumed for purposes of calculating the radar look angle.

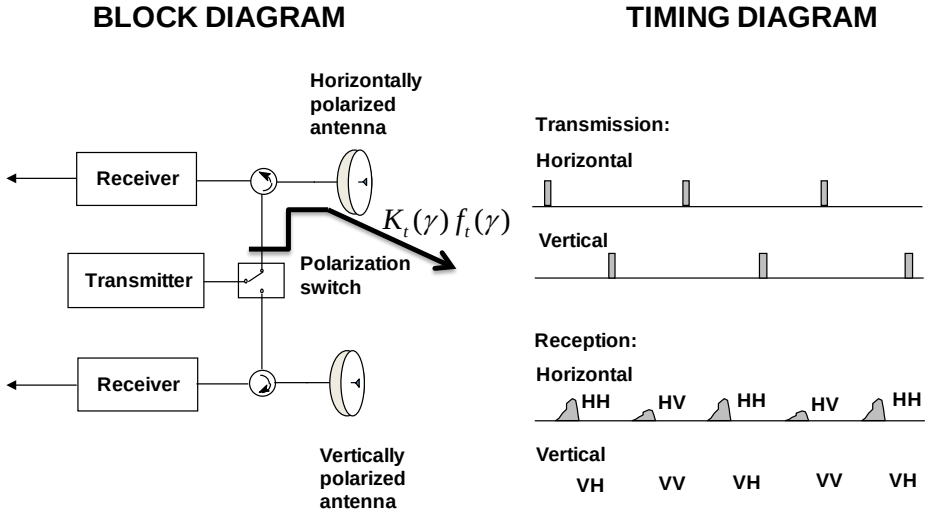


Fig. 4-1. Calibration of polarimetric radar. A polarimetric radar is implemented by alternatively transmitting signals out of horizontally and vertically polarized antennas and receiving at both polarizations simultaneously. Two pulses are needed to measure all the elements in the scattering matrix.

It is then assumed that the radar platform was at an altitude *h* above the Earth reference surface, as shown in Fig. 4-2. Under this flat-Earth assumption in the airborne case, the look angle γ and the incidence angle η for a given pixel are the same, as shown in Fig. 4-2. In the spaceborne case, the look angle and the incidence angle are not the same, but are related by

$$R_t \sin\eta = R_s \sin\gamma\,, \tag{4.1-2}$$

with R_s and R_t the magnitude of the spacecraft and target position vectors relative to the center of the Earth, respectively.

The slant range R to the scattering area is measured through the time delay between the transmission of the pulse and the receipt of the scattered power. To know the slant range accurately, however, we must calibrate possible time delays in the radar system, such as the delay between the transmit event as issued by the timing system and the actual transmit event. One could calibrate the slant range by imaging a calibration site with a number of point targets (discussed in more detail below) of accurately surveyed positions. From the relative positions of these objects in the image, one can accurately calculate the absolute slant range to each, from which the system delays can be estimated. The time interval between these calibrations depends on how stable the radar system is in time. For most radar systems, this measurement is only performed a few times per year.

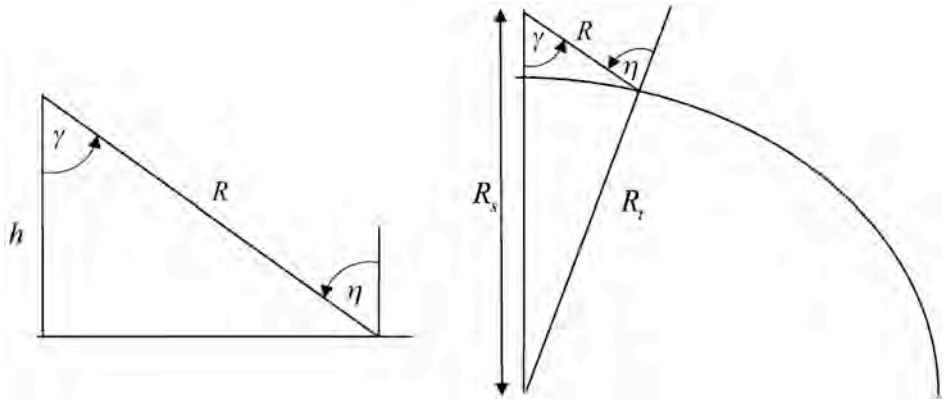


Fig. 4-2. Airborne (left) and spaceborne (right) imaging geometry usually assumed in SAR processing. In both cases, the radar platform is assumed to be at an elevation *h* above the reference plane.

Unless the radar is operated in an interferometric mode, however, the look angle γ is not measured directly and must be inferred indirectly. In most SAR processors, including the National Aeronautics and Space Administration/Jet Propulsion Laboratory (NASA/JPL) Airborne Synthetic Aperture Radar (AIRSAR) processors, this angle is calculated using the altitude of the radar platform, *h*, above the reference plane, the slant range R, and assuming a flat Earth. From Fig. 4-2 it follows that if this assumption is correct, one can write

$$\gamma = \cos^{-1}\left(h/R\right). \tag{4.1-3}$$

The equivalent expression for the spaceborne case assuming a spherical Earth is as follows:

$$\gamma = \cos^{-1}\left(\frac{R^2 + R_s^2 - R_t^2}{2RR_s}\right). \tag{4.1-4}$$

The antenna radiation pattern that is of interest here is the so-called elevation pattern. This is measured in a plane that is orthogonal to the direction in which the radar is traveling. Returning to Eq. (4.1-1), we note that the cross-talk coefficients denoted by the δ are not only a function of the actual antenna; that is, if the radar platform flies with a pitch that is not zero, the antenna coordinates are rotated by the pitch angle. This will then appear as a cross-talk in the transmitted signal.

The quantity f_t in Eq. (4.1-1) is the result of differences in antenna patterns for the two polarizations, as well as potential gain and path length differences in the two transmit paths, including the polarization switch. The latter effects

result in an overall complex constant multiplier f_t, that is a function of the look angle.

This transmitted signal now propagates to the terrain being imaged. Unless the signal propagates through a plasma, we can consider this propagation to be represented by an identity matrix. In the case of a plasma, the signal will suffer Faraday rotation, which we can represent by a rotation matrix. For now, we shall assume no rotation. After reflection from the terrain, the signal propagates back to the radar. This reflection is represented by the complex scattering matrix. Upon reception, the system might again introduce distortions similar to those shown in Eq. (4.1-1). Putting all these together and ignoring Faraday rotation, the measured scattering matrix is then

$$\begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix}_{measured} = K(\gamma)[\mathbf{R}]\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual}[\mathbf{T}], \tag{4.1-5}$$

with

$$\begin{aligned}[\mathbf{R}] &= \begin{pmatrix} f_r(\gamma) & \delta_1^r(\gamma) \\ \delta_2^r(\gamma) & 1 \end{pmatrix} \\ [\mathbf{T}] &= \begin{pmatrix} f_t(\gamma) & \delta_1^t(\gamma) \\ \delta_2^t(\gamma) & 1 \end{pmatrix}\end{aligned}. \tag{4.1-6}$$

The quantity f_r in Eq. (4.1-6) includes not only differences in the antenna patterns for the two polarizations but also gain and path length differences in the two receiver chains.

Note that we can write the transmit and receive system distortion matrices as the product of two matrices

$$\begin{aligned}[\mathbf{R}] &= \begin{pmatrix} f_r(\gamma) & \delta_1^r(\gamma) \\ \delta_2^r(\gamma) & 1 \end{pmatrix} = \begin{pmatrix} 1 & \delta_1^r(\gamma) \\ \delta_2^r(\gamma)/f_r(\gamma) & 1 \end{pmatrix}\begin{pmatrix} f_r(\gamma) & 0 \\ 0 & 1 \end{pmatrix} = [\mathbf{R}_x][\mathbf{R}_c] \\ [\mathbf{T}] &= \begin{pmatrix} f_t(\gamma) & \delta_1^t(\gamma) \\ \delta_2^t(\gamma) & 1 \end{pmatrix} = \begin{pmatrix} f_r(\gamma) & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \alpha & \alpha\left[\delta_1^t(\gamma)/f_t(\gamma)\right] \\ \delta_2^t(\gamma) & 1 \end{pmatrix} = [\mathbf{R}_c][\mathbf{T}_x] \\ \alpha &= f_t(\gamma)/f_r(\gamma)\end{aligned}. \tag{4.1-7}$$

Following the nomenclature used by Quegan (1994) [1], we introduce the following symbols:

$$
\begin{aligned}
[\mathbf{R}_x][\mathbf{R}_c] &= \begin{pmatrix} 1 & w \\ u & 1 \end{pmatrix}\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} \\
[\mathbf{R}_c][\mathbf{T}_x] &= \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \alpha & \alpha z \\ v & 1 \end{pmatrix}. \\
\alpha &= f_t(\gamma)/f_r(\gamma)
\end{aligned}
\tag{4.1-8}
$$

Therefore, using Eq. (4.1-7) in Eq. (4.1-5), we can write

$$
\begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix}_{measured} = K(\gamma)[\mathbf{R}_x][\mathbf{R}_c]\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual}[\mathbf{R}_c][\mathbf{T}_x]. \tag{4.1-9}
$$

Strictly speaking, we should also add a noise matrix to the right-hand side of Eq. (4.1-9) to account for thermal system noise; we shall, however, ignore that term for now. This system model was derived in slightly different forms by Klein and Freeman (1991) [2] and by Quegan (1994) [1]. Ignoring the additive noise, this is the most general form of the system model to include system distortions in the measurements.

The scattering matrix for linear, reciprocal media is symmetrical. The vast majority of cases encountered in practical remote sensing fall in this category. Recognizing this, Raney (1988) [3] suggested that one could gain 3 decibels (dB) in signal-to-noise ratio (SNR) in the cross-polarized channel by coherently averaging the two measured cross-polarized terms in the scattering matrix. The 3-dB increase results from the fact that the cross-polarized signals add coherently, but the noise in the two channels add incoherently because these measurements are made at different times and through different channels. We can write this symmetrized matrix as

$$
\begin{aligned}
\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{sym} &= \frac{1}{2}\left\{[\mathbf{S}]_{measured} + \left[\mathbf{S}^T\right]_{measured}\right\} \\
&= K(\gamma)\frac{1}{2}\left\{[\mathbf{R}] + \left[\mathbf{T}^T\right]\right\}\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual}\left\{[\mathbf{T}] + \left[\mathbf{R}^T\right]\right\}
\end{aligned}
. \tag{4.1-10}
$$

In this equation, the superscript *T* refers to the transpose of the matrix. A close examination of this equation shows that we can write this as

$$
\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{sym} = K_c(\gamma)\left[\mathbf{D}^T\right]\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual}[\mathbf{D}] \tag{4.1-11}
$$

and we can write the distortion matrix $[\mathbf{D}]$ in the form

$$[\mathbf{D}] = \begin{pmatrix} 1 & \delta_1(\gamma) \\ \delta_2(\gamma) & f(\gamma) \end{pmatrix}. \tag{4.1-12}$$

It is important to realize that once we have symmetrized the matrix to take advantage of the increase in SNR, we do not need to know the individual components of the transmit-and-receive system distortion matrices. It is sufficient to estimate the elements of the equivalent distortion matrix in Eq. (4.1-12) to calibrate the elements of the scattering matrix.

The system model in Eq. (4.1-11) can also be written in a form similar to Eq. (4.1-9) as follows

$$\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{sym} = K_c(\gamma)\left[\mathbf{D}_x^T\right]\left[\mathbf{D}_c^T\right]\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual} [\mathbf{D}_c][\mathbf{D}_x], \tag{4.1-13}$$

with

$$[\mathbf{D}] = \begin{pmatrix} 1 & \delta_1(\gamma) \\ \delta_2(\gamma) & f(\gamma) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & f(\gamma) \end{pmatrix}\begin{pmatrix} 1 & \delta_1(\gamma) \\ \delta_2(\gamma)/f(\gamma) & 1 \end{pmatrix} = [\mathbf{D}_c][\mathbf{D}_x]. \tag{4.1-14}$$

The system model in Eq. (4.1-13) was derived by van Zyl (1990) [4] and applied to the calibration of the NASA/JPL AIRSAR polarimetric data. This model is only applicable to data that have been symmetrized and is, therefore, less general than the model in Eq. (4.1-9). On the other hand, studies (for example, Cordey, 1993 [5]) has shown that this algorithm is applicable to most terrain types. The only exceptions found were certain man-made structures, including villages and one agricultural field. This conclusion is not surprising; the derivation of Eq. (4.1-13) follows directly from Eq. (4.1-9) and only relies on reciprocity of the scattering process.

Before discussing in detail how to calibrate polarimetric radars using the system models discussed so far, let us look at a few examples of the effects of these errors. We shall use the polarization responses of a trihedral corner reflector to illustrate the effects of various errors. To simplify our discussion, we shall use the system model in Eq. (4.1-11). Figure 4-3 shows the effects of different amounts of amplitude errors in the co-polarized channel imbalance.

It is also instructive to look at these figures using the three-dimensional display of the polarization responses we introduced in Chapter 2. These are shown in Fig. 4-4. From Fig. 4-4 it is clear that the co-polarized channel amplitude imbalance causes the original figure to be distorted by moving the maximum along the S_1 axis on the Poincaré sphere. Values of $f(\gamma) > 1$ move the maximum to the $-S_1$ axis, while values of $f(\gamma) < 1$ move the maximum to the

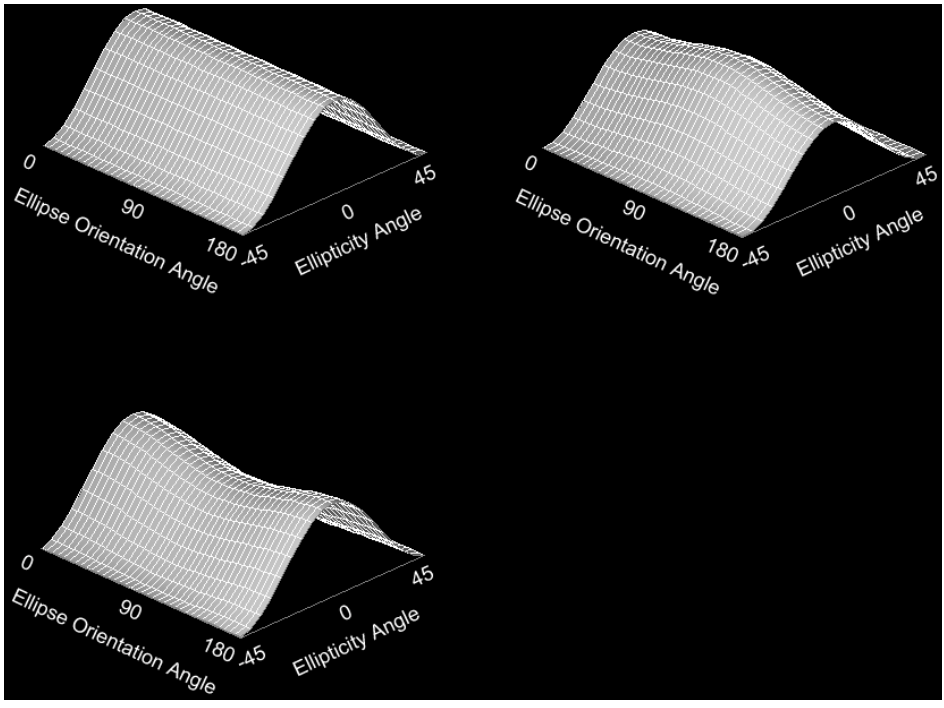


Fig. 4-3. Examples of polarization responses of a trihedral corner reflector. The top left image shows the theoretical response with no calibration errors. The top right image has a co-polarized channel imbalance error of +0.5 dB. The bottom image has an error of -0.5 dB.

$+S_1$ axis. The further the value of $f(\gamma)$ is from 1, the more distorted the figure becomes.

The responses in Fig. 4-3 and Fig. 4-4 were calculated assuming that the phase of $f(\gamma)$ is zero. Fig. 4-5 and Fig. 4-6 show the responses assuming that $|f(\gamma)|=1$; we will now vary the phase angle of $f(\gamma)$. These figures, particularly the three-dimensional displays, show that the main effect of a phase-angle error in the co-polarized channels is to rotate the polarization response about the S_1 axis.

Finally, Fig. 4-7 and Fig. 4-8 show the effects of the cross-talk errors. When the phase angle of the cross-talk parameter is near zero or 180 deg, the main effect is to distort the three-dimensional figure by shifting along the S_2 axis. This can be seen in the images on the right in Fig. 4-8. The standard co-polarized signatures for these cases look like a co-polarized amplitude imbalance, but the maximum in these figures are shifted from either vertical or horizontal (see the two signatures on the right in Fig. 4-7 compared to the top right signature in

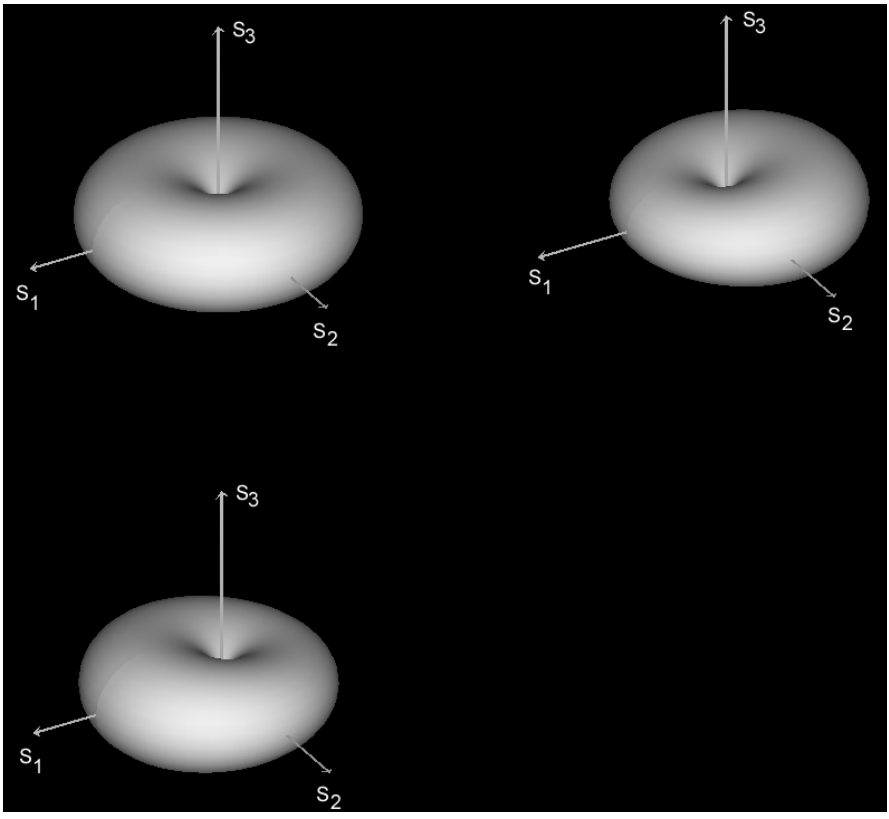


Fig. 4-4. The same polarization responses as in Fig. 4-3, but displayed as three-dimension images as explained in Chapter 2.

Fig. 4-3). In fact, the magnitude of this shift is related to the strength of the cross-talk parameter. When the phase of the cross-talk parameter is different than zero or 180 deg, however, the distortion is more complicated (as indicated in the figures).

With this background, we now return to the process of calibrating polarimetric radar image data. The general calibration process follows the following steps:

1. First, the cross-talk parameters are estimated and corrected.
2. Then the relative co-polarized phase and amplitude are estimated and corrected.
3. Finally, the absolute radiometric factor is estimated to turn the measured powers into normalized radar cross-sections.

In the next few sections we shall discuss these steps in detail.

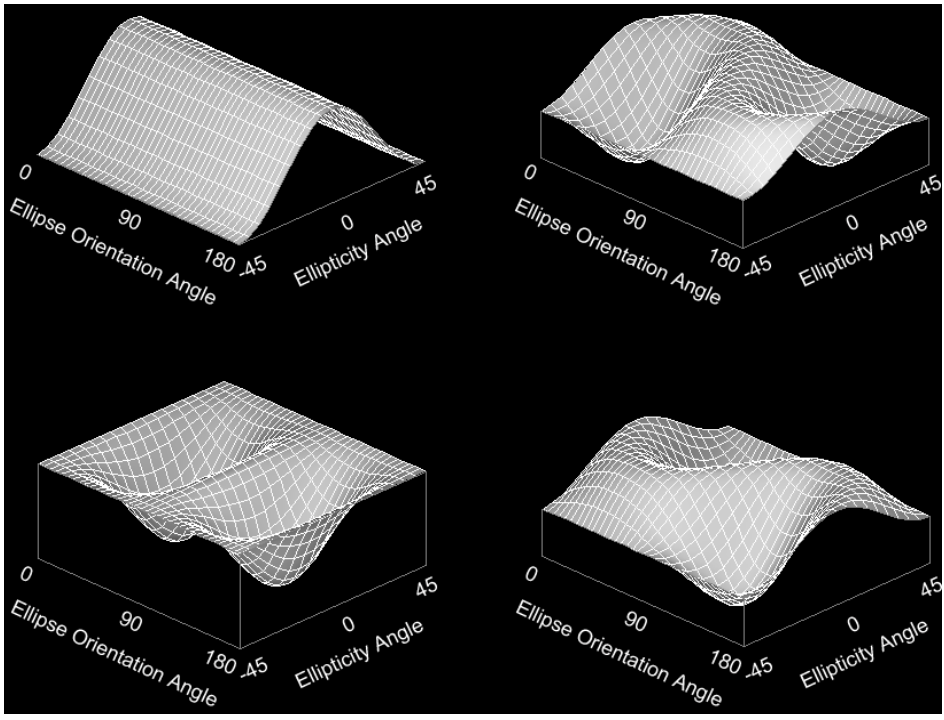


Fig. 4-5. Examples of polarization responses of a trihedral corner reflector with different amounts of co-polarization phase errors. The top left image shows the theoretical response with no calibration errors. The top right image has a phase error of 90 deg. The bottom left image has a phase error of 180 deg. The bottom right image has a phase error of 270 deg.

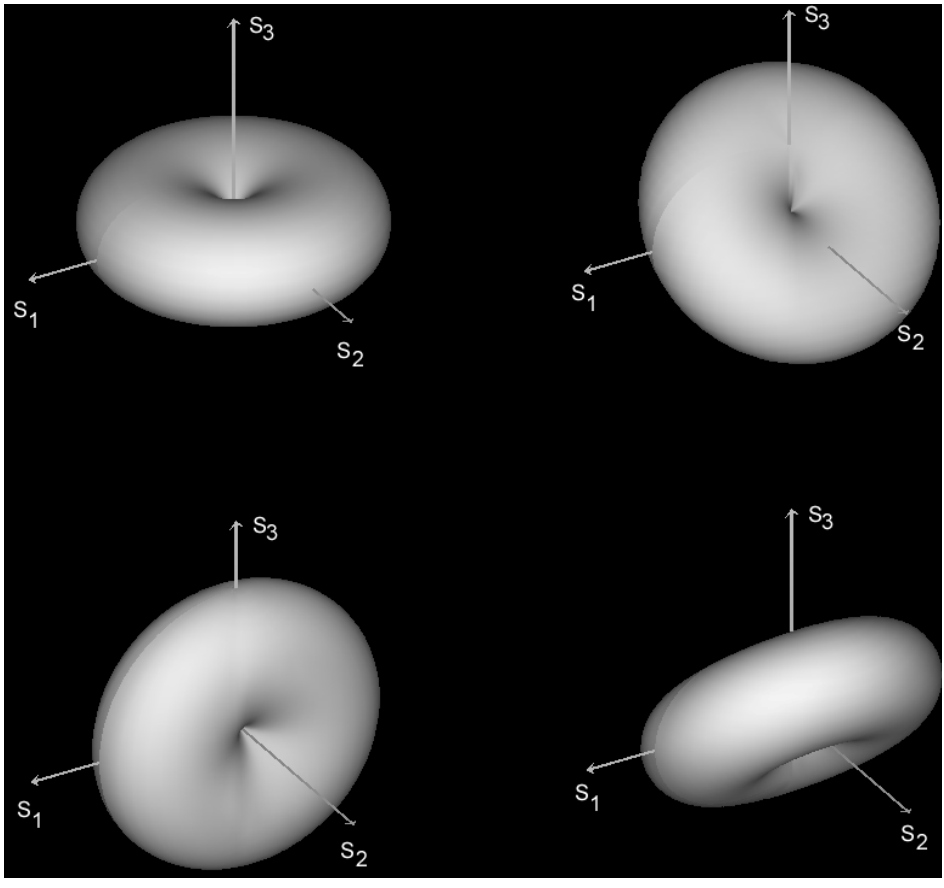


Fig. 4-6. The same polarization responses as in Fig. 4-5, but displayed as three-dimensional images as explained in Chapter 2.

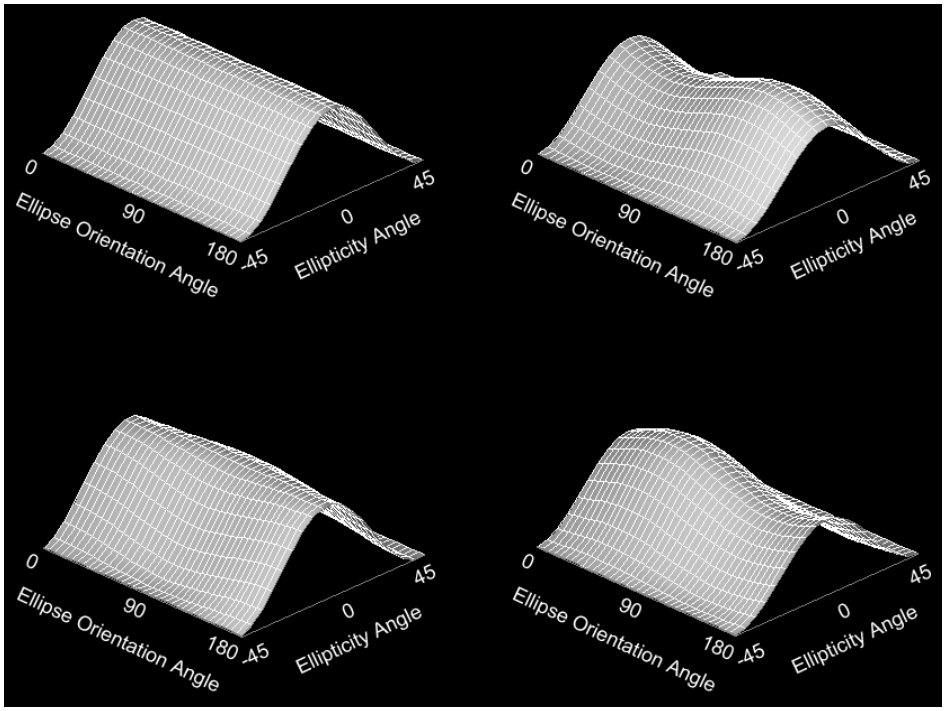


Fig. 4-7. Examples of polarization responses of a trihedral corner reflector with different amounts of cross-talk errors. The top left image shows the theoretical response with no calibration errors. In the other three cases, we assumed that the magnitude of the cross-talk error is –20 dB, and that $\delta_1 = \delta_2$. The top right image has a phase error of 0 deg. The bottom left image has a phase error of 90 deg. The bottom right image has a phase error of 180 deg.

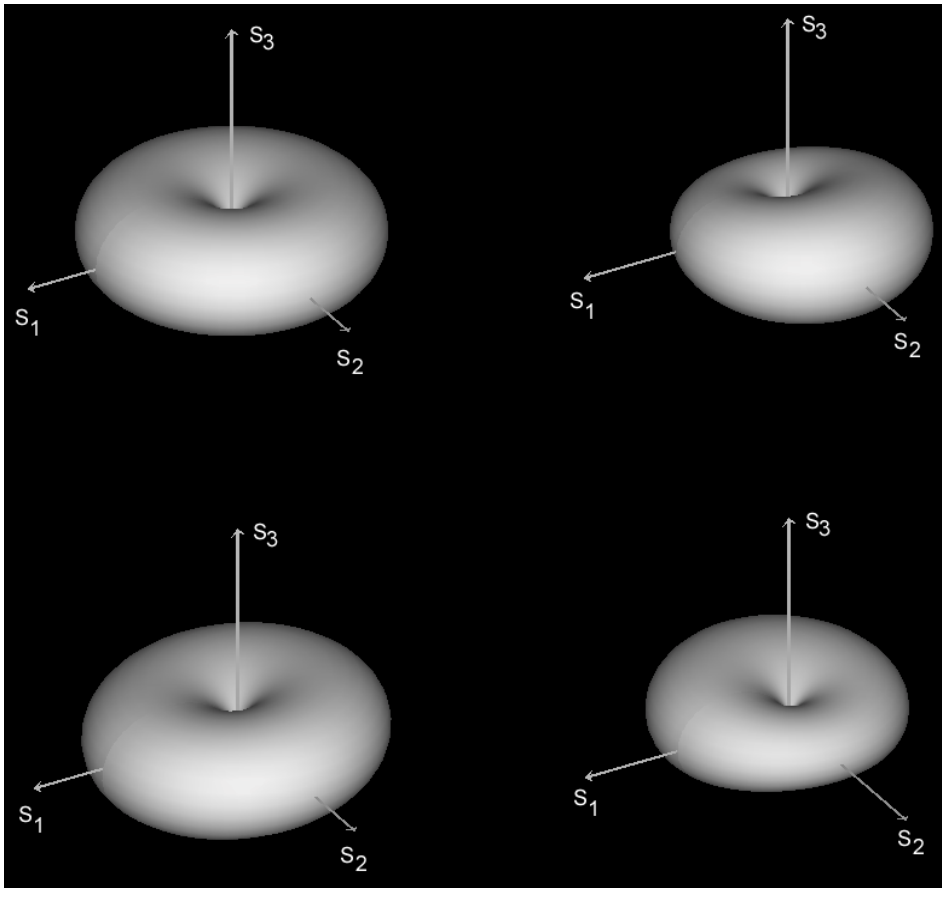


Fig. 4-8. The same polarization responses as in Fig. 4-7, but displayed as three-dimensional images as explained in Chapter 2.

4.2 Cross-Talk Estimation and Removal

The first step in the calibration is to estimate the cross-talk parameters. To do this, we rewrite Eq. (4.1-9) in vector form, as follows

$$\begin{pmatrix} S_{hh} \\ S_{vh} \\ S_{hv} \\ S_{vv} \end{pmatrix}_{measured} = K(\gamma) \begin{pmatrix} \alpha & v+\alpha w & vw \\ \alpha u & \alpha & v \\ \alpha z & 1 & w \\ \alpha uz & u+\alpha z & 1 \end{pmatrix} \begin{pmatrix} k^2 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vv} \end{pmatrix}_{actual}, \quad (4.2\text{-}1)$$

which is of the form

$$\begin{pmatrix} S_{hh} \\ S_{vh} \\ S_{hv} \\ S_{vv} \end{pmatrix}_{measured} = K(\gamma)[\mathbf{X}][\mathbf{Q}] \begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vv} \end{pmatrix}_{actual} . \quad (4.2\text{-}2)$$

Forming the covariance matrix of the measured vector in Eq. (4.2-2), we find

$$[\mathbf{C}]_{measured} = |K(\gamma)|^2 [\mathbf{X}][\mathbf{Q}][\mathbf{C}]_{actual} \left[\mathbf{Q}^{\dagger}\right]\left[\mathbf{X}^{\dagger}\right]. \quad (4.2\text{-}3)$$

The superscript † denotes the *adjoint* (transpose complex conjugate) of the matrix. To estimate the distortion matrices, particularly the matrix $[\mathbf{X}]$, we make the crucial assumption that the terrain being imaged exhibits reflection symmetry. This idea was first introduced in the context of calibration by van Zyl [4] and used by Klein and Freeman [2] and Quegan [1]. Under this assumption, the averaged actual covariance matrix is of the form (Borgeaud et al., 1987 [6]; Nghiem et al., 1992 [7])

$$[\mathbf{C}]_{actual} = \begin{pmatrix} \langle S_{hh}S_{hh}^* \rangle & 0 & \langle S_{hh}S_{vv}^* \rangle \\ 0 & \langle S_{hv}S_{hv}^* \rangle & 0 \\ \langle S_{hh}^*S_{vv} \rangle & 0 & \langle S_{vv}S_{vv}^* \rangle \end{pmatrix}. \quad (4.2\text{-}4)$$

In this expression, the angular brackets $\langle\ \rangle$ denote spatial averaging of the image data. Performing the inner multiplications in Eq. (4.2-3), we find that we can write

$$[\mathbf{O}] = [\mathbf{Q}][\mathbf{C}]_{actual}\left[\mathbf{Q}^{\dagger}\right] = \begin{pmatrix} |k^2|^2 \langle S_{hh}S_{hh}^* \rangle & 0 & k^2 \langle S_{hh}S_{vv}^* \rangle \\ 0 & |k|^2 \langle S_{hv}S_{hv}^* \rangle & 0 \\ k^{2*} \langle S_{hh}^*S_{vv} \rangle & 0 & \langle S_{vv}S_{vv}^* \rangle \end{pmatrix}. \quad (4.2\text{-}5)$$

This matrix still has the same basic form as the original actual covariance matrix in the sense that several of the matrix elements are zero. To estimate the cross-talk parameters, we perform the multiplication

$$[\mathbf{T}] = [\mathbf{X}][\mathbf{O}]\left[\mathbf{X}^{\dagger}\right]. \quad (4.2\text{-}6)$$

In doing so, however, we neglect those terms that are of second order *compared to the other terms in the same element.* The results are

$$\begin{aligned}
T_{11} &\approx |\alpha|^2 O_{11} \\
T_{21} &\approx |\alpha|^2 uO_{11} + \alpha^* vO_{13}^* + \alpha\left(v^* + \alpha^* w^*\right)O_{22} \\
T_{31} &\approx |\alpha|^2 zO_{11} + \alpha^* wO_{13}^* + \left(v^* + \alpha^* w^*\right)O_{22} \\
T_{41} &\approx \alpha^* O_{13}^* \\
T_{22} &\approx |\alpha|^2 |u|^2 O_{11} + \alpha^* u^* vO_{13}^* + \alpha uv^* O_{13} + |\alpha|^2 O_{22} + |v|^2 O_{33} \\
T_{32} &\approx |\alpha|^2 u^* zO_{11} + \alpha^* u^* wO_{13}^* + \alpha zv^* O_{13} + \alpha O_{22} + wv^* O_{33} \\
T_{42} &\approx \alpha^* u^* O_{13}^* + v^* O_{33} + \alpha^*\left(u + \alpha z\right)O_{22} \\
T_{33} &\approx |\alpha|^2 |z|^2 O_{11} + \alpha^* z^* wO_{13}^* + \alpha zw^* O_{13} + O_{22} + |w|^2 O_{33} \\
T_{43} &\approx \alpha^* z^* O_{13}^* + w^* O_{33} + \left(u + \alpha z\right)O_{22} \\
T_{44} &\approx O_{33}
\end{aligned} \tag{4.2-7}$$

From these expressions, we can form the following four simplified expressions

$$\begin{aligned}
T_{21} &\approx uT_{11} + vT_{41} + \alpha\left(v^* + \alpha^* w^*\right)O_{22} \\
T_{31} &\approx zT_{11} + wT_{41} + \left(v^* + \alpha^* w^*\right)O_{22} \\
T_{24} &\approx uT_{14} + vO_{33} + \alpha\left(u^* + \alpha^* z^*\right)O_{22} \\
T_{34} &\approx zT_{14} + wO_{33} + \left(u^* + \alpha^* z^*\right)O_{22}
\end{aligned} \tag{4.2-8}$$

In this expression, we used the fact that $[\mathbf{T}]$ is a Hermitian matrix. From these four expressions, we can solve for *u*, *v*, *z*, and *w* as follows

$$\begin{pmatrix} T_{21} - \alpha T_{31} \\ T_{24} - \alpha T_{34} \end{pmatrix} = \begin{pmatrix} T_{11} & T_{41} \\ T_{14} & T_{44} \end{pmatrix} \begin{pmatrix} u - \alpha z \\ v - \alpha w \end{pmatrix}. \tag{4.2-9}$$

From which we find

$$\begin{aligned}
u - \alpha z &= \left(T_{44}T_{21} - T_{41}T_{24}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right) - \alpha\left(T_{11}T_{34} - T_{31}T_{14}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right) \\
v - \alpha w &= \left(T_{11}T_{24} - T_{21}T_{14}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right) - \alpha\left(T_{44}T_{31} - T_{41}T_{34}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right)
\end{aligned} \tag{4.2-10}$$

Simple inspection shows that

$$\begin{aligned} u &= \left(T_{44}T_{21} - T_{41}T_{24}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right) \\ v &= \left(T_{11}T_{24} - T_{21}T_{14}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right) \\ w &= \left(T_{44}T_{31} - T_{41}T_{34}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right) \\ z &= \left(T_{11}T_{34} - T_{31}T_{14}\right)/\left(T_{11}T_{44} - T_{14}T_{41}\right) \end{aligned} \quad . \tag{4.2-11}$$

Note that knowing these four parameters does not yet allow us to recover $[\mathbf{O}]$ from $[\mathbf{T}]$. We also need to know the value of α. We can solve for α by writing

$$\begin{aligned} T_{32} &= u^* z T_{11} + u^* w T_{41} + z v^* T_{14} + w v^* T_{44} + \alpha O_{22} \\ T_{33} &= |z|^2 T_{11} + z^* w T_{41} + z w^* T_{14} + |w|^2 T_{44} + O_{22} \end{aligned} \quad . \tag{4.2-12}$$

Combining these two equations, we find

$$\alpha = \frac{T_{32} - u^*\left(zT_{11} + wT_{41}\right) - v^*\left(zT_{14} + wT_{44}\right)}{T_{33} - z^*\left(zT_{11} + wT_{41}\right) - w^*\left(zT_{14} + wT_{44}\right)} . \tag{4.2-13}$$

With the values in Eq. (4.2-11) and Eq. (4.2-13), we can now construct the matrix $[\mathbf{X}]$ in Eq. (4.2-6), which can then be used to solve for an estimate of the matrix $[\mathbf{O}]$, as follows:

$$\left[\hat{\mathbf{O}}\right] = \left[\mathbf{X}^\dagger\mathbf{X}\right]^{-1}\left[\mathbf{X}^\dagger\right][\mathbf{T}][\mathbf{X}]\left[\mathbf{X}^\dagger\mathbf{X}\right]^{-1} . \tag{4.2-14}$$

If we started with scattering matrix data, the equivalent solution is

$$\begin{pmatrix} k^2 S_{hh} \\ k S_{hv} \\ S_{vv} \end{pmatrix}_{actual} = \left[\mathbf{X}^\dagger\mathbf{X}\right]^{-1}\left[\mathbf{X}^\dagger\right]\begin{pmatrix} S_{hh} \\ S_{vh} \\ S_{hv} \\ S_{vv} \end{pmatrix}_{measured} . \tag{4.2-15}$$

We still need to estimate the channel imbalance to recover the actual scattering matrix. We shall discuss this process in the next section. Before we do that, however, let us also look at the case where the measured scattering matrix was symmetrized to gain the extra 3 dB in SNR in the cross-polarized channel.

Once we symmetrized the scattering matrix to gain the 3 dB in SNR in the cross-polarized channel, the system model can be written as pointed out by van Zyl (1990) [4] and derived in Eq. (4.1-13):

$$\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{sym} = K_c(\gamma) \begin{pmatrix} 1 & b \\ a & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & f \end{pmatrix} \begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual} \begin{pmatrix} 1 & 0 \\ 0 & f \end{pmatrix} \begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix}. \quad (4.2\text{-}16)$$

Performing the multiplications, we find that we can write

$$\begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vv} \end{pmatrix}_{sym} = \begin{pmatrix} 1 & 2b & b^2 \\ a & 1+ab & b \\ a^2 & 2a & 1 \end{pmatrix} \begin{pmatrix} W_{hh} \\ W_{hv} \\ W_{vv} \end{pmatrix}$$
$$\begin{pmatrix} W_{hh} \\ W_{hv} \\ W_{vv} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & f^2 \end{pmatrix} K_c(\gamma) \begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vv} \end{pmatrix}_{actual} \quad . \quad (4.2\text{-}17)$$

In this way, we have separated the cross-talk corrections from the co-polarized channel imbalance and absolute radiometric calibrations. Forming the covariance matrices, we define

$$[\mathbf{T}] = \begin{pmatrix} 1 & 2b & b^2 \\ a & 1+ab & b \\ a^2 & 2a & 1 \end{pmatrix} \left\langle \begin{pmatrix} W_{hh} \\ W_{hv} \\ W_{vv} \end{pmatrix} \begin{pmatrix} W_{hh} & W_{hv} & W_{vv} \end{pmatrix}^* \right\rangle \begin{pmatrix} 1 & a & a^2 \\ 2b & 1+ab & 2a \\ b^2 & b & 1 \end{pmatrix}^* . \quad (4.2\text{-}18)$$

Next, we make the assumption that the covariance matrix of the terrain satisfies Eq. (4.2-4) as introduced in the context of calibration by van Zyl (1990) [4]. Under this assumption, we find

$$\begin{aligned} T_{11} &\approx \left\langle W_{hh} W_{hh}^* \right\rangle \\ T_{21} &\approx a \left\langle W_{hh} W_{hh}^* \right\rangle + b \left\langle W_{vv} W_{hh}^* \right\rangle + 2b \left\langle W_{hv} W_{hv}^* \right\rangle \\ T_{31} &\approx \left\langle W_{vv} W_{hh}^* \right\rangle \\ T_{22} &\approx \left\langle W_{hv} W_{hv}^* \right\rangle + |a|^2 \left\langle W_{hh} W_{hh}^* \right\rangle + |b|^2 \left\langle W_{vv} W_{vv}^* \right\rangle + ab^* \left\langle W_{hh} W_{vv}^* \right\rangle + a^* b \left\langle W_{vv} W_{hh}^* \right\rangle \\ T_{32} &\approx a^* \left\langle W_{vv} W_{hh}^* \right\rangle + b^* \left\langle W_{vv} W_{vv}^* \right\rangle + 2a \left\langle W_{hv} W_{hv}^* \right\rangle \\ T_{33} &\approx \left\langle W_{vv} W_{vv}^* \right\rangle \end{aligned} \quad . \quad (4.2\text{-}19)$$

Therefore, we can write

$$\begin{aligned} T_{21} &\approx aT_{11} + bT_{31} + 2b^* \left\langle W_{hv} W_{hv}^* \right\rangle \\ T_{23} &\approx aT_{13} + bT_{33} + 2a^* \left\langle W_{hv} W_{hv}^* \right\rangle \end{aligned} . \tag{4.2-20}$$

We can now solve for the two complex cross-talk parameters. First, from the second equation in Eq. (4.2-20) we find

$$bT_{33} = T_{23} - aT_{13} - 2a^* \left\langle W_{hv} W_{hv}^* \right\rangle . \tag{4.2-21}$$

Multiplying the first expression in Eq. (4.2-20) by T_{33}, we find

$$\begin{aligned} T_{21}T_{33} &= aT_{11}T_{33} + bT_{31}T_{33} + 2b^* T_{33} \left\langle W_{hv} W_{hv}^* \right\rangle \\ &= aT_{11}T_{33} + T_{31}\left(T_{23} - aT_{13} - 2a^* \left\langle W_{hv} W_{hv}^* \right\rangle\right) + 2\left\langle W_{hv} W_{hv}^* \right\rangle \left(T_{23} - aT_{13} - 2a^* \left\langle W_{hv} W_{hv}^* \right\rangle\right)^* \end{aligned} , \tag{4.2-22}$$

where we have now made use of (4.2-21) in the second line of this expression. Gathering terms, we can write this expression as

$$a\left(T_{11}T_{33} - T_{31}T_{13} - 4\left\langle W_{hv} W_{hv}^* \right\rangle^2\right) - 4a^* T_{31} \left\langle W_{hv} W_{hv}^* \right\rangle = T_{21}T_{33} - T_{31}T_{23} - T_{23}^* \left\langle W_{hv} W_{hv}^* \right\rangle . \tag{4.2-23}$$

To solve for *a* from this expression, we write the real and imaginary parts of this expression and then solve for the real and imaginary parts of *a*. The result is

$$\begin{aligned} a &= \frac{\Delta P + 4P^* T_{31} \left\langle W_{hv} W_{hv}^* \right\rangle}{\Delta^2 - 16T_{13}T_{31} \left\langle W_{hv} W_{hv}^* \right\rangle^2} \\ P &= T_{21}T_{33} - T_{31}T_{23} - 2T_{23}^* \left\langle W_{hv} W_{hv}^* \right\rangle . \\ \Delta &= T_{11}T_{33} - T_{31}T_{13} - 4\left\langle W_{hv} W_{hv}^* \right\rangle^2 \end{aligned} \tag{4.2-24}$$

In a similar way, we can solve for *b* to find

$$\begin{aligned} b &= \frac{\Delta Q + 4Q^* T_{13} \left\langle W_{hv} W_{hv}^* \right\rangle}{\Delta^2 - 16T_{13}T_{31} \left\langle W_{hv} W_{hv}^* \right\rangle^2} \\ Q &= T_{11}T_{23} - T_{13}T_{21} - 2T_{21}^* \left\langle W_{hv} W_{hv}^* \right\rangle \end{aligned} . \tag{4.2-25}$$

Note that the original equations in van Zyl (1990) [4] contain algebraic errors, as pointed out by Cordey (1993) [5]. The expressions reported here have been corrected. Note that Cordey's expressions contain a typographical error, where the square was left off the $\left\langle W_{hv}W_{hv}^{*}\right\rangle$ term in the second term of the denominators.

Once we know the cross-talk parameters, their effect can be removed by inverting Eq. (4.2-18), as follows:

$$\left\langle \begin{pmatrix} W_{hh} \\ W_{hv} \\ W_{vv} \end{pmatrix} \begin{pmatrix} W_{hh} & W_{hv} & W_{vv} \end{pmatrix}^{*} \right\rangle = \begin{pmatrix} 1 & 2b & b^2 \\ a & 1+ab & b \\ a^2 & 2a & 1 \end{pmatrix}^{-1} [\mathbf{T}] \begin{pmatrix} 1 & a & a^2 \\ 2b & 1+ab & 2a \\ b^2 & b & 1 \end{pmatrix}^{*-1} . \quad (4.2\text{-}26)$$

The expressions for a and b cannot be solved in a simple way; specifically, to solve for the cross-talk parameters, we need to know the cross-polarized return $\left\langle W_{hv}W_{hv}^{*}\right\rangle$. Note that this quantity is the cross-polarized return after the cross-talk effects have been removed. To solve this dilemma, van Zyl (1990) [4] proposed an iterative scheme. First, use Eq. (4.2-26) with $a = b = 0$ to estimate an initial guess of $\left\langle W_{hv}W_{hv}^{*}\right\rangle$, which is then used to estimate an initial guess for the cross-talk parameters using Eqs. (4.2-24) and (4.2-25). This initial guess is then used to improve our estimate of $\left\langle W_{hv}W_{hv}^{*}\right\rangle$ using Eq. (4.2-26). This improved estimate of $\left\langle W_{hv}W_{hv}^{*}\right\rangle$ is then used to improve our estimate of the cross-talk parameters. This procedure is repeated until the values of $\left\langle W_{hv}W_{hv}^{*}\right\rangle$, a, and b have reached stable solutions.

These estimations all rely on calculating average values for the measured covariance matrix. A natural question is as follow: How many values need to be averaged to estimate these parameters? The practical answer ties back to the fundamental assumption about the actual covariance matrix as stated in Eq. (4.2-4). At this point, it is instructive to discuss a subtle but important point behind this assumption. The form of the covariance matrix as expressed in Eq. (4.2-4) assumes reflection symmetry, as is shown in the references cited earlier. One way this reflection symmetry can be broken is if there are local tilts in the azimuth direction. This can be modeled as a local rotation of the polarization basis

$$\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{tilt} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} = \cos^2\beta \begin{pmatrix} 1 & \tan\beta \\ -\tan\beta & 1 \end{pmatrix} \begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual} \begin{pmatrix} 1 & -\tan\beta \\ \tan\beta & 1 \end{pmatrix} \quad . \quad (4.2\text{-}27)$$

Comparing this expression with Eq. (4.2-16), it is clear that the effect of the tilt will appear as a cross-talk with $a = -b = -\tan\beta$. Recognizing this, Schuler et al. (1996) [8] have proposed an algorithm to use the observed correlation between co- and cross-polarized returns to estimate azimuthal slopes in polarimetric radar images. To avoid situations where preferentially tilted surfaces are present in the images, care should be exercised when estimating the cross-talk parameters.

In practice, the cross-talk parameters are functions of the antenna angle, which translates to different values of the slant range. Therefore, the cross-talk parameters are estimated by averaging the covariance matrix of an entire line of constant slant range. In this way, local tilts in the azimuth direction are usually averaged out, and a better estimate for the cross-talk parameters is obtained.

4.3 Co-Polarized Channel Imbalance Calibration

The second step in the polarimetric calibration process is to correct for any co-polarized channel imbalances, both in amplitude and phase, that might exist. There a number of different ways to accomplish this step; each requires deployment of some external calibration devices and, potentially, the use of some calibration signals internal to the radar system. If the radar hardware is very stable over time, one could rely more on the internal signal to perform the calibration. In that case, the external devices will be used only periodically to verify the calibration of the system. Conversely, if there are no internal signals, or if the hardware stability is only short term, one must rely on external devices to perform the calibration.

The ideal calibration device for this purpose is a trihedral corner reflector. In Chapter 2 we showed that these devices have the following scattering matrix

$$[\mathrm{S}] = \frac{k_0 l^2}{\sqrt{12}\pi} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (4.3\text{-}1)$$

In this equation, the parameter l is the length of one of the sides of the corner reflector, as shown in Fig. 4-9. From a calibration point of view, the desirable characteristics of a trihedral corner reflector are:

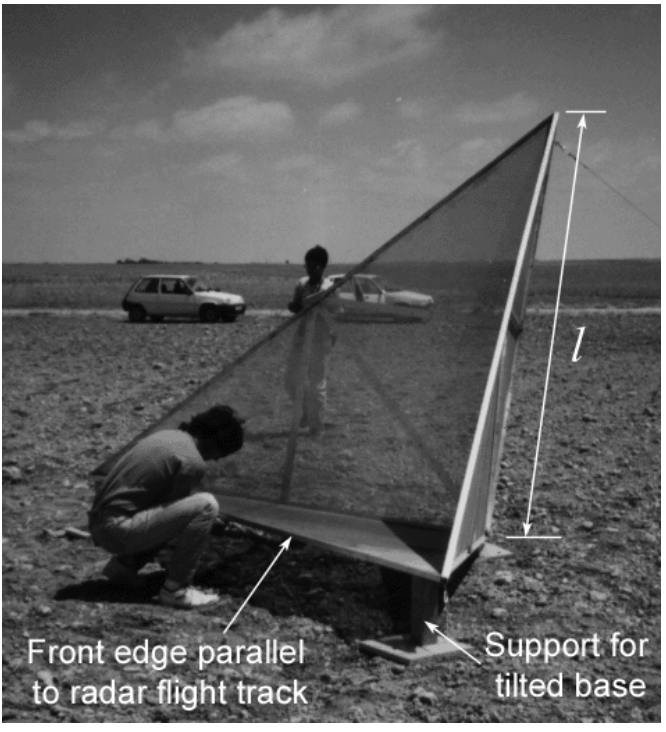


Fig. 4-9. Trihedral corner reflectors are excellent calibration targets for polarimetric radar systems. This photograph shows one of the authors measuring the orientation of a trihedral corner reflector used for calibration. Wooden supports are used to lift the base of the corner reflector so that the boresight points at the expected flight elevation of the radar. The front edge of the corner reflector is placed parallel to the expected flight track of the radar.

1. No cross polarization components are generated ($\sigma_{hv} = \sigma_{vh} = 0$) for the linear polarization case.
2. Horizontal and vertical backscattering cross sections are identical ($\sigma_{hh} = \sigma_{vv}$).
3. Horizontal and vertical co-polarized components are in phase.
4. The device is entirely passive; that is, it requires no power to operate in the field.

These devices are easy to manufacture and subsequently deploy in the field. The basic construction uses three identical flat triangular panels, typically constructed using an aluminum frame and covered with an aluminum mesh to reduce weight. These panels are bolted together at the calibration site to form the trihedral corner reflector. The device is then oriented such that the front

edge of the base is parallel to the expected radar flight direction. The base is tilted to place the boresight of the corner reflector reflection pattern as close to the radar direction as possible. Exact orientation is nearly impossible, so careful measurements must be made of the orientations so as to allow later corrections for the pointing of the reflector relative to the radar look direction. These will be discussed in the next section.

Since the co-polarized channel imbalance might be a function of the radar look direction, we typically deploy several of these devices in any image to be calibrated spaced relatively uniformly through the image at different positions across the radar track. Since we have to estimate the radar cross-sections of these devices from the radar image data, we typically look for smooth bare surfaces on which to place the corner reflectors, thereby keeping the background radar signals as low as possible in the immediate vicinity of the calibration devices.

During routine polarimetric SAR processing, the images are formed and the antenna pattern and range corrections described in the next section are applied. The next step is usually to estimate and remove the cross-talk as described in the previous section. Now we are ready to calibrate the co-polarized channel imbalance. Typically, this is done in two steps: correct the phase, then correct the amplitude.

To understand the co-polarized phase calibration, we consider the system model shown in Fig. 4-10, which is based on Lou and van Zyl (1992) [9] (note that "Pol. Switch" stands for Polarization Switch). This model includes a calibration signal to help measure some of the path delays in the system. We

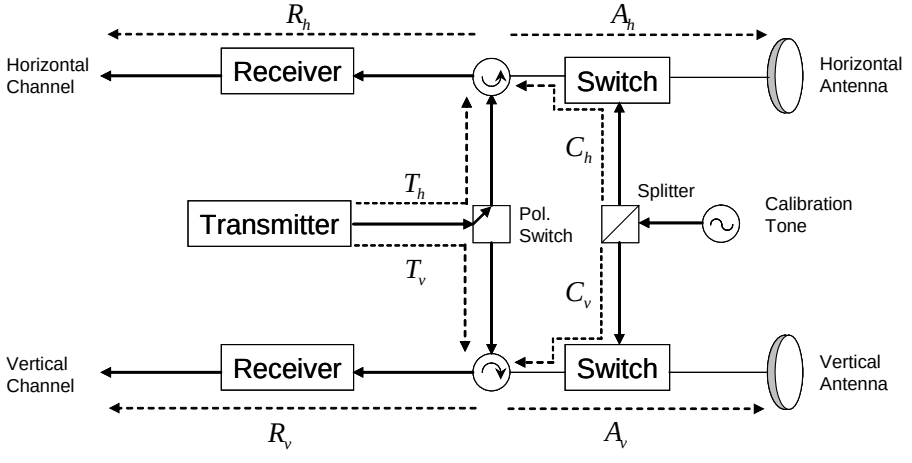


Fig. 4-10. The system model used for co-polarized phase calibration. The instrumentally induced phase paths for the different paths are indicated with dashed lines.

shall first discuss the use of this signal to calibrate the phase difference because it has the advantage that we do not need to make any assumptions about the scattering in the image itself. We shall later discuss the case where this signal is not present.

After cross-talk removal, we can write the scattering matrix as

$$[\mathbf{R}] = \begin{pmatrix} S_{hh}\exp\{i(\phi_{Th}+2\phi_{Ah}+\phi_{Rh})\} & |f|S_{hv}\exp\{i(\phi_{Tv}+\phi_{Ah}+\phi_{Av}+\phi_{Rh})\} \\ |f|S_{hv}\exp\{i(\phi_{Th}+\phi_{Ah}+\phi_{Av}+\phi_{Rv})\} & |f|^2 S_{vv}\exp\{i(\phi_{Tv}+2\phi_{Av}+\phi_{Rv})\} \end{pmatrix}. \quad (4.3\text{-}2)$$

Remember that $f = f(\gamma)$ as shown in Eq. (4.1-12). In each case, terms such as ϕ_{Th} refer to the phase of the signal because of the path along the dotted line labeled T_h in Fig. 4-10. For example, in the case of R_{hh} the signal travels along the path T_h from the transmitter to the circulator then along path A_h between the circulator and the antenna before being radiated. Upon reception, the signal again travels along path A_h, now from the antenna to the circulator, and then along path R_h between the circulator and the rest of the horizontal channel receiving chain. We can rewrite Eq. (4.3-2) by extracting an overall absolute phase as follows

$$[\mathbf{R}] = \exp\{i(\phi_{Tv}+2\phi_{Av}+\phi_{Rv})\}\begin{pmatrix} S_{hh}\exp\{i(\phi_T+2\phi_A+\phi_R)\} & |f|S_{hv}\exp\{i(\phi_A+\phi_R)\} \\ |f|S_{hv}\exp\{i(\phi_T+\phi_A)\} & |f|^2 S_{vv} \end{pmatrix}, \quad (4.3\text{-}3)$$

where $\phi_T = \phi_{Th} - \phi_{Tv}$; $\phi_A = \phi_{Ah} - \phi_{Av}$; $\phi_R = \phi_{Rh} - \phi_{Rv}$. The absolute phase outside of the matrix in Eq. (4.3-3) is of little importance in polarimetry; this information is lost when the various cross-products are formed in the covariance matrix. The elements of the covariance matrix are

$$
\begin{aligned}
R_{hh}R_{hh}^{*} &= S_{hh}S_{hh}^{*} \\
R_{hv}R_{hv}^{*} &= R_{vh}R_{vh}^{*} = |f|^{2} S_{hv}S_{hv}^{*} \\
R_{vv}R_{vv}^{*} &= |f|^{4} S_{vv}S_{vv}^{*} \\
R_{hh}R_{hv}^{*} &= |f|^{2} S_{hh}S_{hv}^{*} \exp\{i(\phi_T + \phi_A)\} \\
R_{hh}R_{vh}^{*} &= |f|^{2} S_{hh}S_{hv}^{*} \exp\{i(\phi_R + \phi_A)\} \\
R_{hh}R_{vv}^{*} &= |f|^{2} S_{hh}S_{vv}^{*} \exp\{i(\phi_T + 2\phi_A + \phi_R)\} \\
R_{hv}R_{vh}^{*} &= |f|^{2} S_{hv}S_{hv}^{*} \exp\{i(\phi_R - \phi_T)\} \\
R_{hv}R_{vv}^{*} &= |f|^{3} S_{hv}S_{vv}^{*} \exp\{i(\phi_R + \phi_A)\} \\
R_{vh}R_{vv}^{*} &= |f|^{3} S_{hv}S_{vv}^{*} \exp\{i(\phi_T + \phi_A)\}
\end{aligned}
\quad . \qquad (4.3\text{-}4)
$$

There are three unknown phases to calibrate: ϕ_T, ϕ_A and ϕ_R. We can measure the antenna path phase difference ϕ_A when the radar system is assembled on the spacecraft or aircraft. This path involves only cables (or possibly waveguides) and the antennas and should remain stable over long periods of time as long as the antennas and cables are not disturbed. To measure the receiver path phase difference, we use the calibration tone that is injected into the two channels via a switch network. This tone has a common source. Therefore, any phase difference measured at the output of the receiver chains (for example, ϕ_{cal}) is related to the differences in path length that the common signal traveled through the system. Referring to Fig. 4-10, we can write this

$$\phi_{cal} = \phi_R + \phi_C \quad \Rightarrow \quad \phi_R = \phi_{cal} - \phi_C \,. \qquad (4.3\text{-}5)$$

The term ϕ_C can be minimized by careful design and ϕ_C should stay stable over long periods because it involves path length differences through cables, splitters, and couplers. We can measure this after radar installation; it should then remain stable as long as these devices are not changed or disturbed. We can, therefore, recover the receiver path length phase difference from Eq. (4.3-5) by monitoring the phase difference between the calibration signals in the two receiver chains.

There is now only one path length phase difference that is not known: ϕ_T This phase difference can be recovered by using the fact that reciprocity dictates that $S_{hv} = S_{vh}$. From Eq. (4.3-4), we find that

$$R_{hv}R_{vh}^{*} = \left| fS_{hv} \right|^2 \exp\left\{ i(\phi_R - \phi_T) \right\} . \tag{4.3-6}$$

The phase of the complex product $R_{hv}R_{vh}^{*}$ will vary slightly for each pixel in the image due to system noise. To get a good estimate of $\phi_R - \phi_T$, therefore, we can average this complex product over the entire scene. We can then recover the transmit path length phase difference from

$$\phi_T = \phi_R - \arg\left\{ \left\langle R_{hv}R_{vh}^{*} \right\rangle \right\} . \tag{4.3-7}$$

With these relative phases known, the scattering matrix elements can be calibrated. This discussion makes only one assumption about the scattering: that reciprocity holds. No further assumptions are necessary.

But what if we do not have a calibration tone built into the radar system? We shall now show that with a relatively simple assumption about the scattering, we can still calibrate that relative co-polarization phase, as discussed by Zebker and Lou (1990) [10]. The basic assumption is that there is some extended area in the image that we know what the co-polarized phase difference should be. Let us denote this expected phase difference by $\hat{\phi}_{co}$. From the image data for the area where we expect to know the co-polarized phase difference we extract the measured co-polarized phase as

$$\arg\left\{ \left\langle R_{hh}R_{vv}^{*} \right\rangle \right\} = \phi_T + 2\phi_A + \phi_R = \hat{\phi}_{co} \Rightarrow \phi_T + \phi_R = \hat{\phi}_{co} - 2\phi_A . \tag{4.3-8}$$

Using Eq. (4.3-7) and Eq. (4.3-8) we can now solve for the values of ϕ_T and ϕ_R. So what type of surfaces can be used with this method to calibrate the phases? Models of scattering from slightly rough surfaces predict an average co-polarized phase that is near zero. Of course, if there are trihedral corner reflectors in the image, one could use their measured co-polarized phases in the same way.

Once the phases have been calibrated, we need to correct the effects of the co-polarized amplitude imbalance. For this, we use the measured image power from trihedral corner reflectors. Recall that for these devices we expect the co-polarized responses to be identical. After phase calibration, we end with covariance matrix elements that are

$$
\begin{aligned}
W_{hh}W_{hh}^{*} &= R_{hh}R_{hh}^{*} = S_{hh}S_{hh}^{*} \\
W_{hv}W_{hv}^{*} &= R_{hv}R_{hv}^{*} = R_{vh}R_{vh}^{*} = |f|^{2} S_{hv}S_{hv}^{*} \\
W_{vv}W_{vv}^{*} &= R_{vv}R_{vv}^{*} = |f|^{4} S_{vv}S_{vv}^{*} \\
W_{hh}W_{hv}^{*} &= R_{hh}R_{hv}^{*} \exp\{-i(\phi_T + \phi_A)\} = |f|^{2} S_{hh}S_{hv}^{*} \\
W_{hh}W_{vh}^{*} &= R_{hh}R_{vh}^{*} \exp\{-i(\phi_R + \phi_A)\} = |f|^{2} S_{hh}S_{hv}^{*} \\
W_{hh}W_{vv}^{*} &= R_{hh}R_{vv}^{*} \exp\{-i(\phi_T + 2\phi_A + \phi_R)\} = |f|^{2} S_{hh}S_{vv}^{*} \\
W_{hv}W_{vh}^{*} &= R_{hv}R_{vh}^{*} \exp\{-i(\phi_R - \phi_T)\} = |f|^{2} S_{hv}S_{hv}^{*} \\
W_{hv}W_{vv}^{*} &= R_{hv}R_{vv}^{*} \exp\{-i(\phi_R + \phi_A)\} = |f|^{3} S_{hv}S_{vv}^{*} \\
W_{vh}W_{vv}^{*} &= R_{vh}R_{vv}^{*} \exp\{-i(\phi_T + \phi_A)\} = |f|^{3} S_{hv}S_{vv}^{*}
\end{aligned}
\quad . \qquad (4.3\text{-}9)
$$

We measure the co-polarized channel amplitude imbalance by extracting the horizontal-horizontal (HH) and vertical-vertical (VV) image power in a small box centered on the corner reflector. Generally, we add the power in this small box so as to find the integrated power under the point response of the corner reflector. This method is less sensitive to processor focusing than using the peak value in the point response. For co-polarized channel imbalance, the absolute value of this integrated power is not important; only the relative values in the HH and VV images matter.

The square root of the square root of the measured ratios of the power in the HH and VV images as a function of angle of incidence is then used to derive a correction curve to apply to the images. This curve represents the magnitude of the quantity $f(\gamma)$ in Eq. (4.1-14) or $f_r(\gamma)$ in Eq. (4.1-7). From the curve we can then calculate a value of the co-polarized amplitude correction for each range line. Once we apply this correction to the data, we have calibrated all the polarimetric channels relative to each other. What remains is to perform an absolute calibration, which will enable us to compare the measured image radar cross-sections to theoretical models. We shall discuss this process next.

4.4 Absolute Radiometric Calibration

To understand the absolute radiometric corrections performed during SAR processing, let us examine the radar equation we first encountered in Chapter 1 (see Fig. 1-6). The received power from a scattering area A is given by

$$P_r = \frac{P_t G_t(\gamma)}{4\pi R^2} A\sigma_0 \frac{\lambda^2 G_r(\gamma)}{(4\pi R)^2}, \tag{4.4-1}$$

Where

P_t	=	transmitted power	
λ	=	radar wavelength	
R	=	distance to scattering area	(4.4-2)
$G_t(\gamma)$	=	transmit antenna gain at angle γ	
$G_r(\gamma)$	=	radar look angle γ	
σ_0	=	normalized radar cross-section for area A	

Note here that A is the area *on the ground* responsible for the scattering.

Since the radar signals are sampled in the time domain, which corresponds to distances in the line-of-sight (the slant range) direction, the following approximation is usually made during SAR calibration when calculating the scattering area

$$A = \frac{\delta_r \delta_a}{\sin\eta}. \tag{4.4-3}$$

Here δ_r and δ_a are the slant range (cross-track) and azimuth (along-track) pixel spacings, respectively. Note that the scattering area is determined by the pixel spacing; that is the way in which the radar signals are sampled, not the resolution of the radar. Typically, the radar pixel spacing is smaller than the radar resolution so that a point target typically occupies more than one pixel in a radar image.

The goal of SAR imaging is to measure σ_0 for each pixel. Rewriting Eq. (4.4-1), one finds that

$$\sigma_0 = K(R,\gamma)/A, \tag{4.4-4}$$

Where

$$K(R,\gamma) = \frac{P_r}{P_t} \frac{(4\pi)^3 R^4}{\lambda^2 G_t(\gamma) G_r(\gamma)}. \tag{4.4-5}$$

We note that, apart from the measured transmit and received power levels, two parameters, R and γ, must be estimated to determine $K(R,\gamma)$. We have already discussed how these parameters are estimated for the airborne and spaceborne cases. For each line of constant range in the data, one can estimate the corresponding slant range and look angle. The look angle is then used to determine the value of the antenna gain to apply for that line. These gains, together with the other parameters in Eq. (4.4-5), are then used to determine the value of $K(R,\gamma)$ for that particular line.

In most SAR processors, the look angle is determined assuming a reference surface that contains no relief. In reality the local topography raises or lowers the pixel so that the actual look angle is different from the one calculated from a reference surface. If the topography is known in digital form, one can actually take this effect into account during processing to provide more accurate calibration of the data. Radar interferometry is especially useful in this regard: the interferometer provides the topography that is needed to more accurately calibrate the data.

Finally, to complete the calibration, one has to remove the scattering area A. Here again most SAR post-processors use the flat or spherical Earth assumptions and approximate the area A by Eq. (4.4-3). We note that, as in the case of the look angle, the incidence angle is not measured directly and must be inferred indirectly. If digital topographic data are available, one could estimate the actual scattering area much more accurately and perform a more accurate calibration. We shall discuss this in more detail later.

Even with excellent knowledge of antenna-gain patterns and system parameters, the absolute calibration of the image is best achieved using external calibration targets placed in an image to verify the overall calibration. Trihedral corner reflectors are excellent devices to use for this purpose. The basic idea is to deploy the corner reflectors so that they are spaced across the image swath, covering most of the range extent of the image. For a corner reflector constructed with triangular panels, such as the one shown in Fig. 4-9, the radar cross-section is, per Ruck et al. (1970) [11], as follows

$$\sigma(\theta,\phi)=\frac{4\pi l^4}{\lambda^2}\left[\cos\theta+\sin\theta(\sin\phi+\cos\phi)-\frac{2}{\cos\theta+\sin\theta(\sin\phi+\cos\phi)}\right]^2. \quad (4.4\text{-}6)$$

The angles θ and ϕ are shown in Fig. 4-11. The peak of the response happens for $\theta=54.74°;\ \phi=45°,$ and the half-power response is approximately a cone of 40 deg centered on this axis of symmetry. Therefore, if the corner reflector is

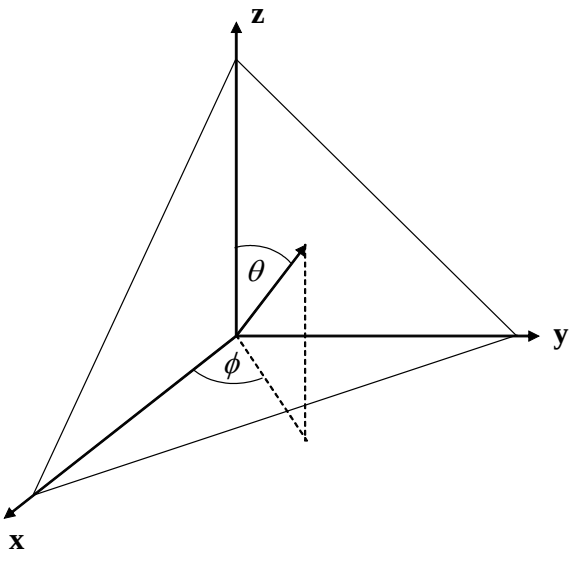


Fig. 4-11. Diagram showing the spherical angles relative to the panels of the trihedral corner reflector. The bottom panel is in the x-y plane.

placed in that portion of the image where the angle of incidence is 54.74 deg, the radar signal will be reflected with maximum intensity.

To use the corner reflector as a calibration device, we try to point it such that the radar signal will be reflected optimally. This means that if the expected angle of incidence is $\theta_i \leq 54.74\,\text{deg},$ we need to raise the bottom panel by an angle $\alpha_{CR} = 54.74\,\text{deg} - \theta_i$. At the same time, we need to ensure that the front edge of the bottom panel is parallel to the expected flight line. Careful measurements of the final orientation angles must be made for later use in the calibration process. These angles are typically measured as close in time to the actual overflight as possible. Figure 4-9 shows one of the authors measuring the angles of such a corner reflector deployed for calibration purposes. To ensure accuracy of the measurements, the corner reflector signal must be much brighter than the background in the image. This is accomplished by placing the corner reflectors on as smooth a surface as possible, preferably with no vegetation cover. Corner reflectors on the order of 2.4-m panels are typically sufficient for L-band calibration; for C-band the typical size is about 1 m.

To find the absolute calibration factor to apply to the radar image, we compare the measured corner reflector signal in the image to the theoretical value, keeping in mind that radar images are reported as normalized radar cross-sections (that is, the power is divided by the scattering area). During this step, we take into account the actual orientation angles as measured for each

individual corner reflector. To get a more accurate measurement, we typically deploy as many corner reflectors at the calibration site as is practical. The individual measurements are then averaged together to arrive at the final calibration value. The overall normalized radar cross-section of each pixel in the image is then scaled by this value to ensure that the average of the corner reflector values will be the correct theoretical value. Figure 4-12 shows such a calibration site in California used to calibrate the NASA/JPL AIRSAR system. The corner reflectors are spaced across the Rosamond dry lake to cover as much of the imaging swath as possible.

For most SAR processors, this will conclude the calibration process. In the next two sections, we shall examine a bit more closely some of the assumptions made during the calibration process. In particular, we shall look at the accuracy of approximating the scattering area by a flat facet. We shall follow that discussion by looking at the effect of local topography on antenna pattern correction when the antenna look angle is calculated assuming a flat or spherical reference plane.

4.4.1 Effect of Topography on the Scattering Area

Figure 4-13 shows a surface tilted in an arbitrary direction being illuminated by a plane wave. It follows from simple geometry that the pixel area $\delta_r \delta_a$ is the projection of the actual surface area A' on the image plane. Therefore,

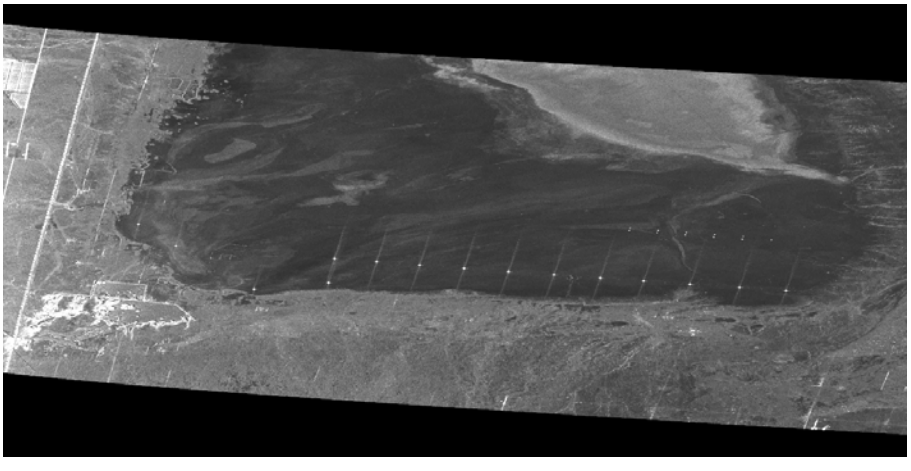


Fig. 4-12. Image of the Rosamond dry lake in California showing the array of corner reflectors used to calibrate the NASA/JPL AIRSAR system. The dry lake bed is the large dark feature forming most of the image. The corner reflectors are visible as the line of bright objects near the bottom edge of the lake. The image is an L-band total-power image, and the radar illumination is from the left in the image. The corner reflectors are spaced to cover as much of the range swath as possible.

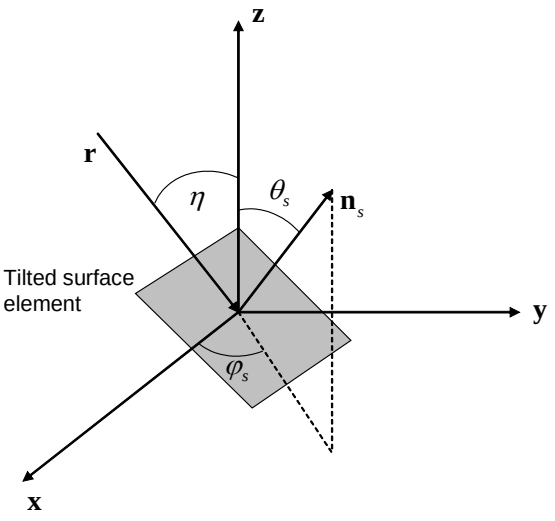


Fig. 4-13. Diagram of a tilted ground element in relation to the image plane that is defined by the x axis (assumed to be the azimuth direction) and the unit vector that points in the slant range direction.

$$\delta_r \delta_a = A' \cos\psi \,, \tag{4.4-7}$$

where ψ is the angle between the actual surface normal and the image plane normal. Using spherical coordinates for the surface slope tilt angle θ_S and the slope aspect angle φ_S, we can write the surface normal as

$$\mathbf{n}_S = \sin\theta_S \cos\varphi_S \mathbf{x} + \sin\theta_S \sin\varphi_S \mathbf{y} + \cos\theta_S \mathbf{z} \,. \tag{4.4-8}$$

Here it is assumed that the unit vector $\mathbf{x}$ points in the along-track direction. Denoting the radar line of site direction by the unit vector $\mathbf{r}$, the normal to the image plane is

$$\mathbf{n}_i = \mathbf{x} \times \mathbf{r} = \cos\eta \mathbf{y} + \sin\eta \mathbf{z} \,. \tag{4.4-9}$$

Therefore,

$$\cos\psi = \mathbf{n}_S \cdot \mathbf{n}_i = \sin\eta \cos\theta_S + \cos\eta \sin\theta_S \sin\varphi_S \,. \tag{4.4-10}$$

The actual scattering area is, therefore,

$$A' = \frac{\delta_r \delta_a}{\sin\eta \cos\theta_S + \cos\eta \sin\theta_S \sin\varphi_S} \,. \tag{4.4-11}$$

This expression reduces to Eq. (4.4-3) if the surface is not tilted (that is, $\theta_S = 0$).

To quantify the errors introduced by using Eq. (4.4-3) instead of Eq. (4.4-11) for the actual scattering area, we calculate the ratio of the two areas:

$$Error\left(dB\right)=10\log\left(\frac{A'}{A}\right)=10\log\left\{\frac{\sin\eta_0}{\sin\eta\cos\theta_s+\cos\eta\sin\theta_s\sin\varphi_s}\right\}. \tag{4.4-12}$$

We used η_0 in the numerator of Eq. (4.4-12) as a reminder that this angle represents the angle of incidence calculated for a reference plane, either a flat surface for airborne cases or a spherical surface for spaceborne SARs. Due to local relief, this angle might be different than the local angle of incidence for an elevated but untilted surface element. This local angle of incidence can be calculated from the slant range R, the height of the radar platform h, and the elevation of the image pixel h_p above the reference plane, respectively. These expressions are

$$\eta=\gamma=\cos^{-1}\left\{\left(h-h_p\right)\Big/R\right\} \tag{4.4-13}$$

for the airborne case and

$$\eta=\sin^{-1}\left\{R_s\sin\gamma\Big/\left(R_t+h_p\right)\right\} \tag{4.4-14}$$

for the spaceborne case. In this expression

$$\gamma=\cos^{-1}\left(\frac{R^2+R_s^2-\left(R_t+h_p\right)^2}{2RR_s}\right). \tag{4.4-15}$$

Note that under the flat or spherical Earth assumption for the reference plane h_p is assumed to be zero, or at least the same value for all pixels in the image. Also, in Eq. (4.4-13) - (4.4-15) a positive value of h_p means that the pixel is above the reference plane.

Figure 4-14 shows the magnitude of the calibration error introduced by making the assumption shown in Eq. (4.4-3) when calibrating the effect of the scattering area. The results are shown for different values of the angle of incidence and assume that the radar altitude above the reference plane and the pixel elevation are accurately known.

These results show the following important points:

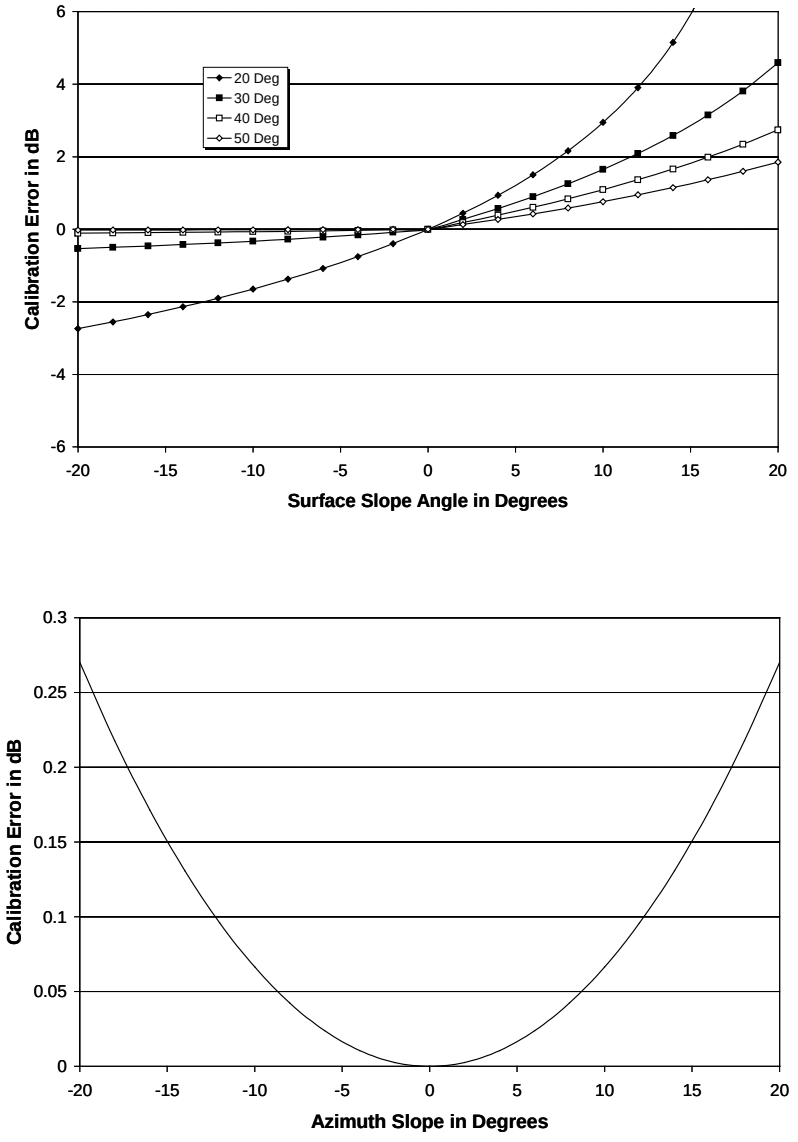


Fig. 4-14. These graphs show the calibration error that could be introduced if local surface elements are tilted as a result of topographical variations in elevation. The graph on top assumes that the tilts are only in the range direction, with positive slopes towards the radar. The errors are plotted for different values of the radar look angle. The graph on the bottom shows the effect of purely azimuthal tilts. In this case, the results are independent of the angle of incidence and the errors are significantly smaller than the case of tilts in the range direction.

1. The error due to the scattering area is largest when the slope of the surface is such that the range slope equals the angle of incidence. When this happens, the local angle of incidence approaches zero deg.
2. The smaller the look angle, the larger the errors due to the surface slope. In these cases, a smaller surface tilt leads to the situation described in 1.
3. The effects of azimuth tilts are much smaller than those of range tilts.

To better quantify what magnitude errors can be expected in practice, van Zyl et al. (1992) [12] calculated, using digital elevation model (DEM) data, the distribution of surface slopes for a moderate relief area near Tombstone, Arizona, and a high-relief area near Oetztal in Austrian Alps. In both cases, DEM data were used to calculate the slope in two orthogonal directions using the difference in height of two adjacent pixels and dividing by the horizontal separation. Using the information in the slope images, they calculated the expected errors for each area for an airborne case (using the nominal parameters of the NASA/JPL AIRSAR system flown on the NASA DC-8 aircraft) and a spaceborne case (using Earth Resource Satellite 1 [ERS-1] parameters). In the airborne case, the flat-Earth incidence angle typically varies between approximately 15 deg and 60 deg, while for the spaceborne case, the incidence angle at the center of the image is assumed to be 23 deg. For the moderate relief area, it was found that the airborne case shows relatively small errors, although some leading slopes of the hills exhibit errors exceeding 3 dB. The errors in the spaceborne case are larger, mainly because of the smaller incidence angle. Most of the errors were less than 1 dB, even in the spaceborne case. It was found, however, that the situation for the high-relief area was dramatically different. In that case, most of the leading slopes show errors on the order of 5 dB or larger, even in the airborne case. In the spaceborne example of the Oetztal area, most of the errors exceed 1 dB, with large areas showing errors exceeding 5 dB in magnitude.

It also follows from Eq. (4.4-11) that the calibration error will be influenced by how accurately the angle of incidence is inferred. In the airborne case, one has to rely on some measurement of the platform altitude to calculate the incidence angle as shown in Eq. (4.4-13). It is important to note (see Fig. 4-15) that even if one could measure the altitude of the plane above the terrain directly underneath it very accurately, it still does not mean that the angle of incidence is calculated correctly in the swath to the side of the plane. This, of course, is a consequence of the changing topography. As an example, a 500-m error in the estimation of $h - h_p$ will lead to an additional error of about 2 dB at a 30-deg look angle in the case of a plane with nominal $h - h_p$ of 8000 m. Note that the 500-m number is not the accuracy with which the radar altitude can be measured. It is the difference in elevation of the terrain directly underneath the

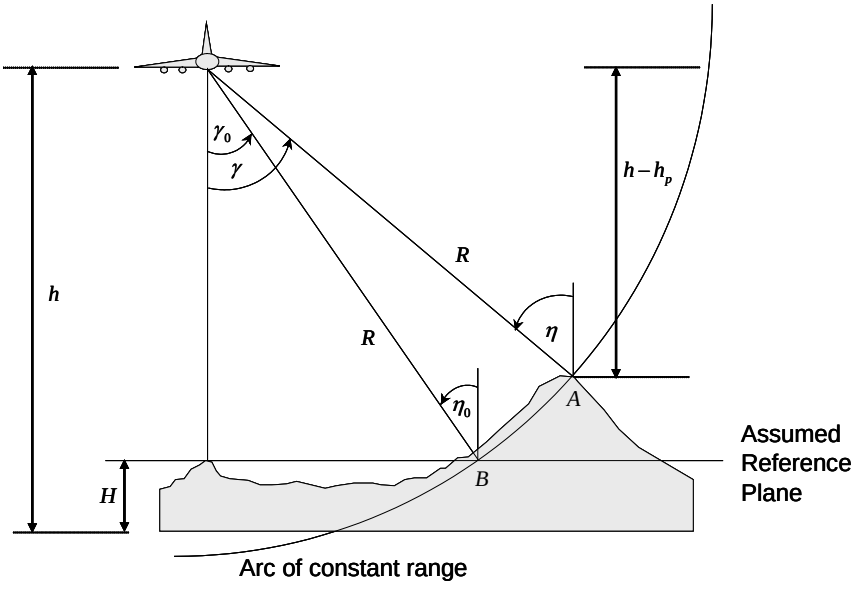


Fig. 4-15. Imaging geometry where the elevation of the terrain being imaged varies. Not taking the varying topography into account will result in the incorrect look angle being used during antenna pattern removal.

plane and that of the area being imaged off to the side. In our experience with the NASA/JPL AIRSAR system flying on a DC-8 plane, we often image terrain in which we have more than 500-m variation in elevation. Of course, this error can be minimized by a proper choice of the radar platform altitude during processing. In the spaceborne case, the terrain variations are much smaller than the platform altitude, which means that the additional errors introduced by misestimating the angle of incidence are negligible.

In summary, we have shown in this section that the error due to the flat surface approximation when calculating the scattering area is significant in both airborne and spaceborne SAR data.

4.4.2 Effect of Topography on Antenna Pattern Corrections

In removing the antenna patterns during radiometric calibration, one has to estimate where in the antenna beam the pixel was during data collection. Typically the antenna is mounted on the platform such that the boresight of the antenna points at a specific look angle. For electronically or mechanically steerable antennas, this boresight angle can be changed for each data collection. To remove the antenna gain pattern from the image, therefore, we need to know the radar look angle as well as the roll angle of the radar platform. If the radar platform roll stays constant, the only remaining unknown is the radar look angle for each pixel. In most cases, the radar platform roll changes with time.

Typically three-axis gyroscopes are used to measure radar platform attitude. If the accuracy of these measurements is not sufficient, other techniques, such as those discussed by Curlander and McDonough (1991) [13], may be employed to estimate platform roll more accurately.

To evaluate the effect of the topography on the antenna pattern removal, let us assume that the roll angle is known and discuss only the additional effect of varying topography. Key to antenna pattern removal is the estimation of the actual radar look angle γ for each pixel. In the spaceborne case, the look angle in the presence of varying terrain is given by Eq. (4.4-15). Typically h_p is much smaller than the other quantities in this expression, meaning that the effect of the varying topography on the estimation of the look angle can be neglected. The limiting factor in the spaceborne case is much more likely to be knowledge of the antenna pointing.

In the airborne case, the local terrain variation might be a significant fraction of the plane altitude, as illustrated in Fig. 4-15. Suppose the height of the radar platform above some reference plane is h. If now the elevation of the terrain directly underneath the plane with respect to the same reference plane is H and the elevation of the pixel being imaged is h_p, it follows that the actual look angle is given by Eq. (4.4-13) (see Fig. 4-14). If one uses the measured plane elevation above the terrain directly underneath it, the incorrectly inferred look angle will be given by

$$\eta = \gamma = \cos^{-1}\left\{\left(h - H - h_p\right)\Big/ R\right\}. \tag{4.4-16}$$

The error introduced in the antenna pattern removal can be written [see Eq. (4.4-5)] as

$$\frac{K\left(R,\gamma_0\right)}{K\left(R,\gamma\right)} = \frac{G_t\left(\gamma\right)G_r\left(\gamma\right)}{G_t\left(\gamma_0\right)G_r\left(\gamma_0\right)}. \tag{4.4-17}$$

Clearly, the magnitude of the error depends on the exact values of the antenna patterns G_t and G_r. From Fig. 4-14, however, it is clear that if the terrain elevation is a significant fraction of the altitude at which the plane flies, the actual and assumed look angles might be significantly different, even if we use the correct altitude for the plane.

These errors were quantified by van Zyl et al. (1993) [12]. They used the C-band antenna patterns for the NASA/JPL AIRSAR system and evaluated the antenna pattern removal errors for a high-relief area in the Austrian Alps. Their calculations show that extremely large errors, some much larger than 10 dB,

can be expected in the near range. In the case of the AIRSAR system, the antenna boresight points at a 50-deg look angle. The near range of the image, typically less than 20-deg look angle, is imaged by that part of the antenna pattern where the gain changes rapidly with angle, Therefore, a relatively small error in look angle can cause a significant error in antenna pattern correction. Their result indicates that one can expect that any time the SAR system acquires data over that portion of the antenna beam where the gain changes rapidly with angle away from boresight, the effect of topography might lead to significant calibration errors.

There is another subtle point to realize regarding the antenna pattern correction errors discussed here. Since the antenna gain patterns for the different polarizations are typically slightly different, the errors due to the topography might be different for the different polarizations. This, in turn, means that the polarimetric calibration (that is, the relative calibration between the different polarization channels) will be affected by the topography. This is especially important for those geophysical algorithms that are designed to use polarization ratios to infer some geophysical parameter from the radar data. Note, however, that this error results from using the wrong look angle in the calibration; that is, it is an antenna pattern removal error. As argued before, this type of error is not expected to be a major problem for spaceborne SAR systems.

4.4.3 AIRSAR Image Example

Figure 4-16 shows an example of an image acquired with the NASA/JPL AIRSAR system at C-band calibrated with and without taking the topography into account. This image was acquired of over the Tennessee Valley area of Marin County near San Francisco in California. The radar was operated in the TOPSAR interferometric mode; consequently, a co-registered digital elevation model was available from this mode with which to correct the SAR image data.

The image calibrated assuming a flat-Earth reference plane shows distinct variations in radar cross-section associated with the topography. Slopes facing the radar (at the top of the image) are consistently brighter than those facing away from the radar. This effect gives the image a three-dimensional appearance. This effect was exploited by Guindon (1990) [14] to develop an algorithm to estimate topography from the brightness variations in radar images. When the calibration is properly done taking into account the topography, however, the image on the left shows very little of these effects and appears much flatter in brightness.

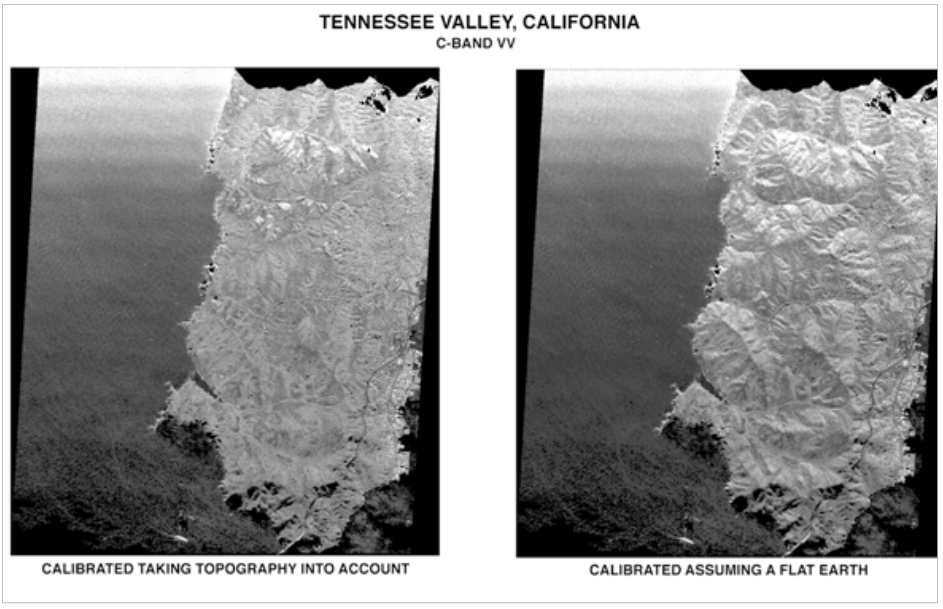


Fig. 4-16. Two images showing the effects of taking the topography into account when calibrating SAR data acquired over terrain with varying topography. The image on the left was calibrated taking topographical effects into account, while the image in the right was calibrated using the standard smooth (flat) Earth approximation. The radar illumination is from the top. Note that the image on the left appears to have much less variation in brightness that can be attributed to the topography.

These images strikingly show the differences between calibrations that take the topography into account and calibrations assuming a flat reference plane. It is clear that extra care must be taken when attempting to derive quantitative information from SAR data acquired in areas with significant relief, especially with airborne systems, if the calibration does not take the topographical effects into account.

4.5 Faraday Rotation

Radar signals transmitted from most orbiting spacecraft have to propagate through the ionosphere before on their way to the surface of the earth and again after being scattered by the surface. Linearly polarized electromagnetic waves suffer a rotation of the polarization, known as Faraday rotation, when propagating through a plasma like the ionosphere. This rotation, of course, could introduce significant errors into the measured scattering matrix if not removed properly during calibration.

The amount of rotation that an electromagnetic wave polarization suffers as a result of Faraday rotation is proportional to the total electron content (TEC) of the ionosphere [15]. Therefore, this effect varies with solar activity, reaching its

maximum during solar maximum. The amount of rotation is also inversely proportional to the square of the frequency, which means that low frequency radars will be affected more than higher frequencies. At solar maximum, L-band radars could see as much as 40 deg of rotation [16], while a P-band radar could see as much as 321 deg of rotation. For C-band and higher frequencies, the rotation is generally negligible. It should be pointed out that these numbers are the maximum expected. Rignot [17], for example, estimated that a Faraday rotation of approximately 30 deg would explain some anomalous scattering observed with HH polarization using the Japanese JERS-1 SAR system. In a different study, Wright et al. [18] argue that the upper limit on the Faraday rotation at L-band is on the order of 30 deg.

The major effect of Faraday rotation, since it involves a rotation of a linear polarization, is to create an artificial correlation between the co- and cross-polarized elements of the scattering matrix. As discussed previously in Section 4.2, polarimetric cross-talk removal algorithms use the fact that for terrain with reflection symmetry there should be no correlation between these terms. Any inferred correlation is then estimated as cross-talk in the antennas and removed. In the presence of Faraday rotation, this additional artificial correlation would lead to large errors in the estimation of the cross-talk. Recent experience with spaceborne SAR systems, however, shows that antennas can be implemented with adequate cross-talk performance, so that the effect of the antenna cross-talk is generally negligible. In that case, the perceived correlation between co- and cross-polarized terms in the scattering matrix can be attributed to Faraday rotation only.

The problem of estimating Faraday rotation directly from polarimetric data has received considerable attention recently [16, 18, 19, 20, 21]. In the presence of Faraday rotation, the radar system model given in Eq. (4.1-9) needs to be modified in the following way

$$\begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix}_{measured} = K(\gamma)[\mathbf{R}_x][\mathbf{R}_c][\mathbf{F}_\Omega]\begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix}_{actual}[\mathbf{F}_\Omega][\mathbf{R}_c][\mathbf{T}_x] \tag{4.5-1}$$

Where

$$[\mathbf{F}_\Omega] = \begin{pmatrix} \cos\Omega & \sin\Omega \\ -\sin\Omega & \cos\Omega \end{pmatrix} \tag{4.5-2}$$

represents the one-way Faraday rotation through the ionosphere given by an angle Ω. In writing Eq. (4.5-1), we have neglected the effects of thermal noise in the system, which will be additive. Freeman [16] shows various ways to

estimate the amount of Faraday rotation. If we can assume that the radar system is well calibrated, except for the Faraday rotation, we can write the measured scattering matrix as [19, 21] as

$$\begin{pmatrix} M_{hh} & M_{hv} \\ M_{hv} & M_{vv} \end{pmatrix} = \begin{pmatrix} \cos\Omega & \sin\Omega \\ -\sin\Omega & \cos\Omega \end{pmatrix} \begin{pmatrix} S_{hh} & S_{hv} \\ S_{hv} & S_{vv} \end{pmatrix} \begin{pmatrix} \cos\Omega & \sin\Omega \\ -\sin\Omega & \cos\Omega \end{pmatrix} \tag{4.5-3}$$

Here we assumed that reciprocity holds for the scattering matrix. Bickel and Bates [22] actually proposed an estimator of the Faraday rotation angle derived from expressing the measured scattering matrix in a circular polarization basis:

$$\hat{\Omega} = \frac{1}{4}\arg\left(Z_{12}Z_{21}^{*}\right) \tag{4.5-4}$$

In this expression, $Z_{12} = (M_{hv} - M_{vh}) + j(M_{hh} + M_{vv})$ and $Z_{21} = Z_{12} + 2(M_{vh} - M_{hv})$. Expanding Eq. (4.5-3), we find

$$\begin{aligned} M_{hh} &= S_{hh}\cos^2\Omega - S_{vv}\sin^2\Omega \\ M_{hv} &= S_{hv} - \frac{1}{2}\left(S_{hh} + S_{vv}\right)\sin 2\Omega \\ M_{vh} &= S_{hv} + \frac{1}{2}\left(S_{hh} + S_{vv}\right)\sin 2\Omega \\ M_{vv} &= S_{vv}\cos^2\Omega - S_{hh}\sin^2\Omega \end{aligned} \tag{4.5-5}$$

Note that when Faraday rotation is present, $M_{hv} \neq M_{vh}$, even if the scattering matrix satisfies reciprocity. From this expression, we find

$$\begin{aligned} M_{hh} + M_{vv} &= \left(S_{hh} + S_{vv}\right)\cos 2\Omega \\ M_{vh} - M_{hv} &= \left(S_{hh} + S_{vv}\right)\sin 2\Omega \end{aligned} \tag{4.5-6}$$

Using these relations, Freeman [16] showed that one can estimate the Faraday rotation angle via

$$\hat{\Omega} = \frac{1}{2}\tan^{-1}\left\{\frac{M_{vh} - M_{hv}}{M_{hh} + M_{vv}}\right\} \tag{4.5-7}$$

While this estimator appears straightforward to implement, Freeman [16] points out that if applied on a single pixel basis, this estimator could result in large errors. Instead, he proposes to use an averaged estimator given by

$$\hat{\Omega} = \pm\frac{1}{2}\tan^{-1}\sqrt{\frac{\left\langle \left|M_{vh} - M_{hv}\right|^2 \right\rangle}{\left\langle \left|M_{hh} + M_{vv}\right|^2 \right\rangle}} \tag{4.5-8}$$

Freeman [16] reports that in the presence of residual calibration errors, the estimator given in Eq. (4.5-4) performs better than the one in Eq. (4.5-8).

4.6 Summary

As discussed in this chapter, polarimetric radars are significantly more challenging to calibrate than single-channel radars. However, the techniques discussed here have been tested thoroughly calibrating a wide range of airborne and spaceborne SAR systems. It is safe to say that these techniques have proven so successful that calibrated polarimetric data are now routinely available from all modern SAR sensors.

The estimation and removal of Faraday rotation is a significant topic for lower frequency radar systems. This topic will be especially important for lower frequency radars (L-band and particularly P-band) proposed to study the Earth's ecosystems. It seems, however, that when the techniques discussed in the previous section are applied to L-band data, reasonable results are obtained. This removes a major obstacle towards implementing future low-frequency radar systems.

References

[1] S. Quegan, "A Unified Algorithm for Phase and Cross-Talk Calibration of Polarimetric Data-Theory and Observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 89–99, 1994.

[2] J. D. Klein and A. Freeman, "Quadpolarisation SAR Calibration Using Target Reciprocity," *Journal of Electromagnetic Waves and Applications*, vol. 5, pp. 735–751, 1991.

[3] R. K. Raney, "A Free 3-dB in Cross-Polarized SAR Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, pp. 700–702, 1988.

[4] J. J. van Zyl, "A Technique to Calibrate Polarimetric Radar Images Using Only Image Parameters and Trihedral Corner Reflectors," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 337-348, 1990.

[5] R. A. Cordey, "On the Accuracy of Cross-Talk Calibration of Polarimetric SAR Using Natural Clutter Statistics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, pp. 447–454, 1993.

[6] M. Borgeaud, R. T. Shin, and J. A. Kong, "Theoretical Models for Polarimetric Radar Clutter," *Journal of Electromagnetic Waves and Applications*, vol. 1, pp. 67–86, 1987.

[7] S. V. Nghiem, S. H. Yueh, R. Kwok, and F. K. Li, "Symmetry Properties in Polarimetric Remote Sensing," *Radio Science*, vol. 27, pp. 693–711, 1992.

[8] D. L. Schuler, J. S. Lee, and G. de Grandi, "Measurement of Topography using Polarimetric SAR Images," *IEEE Transactions on Geoscience Remote Sensing*, vol. 34, no. 5, pp. 1266-1277, 1996.

[9] Y. Lou and J. J. van Zyl, "Relative Phase Calibration of the NASA/DC-8 Three-Frequency Polarimetric SAR System," *Proceedings of IGARSS1992*, Houston, Texas, May 26–29, pp. 104–105, 1992.

[10] H. A. Zebker and Y. Lou, "Phase Calibration of Imaging Radar Polarimeter Stokes Matrices," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 246–252, 1990.

[11] G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, *Radar Cross Section Handbook*, Chapter 8, Volume 2, Plenum Press, New York, New York, 1970.

[12] J. J. van Zyl, B. D. Chapman, P. Dubois, and J. Shi, "The Effect of Topography on SAR Calibration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, pp. 1036–1043, 1993.

[13] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar Systems and Signal Processing*, Wiley and Sons, New York, New York, 1991.

[14] B. Guindon, "Development of a Shape-from-Shading Technique for the Extraction of Topographic Models from Individual Spaceborne SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 654–661, 1990.

[15] O. K. Garriott, F. L. Smith, and P. C. Yuen, "Observations of Ionospheric Electron Content Using a Geostationary Satellite," *Planetary and Space Science*, vol. 13, pp. 829–838, 1965.

[16] A. Freeman and S. Saatchi, "On the Detection of Faraday Rotation in Linearly Polarized L-Band SAR Backscatter Signatures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 1607–1616, 2004.

[17] E. Rignot, "Effect of Faraday Rotation on L-Band Interferometric and Polarimetric Synthetic-Aperture Radar Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 383–390, 2000.

[18] P. Wright, S. Quegan, N. Wheadon, and D. Hall, "Faraday rotation effects on L-band spaceborne SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 2735–2744, December 2003.

[19] W. B. Gail, “Effect of Faraday Rotation on Polarimetric SAR,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 301–308, 1998.

[20] A. Freeman and S. S. Saatchi, “On the Detection of Faraday Rotation in Linearly Polarized L-Band SAR Backscatter Signatures,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1607–1616, Aug. 2004.

[21] J. Chen and S. Quegan, “Improved Estimators of Faraday Rotation in Spaceborne Polarimetric SAR Data,” *IEEE Transactions on Geoscience and Remote Sensing. Letters*, vol. PP, no. 99, pp. 846–850, June 2010.

[22] S. H. Bickel and R. H. T. Bates, “Effects of Magneto-Ionic Propagation on the Polarization Scattering Matrix,” *Proceedings of the IEEE*, vol. 53, no. 8, pp. 1089–1091, August 1965.

Chapter 5
Applications: Measurement of Surface Soil Moisture

Soil moisture is a key parameter in numerous environmental studies, including hydrology, meteorology, and agriculture. Soil moisture plays an important role in the interaction between the land surface and the atmosphere, as well as in the partitioning of precipitation into runoff and ground water storage. In spite of its importance, soil moisture has not found a widespread application in the modeling of hydrological and biogeochemical processes and related ecosystem dynamics, in part because soil moisture is a difficult parameter to measure on a large-area, cost-effective, and routine basis.

It is well known that return signals from synthetic aperture radar (SAR) are affected by surface characteristics, such as the roughness, the correlation length, and the dielectric constant of the soil. Some earlier studies (Wang, et al., 1986 [1]; Dobson and Ulaby, 1986 [2]) using single frequency and single polarization Shuttle Imaging Radar-B (SIR-B) imagery could only describe the dependence of backscattering coefficient σ_0 on these surface parameters separately. Estimation of surface soil wetness was usually obtained by employing an empirical relationship to convert the measured σ_0 into volumetric soil moisture m_v (Jackson, 1993 [3]). Several research groups have, for example, reported a linear relationship between the observed radar backscatter expressed in decibels (dB) and the volumetric soil moisture of the surface (Schneider and Oppelt, 1998 [4]; Quesney et al., 2000 [5]). It is commonly reported that the slope of this linear relationship is a function of the vegetation cover, with decreasing slopes as the amount of vegetation increases indicating decreasing sensitivity to soil moisture in the presence of increasing

amounts of vegetation. As an aside, it is often found that the offset in the linear relationship is correlated with the surface roughness.

The inversion of soil moisture information from radar backscatter became more rigorous after the availability of polarimetric radar data. Basically, the additional measurements allow us to separate the effects of surface roughness and soil moisture on the observed radar backscatter. Several algorithms have been developed for measuring bare surface soil moisture quantitatively using dual-polarization L-band SAR image data (Dubois et al., 1995 [6]; Shi et al., 1997 [7]) or three-polarization SAR measurements (Oh et al., 1992 [8]).

The algorithm proposed by Oh et al. (1992) [8] was derived empirically from data measured with a truck-mounted scatterometer and the algorithm involves ratios of various polarization combinations. Dubois et al. (1995) [6] used the same data, plus data from truck-mounted scatterometers measured by the University of Berne (Wegmuller, 1993 [9]) over bare surfaces with a wide range of surface roughnesses to derive an empirical algorithm that uses horizontal-horizontal (HH) and vertical-vertical (VV) polarization combinations at L-band. Shi et al. (1997) [7] generated a synthetic data set using the Integral Equation Method (IEM) model (Fung et al.,1992 [10]) and then derived a set of coefficients to parameterize their synthetic data set. These parameters were then used to invert Airborne Synthetic Aperture Radar (AIRSAR) and SIR-C data. As in the case of Dubois et al., (1995) [6], the algorithm proposed by Shi et al. (1997) [7] uses a pair of measured radar cross-sections to estimate the surface dielectric constant and a roughness parameter. Both these algorithms have been applied to data acquired from space during the SIR-C mission and have shown accuracies on the order of 4 percent when estimating volumetric soil moisture from SAR data.

Since these early studies, many more results have been reported. We will discuss these results in more detail in this chapter. This chapter will also include a detailed of examination of the related algorithms. The chapter also includes a discussion of the algorithm proposed by Kim and van Zyl (2009) [11] that utilizes time series data to track changes in soil moisture. Finally, we shall look at the effects of vegetation on these algorithms and show that the time series algorithm shows some promise in providing the capability to measure soil moisture even in the presence of vegetation.

5.1 Surface Electrical and Geometrical Properties

Before describing the details of the individual models and their performance, we briefly review the properties, both electrical and geometrical, of rough surfaces. For the moment, we shall focus our attention on bare surfaces. The case of vegetated surfaces will be discussed later in the chapter.

5.1.1 Geometrical Properties

We can describe the local surface height of any rough surface by a two-dimensional function $\xi(x,y)$, where x and y represent the horizontal coordinates of the surface. It is typically assumed that the statistical distribution of the surface height is Gaussian with zero mean. The so-called surface roughness, or root mean square (rms) height of the surface, h is defined as

$$h^2 = \left\langle \xi^2(x,y) \right\rangle, \tag{5.1-1}$$

where $\langle x \rangle$ means the average of x Simply knowing the rms height of the surface is not yet a complete description of the geometrical properties of the surface. One also has to know how the surface height at one point on the surface is related to the surface height at a different point. This is described by the surface-height correlation function. For an isotropically rough surface, the surface-height correlation function is a function only of the separation between the two points on the surface, r. The surface-height correlation function is mathematically described by

$$\rho(r) = \frac{\left\langle \xi(x,y)\xi(x',y') \right\rangle}{h^2}. \tag{5.1-2}$$

The two most commonly used surface-height correlation functions are the Gaussian and exponential correlation functions. The Gaussian correlation function for an isotropically rough surface is

$$\rho_g(r) = e^{-r^2/l^2}. \tag{5.1-3}$$

The quantity l is known as the surface correlation length. The exponential surface-height correlation function is given by

$$\rho_e(r) = e^{-|r|/l}. \tag{5.1-4}$$

Instead of the correlation function, surfaces are often characterized in terms of their roughness spectral density, or roughness spectrum. This function is the Fourier transform of the surface autocorrelation function. For isotropically rough surfaces, the roughness spectrum is

$$W(k_x,k_y) = \frac{2}{\pi}\int_0^\infty r\rho(r)J_o(kr)dr. \tag{5.1-5}$$

The roughness spectrum functions for the Gaussian and the exponential correlation functions are

$$W_g\left(k_x,k_y\right)=\frac{l^2}{\pi}\exp\left[\frac{-\left(k_x^2+k_y^2\right)l^2}{4}\right] \tag{5.1-6}$$

and

$$W_e\left(k_x,k_y\right)=\frac{2l^2}{\pi\left[1+\left(k_x^2+k_y^2\right)l^2\right]^{3/2}}. \tag{5.1-7}$$

To understand the difference between surfaces with these different correlation functions, we show a comparison of the correlation functions in Fig. 5-1. We note that the Gaussian correlation function is larger for small separations than the exponential function, but rapidly decreases to become smaller than the exponential correlation function for larger values of r/l. This means that one would expect a surface with an exponential correlation function to appear to have larger slopes at the micro scale than a surface with an equivalent

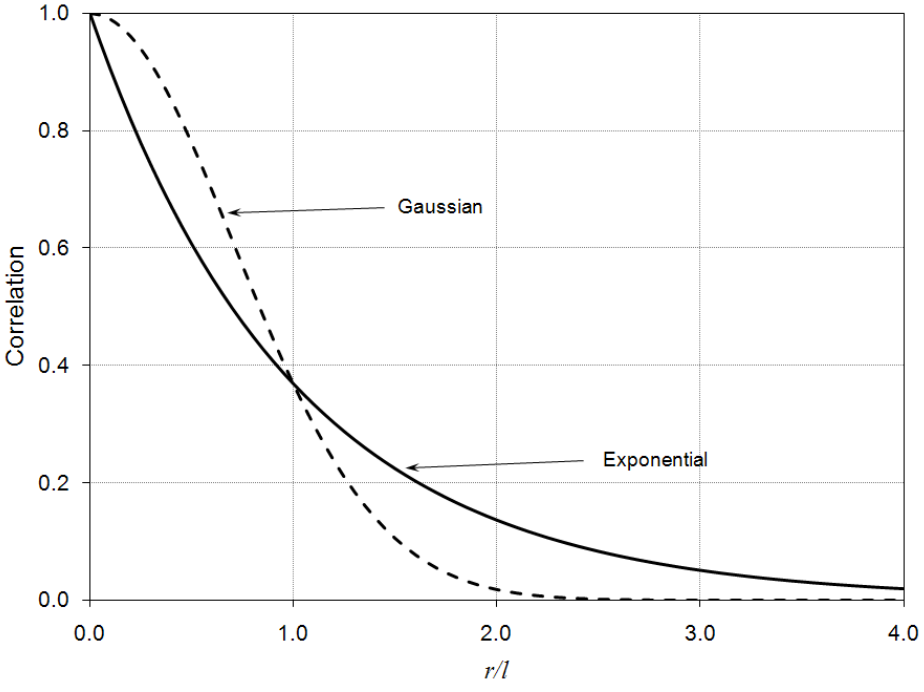


Fig. 5-1. Comparison between the Gaussian and the exponential correlation functions.

correlation length and a Gaussian correlation function. Looking at the slopes of the two types of surfaces, we note that the total mean-square slope of a surface with a Gaussian correlation function is

$$s_g^2 = \left\langle \left(\frac{\partial \xi}{\partial x} \right)^2 + \left(\frac{\partial \xi}{\partial y} \right)^2 \right\rangle = \frac{4h^2}{l^2} . \tag{5.1-8}$$

The surface with the exponential correlation function, on the other hand, has surface slopes and all higher surface derivatives that are infinite; that is,

$$s_e^2 = \infty . \tag{5.1-9}$$

To illustrate the differences between these two surface types more graphically, Figs. 5-2 and 5-3 represent two such synthetic surfaces with the same rms height and correlation length. The surface with Gaussian correlation function (shown in Fig. 5-2) appears to have less high frequency roughness than the surface with the exponential correlation function (shown in Fig. 5-3). Figures 5-3 and 5-4 show the effect of the surface correlation length on the appearance of the rough surface. Not surprisingly, the surface with the shorter correlation length appears to have more high-frequency roughness than the one with the longer correlation length.

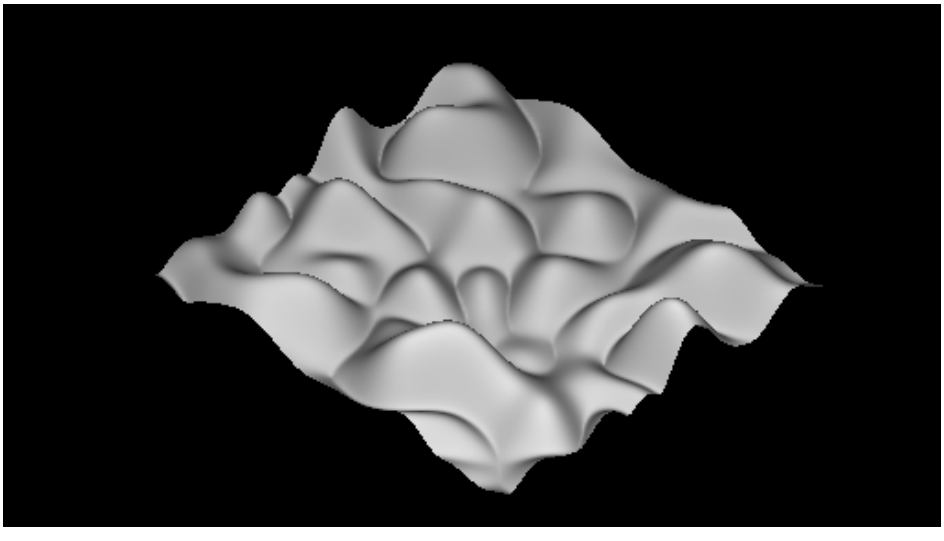


Fig. 5-2. Synthetic rough surface generated using a Gaussian surface height correlation function. Note the smooth appearance of the surface compared to that generated with the exponential surface height correlation function shown in Fig. 5-3.

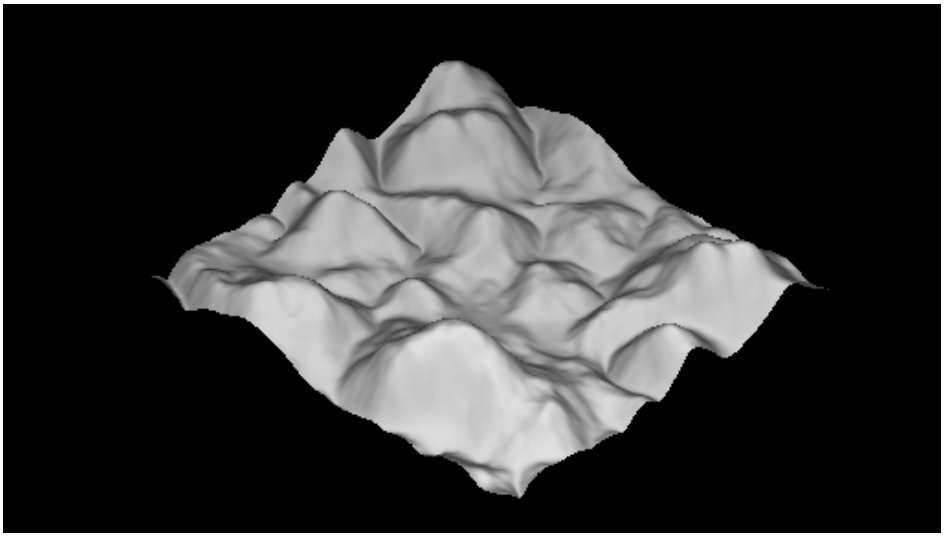


Fig. 5-3. Synthetic rough surface generated with an exponential surface height correlation function. The surface has the same rms height and correlation length as the surface generated with a Gaussian correlation functions shown in Fig. 5-2.

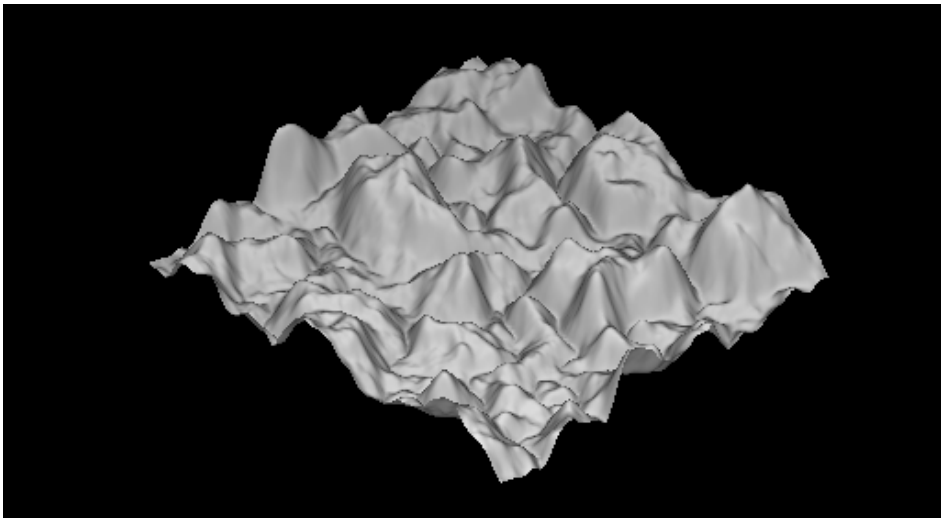


Fig. 5-4. Synthetic rough surface generated with an exponential surface height correlation function. The surface has half the correlation length of the surface shown in Fig. 5-3.

Several researchers have measured profiles of micro-topography in order to better describe the roughness characteristics of natural surfaces. The profiles of micro-topography are measured using various approaches. The simplest approach utilizes a large board with a grid painted on it. The board is then pushed into the surface to the lowest point on the surface, and a photograph is taken of the board covered with the surface profile. The profile is subsequently digitized from the photograph. The advantages offered by this approach are that it is easy to make the measurement and the equipment is relatively cheap and easily operated in the field. Disadvantages include the fact that only relatively short profiles can be measured (typically a meter or two at best) and that the soil has to be soft enough to permit insertion of the board.

A second approach utilizes a horizontal bar with an array of vertical rods of equal length that are dropped to the surface. The heights of the top of the rods above a known level surface are then measured and recorded. While relatively easier to operate than the boards described above, especially over rocky or hard surfaces, the disadvantage of this method is the limited length of the profiles that can be measured with one instrument; consequently, a large amount of time is required to make measurements of reasonably large areas, especially in areas with difficult access.

Laser profilers are also sometimes used to measure micro-topography. In this case, a laser is mounted on a frame that allows the laser to translate in a raster pattern. Measurements are typically taken every centimeter or so along a particular profile. These instruments obviously require power to operate, limiting their utility to easily accessible areas. An additional drawback is that the size of the frame usually limits the area that can be measured to a meter or so square. Another method is to operate lasers from low-flying aircraft; using this method, however, the measurement density is inadequate for micro-topography studies.

Stereo photography, either close-range or from specially equipped helicopters, seems to provide the best balance between accuracy and coverage. As part of this method, the photographs are digitized and then correlated against each other to reconstruct the micro-topography using the same kind of software developed to construct large-scale, digital-elevation models from stereo cameras flown on either aircraft or satellites. While more expensive to acquire, the full three-dimensional surface can be reconstructed over a relatively large area, providing excellent statistics.

Using stereo photographs acquired from a helicopter, Farr (1992) [12] studied the roughness characteristics of several lava flows in the Mojave Desert in southern California. He found that the power spectra of these natural surfaces exhibit the general form

$$W(k) = bk^m, \tag{5.1-10}$$

with the value of the exponent m between –2 and –3, consistent with known behavior of topography at larger scales. His measurements showed values closer to –2 and that the magnitude of m first seems to increase with lava flow age, but then decreases for surfaces between 0.14 and 0.85 million years old. For surfaces older than 0.85 million years, the magnitude of m seems to increase again.

Shi et al. (1997) [7] reports a different approach to their analysis of surface roughness characteristics. Using 117 roughness profiles measured over various fields in the Washita watershed, they fitted the correlation function of the measured profiles with a general correlation function of the form

$$\rho(r) = \exp\left(-(r/l)^n\right). \tag{5.1-11}$$

Values of $n = 1$ correspond to an exponential correlation function; $n = 2$ corresponds to a Gaussian. Their results indicate that 76 percent of the surfaces could be fitted with values of $n \le 1.4$, leading to the conclusion that the exponential correlation function is the more appropriate description of the surface correlation function.

We note that for values of $kl >> 1$, the roughness spectrum of the exponential correlation function behaves like Eq. (5.1-10) with an exponent –3. The results from the Shi et al. (1997) [7] study seem to indicate that agriculture and pasture fields have roughness spectra that contain more energy at the longer spatial scales than the natural lava flow surfaces studied by Farr [12].

5.1.2 Electrical Properties

The electrical properties of a rough surface are described by the complex dielectric constant, or relative permittivity, of the soil, which is a strong function of the soil moisture. This is the result of the fact that the dielectric resonance of both pure and saline water lies in the microwave portion of the electromagnetic spectrum. Dry soil surfaces have dielectric constants on the order of 2 - 3; water has a dielectric constant of approximately 80 at microwave frequencies. Therefore, adding a relatively small amount of water to the soil drastically changes the value of the dielectric constant.

A wet bare soil consists of a mixture of soil particles, air, and liquid water. Usually, the water contained in the soil is further divided into two parts: so-called bound water and free water. Due to the influence of matric and osmotic forces, the water molecules contained within the first few molecular layers

surrounding the soil particles are tightly held by the soil particles; hence, the term bound water. The amount of bound water is directly proportional to the surface area of the soil particles, which, in turn, is a function of the soil texture and mineralogy. Because of the relatively strong forces acting on it, bound water exhibits an electromagnetic spectrum that is different from that of regular liquid water. Since the matric forces acting on a water molecule decrease rapidly with distance away from the soil particle, water molecules located more than a few molecular layers away from the soil particles are able to move throughout the soil with relative ease; for this reason, this water is known as free water. The complex dielectric constant of both bound and free water is a function of frequency, temperature, and salinity of the soil.

In general, there is a nonlinear relationship between dielectric constant and volumetric soil moisture. Wang and Schmugge (1980) [13] present an empirical soil dielectric mixing model based on various measurements at 1.4 and 5 gigahertz (GHz). If one is concerned only with the real part of the dielectric constant, one can write the expressions as follows:

$$\varepsilon = \begin{cases} 3.25 + 2.2m_v + \dfrac{76.3E_T m_v^2}{M_T} & for \quad m_v \le M_T \\ 3.25 + 76.3M_T\left(E_T - 1\right) + 78.5m_v & for \quad m_v > M_T \end{cases}, \qquad (5.1\text{-}12)$$

where m_v is the volumetric soil moisture (units cm^3/cm^3) $0 \le m_v \le 1$ and M_T is a transition moisture level (units cm^3/cm^3), which is given by

$$M_T = 0.49W_P + 0.165\,. \qquad (5.1\text{-}13)$$

W_P is the wilting point, (in units of cubic centimeter per cubic centimeter (cm^3/cm^3)), which is a function of the soil texture, and is given by

$$W_P = 0.06774 - 0.00064S + 0.00478C\,. \qquad (5.1\text{-}14)$$

In Eq. (5.1-14), S and C represent the percentages of sand and clay, respectively, in the soil. Finally, E_T in Eq. (5.1-12) is an empirical fit coefficient, which is

$$E_T = 0.481 - 0.57W_P\,. \qquad (5.1\text{-}15)$$

For a typical loam soil:

$$W_P = 0.1; \quad M_T = 0.22; \quad E_T = 0.4\,. \qquad (5.1\text{-}16)$$

Hallikainen et al. (1985) [14] report the results of extensive measurements of soil dielectric constants. Hallikainen et al. also found that the dielectric constant can be described as a quadratic function of volumetric soil moisture. Their results show the coefficients of the polynomial to be a function of the soil texture, and the polynomial is of the general form:

$$\varepsilon = (a_0 + a_1 S + a_2 C) + (b_0 + b_1 S + b_2 C) m_v + (c_0 + c_1 S + c_2 C) m_v^2 . \qquad (5.1\text{-}17)$$

This form is applicable to both the real and imaginary parts of the dielectric constant. The coefficients are functions of frequency, with values reported from 1.4 GHz to 18 GHz. At 1.4 GHz, the values for the real part of the dielectric constant are:

$$\begin{array}{lll} a_0 = 2.862 & a_1 = -0.012 & a_2 = 0.001 \\ b_0 = 3.803 & b_1 = 0.462 & b_2 = -0.341 . \\ c_0 = 119.006 & c_1 = -0.500 & c_2 = 0.633 \end{array} \qquad (5.1\text{-}18)$$

For the imaginary part of the dielectric constant, the values are

$$\begin{array}{lll} a_0 = 0.356 & a_1 = -0.003 & a_2 = -0.008 \\ b_0 = 5.507 & b_1 = 0.044 & b_2 = -0.002 . \\ c_0 = 17.753 & c_1 = -0.313 & c_2 = 0.206 \end{array} \qquad (5.1\text{-}19)$$

For a typical loam soil, Eq. (5.1-17) becomes

$$\varepsilon = 2.2575 + 22.9925 m_v + 101.8015 m_v^2 \qquad (5.1\text{-}20)$$

for the real part of the dielectric constant.

It should be pointed out that more recent work by Brisco et al. (1992) [15] concluded that soil texture has no effect on the results, and that a single cubic polynomial with constant coefficients could be used to link soil moisture and the measured surface dielectric constant, as follows:

$$m_v = -0.0278 + 0.0280\varepsilon - 0.000586\varepsilon^2 + 0.00000503\varepsilon^3 . \qquad (5.1\text{-}21)$$

Other work (Dobson et al., 1985) [16] derived a semi-empirical relationship between the real part of the dielectric constant and the volumetric soil moisture of the form.

$$\varepsilon' = \left[1 + \frac{\rho_b}{\rho_s} \varepsilon_s^{\alpha} + m_v^{\beta} \varepsilon_{fw}^{\alpha} - m_v \right]^{1/\alpha} , \qquad (5.1\text{-}22)$$

where $\alpha = 0.65$, $\rho_s = 2.66\ g/cm^3$, ε_s is the dielectric constant of the solid soil (typical value ≈ 4.7), ρ_b is the bulk density of the soil (on the order of $1.1\ g/cm^3$ for sandy loam soils), and β is a parameter that is a function of the soil texture

$$\beta = 1.2748 - 0.00519S - 0.00152C\,. \tag{5.1-23}$$

As before, S and C are the percentage of sand and clay in the soil, respectively. The dielectric constant of free water is a function of temperature and frequency and is given by

$$\varepsilon_{fw} = \varepsilon_{w\infty} + \frac{\varepsilon_{w0} - \varepsilon_{w\infty}}{1 + (2\pi f \tau_w)^2}\,. \tag{5.1-24}$$

In Eq. (5.1-24), τ_w is the relaxation time for water, ε_{w0} is the static dielectric constant for water, $\varepsilon_{w\infty} = 4.9$ is the high frequency limit of the real part of the dielectric constant for water, and f is the frequency. Both τ_w and ε_{w0} are functions of temperature (Ulaby et al., 1986 [17]). At $20°$ C, the values are $2\pi\tau_w = 0.58 \times 10^{-10}$ seconds and $\varepsilon_{w0} = 80.1$.

A comparison of the results of the models described above is shown in Fig. 5-5 for the case of a sandy loam soil at 1.4 GHz. All models show the non-linear relationship between the soil dielectric constant and volumetric soil moisture, especially at low moisture values.

5.1.3 Penetration Depth

The question of how deep a particular radar senses the surface moisture often arises. The answer to this question is unfortunately not straightforward.

Experimental results indicate that at C-band, the radar signal is well correlated with the moisture in the top 2 cm of the soil (Bruckler et al., 1988 [18]). At L-band, good correlation is typically observed with the moisture in the top 5 cm of the soil.

From a theoretical point of view, the penetration depth depends on the complex dielectric constant ε of the surface. This quantity can be written as

$$\varepsilon = \varepsilon' + i\varepsilon''\,, \tag{5.1-25}$$

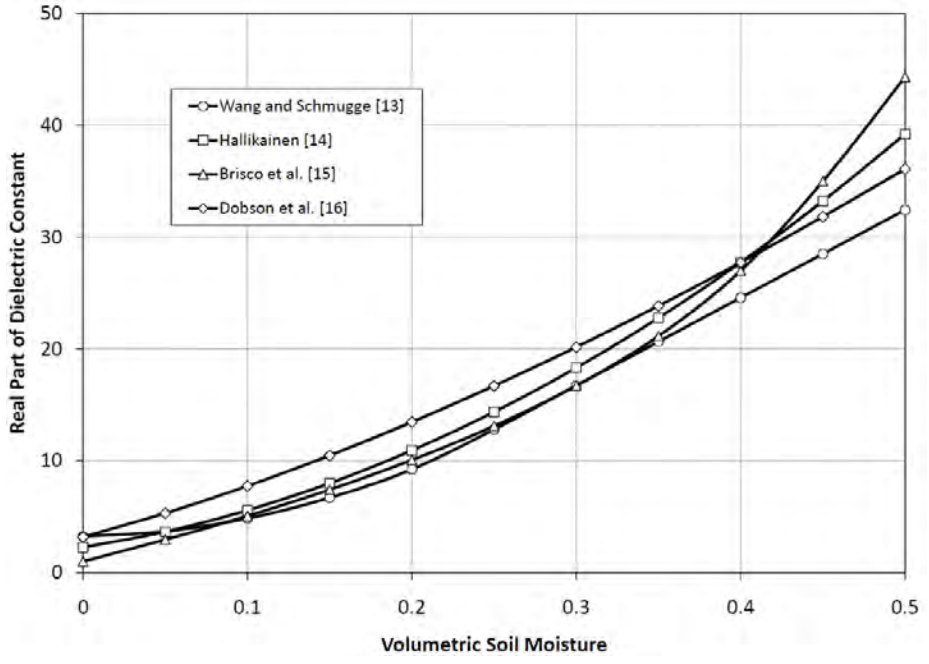


Fig. 5-5. Comparison of the results of several models linking surface dielectric constant to volumetric soil moisture. The case shown is for a sandy loam soil with 51.5 percent sand and 13.5 percent clay. Results are applicable to 1.4 GHz. Note the close comparison between the Hallikainen and Brisco curves.

where the imaginary part corresponds to the ability of the medium to absorb the wave and transform its energy to another type of energy (heat, chemical, etc.). If we consider a wave propagating purely in the x-direction in a homogeneous medium, then the electric field varies as

$$E = E_0 e^{i\sqrt{\varepsilon}kx}.$$

If we can assume that $\varepsilon'' \ll \varepsilon'$, then

$$\sqrt{\varepsilon} = \sqrt{\varepsilon' + i\varepsilon''} \approx \sqrt{\varepsilon'} + i\frac{\varepsilon''}{2\sqrt{\varepsilon'}} \tag{5.1-26}$$

and

$$E = E_0 e^{-\alpha_a x}\ e^{i\sqrt{\varepsilon'}kx}, \tag{5.1-27}$$

where

$$\alpha_a = \frac{\varepsilon'' k}{2\sqrt{\varepsilon'}} = \frac{\pi \varepsilon''}{\lambda \sqrt{\varepsilon'}}.$$

The observed radar cross-section is proportional to power. The power of the wave as a function of x can be written as:

$$P(x) = P(0)\, e^{-2\alpha_a x}. \tag{5.1-28}$$

If α_a is also a function of x, such as when the moisture profile varies with depth in the soil, then the above equation will become:

$$P(x) = P(0)\, \exp\left(-2 \int_0^x \alpha_a(\xi)\, d\xi \right). \tag{5.1-29}$$

The penetration depth L_p is defined as the depth at which the power decreases to $P(0)\, e^{-1}$ (i.e., 4.3-dB loss). Thus:

$$L_\mathrm{p} = \frac{1}{2\alpha_a} = \frac{\lambda \sqrt{\varepsilon'}}{2\pi \varepsilon''}. \tag{5.1-30}$$

This can also be expressed as a function of the medium loss tangent $(\tan \delta = \varepsilon'' / \varepsilon')$:

$$L_\mathrm{p} = \frac{\lambda}{2\pi \sqrt{\varepsilon'} \tan \delta}. \tag{5.1-31}$$

At this depth, the incident power is 4.3 dB weaker than at the surface. In the case of radar observation, if the interface is covered by a layer of thickness L_p, the attenuation due to absorption alone will be $2 \times 4.3 = 8.6$ dB for a normally incident wave. The factor 2 represents the fact that absorption affects both the incident and the scattered waves.

These expressions show that the penetration depth is a function of actual value of the dielectric constant. Based on the discussion in the previous section, this means that the penetration depth of the radar signals will then be a function of the soil moisture, with deeper penetration into dry soils and less penetration into wet soils.

Note that the penetration depth does not necessarily tell us from how deep into the soil most of the radar signal was received. As was pointed out above, the signal returned from a depth equal to the penetration depth suffered 8.6 dB of attenuation. Walker et al. (1997) [19] argue that it is unlikely that scattering

from this depth would add significantly to that arising from the soil-air interface. They suggest a modified depth based on a comparison of signals that would be received from a buried layer to that from the surface. Their results suggest a somewhat smaller sensing depth than what is normally quoted, with a reduction on the order of a factor of 2, depending on which polarization is used. A similar approach was used by Le Morvan et al. (2008) [20], who used a three-layer model to explain Advanced Synthetic Aperture Radar (ASAR) observation of soil moisture in France. As an aside, their results show significant correlations between the observed C-band backscatter and soil moisture measurements in a layer 2 - 5 cm beneath the surface. As is pointed out by these studies, however, to calculate the actual sensing depth from a theoretical point of view for anything but the simplest scattering geometries is a daunting task. For our purposes, we shall assume that, on the basis of observations and these studies, radars sense soil moisture in the top few centimeters of the soil, keeping in mind that the sensing depth is, indeed, shallower for wetter surfaces than for dry surfaces.

5.1.4 Soil Moisture Profile

When the penetration depth of a soil surface is evaluated, it is assume that the dielectric constant of the surface is uniform. However, soil moisture varies with depth [21, 22] depending upon the temperature profile, the soil type, and the surface evaporation. The soil moisture profile changes significantly as a soil surface dries after a precipitation event. The effect of the soil moisture profile on radar measurements can be evaluated by calculating the backscattering cross section from a rough surface interface on top of a layered medium with varying permittivity profile.

The reflection coefficient from a layered medium with a flat interface can be calculated for a given permittivity profile ($\varepsilon(z)$) as [23]

$$\frac{dR_n}{dz} = -i2\beta(z)R_n + \frac{d\beta(z)/dz}{2\beta(z)}\left(1 - R_n^2\right)$$

for horizontal polarization (5.1-32)

$$\frac{dR_n}{dz} = -i2\beta(z)R_n + \frac{d\left(\frac{\beta(z)}{\varepsilon(z)}\right)}{dz}\frac{\varepsilon(z)}{2\beta(z)}\left(1 - R_n^2\right)$$

for vertical polarization (5.1-33)

where

$$\beta(z) = \frac{2\pi}{\lambda}\sqrt{\varepsilon(z)}\cos\theta \tag{5.1-34}$$

Equations (5.1-32) and (5.1-33) can be solved exactly using numerical techniques. When the permittivity profile changes slowly, the reflection coefficient of the layered medium is smaller. Therefore, it is expected that the reflection coefficient will be smaller as the surface dries since the impedance changes gradually from the surface to the subsurface in addition to the lower surface dielectric constant. It is important to point out that the reflection coefficient of a soil surface with a varying moisture profile may not be the same as that calculated assuming a homogenous surface characterized by the average value of soil moisture within top 5 cm. We believe that the moisture profile may play an important role, especially for a dry surface. Notice that R_n in Eq. (5.1-32) and Eq. (5.1.33) is the specular reflection coefficient. Although the specular reflection coefficient is useful for calculating the effective impedance at the reflection interface, we must study the effects of a rough interface to realistically model backscatter from such a surface.

The backscattering cross section from a layered medium with a rough interface has been studied in [24, 25, 26]. An analytic method was derived to calculate the bistatic-scattering coefficient from a layered medium (less than three layers) with slightly rough interfaces [24]. In [24], the electromagnetic waves in each layer are expressed in terms of an infinite number of angular spectral components. A small perturbation approach is used to calculate the scattered wave, which is appropriate for the scattering calculation for a low-frequency radar system. In addition, Kuo and Moghaddam [25] used field-collected soil moisture data to understand the penetration capability of a low frequency radar. The simulation results in [25] demonstrated that both the backscattering coefficient and the co-polarization phase difference are sensitive to deep soil moisture. The inversion of the model parameters of a two-layer dielectric medium was investigated using the method of simulated annealing [26]. Although these approximate methods [24, 25] have provided the useful results, more accurate results can be obtained using a finite element method (FEM) [27, 28].

In the previous section we showed that the penetration depth is a function of the wavelength. The longer wavelength signals are expected to penetrate deeper into the surface and, depending on the moisture profile, may then sense either a drier or wetter surface. Figure 5-6 shows an example of the inferred dielectric constant of Lunar Dry Lake in Nevada using the algorithm proposed by Dubois et al. which we will discuss in detail a bit later. The oval shaped feature is the dry lake bed, which at the time of data acquisition was dry enough on the surface so that the author was able to drive a truck across the lake bed. The

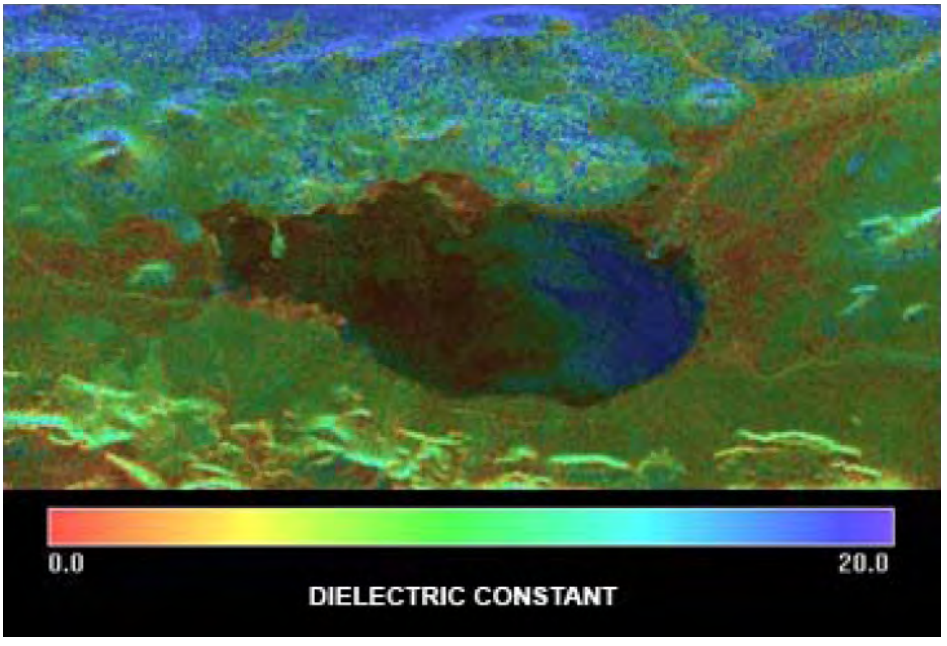


Fig. 5-6. Dielectric constant map derived from an L-Band AIRSAR image of Lunar Lake playa in Nevada. The oval shaped feature is a dry lake bed, which at the time of data acquisition was dry at the surface. The higher dielectric constant on the right portion of the lake is due to subsurface moisture. [6]

radar signals penetrated deep enough into the surface to sense subsurface moisture on the right hand side of the lake in the picture. The moisture was the remnants of runoff from light rains about a week earlier that entered the lake surface from the top right and covered only the right portion of the lake bed. The water had dried away before the data collection to the point that the surface was completely dry.

Figure 5-7 shows another example, this time showing the difference between two frequencies. The image shows a feature known as Cottonball Basin in Death Valley, California. Cottonball Basin is at the northern end of the larger Death Valley salt pan, and receives most of its inflow from groundwater. The southern edge of Cottonball Basin are covered with mud flats resulting from seeping salty water. These areas stay wet the longest after an inflow event.

Figure 5-7 shows two dielectric constant maps, the one inferred from L-band data on the left, and from P-band data on the right. There are many similarities between the two maps. First, both show relatively large dielectric constants in the mudflats, and lower dielectrics in the rest of Cottonball Basin. There are also some important differences, however. When looking closer at the mud flat areas, we note that the P-band dielectric constants show higher values over

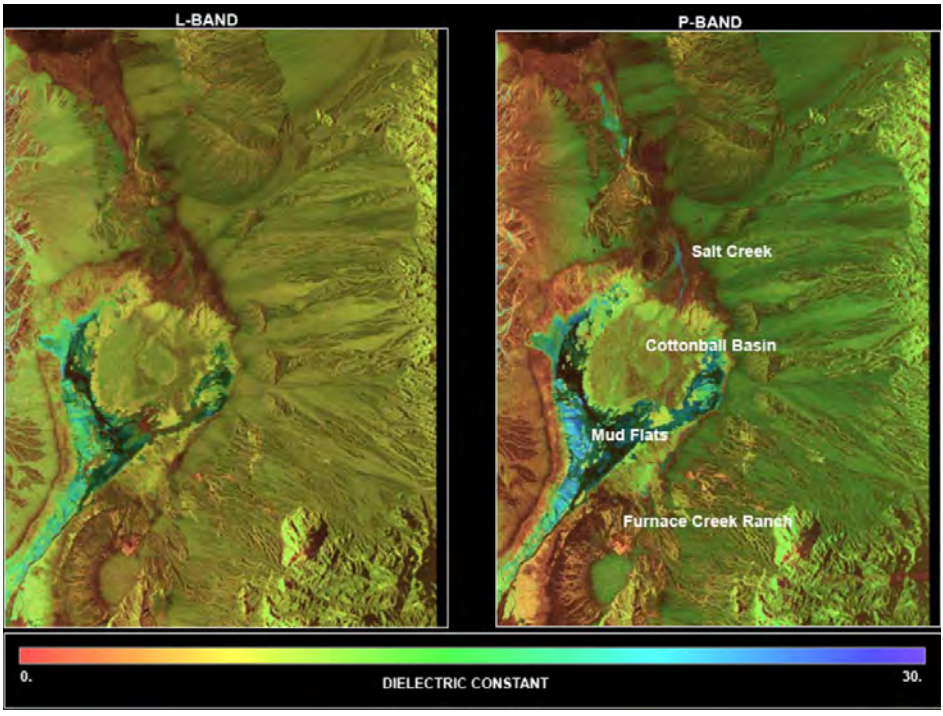


Fig. 5-7. Dielectric constant maps inferred from L-band (left) and P-band (right) AIRSAR data of Cottonball Basin in Death Valley, California. See text for discussion.

larger areas towards the edges of the mudflats. These are the shallower areas of the mud flats. As the mud flats dry out, the surface of the shallower areas dry first, while the sub-surface in these areas can stay wetter longer. The longer wavelength P-band signals more than likely penetrate deeper into these surfaces and sense more of the wetter subsurface than the L-band signals. We also note the same behavior to the northern part of the Basin, next to the word “Salt Creek” in the image. There is an area, located in the Salt Creek, that shows a higher dielectric constant at P-Band. Another area is also visible further north following the Salt Creek further up in the image. Both these areas more than likely represent subsurface moisture. Unfortunately, no actual ground measurements were made during the data collection, so these explanations cannot be verified. But given that the Dubois et al. algorithm consistently infers higher moistures at the longer wavelength in only some areas, and similar values in others, supports this conjecture.

5.2 Scattering from Bare Rough Surfaces

Many excellent texts describing the details of scattering from rough surfaces are available. Here, we will summarize some of the better known models only in

enough detail to set the stage for a better understanding of the inversion models that we will describe later. We will not, however, describe the detailed electromagnetic modeling aspects of these models.

Consider first the case of a perfectly smooth surface of infinite extent that is uniformly illuminated by a plane wave. This surface will reflect the incident wave into the specular direction with scattering amplitudes equal to the well-known Fresnel reflection coefficients. In this case, no scattered energy will be received in any other direction. If now the surface is made finite in extent, or the infinite surface is illuminated by a finite extent uniform plane wave, the situation changes. In this case, the far-field power will decrease proportional to the distance squared (the well-known R-squared law). The maximum amount of reflected power still appears in the specular direction, but a lobe structure, similar to an "antenna pattern," appears around the specular direction. The exact shape of the lobe structure depends on the size and shape of the finite illuminated area, and the pattern is adequately predicted using physical optics calculations. This component of the scattered power is often referred to as the coherent component of the scattered field. For angles far away from the specular direction, there will be very little scattered power in the coherent component.

The next step is to add some roughness to the finite surface such that the mean-square height of the surface is still much less than the wavelength of the illuminating source. The first effect is that some of the incident energy will now be scattered in directions other than the specular direction. The net effect of this scattered energy is to fill the nulls in the "antenna pattern" of the surface described before. The component of the scattered power that is the result of the presence of surface roughness is referred to as the incoherent component of the scattered field. At angles significantly away from the specular direction, such as the backscatter direction at larger incidence angles, the incoherent part of the scattered field usually dominates.

As the roughness of the surface increases, less power is contained in the coherent component and more power is contained in the incoherent component. In the limit where the rms height becomes larger than the wavelength, the coherent component is typically no longer distinguishable, and the incoherent power dominates in all directions. In this limit, the strength of the scattering in any given direction is related to the number of surface facets that are oriented such that they reflect specularly in that direction. This is the same phenomenon that causes the shimmering of the moon on a roughened water surface.

Several different criteria exist to decide if a surface is "smooth" or "rough." The most commonly used one is the so-called Rayleigh criterion that classifies a surface as rough if the rms height satisfies

$$h > \lambda/8\cos\theta\,. \tag{5.2-1}$$

In this criterion, θ is the angle at which the radar wave is incident on the surface. A more accurate approximation of surface roughness was introduced by Peake and Oliver (1971) [29]. According to this approximation, a surface is considered smooth if

$$h < \lambda/25\cos\theta\,. \tag{5.2-2}$$

A surface is considered rough if

$$h > \lambda/4\cos\theta\,. \tag{5.2-3}$$

Any surface that falls in between these two values is considered to have intermediate roughness.

Depending on the angle of incidence, two different approaches are used to model radar scattering from rough surfaces. For small angles of incidence, scattering is dominated by reflections from appropriately oriented facets on the surface. In this regime, physical optics principles are used to derive the scattering equations. As a rule of thumb, facet scattering dominates for angles of incidence less than 20 - 30 degrees (deg). For the larger angles of incidence, scattering from the small scale roughness dominates. The best known model for describing this type of scattering is the small perturbation model. This model, as its name suggests, treats the surface roughness as a small perturbation from a flat surface. More recently, Fung et al. (1992) [10] proposed a model, based on an integral equation method (IEM) solution to the scattering problem, that seems to describe the scattering adequately in both limits. Some refinements to this IEM model have been suggested by Wu et al. (2001) [30].

All models of rough surface scattering assume that the average surface is horizontal. In practice, the scattering is modulated by the local topography; this assumption, therefore, is not necessarily valid. The effect of the local topography can be described as a local tilt to the surface. This tilt is adequately described by the slope of the surface in two orthogonal directions. Note that, in this case, we mean large scale tilts, the scale of which is larger than or equal to the size of the radar pixels. The effect of such a local tilt on the scattering is two-fold. First, the wave impinges on the surface with a different local angle of incidence than it would have if the surface was not tilted. Second, the local tilt rotates the local coordinate system of the surface element with respect to that of the global system in which the radar makes its measurement. This rotation will modify the relative strengths of the scattering coefficients at the different polarizations. We will look at this effect in more detail in the next section.

Several algorithms have been proposed to invert measured radar signals to infer soil moisture. In general, the radar backscatter cross-section is a function of both the surface roughness and the surface dielectric constant (or moisture). Therefore, in order to measure either of these quantities, we must be able to separate their effects on the measured radar backscatter cross-section. Since we are trying to solve for more than one unknown, it follows that more than one measurement is needed. Most algorithms use multi-polarization measurements to accomplish the goal of separating the effects of surface roughness and dielectric constant. The sections that follow briefly discuss several of these models.

5.2.1 First-Order Small Perturbation Model

The use of the first-order small perturbation model to describe scattering from slightly rough surfaces dates back to Rice (1951) [31, 32]. Rice used a perturbation technique to show that, to first order, the scattering cross-sections of a slightly rough surface can be written as

$$\sigma_{xx} = 4\pi k^4 h^2 \cos^4\theta \left|\alpha_{xx}\right|^2 W\left(2k\sin\theta\right); \quad xx = hh \text{ or } vv, \tag{5.2-4}$$

where $k = 2\pi/\lambda$, is the wavenumber, λ is the wavelength, and θ is the local incidence angle at which the radar waves impinge on the surface. The roughness characteristics of the soil are described by two parameters: h is the surface root mean square (rms) height and $W\left(\xi_x, \xi_y\right)$ is the two-dimensional normalized surface roughness spectrum, which is the Fourier transform of the two-dimensional normalized surface autocorrelation function previously discussed in subsection 5.1.1. We note that the surface rms height should be calculated after local slopes have been removed from the surface profile; the slope of the surface changes the radar cross-section because of the change in the local incidence angle. Local slopes that tilt towards or away from the radar do not change the surface roughness; instead, they affect the local incidence angle. This is a frequent source of error in the interpretation of the results from laboratory and field experiments.

The surface electrical properties are contained in the variable α_{xx}, which is given by

$$\alpha_{hh} = \frac{(\varepsilon - 1)}{(\cos\theta + \sqrt{\varepsilon - \sin^2\theta})^2} \tag{5.2-5}$$

and

$$\alpha_{vv} = \frac{(\varepsilon - 1)\left[(\varepsilon - 1)\sin^2\theta + \varepsilon\right]}{(\varepsilon\cos\theta + \sqrt{\varepsilon - \sin^2\theta})^2}. \tag{5.2-6}$$

In these equations, ε is the dielectric constant, or relative permittivity, of the soil. We note that the small perturbation model as described here is applicable only to smooth surfaces. The usual assumptions are that the roughness is small compared to the wavelength (i.e., $kh < 0.3$) and that the rms slope (*s*) satisfies $s < 0.3$.

5.2.2 The Integral Equation Model

Fung et al. (1992) [10] showed that the expressions for the tangential surface fields on a rough dielectric surface can be written as a pair of integral equations. The scattered fields, in turn, are written in terms of these tangential surface fields. Using this formulation, and standard approximations, Fung et al. [10] showed that the scattered field can be interpreted as a single scattering term and a multiple scattering term. When the surface is smooth enough, the single scattering term reduces to the well-known small perturbation model described above and the cross-polarized terms reduce numerically to the second-order small perturbation result. Their results also show that in the high-frequency limit, only the well-known Kirchoff term described by the physical optics model remains significant for surfaces with small rms slopes. When the surface rms slopes are large, however, the multiple scattering terms are important.

Fung et al. (1992) [10] showed that the single scattering backscatter cross-sections can be written as

$$\sigma_{xy} = \frac{k^2}{2}\exp\left(-2k^2h^2\cos^2\theta\right)\sum_{n=1}^{\infty} h^{2n}\left|I_{xy}^n\right|^2 \frac{W^n(-2k\sin\theta, 0)}{n!}, \tag{5.2-7}$$

with

$$I_{xy}^n = (2k\cos\theta)^n f_{xy}\exp\left(-k^2h^2\cos^2\theta\right) + \frac{k^n\cos^n\theta\left[F_{xy}(-k\sin\theta, 0) + F_{xy}(k\sin\theta, 0)\right]}{2}. \tag{5.2-8}$$

The term W^n is the Fourier transform of the *n*th power of the surface correlation function, which can be calculated using

$$W^n(k) = \frac{2}{\pi}\int_0^{\infty} r\rho^n(r) J_o(kr)\, dr, \tag{5.2-9}$$

where $J_0(x)$ is the Bessel function of the first kind and order zero. Also,

$$f_{hh} = \frac{-2R_h}{\cos\theta}; \quad f_{vv} = \frac{2R_v}{\cos\theta}; \quad f_{hv} = f_{vh} = 0, \tag{5.2-10}$$

with R_h and R_v the well-known Fresnel reflection coefficients for horizontal and vertical polarization, respectively. Finally,

$$F_{hh}(-k\sin\theta,0) + F_{hh}(k\sin\theta,0) = \frac{-2\sin^2\theta(1+R_h)^2}{\cos\theta}\left[\left(1-\frac{1}{\mu}\right) + \frac{\mu\varepsilon - \sin^2\theta - \mu\cos^2\theta}{\mu^2\cos^2\theta}\right] \tag{5.2-11}$$

$$F_{vv}(-k\sin\theta,0) + F_{vv}(k\sin\theta,0) = \frac{2\sin^2\theta(1+R_v)^2}{\cos\theta}\left[\left(1-\frac{1}{\varepsilon}\right) + \frac{\mu\varepsilon - \sin^2\theta - \varepsilon\cos^2\theta}{\varepsilon^2\cos^2\theta}\right] \tag{5.2-12}$$

and

$$F_{hv}(-k\sin\theta,0) + F_{hv}(k\sin\theta,0) = 0, \tag{5.2-13}$$

where μ is the relative permeability of the surface and ε is the relative permittivity, or dielectric constant. Note again that the single scattering term does not predict any depolarization. The cross-polarized return is predicted by the multiple scattering term. The expressions are quite complicated and are given in Fung et al. (1992) [10].

Figure 5-8 shows the predicted backscatter cross-section as a function of incidence angle for different surface roughness values and different dielectric constants. The plot on the left shows that increasing the surface roughness generally causes an increase in the radar cross-sections for all polarization combinations. Notice how the difference between the HH and VV cross-sections becomes smaller as the surface gets rougher. The plot on the right shows that increasing the dielectric constant (or soil moisture) also increases the radar cross-sections for all polarizations. In this case, however, increasing the dielectric constant also increases the difference between the HH and VV cross-sections.

5.3 Example Bare Surface Soil Moisture Inversion Models

In this section, we discuss some examples of previously reported soil moisture inversion models. All of these examples try to invert for soil moisture using

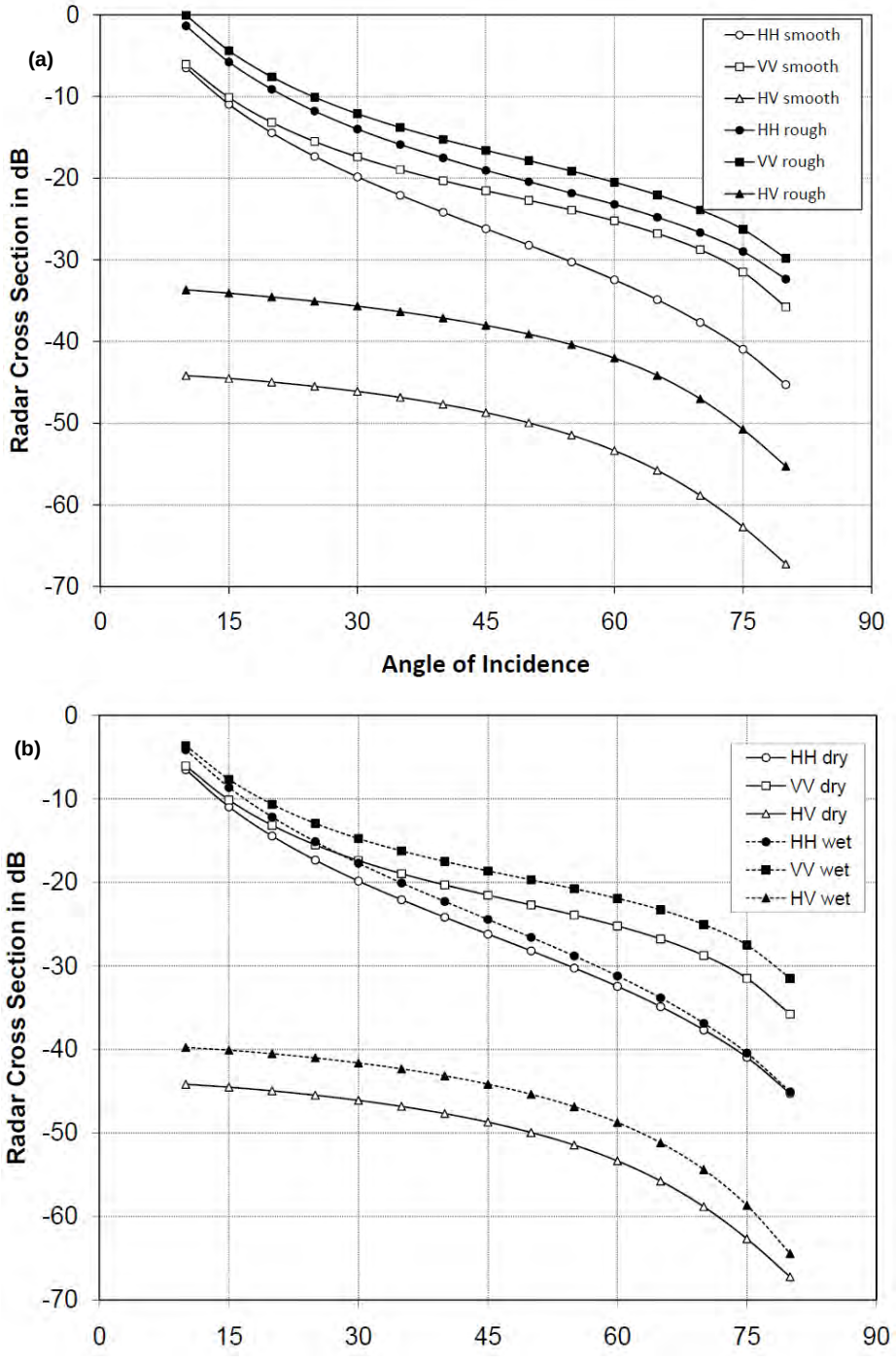


Fig. 5-8. The predicted radar cross-sections for a slightly rough surface, assuming an exponential correlation function. Chart (a) on the top shows the effect of changing surface roughness for constant dielectric constant; chart (b) on the bottom shows the effect of changing dielectric constant for constant roughness.

radar measurements made at one instant in time. As such, they can be considered to attempt to provide a "snap-shot" of soil moisture at the time the radar measurements were made. They commonly use only the radar measurements from that time instant and no other measurements from other times or other sensors.

5.3.1 The First-Order Small Perturbation Model

Returning to the expression for the small perturbation radar cross-section as given in Eq. (5.2-4), we note that the copolarized ratio

$$\frac{\sigma_{hh}}{\sigma_{vv}} = \left|\frac{\alpha_{hh}}{\alpha_{vv}}\right|^2 = \left|\frac{1}{(\varepsilon-1)\sin^2\theta+\varepsilon}\,\frac{(\varepsilon\cos\theta+\sqrt{\varepsilon-\sin^2\theta})^2}{(\cos\theta+\sqrt{\varepsilon-\sin^2\theta})^2}\right|^2 \qquad (5.3\text{-}1)$$

is only a function of the surface dielectric properties, and not of the surface roughness. Therefore, one should be able to use the measure ratio of radar cross-sections to infer the surface dielectric constant. Figure 5-9 shows how this ratio changes smoothly as the dielectric constant increases or as the incidence angle changes.

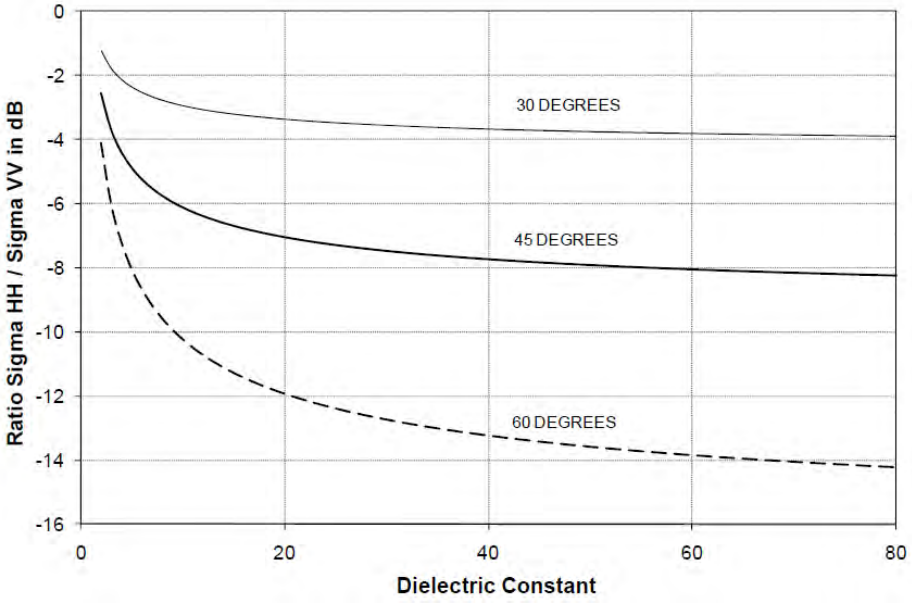


Fig. 5-9. The ratio of σ_{HH} to σ_{VV} as predicted by the small perturbation model. This ratio monotonically decreases as the dielectric constant increases. It also decreases with increasing incidence angle.

If we assume the dielectric constant to be a real number, it is possible to rewrite this expression into a fourth-order polynomial in ε as follows:

$$a_4\varepsilon^4 + a_3\varepsilon^3 + a_2\varepsilon^2 + a_1\varepsilon + a_0 = 0\,, \tag{5.3-2}$$

where

$$\begin{aligned}
a_4 &= \left[1+\sin^2\theta+\sqrt{R}\cos^2\theta\right]^2 \\
a_3 &= -2\left[2\sin^2\theta-\cos^2\theta\left(1+\sin^2\theta\right)+\sin^4\theta-\sqrt{R}\right]\left[1+\sin^2\theta+\sqrt{R}\cos^2\theta\right] \\
&\quad -4\cos^2\theta\left[1+\sqrt{R}+\sin^2\theta\right]^2 \\
a_2 &= \left[2\sin^2\theta-\cos^2\theta\left(1+\sin^2\theta\right)+\sin^4\theta-\sqrt{R}\right]^2 \\
&\quad -2\sin^2\theta\left[1+\sin^2\theta+\sqrt{R}\cos^2\theta\right]\left[\cos 2\theta+\sqrt{R}\right] \\
&\quad +\sin^2 2\theta\left[1+\sqrt{R}+\sin^2\theta\right]\left[3+\sqrt{R}+\sin^2\theta\right] \\
a_1 &= 2\left[2\sin^2\theta-\cos^2\theta\left(1+\sin^2\theta\right)+\sin^4\theta-\sqrt{R}\right]\left[\cos 2\theta+\sqrt{R}\right] \\
&\quad -2\sin^2\theta\sin^2 2\theta\left[1.5+\sqrt{R}+\sin^2\theta\right] \\
a_0 &= \sin^4\theta\left[1+2\sqrt{R}\cos 2\theta+R\right]
\end{aligned} \tag{5.3-3}$$

and

$$R = \frac{\sigma_{vv}}{\sigma_{hh}}\,. \tag{5.3-4}$$

To invert the radar measurements for surface dielectric constant, one then has to find the roots of this fourth-order polynomial. As an example, consider the case of a radar wave incident at 45 deg on a surface with $\varepsilon = 10$. The polynomial in this case is

$$0.240270\varepsilon^4 - 3.01021\varepsilon^3 + 6.56937\varepsilon^2 - 5.06918\varepsilon + 1.26975 = 0\,,$$

with roots

$$\varepsilon = 0.528;\quad 0.996;\quad 1.004;\quad 10.000\,.$$

Solving for the roots of a fourth-order polynomial is computationally quite intensive. Looking again at Fig. 5-9, we note that the co-polarized ratio decreases monotonically as the dielectric constant increases. This type of function is ideally suited to be inverted using a look-up table approach. This approach uses a table with one row for every 1 deg in incidence angle and dielectric constants as integers. Once the four table entries are identified that

bracket the measured value in incidence angle and the dielectric constant, bilinear interpolation is used between these values to find the final dielectric constant. This method provides an increase of approximately a factor 20 in computational speed over solving for the roots of the polynomial with a negligible decrease in accuracy.

We can relax the requirement that the dielectric constant must be real by performing the look-up table calculation using the soil moisture and one of the models that relate soil moisture to the complex dielectric constant. This approach would require making some assumptions about the type of soil: specifically, the texture and salinity. These assumptions must be made in any case if one wants to translate the measured dielectric constant to soil moisture.

As pointed out before, the small perturbation model is only applicable to surfaces that are smooth. While this model predicts the ratio (HH/VV) of the co-polarized radar cross-sections to be independent of the surface roughness, observations show this ratio to increase with increasing roughness. Extending the perturbation model to include second-order terms, one finds that this ratio is, indeed, affected by the surface roughness and that the effect of the roughness is to increase this ratio; that is, to make the VV cross-section closer to that at HH polarization. The net effect of this is that the measured value appears to shift to the left on the curves shown in Fig. 5-9, with the result that a first-order small perturbation inversion will tend to underestimate the surface dielectric constant in the presence of significant roughness. (We will illustrate this more clearly in the next section.) As a result of this, one could argue that the actual dielectric constant of a rough surface will be larger than or equal to the result of a first-order small perturbation inversion.

5.3.2 Algorithm Proposed by Oh et al. (1992)

Based on the scattering behavior in limiting cases and experimental data, Oh et al. (1992) [8] have developed an empirical model in terms of the rms surface height, the wave number, and the relative dielectric constant. The key to this approach is the co-polarization ratio p and cross-polarization ratio q, which are given explicitly in terms of the roughness and the soil dielectric constant. The parameters p and q from the Oh's algorithm are derived using an empirical fit to the data collected by their truck-mounted scatterometer system over bare soils of different roughness and moisture conditions. The explicit expressions for p and q are:

$$p = \frac{\sigma_{hh}}{\sigma_{vv}} = \left\{1 - \left(\frac{2\theta}{\pi}\right)^{1/3\Gamma_0} \cdot \exp(-kh)\right\}^2 \tag{5.3-5}$$

and

$$q = \frac{\sigma_{hv}}{\sigma_{vv}} = 0.23\sqrt{\Gamma_0}\cdot[1-\exp(-kh)], \tag{5.3-6}$$

where

$$\Gamma_0 = \left|\frac{1-\sqrt{\varepsilon}}{1+\sqrt{\varepsilon}}\right|^2. \tag{5.3-7}$$

As before, $k = 2\pi/\lambda$, is the wavenumber, λ is the wavelength, and ε is the complex dielectric permittivity.

Figure 5-10 shows how the co- and cross-polarized ratios are predicted to vary with changing surface roughness and dielectric constant. For large surface roughness, the co-polarized ratio approaches 1, independent of the surface dielectric constant, consistent with experimental observations. For lower dielectric constants, the co-polarized ratio is more sensitive to variations in dielectric constant than in roughness; note, however, that this is not the case for

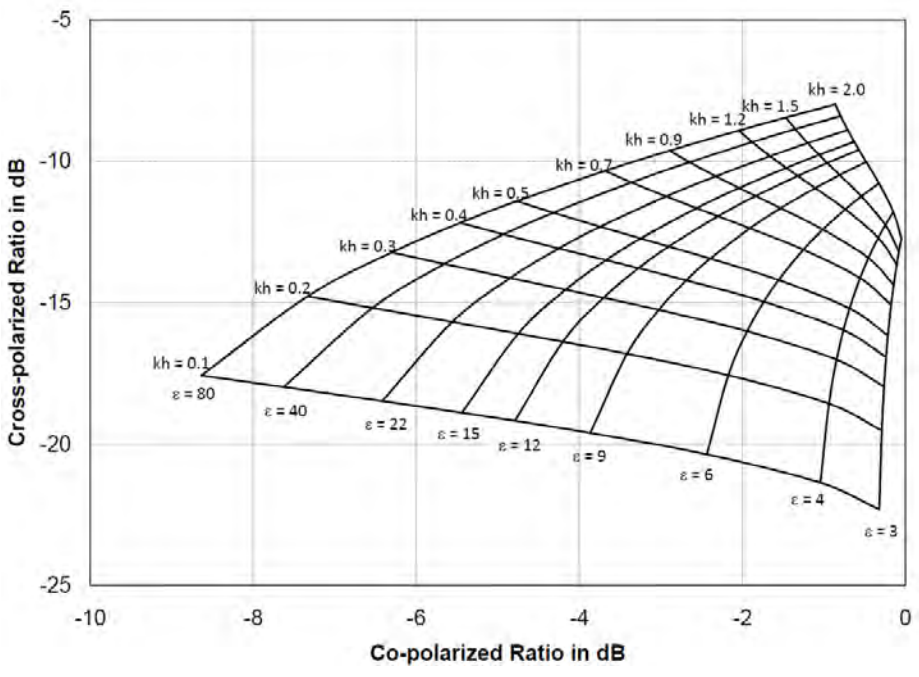


Fig. 5-10. Variation of the co- and cross-polarized ratios as a function of dielectric constant and surface roughness, as predicted by the model proposed by Oh et al. (1992) [8]. The results shown are for a 45-deg incidence angle.

high dielectric constant surfaces. The cross-polarized ratio is always a strong function of surface roughness. For rougher surfaces, this ratio is also quite sensitive to changes in dielectric constant and, in the limit of very rough surfaces, becomes a function of dielectric constant only. We also note that as the surface roughness increases, all the values bunch together more, meaning that we will have less sensitivity to moisture (or roughness) in our measurements, and larger uncertainties in the inferred quantities will likely result in the presence of measurement errors.

To invert the expressions for the soil dielectric constant, we note that we can rewrite Eq. (5.3-5) and Eq. (5.3-6) in terms of only the surface dielectric constant as

$$\sqrt{p} = 1 - \left(\frac{2\theta}{\pi}\right)^{1/3\Gamma_0} \left(1 - \frac{q}{0.23\sqrt{\Gamma_0}}\right). \tag{5.3-8}$$

Two approaches can be followed to solve this nonlinear equation in the dielectric constant: the first is to use an iterative scheme; the second is to use a look-up table approach. We found that the most efficient way (in terms of computational speed) to solve this expression for the dielectric constant is to use a look-up table inversion.

We can derive a similar expression to solve for the surface roughness from Eq. (5.3-5) and Eq. (5.3-6) as

$$\frac{0.0529\left(1 - e^{-kh}\right)^2}{q^2} \log\left(\frac{2\theta}{\pi}\right) - kh = \log\left(1 - \sqrt{p}\right). \tag{5.3-9}$$

As in the case of the dielectric constant, a look-up table approach is used to solve this expression for the rms height from the measured values of the co- and cross-polarized ratios.

One potential practical problem of the algorithm proposed by Oh et al. (1992) [8] is the fact that the cross-polarized ratio is strongly affected by the presence of vegetation. Several studies have shown that the cross-polarized return is strongly correlated with vegetation biomass; and as a result, even a relatively small amount of vegetation increases the cross-polarized ratio significantly over that of a bare surface. The presence of natural vegetation also tends to increase the co-polarized ratio, but to a lesser extent. The result is to artificially move the measurement to the upper right in Fig. 5-8, which means that the dielectric constant of the surface will be underestimated and the surface roughness will be overestimated.

5.3.3 Algorithm Proposed by Dubois et al.

Partly to avoid the difficulties that the Oh et al. (1992) [8] algorithm has when vegetation is present, Dubois et al. (1995) [6] developed an empirical model that only requires measurements of σ_{hh} and σ_{vv} at a frequency between 1.5 and 11 GHz to retrieve both the surface rms height *h* and soil dielectric constant ε from bare soils. They used two sets of ground-based scatterometer data collected by Oh et al. (1992) [8] and by the University of Berne's Radiometer Scatterometer (RASAM) (Wegmuller, 1993 [9]) system to develop two equations that relate the measured co-polarized cross sections to surface roughness and dielectric constant. Those equations are:

$$\sigma_{hh} = 10^{-2.75} \cdot \left(\frac{\cos^{1.5}\theta}{\sin^{5}\theta}\right) \cdot 10^{0.028 \cdot \varepsilon \cdot \tan\theta} \cdot (kh \cdot \sin\theta)^{1.4} \cdot \lambda^{0.7} \qquad (5.3\text{-}10)$$

and

$$\sigma_{vv} = 10^{-2.35} \cdot \left(\frac{\cos^{3}\theta}{\sin^{3}\theta}\right) \cdot 10^{0.046 \cdot \varepsilon \cdot \tan\theta} \cdot (kh \cdot \sin\theta)^{1.1} \cdot \lambda^{0.7}. \qquad (5.3\text{-}11)$$

Note that the wavelength used in these expressions must be in centimeters. These equations have been applied to a number of AIRSAR images of the Little Washita watershed and SIR-C measurements over a bare field by Dubois et al. (1995) [6], and the estimated m_v (using Hallikainen's equations to relate dielectric constant and soil moisture) values were found to agree well with those measured by ground sampling, with a resulting rms error on the order of 4–6 percent [6].

We note that the expressions derived by Dubois et al. [6] predict that the co-polarized ratio will exceed 1 in the limit of surfaces with large rms heights. Experimental results suggest that this ratio will asymptotically approach one, but will not exceed 1. This deficiency in the Dubois et al. model makes their results for rougher surfaces suspect.

Figure 5-11 shows how the VV cross-section and the co-polarized ratio are predicted to vary with changing surface roughness and dielectric constant. In contrast to the model proposed by Oh et al. [8], the co-polarized ratio does not approach 1 for large surface roughness. As mentioned before, this represents one shortcoming in the Dubois et al. model, and this failure to approach 1 is a consequence of the linear approximation they applied to the measured data in the logarithmic domain. All the curves run parallel in both directions, suggesting that the sensitivity of the cross-sections to roughness is independent of the actual value of the dielectric constant and vice versa. This, again, is a consequence of the linear approximation assumed.

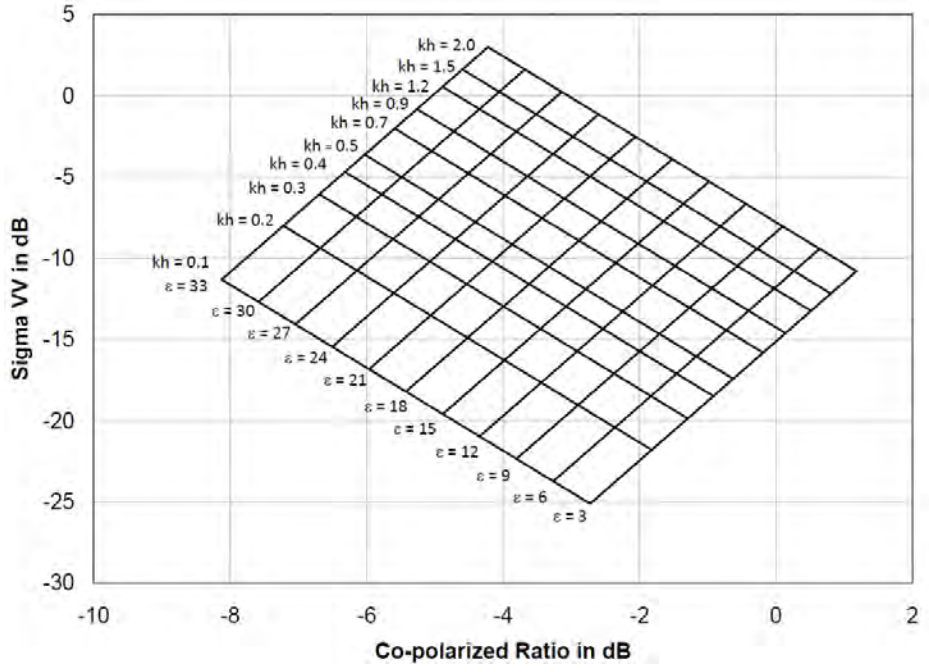


Fig. 5-11. Variation of the VV cross-section and the co-polarized ratio as a function of dielectric constant and surface roughness as predicted by the model proposed by Dubois et al. (1995) [6]. The results shown are for 45 deg incidence angle and L-band.

The Dubois et al. [6] expressions can be rewritten by expressing the radar cross-sections in decibels as follows

$$\sigma_{xx}\left(dB\right) = A_{xx}\varepsilon + B_{xx}\log_{10}(kh) + C_{xx}; \quad xx = hh \; or \; vv \,. \tag{5.3-12}$$

Using the expressions in Eq. (5.3-10) and Eq. (5.3-11), we find that

$$\begin{aligned} A_{hh} &= 0.28\tan\theta \\ B_{hh} &= 14 \\ C_{hh} &= -27.5 + 15\log_{10}\left(\cos\theta\right) - 36\log_{10}\left(\sin\theta\right) + 7\log_{10}\left(\lambda\right) \\ A_{vv} &= 0.46\tan\theta \\ B_{vv} &= 11 \\ C_{vv} &= -23.5 + 30\log_{10}\left(\cos\theta\right) - 19\log_{10}\left(\sin\theta\right) + 7\log_{10}\left(\lambda\right) \end{aligned} \,. \tag{5.3-13}$$

Equation (5.3-12) can be written in matrix form as follows

$$\begin{bmatrix} A_{hh} & B_{hh} \\ A_{vv} & B_{vv} \end{bmatrix} \begin{pmatrix} \varepsilon \\ \log_{10}(kh\cdot) \end{pmatrix} = \begin{pmatrix} \sigma_{hh} - C_{hh} \\ \sigma_{vv} - C_{vv} \end{pmatrix}. \tag{5.3-14}$$

Inverting this expression, we find that the solution for the Dubois et al. algorithm [6] is

$$\begin{pmatrix} \varepsilon \\ \log_{10}(kh\cdot) \end{pmatrix} = \frac{1}{A_{hh}B_{vv} - A_{vv}B_{hh}} \begin{bmatrix} B_{vv} & -B_{hh} \\ -A_{vv} & A_{hh} \end{bmatrix} \begin{pmatrix} \sigma_{hh} - C_{hh} \\ \sigma_{vv} - C_{vv} \end{pmatrix}. \tag{5.3-15}$$

The explicit inversion for the dielectric constant is

$$\begin{aligned} \varepsilon = \frac{1}{3.36\tan\theta}\{&14\sigma_{vv}(dB) - 11\sigma_{hh}(dB) + 26.5 - 255\log_{10}(\cos\theta) \\ &- 130\log_{10}(\sin\theta) - 21\log_{10}(\lambda)\} \end{aligned}. \tag{5.3-16}$$

According to Dubois et al. (1995) [6], their algorithms is applicable to surfaces with $kh < 3.0$ and $30^0 \le \theta \le 70^0$. The range of experimental values available during algorithm development is what places the upper limit on the surface roughness. From Eq. (5.3-15), we find the inversion for the surface roughness to be

$$\begin{aligned} \log(kh) = &-0.083\sigma_{vv}(dB) + 0.137\sigma_{hh}(dB) + 1.807 - 0.446\log_{10}(\cos\theta) \\ &- 3.345\log_{10}(\sin\theta) - 0.375\log_{10}(\lambda) \end{aligned}. \tag{5.3-17}$$

In the presence of vegetation, the co-polarized ratio is affected more than the absolute radar cross section. This means that we can expect the presence of vegetation to move a measurement mostly to the right in Fig. 5-11. The result is that the dielectric constant will be underestimated and the roughness will be overestimated. An obvious advantage of the Dubois et al. model is the simple inversion equations that are easily and efficiently implemented [6].

5.3.4 Algorithm Proposed by Shi et al. (1997) [7]

A concern about the empirical approaches described so far is that these models do not take into account the shape of the surface power spectrum that is related to the surface roughness correlation function and correlation length. This is not consistent with theoretical surface backscattering model predictions: i.e., the backscattering coefficients are sensitive not only to soil moisture m_v and surface root mean square (rms). height *h*, but also to the shape of the surface roughness power spectrum. In addition, any empirical model developed from a limited number of observations might give site-specific results because of the nonlinear response of backscattering to the soil moisture and surface roughness parameters. This drawback might be reduced by using data from many different

sites, but it is very difficult to acquire data experimentally that would cover all possible types of surfaces and include the entire range of expected roughness and moisture conditions.

Progress in theoretical modeling, such as the Integral Equation Method (IEM) (Fung et al., 1992 [10]), offers an alternative approach for the retrieval of soil moisture from radar data. Although the IEM model is valid for a wider range of surface roughness conditions when compared to other earlier theoretical models, the complexity of this model makes its application directly to the radar data to infer soil moisture and roughness parameters rather difficult. Since the number of independent radar measurements is usually limited, Shi et al. (1997) [7] developed a model by parameterizing IEM model-based numerical simulations for a wide range of surface roughness and soil moisture conditions.

Shi et al. (1997) [7] examined many different combinations of polarizations of AIRSAR and SIR-C measurements at L-band to evaluate their effectiveness in the estimation of m_v and h. Several pairs of the measurements were found to be nearly equally effective when the AIRSAR and SIR-C quad-polarized data acquired over the Little Washita watershed were used. Two of these combinations, $\sqrt{\sigma_{hh}\sigma_{vv}}$ and $\sigma_{hh}+\sigma_{vv}$, provided the best fit with three different correlation functions used in the simulation. The following two equations were used to estimate soil moisture and a surface roughness parameter S_R from SIR-C and AIRSAR data (Shi et al., 1997 [7]):

$$10\log_{10}\left[\frac{|\alpha_{vv}|^2}{\sigma_{vv}}\right]=a_{vv}(\theta)+b_{vv}(\theta)10\log_{10}\left[\frac{1}{S_R}\right] \tag{5.3-18}$$

and

$$10\log_{10}\left[\frac{|\alpha_{vv}|^2+|\alpha_{hh}|^2}{\sigma_{vv}+\sigma_{hh}}\right]=a_{vh}(\theta)+b_{vh}(\theta)10\log_{10}\left[\frac{|\alpha_{vv}||\alpha_{hh}|}{\sqrt{\sigma_{vv}\sigma_{hh}}}\right]. \tag{5.3-19}$$

In these expressions, the surface roughness parameter $S_R=(kh)^2W$, k is the wave number, and h is the surface rms height. W is the roughness spectrum of the surface, α_{hh} and α_{vv} are the polarization amplitudes for HH- and VV-polarization as in the small perturbation model that depends only on the dielectric constant of the soil ε and the incidence angle θ, as was previously demonstrated. The coefficients that are functions of the angle of incidence in Eq. (5.3-6) and Eq. (5.3-7) are

$$\begin{aligned} a_{vv}(\theta) &= -6.901 + 5.492\tan\theta - 1.051\log(\sin\theta) \\ b_{vv}(\theta) &= 0.515 + 0.896\sin\theta - 0.475\sin^2\theta \\ a_{vh}(\theta) &= \exp\left[-12.37 + 37.206\sin\theta - 41.187\sin^2\theta + 18.898\sin^3\theta\right] \\ b_{vh}(\theta) &= 0.649 + 0.659\cos\theta - 0.306\cos^2\theta \end{aligned} \quad . \qquad (5.3\text{-}20)$$

To understand the inversion of this model better, we note that it is possible to rewrite Eq. (5.3-19) in the following form:

$$\frac{|\alpha_{vv}|^2 + |\alpha_{hh}|^2}{\left(|\alpha_{vv}||\alpha_{hh}|\right)^{b_{vh}(\theta)}} 10^{-a_{vh}(\theta)/10} = \frac{\sigma_{vv} + \sigma_{hh}}{\left(\sigma_{vv}\sigma_{hh}\right)^{b_{vh}(\theta)/2}} . \qquad (5.3\text{-}21)$$

Figure 5-12 shows the behavior of the left-hand side of Eq. (5.3-21) as a function of dielectric constant.

Conceptually, the inversion algorithm for the Shi et al. algorithm [7] works as follows. The measured co-polarized returns are used to calculate the right side of Eq. (5.3-21). This value is then used in Fig. 5-12, at the appropriate incidence angle, to "read" the dielectric constant. We found a look-up table approach to be the most efficient way to invert Eq. (5.3-21) for the surface dielectric constant.

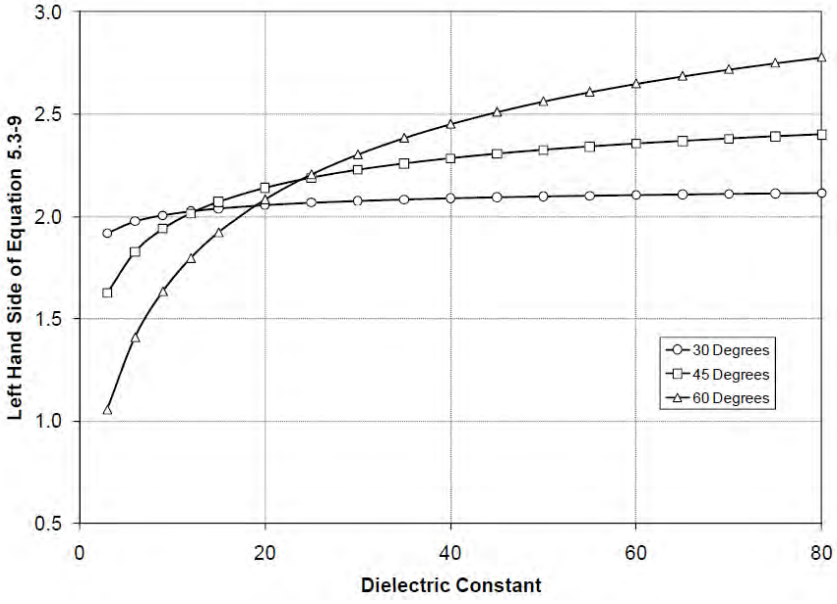


Fig. 5-12. The relationship between dielectric constant and the left hand side of Eq. (5.3-9) is a monotonically increasing function. The sensitivity of the Shi et al. algorithm is better at larger incidence angles.

5.4 Comparison of the Performance of Bare Surface Inversion Models

To investigate the ability of the individual algorithms to take the effect of surface roughness into account during the inversion, we follow this assessment by applying all the algorithms to the same AIRSAR L-band data set of a sparsely vegetated (less than 5% cover) alluvial fan in Death Valley, California, acquired during the summer of 1993. The moisture content of the soil is extremely low this time of the year, with the result that one expects similar low soil moisture throughout the scene. If we invert for the surface dielectric constant, we would expect those areas covered with more rocks to have a slightly higher dielectric constant and the valley floor where salt deposits causes the dielectric constant to be quite high (typically > 40) because of the increase in salinity. The L-band total power image for this data set is shown in Fig. 5-13.

The large alluvial fan is the Trail Canyon fan on the west side of Death Valley. The older, smoother surfaces are shown as the darker features; the rougher,

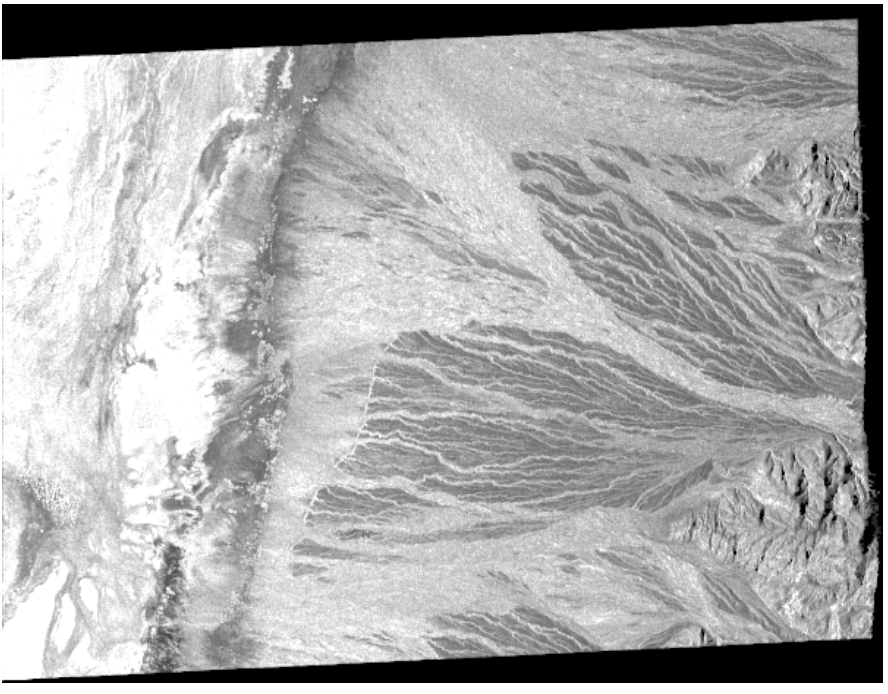


Fig. 5-13. L-band total power image of the Trail Canyon alluvial fan on the west side of Death Valley, California. The radar illumination is from the left. The bright areas on the left in the images are salty soils in the valley floor. The darker areas on the fan surface represent older, smoother fan surfaces. Active stream channels are rougher, appearing bright in the image. The image was acquired in summer when the surfaces are uniformly dry and low dielectric constants are expected.

younger surfaces are brighter, indicating more backscatter. The valley floor is covered with a rougher, salty soil and appears very bright in the image, as does the active stream channels that are covered with rocks and small boulders.

Figure 5-14 shows the inversion result for the dielectric constant for the small-perturbation algorithm described in the previous section. The first obvious observation is that this result does not appear uniform on the alluvial fan, as expected. Instead, the darker areas in the total power image, corresponding to the smoother surfaces, have higher dielectric constants than the brighter, rougher areas. As mentioned before, the ratio of the cross-sections at *hh* and *vv* is a function of the surface roughness. The first-order small-perturbation model fails to take this into account and underestimates the dielectric constant of the rougher surfaces. Note that we have applied the small-perturbation model to the

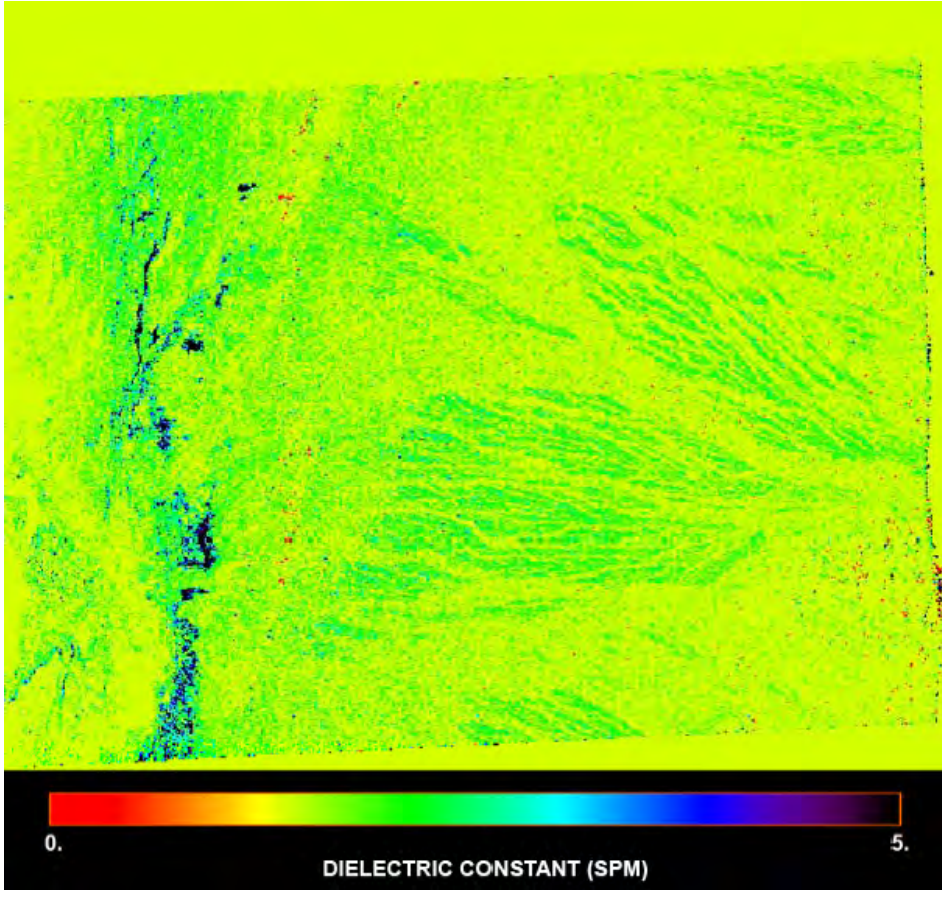


Fig. 5-14. Inversion results using the small-perturbation model algorithm. The algorithm results, as expected, are affected significantly by the changes in surface roughness, and consistently estimates lower values of dielectric constant for the rougher surfaces. Note the scale difference with the other figures following.

entire image even if the surface is rougher than that for which the small-perturbation model is supposed to be applicable. This was done only to illustrate the shortcomings of the small perturbation model in terms of not taking into account the effects of surface roughness on the HH/VV ratio.

Figure 5-15 shows the inversion result for the algorithm proposed by Oh et al. (1992) [8]. The algorithm also estimates higher dielectric constants on the smoother surfaces, but the difference is somewhat less pronounced than in the case of the small perturbation model. The inferred dielectric constant values are higher than that of the small-perturbation model (note the different scales), indicating that the Oh et al. algorithm does a better job of taking the effects of surface roughness into account. However, it is clear from the results that this algorithm still fails to completely separate the effects of surface roughness and dielectric constant.

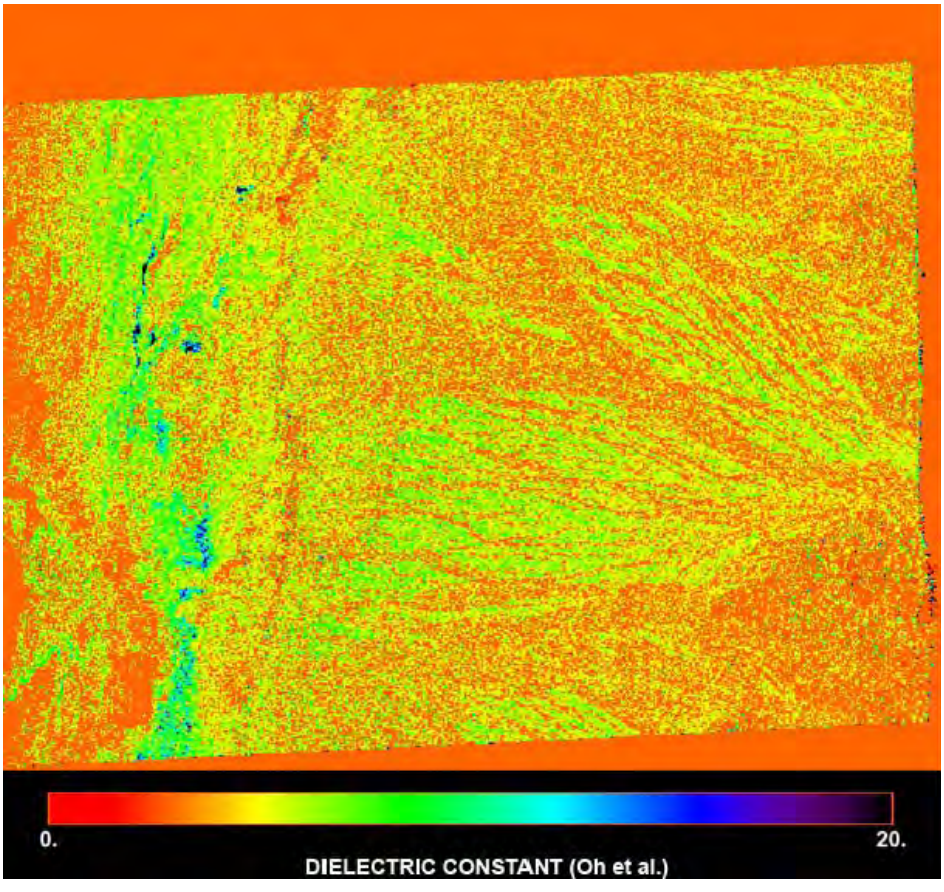


Fig. 5-15. Inversion results using the algorithm proposed by Oh et al. [8]. The algorithm results are affected significantly by the changes in surface roughness, and consistently estimate lower values of dielectric constant for the rougher surfaces.

Figure 5-16 shows the inversion results for the Dubois et al. algorithm [6]. This algorithm clearly does a better job than the others at removing the effects of surface roughness. It also clearly shows that the saline valley floor has a higher dielectric constant than the dry alluvial fan surface. The values of the dielectric constants calculated on the alluvial fan are consistent with dry soils that are covered by desert pavement. Also note that, although quite noisy, the inversion does indicate that the dielectric constant of the active stream channels is slightly higher than that of the fan surface, consistent with having more rocks and boulders and less exposed soil present in the active stream channels.

Figure 5-17 contains the inversion results for the algorithm proposed by Shi et al. (1997) [7]. Note that this algorithm, in contrast to the small-perturbation model, and the Oh et al. [8] algorithm, reports lower dielectric constants on the smoother surfaces than the rougher surfaces. It also fails to report high

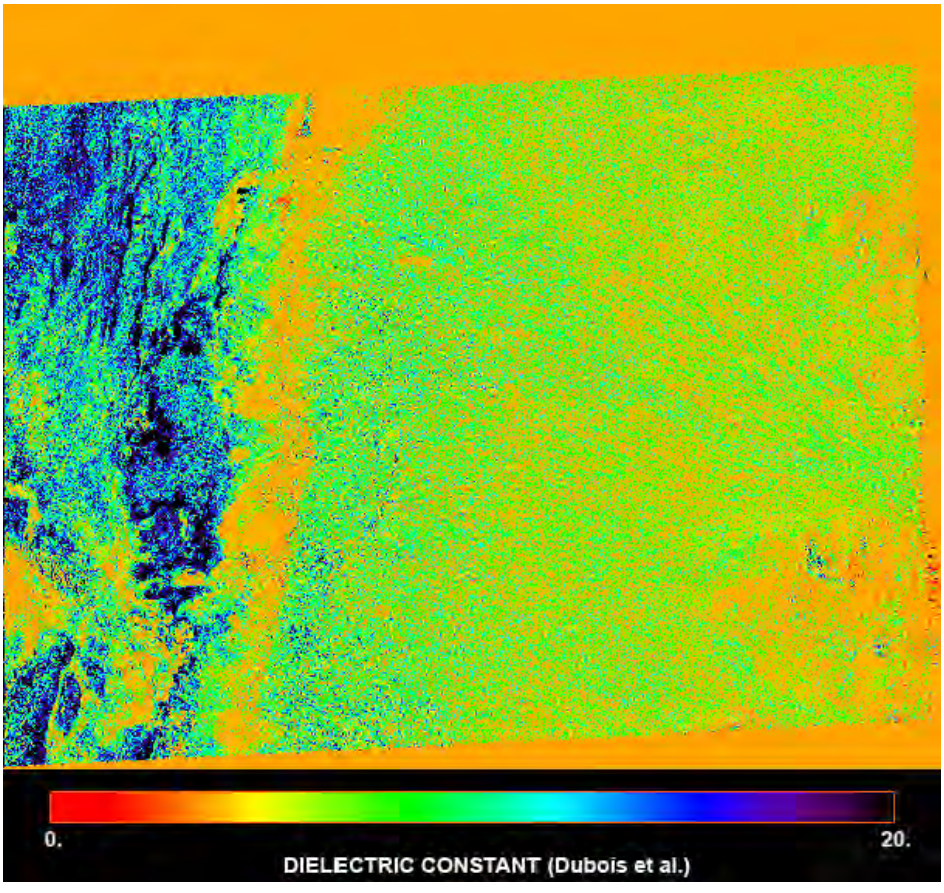


Fig. 5-16. Inversion results using the Dubois et al. algorithm. The algorithm does well at removing the effects of surface roughness. It also does better at indicating the higher dielectric constants of the saline valley floor on the left.

dielectric constants for the saline valley floor. To understand the behavior of this algorithm in the presence of roughness, let us consider the right-hand side of Eq. (5.3-21), repeated below:

$$\frac{|\alpha_{vv}|^2 + |\alpha_{hh}|^2}{(|\alpha_{vv}||\alpha_{hh}|)^{b_{vh}(\theta)}} 10^{-a_{vh}(\theta)/10} = \frac{\sigma^{\circ}_{vv} + \sigma_{hh}}{(\sigma_{vv}\sigma_{hh})^{b_{vh}(\theta)/2}}.$$

Very rough surfaces exhibit co-polarized radar cross-sections that are equal, while slightly rough surfaces exhibit larger VV radar cross-sections than HH. It is, therefore, clear that for these two measurements to become equal, the HH cross-section must increase faster than the VV cross-section with increasing surface roughness in order for the former to "catch up" and become equal to the VV cross-section. This is what the empirical Dubois et al. algorithm [6] shows.

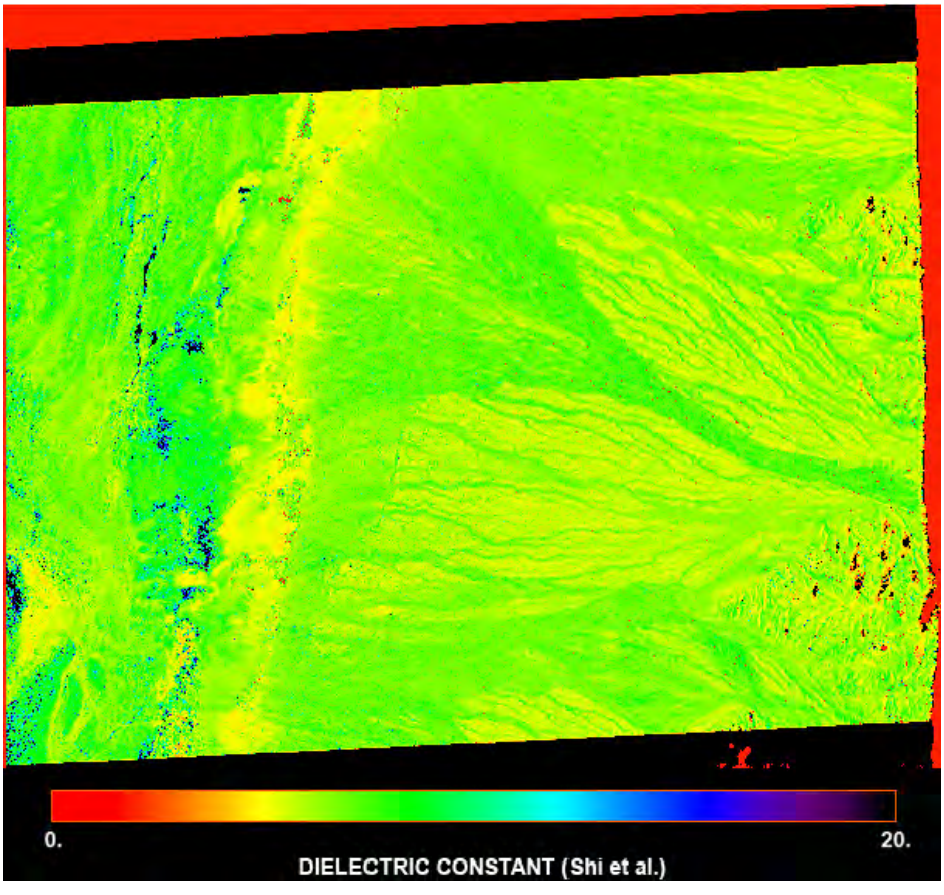


Fig. 5-17. Inversion results for the Shi et al. algorithm [7]. The algorithm seems to over-compensate for roughness, inferring lower values on the smoother surfaces than the rougher surfaces.

To help understand the results in Fig. 5-17, let us assume (as demonstrated by the Dubois et al. results) that

$$\sigma_{hh} \sim (kh)^x; \quad \sigma_{vv} \sim (kh)^y; \quad x \geq y .$$

The right side of Eq. (5.3-21) then becomes

$$\frac{\sigma_{vv} + \sigma_{hh}}{(\sigma_{vv}\sigma_{hh})^{b_{vh}(\theta)/2}} \sim \frac{(kh)^x + (kh)^y}{(kh)^{(x+y)b_{vh}(\theta)/2}} .$$

Neither x nor y are functions of the radar angle of incidence. It follows, therefore, that one could perhaps make the Shi et al. algorithm [7] insensitive to surface roughness at a particular incidence angle; however, this would not be the case at different incidence angles. In practice, however, $b_{vh}(\theta)$ changes slowly with incidence angle, so the sensitivity to roughness does not vary much with incidence angle. If we take the values found by Dubois et al. [6] for x and y and use the value of $b_{vh}(\theta)$ at 45 deg angle of incidence, we find

$$\frac{\sigma_{vv} + \sigma_{hh}}{(\sigma_{vv}\sigma_{hh})^{b_{vh}(\theta)/2}} \sim \frac{(kh)^{1.4} + (kh)^{1.1}}{(kh)^{0.48(1.1+1.4)}} = (kh)^{0.2} + (kh)^{-0.1} .$$

One should, therefore, expect the Shi et al [7] results to be slightly affected by surface roughness.

The next point to investigate is whether this argument results in an over-estimation of the surface dielectric constant in the presence of increased roughness, as Fig. 5-17 indicates. Note that the discussion above shows that the right side of Eq. (5.3-21) increases with increasing surface roughness. The left side of Eq. (5.3-21) is a monotonically increasing function of the surface dielectric constant. Therefore, increasing the right side of Eq. (5.3-21) (artificially) in the presence of surface roughness would, indeed, cause the algorithm to estimate a higher dielectric constant than the actual value.

5.5 Parameterizing Scattering Models

Most scattering models are far too computationally intensive to use in practical inversion algorithms for large images. This is especially the case when the answers are needed on relatively short timescales, such as “near real-time” applications that require soil moisture values as inputs. One way to make this problem more tractable is to use “parameterized” versions of the scattering model with simpler functions that are easily inverted.

To illustrate this process, let us postulate that we can approximate the measured radar cross-section σ_{ij} using $\hat{\sigma}_{ij}$ as follows:

$$10\log\hat{\sigma}_{ij} = A_{ij} + B_{ij} f\left(m_v\right) + C_{ij} g\left(kh\right); \quad i, j = h, v. \tag{5.5-1}$$

Here we use generic functions f and g to denote our hypothesis about the influence of soil moisture and surface roughness on the radar cross-section. The subscript ij refers to the polarization combination used to measure (or in this case simulate) the radar cross-section. We should point out that the coefficients in this function are all functions of the angle of incidence. An example of such a postulate was employed by Dubois et al. (1995) [6], who used the dielectric constant itself as f and the logarithm for g. We typically want to minimize the error between the measured and the estimated radar cross-sections in a least-squares sense. This error for *n* measurements is given by

$$E = \sum_{n=1}^{N}\left(10\log\sigma_{ijn} - 10\log\hat{\sigma}_{ijn}\right)^2 = \sum_{n=1}^{N}\left(10\log\sigma_{ijn} - A_{ij} - B_{ij} f\left(m_{vn}\right) - C_{ij} g\left(kh_n\right)\right)^2. \tag{5.5-2}$$

The solution that minimizes this error is

$$\begin{pmatrix} A_{ij} \\ B_{ij} \\ C_{ij} \end{pmatrix} = \begin{pmatrix} 1 & \langle f \rangle & \langle g \rangle \\ \langle f \rangle & \langle f^2 \rangle & \langle fg \rangle \\ \langle g \rangle & \langle fg \rangle & \langle g^2 \rangle \end{pmatrix}^{-1} \begin{pmatrix} \langle 10\log\sigma_{ij} \rangle \\ \langle 10\log\sigma_{ij} f \rangle \\ \langle 10\log\sigma_{ij} g \rangle \end{pmatrix}, \tag{5.5-3}$$

where the angular brackets $\langle\ \rangle$ denote averaging over all measurements. The question then becomes which functions f and g would provide the best results.

As an example, we use simulated data from the integral equation method at an angle of incidence of 45 deg. We vary the soil moisture from 1 percent to 40 percent and the values of *h* from 0.1 cm to 3 cm. At L-band, this would correspond to values of *kh* that vary from 0.026 to 0.78. We then use different combinations of functions to parameterize the values predicted by the integral equation model.

In the first case, we use $f(m_v) = m_v$. Several other researchers have shown a linear relationship between the logarithm of the radar cross-section and soil moisture (Quesney et al., 2000 [5]). For the function *g*, we follow Dubois et al. (1995) [6] and use the logarithm of *kh*. Using the data described above, we find

$$\begin{aligned} 10\log\hat{\sigma}_{hh} &= -20.17 + 15.33m_v + 13.63\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -18.81 + 25.33m_v + 10.99\log(kh) \end{aligned}. \tag{5.5-4}$$

Note that the HH cross-section increases faster with roughness than the VV, explaining why the difference between HH and VV is observed to be smaller for rougher surfaces than for smoother ones. Conversely, the VV cross-section increases faster with an increase in soil moisture, again explaining why the difference between HH and VV is observed to be larger for wetter surfaces than for drier ones. Also, note that the coefficients for the roughness are very similar to those determined by Dubois et al. (1995) [6], as is shown in Eq. (5.3-13), even though their coefficients were determined from measured data rather than model simulations.

To determine how well these parameterizations approximate the model, we compare the accuracies of the parameterized inversions. We start with simulating 1000 model predictions in which both the surface roughness and the soil moisture are treated as random variables. We then take these 1000 pairs of HH and VV cross-sections and invert them for soil moisture and surface roughness using the expressions for the two parameterizations. We then compare the results of the inversions to the actual random values of soil moisture and surface roughness that went into the simulation.

To invert the parameterizations, we note that (see Eq. (5.5-1)) we can write

$$\begin{aligned} 10\log\hat{\sigma}_{hh} &= \Sigma_{hh} = A_{hh} + B_{hh} f(m_v) + C_{hh} g(kh) \\ 10\log\hat{\sigma}_{vv} &= \Sigma_{vv} = A_{vv} + B_{vv} f(m_v) + C_{vv} g(kh) \end{aligned}. \tag{5.5-5}$$

We can write this in matrix form as follows

$$\begin{pmatrix} B_{hh} & C_{hh} \\ B_{vv} & C_{vv} \end{pmatrix} \begin{pmatrix} f(m_v) \\ g(kh) \end{pmatrix} = \begin{pmatrix} \Sigma_{hh} - A_{hh} \\ \Sigma_{vv} - A_{vv} \end{pmatrix}, \tag{5.5-6}$$

which has the solution

$$\begin{pmatrix} f(m_v) \\ g(kh) \end{pmatrix} = \begin{pmatrix} B_{hh} & C_{hh} \\ B_{vv} & C_{vv} \end{pmatrix}^{-1} \begin{pmatrix} \Sigma_{hh} - A_{hh} \\ \Sigma_{vv} - A_{vv} \end{pmatrix}. \tag{5.5-7}$$

Figure 5-18 shows the results for the inversion. The rms error in estimating the surface roughness is 0.25 cm; the rms error in soil moisture is 5.1 percent.

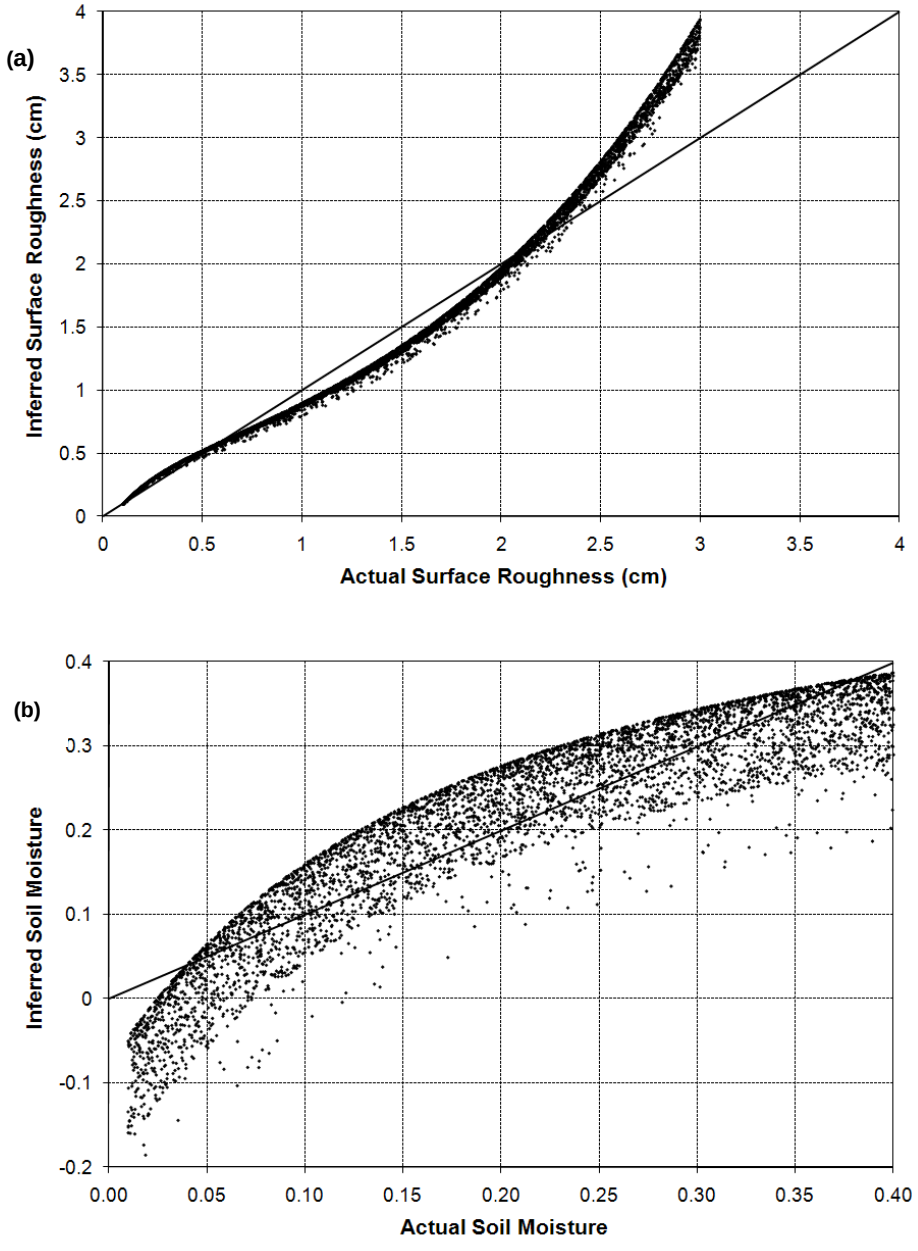


Fig. 5-18. Inversion results for the parameterization shown in Eq. (5.5-4) for (a) surface roughness and (b) soil moisture. The rms error in surface roughness is 0.25 cm; the rms error in soil moisture is 5.1 percent.

Note that particularly the soil moisture is often underestimated. Many of the lower soil moisture cases are inferred to have negative soil moistures. This is a consequence of the linear approximation used for the soil moisture in this parameterization. In fact, the same behavior is observed for the dielectric constant of drier surfaces when using the Dubois et al. (1995) [6] algorithm, as was pointed out by Wang et al. (1995) [33].

As a second case, we use for the function *f* the logarithm of the soil moisture and for *g* the logarithm of *kh*. This choice for the function *f* is an effort to take into account the fact that the radar cross-section is, in fact, a non-linear function of the soil moisture. In this case, we find

$$\begin{aligned} 10\log\hat{\sigma}_{hh} &= -12.88 + 5.23\log(m_v) + 13.63\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -6.87 + 8.51\log(m_v) + 10.99\log(kh) \end{aligned} \quad . \tag{5.5-8}$$

The results for this case are shown in Fig. 5-19. The rms error in surface roughness is 0.26 cm; the rms error in soil moisture is 6.4 percent. While we have removed the negative inferred soil moisture values, there is in the inferred results a larger spread, especially for larger soil moisture values.

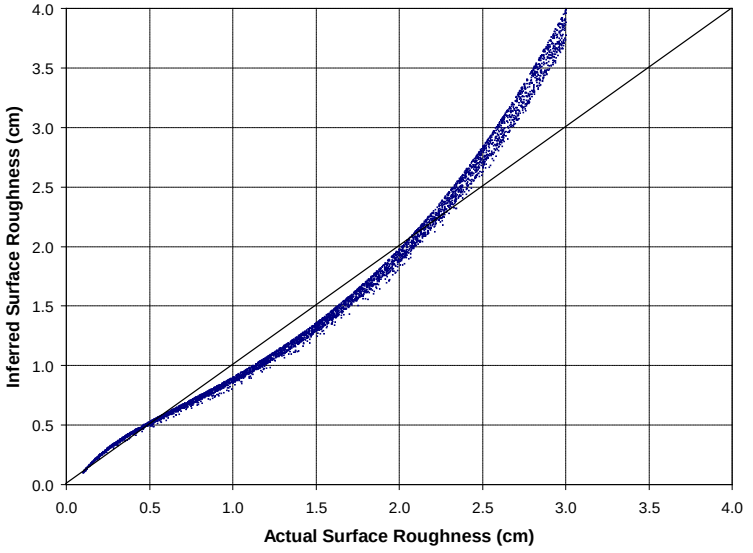
These two examples illustrate the difficulty in parameterizing non-linear functions with simple approximations. But is there a way to use these simple functions (which are easy to invert) to achieve better accuracy? The answer lies in using smaller ranges of moisture and surface roughness when making the approximations. As an illustration, let us assume we want to use the first approximation where we use a linear approximation for the soil moisture and a logarithmic approximation for the surface roughness. We shall assume that we want to have an accuracy of 0.1 cm or better for the surface roughness and an accuracy of 1 percent or better for soil moisture over the ranges we assumed before. We can achieve these accuracies if we are willing to break the ranges up into smaller segments. An example is shown below. For $0.01 \le m_v \le 0.1$:

$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -18.77 + 42.31m_v + 16.46\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -15.86 + 63.18m_v + 16.34\log(kh) \end{aligned}\right\} \text{for } 0.1 \le h \le 0.5$$

$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -23.13 + 42.31m_v + 12.17\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -21.80 + 63.36m_v + 10.12\log(kh) \end{aligned}\right\} \text{for } 0.5 \le h \le 1.5 \quad . \tag{5.5-9}$$

$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -21.72 + 42.31m_v + 15.73\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -22.77 + 64.25m_v + 7.82\log(kh) \end{aligned}\right\} \text{for } 1.5 \le h \le 3.0$$

(a)



(b)

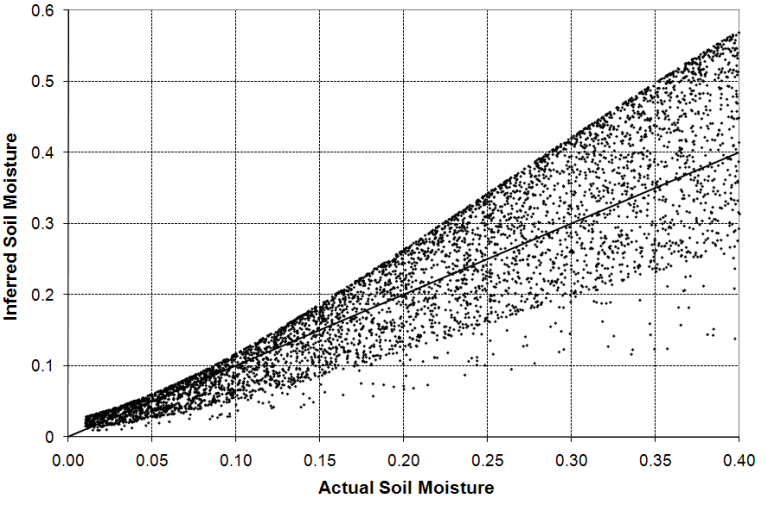


Fig. 5-19. Inversion results for the parameterization shown in Eq. (5.5-8) for (a) actual surface roughness and (b) actual soil moisture. The rms error in surface roughness is 0.26 cm; the rms error in soil moisture is 6.4 percent.

For the range $0.1 \le m_v \le 0.2$:

$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -16.46 + 17.54m_v + 16.46\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -12.64 + 29.08m_v + 16.34\log(kh) \end{aligned}\right\} \text{for } 0.1 \le h \le 0.5$$
$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -20.82 + 17.53m_v + 12.17\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -18.54 + 29.14m_v + 10.17\log(kh) \end{aligned}\right\} \text{for } 0.5 \le h \le 1.5 \cdot \qquad (5.5\text{-}10)$$
$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -19.41 + 17.53m_v + 15.73\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -19.42 + 29.46m_v + 8.11\log(kh) \end{aligned}\right\} \text{for } 1.5 \le h \le 3.0$$

Finally, for the range $0.2 \le m_v \le 0.4$:

$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -14.37 + 7.32m_v + 16.46\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -9.34 + 12.99m_v + 16.34\log(kh) \end{aligned}\right\} \text{for } 0.1 \le h \le 0.5$$
$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -18.74 + 7.32m_v + 12.17\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -15.22 + 13.00m_v + 10.19\log(kh) \end{aligned}\right\} \text{for } 0.5 \le h \le 1.5 \cdot \qquad (5.5\text{-}11)$$
$$\left.\begin{aligned} 10\log\hat{\sigma}_{hh} &= -17.32 + 7.32m_v + 15.73\log(kh) \\ 10\log\hat{\sigma}_{vv} &= -16.04 + 13.10m_v + 8.25\log(kh) \end{aligned}\right\} \text{for } 1.5 \le h \le 3.0$$

Note that these boundaries were not optimized; the actual segments were chosen to illustrate the point. The question now is how to invert the measurements using these nine subspaces. We simply use the original set of expressions in Eq. (5.5-4) as a starting point. Based on the results of the inversion, we then use the subspace wherein the initial inversion result fall. We then use the updated set of expressions to perform the inversion. We find that a small number of iterations are generally required to find the best results. In our test, fewer than 10 iterations are typically required for the result to stabilize. The iterations are performed using the results of the previous calculation to decide on which of the nine expressions to use for the next iteration. Once the results change less than a pre-defined amount, we stop the iterations. The results of this scheme are shown in Fig. 5-20. This clearly represents a significant improvement over the single set of expressions.

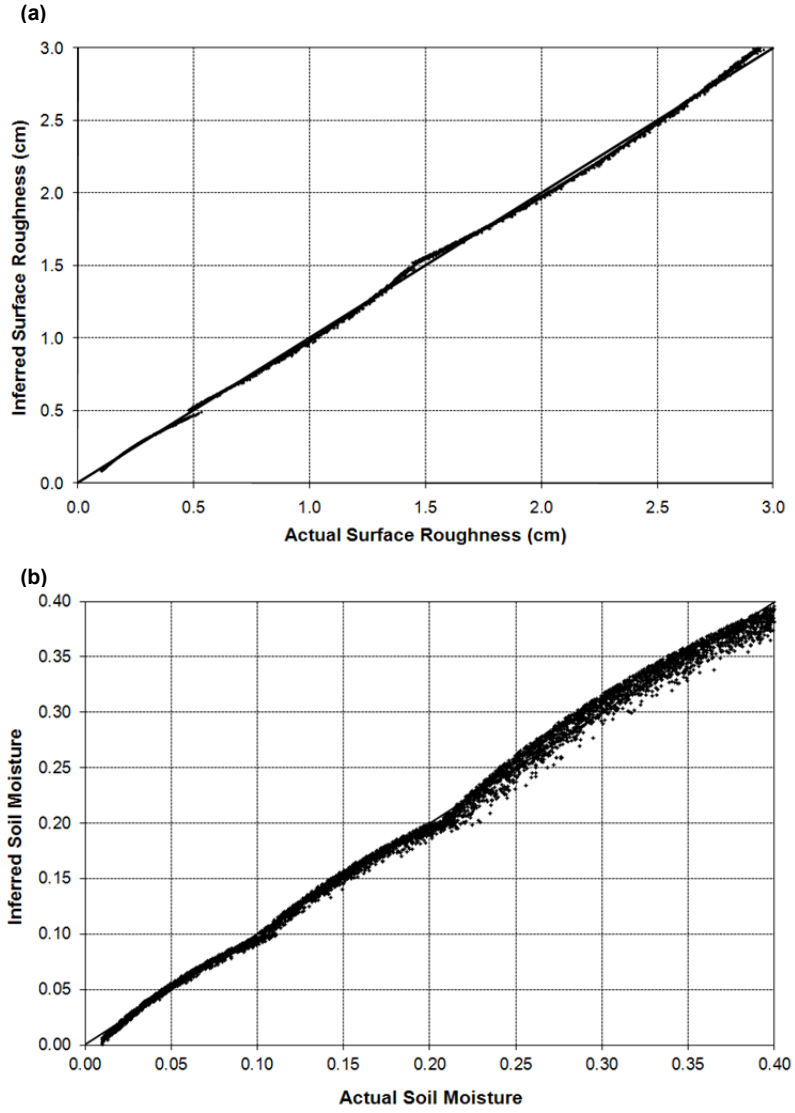


Fig. 5-20. Inversion results for the parameterization that breaks the data space into nine smaller spaces showing (a) inferred surface roughness and (b) inferred soil moisture. The rms error in surface roughness is 0.26 mm, and the rms error in soil moisture is 0.67 percent.

The discussion in this section illustrates the difficulties with parameterizing complicated non-linear expressions with simpler ones. While the inversion might be much more efficiently performed numerically using simple expressions, it is not easy to find a parameterization that is valid over a large range of geophysical variable values. Once we start adding more expressions to

better approximate sub-ranges of the geophysical parameter, the advantages become less clear. With computer capabilities rapidly increasing, direct numerical inversion of the forward radar model is becoming more feasible. We shall explore this in more detail below.

5.6 Inverting the IEM Model

The discussion in the previous section illustrates the difficulty of parameterizing non-linear models, such as the IEM, over a large range of parameters, which, in our case, includes surface roughness, soil moisture, and the angle of incidence. The main advantage of parameterizing such a model is the ease and speed with which the parametrized expressions can be inverted. The price we pay is that the inversion is less accurate.

As computer capabilities grow, however, it might become feasible to "invert" models like the IEM using some merit function to describe how different a simulated and observed result are and to then adjust the input parameters of the simulation to minimize this difference. The obvious advantage is that we no longer have to rely on functions that have a limited range of accuracy. The disadvantage is that the inversion might be much slower than a parameterized case.

To illustrate, we will look at two different approaches. The first approach uses a simple cube of pre-calculated IEM simulations for each polarization combination that is stored on a computer for later use. The inversion then reads the cube of data and uses the values in the cube to estimate the geophysical parameters based on the observed radar cross-sections. The second approach uses the downhill simplex optimization approach to minimize the error between the simulated data and the observed radar cross-sections. These two approaches will be discussed separately below.

For a merit function, we define the rms error between the measured and observed values as follows:

$$E = \sqrt{\left(\sigma_{hh\,obs} - \sigma_{hh\,sim}\right)^2 + \left(\sigma_{vv\,obs} - \sigma_{vv\,sim}\right)^2}\,, \qquad (5.6\text{-}1)$$

where the subscripts *obs* and *sim* refer to the observed and simulated values, respectively. The error function can be extended to include more polarization combinations by simply adding terms inside the square root.

5.6.1 Using a Data Cube

The inversion using a data cube utilizes a previously calculated cube of values for each of the HH and VV radar cross-sections. The three parameters making

up the cube are the surface roughness, the soil moisture, and the angle of incidence. For our illustration, we use a constant ratio of the surface roughness and the surface correlation length. Of course, one could use a four-dimensional data set where the correlation length is explicitly included as one of the parameters. We only have to calculate this data cube once; we can then store the data in a file for later use. For our purposes, we used a cube that is calculated at intervals of 0.5 deg in the angle of incidence, and 512 values in each of the surface roughness and the soil moisture. We allowed the soil moisture and surface roughness values to range over the same values as in the previous section.

When inverting the data, we first calculate the error for the entire data space, using both the HH and VV data cubes. We then select the point where the error is the minimum in the data cube, and interpolate the neighborhood of that point to a finer grid. We then find the minimum error on this finer grid, and report the surface parameters for this point as the inferred ones. The results are shown in Fig. 5-21. The rms error for the surface roughness is 0.0009 cm; and the rms error for the soil moisture is 0.06 percent.

Considering the results in Fig. 5-21, we notice that the soil moisture error is larger for the larger absolute values of the soil moisture. This can be seen by the larger spread of the values in the upper left side of the figure on the right in Fig. 5-21. This is a consequence of the fact that the radar cross-sections saturate as the moisture values become larger. In that case, a small error in roughness estimation leads to a larger error in soil moisture. Everything considered, however, these results are excellent.

The results in Fig. 5-21 were calculated for a fixed angle of incidence of 40 deg. We repeated the experiment by randomly varying the angle of incidence as well as the roughness and soil moisture values. When inverting the simulated data, we assume we know the angle of incidence. We then interpolate linearly between the two closest planes in the HH and VV data cubes to generate two data sets for that angle of incidence. These data sets are then used to invert the simulated data. The results are shown in Fig. 5-22.

Very similar results are found when using the downhill simplex method to invert the IEM model. Both these methods provide excellent results, but they come at the price of increase computational complexity. Even the cube method takes about an order of magnitude longer to invert than the simple parameterized models.

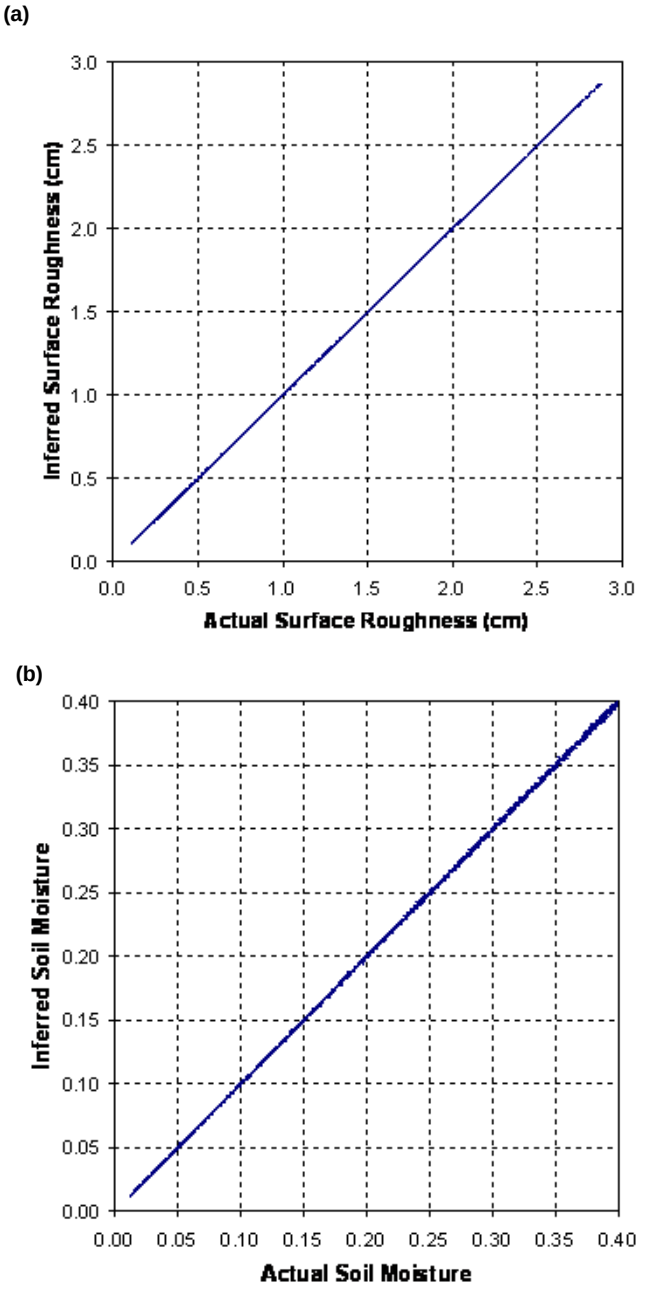


Fig. 5-21. Inversion results using a data space with 512 values for each of the (a) surface roughness and (b) soil moisture. The rms error in surface roughness is 0.009 mm, and the rms error in soil moisture is 0.06 percent. The results are shown for 5000 simulations, with random roughness and moisture values at a fixed angle of incidence of 40 deg.

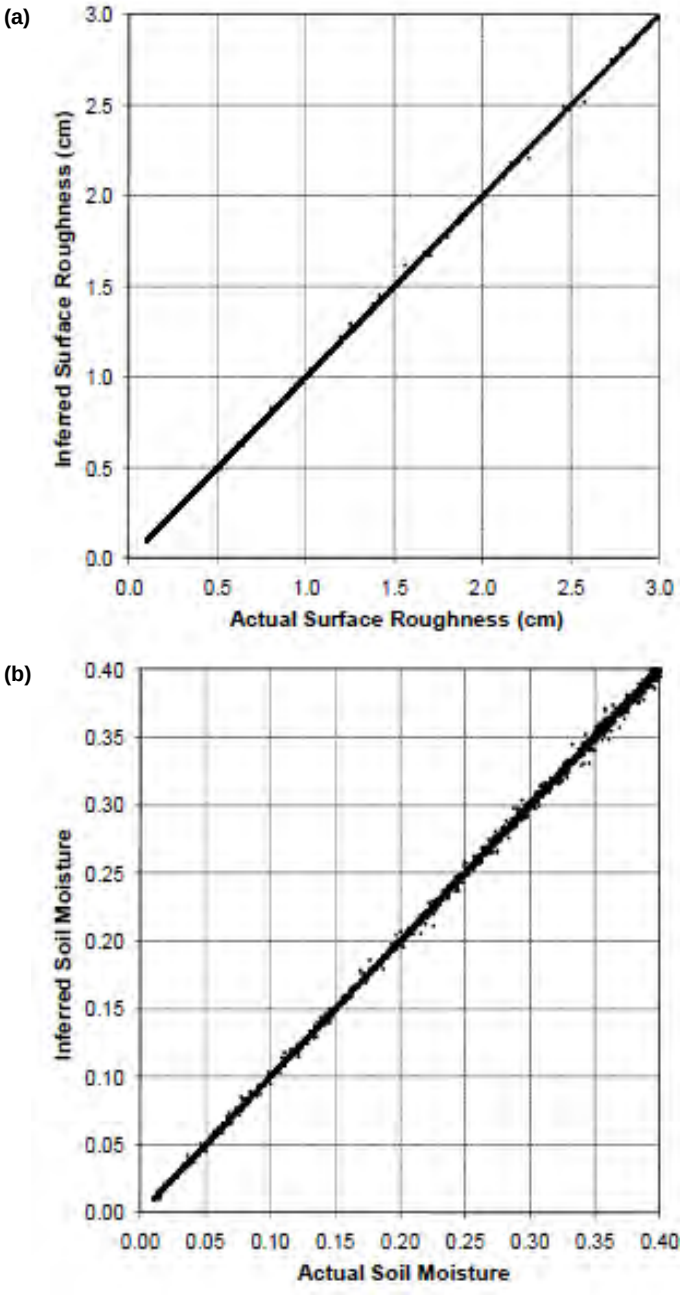


Fig. 5-22. Inversion results using a data space with 512 values for each of the (a) surface roughness and (b) soil moisture. The rms error in surface roughness is 0.03 mm, and the rms error in soil moisture is 0.16%. The results are shown for 5000 simulations with random roughness and moisture values and random incidence angles between 10 and 60 deg.

5.7 Scattering from Vegetated Terrain

Estimating soil moisture from vegetated terrain is significantly more complicated than the bare surface case. Unfortunately, however, much of the earth's surface is covered by some vegetation, so the bare surface approach has limited application. Here we shall discuss briefly how models for backscattering from vegetated terrains are constructed, and then show how these models can aid in estimating soil moisture in vegetated terrain.

Here we shall illustrate the modeling approach using a single layer vegetation model. This type of model would be used to describe scattering from pasture land, grasslands, most agriculture crops and shrubs. To model forest canopies, one or two more layers of vegetation would be added as described by Durden et al. (1989) [34].

The vegetation layer will be modeled assuming that there are two interfaces to consider (see Fig. 5-23): 1) vegetation layer, and 2) the underlying ground surface. The vegetation layer is comprised of primary scatterers, and possibly secondary scatterers. For example, the primary scatterers may be stalks in the case of corn, while the secondary scatterers may be the leaves. In the case of pasture, we will typically use only primary scatterers.

The vegetation layer has a thickness b. Each component of the vegetation (primary and secondary scatterers) is characterized by a dielectric constant $(\varepsilon_p, \varepsilon_s)$, radius (a_p, a_s), length (l_p, l_s), density (ρ_p, ρ_s) and a probability density function ($p_p(\theta_c, \phi_c), p_s(\theta_c, \phi_c)$) describing the statistical distribution of the orientation of the cylinders used to represent the component. The ground surface is characterized by a dielectric constant ε_g, rms. height h, correlation length l_g. The small-scale slopes (s_l) describe the micro-roughness of the

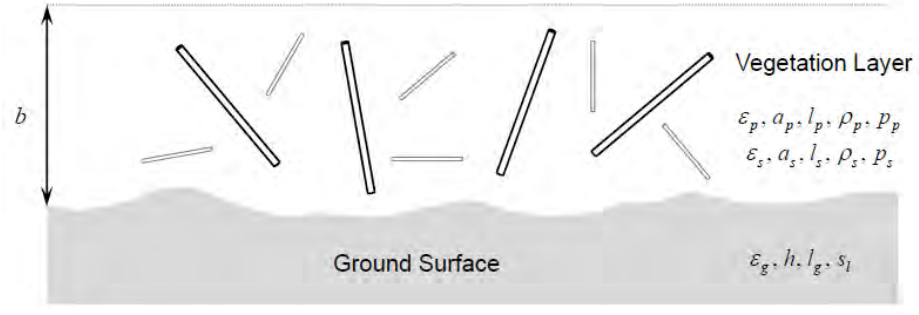


Fig. 5-23. The vegetation is assumed to consist of a layer of randomly oriented scatterers.

surface and are related to the rms. height and the correlation length of the ground surface, which is assumed to have an exponential correlation function.

Central to this type of modeling is the bistatic scattering matrix of a dielectric cylinder. Other types of scatterers often found in vegetation modeling include dielectric disks which are used to model larger flat leaves. Appendix B lists the expressions for scattering from a dielectric cylinder with arbitrary orientation.

There are several possible scattering paths to consider (see Fig. 5-24): 1) direct backscattering from the vegetation elements, 2) specular scattering at the ground surface, followed by bistatic scattering at the vegetation element, 3) bistatic scattering at the vegetation element, followed by specular scattering at the ground surface, and 4) backscattering from the underlying ground surface. For each path, the wave incident on a scatterer in the vegetation traveled through the vegetation between the scatterer and the radar before it reaches the scatterer, and suffered some attenuation as a result. Also, because the vegetation elements are assumed to be longer than the wavelength, the scattering centers in the different paths will in general be separated by several wavelengths. This implies that any phase relationships between the different scattering paths will be uniformly random, meaning that the resulting scattered powers will add incoherently. The exception is cases 2 and 3, which are exactly the same path, but in opposite directions. In that case, there will be a deterministic phase relationship between the two paths, which means the signals will add coherently. We shall now discuss each of these scattering terms separately.

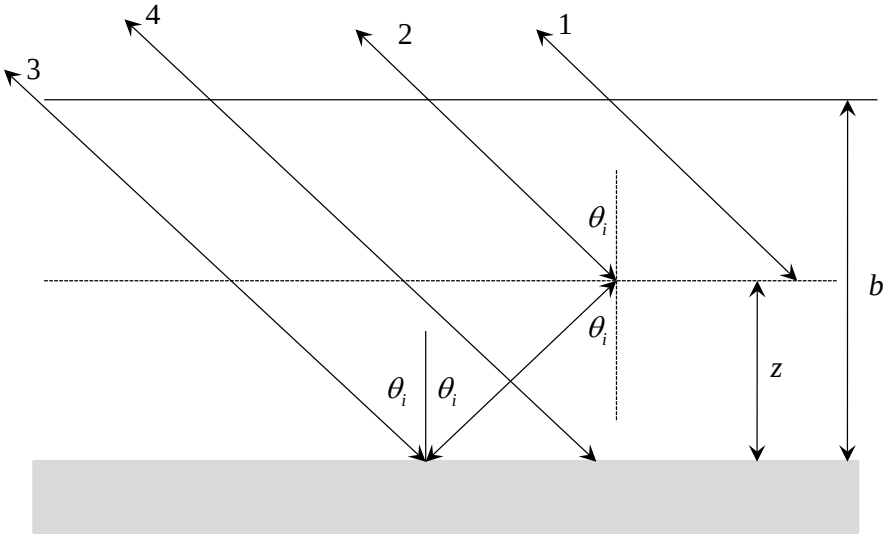


Fig. 5-24. Scattering paths to consider for a single layer vegetation model. See the text for details.

5.7.1 Scattering from the Vegetation Layer (Scattering Path 1)

The next step is to calculate the scattering coefficients from a layer randomly oriented scatterers. We shall model this by assuming that the scatterers are oriented according to a statistical distribution function that is given by $p(\theta_c,\phi_c)$. By definition,

$$\int_0^{2\pi}\int_0^{\pi} p(\theta_c,\phi_c)\sin\theta_c\, d\theta_c\, d\phi_c = 1 \tag{5.7-1}$$

Our interest is in calculating the backscattering radar cross-section of this layer, as well as the transmission coefficient for this layer. The latter is important in calculating the reflection from the underlying ground surface after the signals have propagated through the vegetation layer.

First, let us consider a scatterer that is at a height *z* above the ground surface inside the vegetation layer as shown in Figure 5-24. The wave incident on this scatterer has already propagated a distance

$$d(z) = \frac{b-z}{\cos\theta_i} \tag{5.7-2}$$

through part of the layer before reaching the scatterer. The strength of the electric field incident on this scatterer is therefore

$$\begin{pmatrix} E_h \\ E_v \end{pmatrix}^{isc} = \begin{pmatrix} e^{-\tau_h(b-z)} & 0 \\ 0 & e^{-\tau_v(b-z)} \end{pmatrix} \begin{pmatrix} E_h \\ E_v \end{pmatrix}^{inc} = \left[\boldsymbol{\alpha}(b-z)\right] \begin{pmatrix} E_h \\ E_v \end{pmatrix}^{inc} \tag{5.7-3}$$

Here the superscript *isc* refers to the wave incident on the scatterer, while *inc* refers to the wave incident on the vegetation layer. Also,

$$\tau_p(b-z) = \frac{\kappa_p^e(b-z)}{\cos\theta_i} \tag{5.7-4}$$

The extinction coefficients are average extinction coefficients as defined in Appendix B and takes into account the average extinction between the scatterer and the top of the layer. This wave is now scattered by the scatterer. In the backscatter direction, the scattering matrix of the scatterer is (see Appendix B)

$$\mathbf{S}(\theta_i,\phi_i,\theta_s=\theta_i,\phi_s=\phi_i,\theta_c,\phi_c) = \begin{pmatrix} \mathbf{h}_s\cdot\mathbf{h}_s' & \mathbf{h}_s\cdot\mathbf{v}_s' \\ \mathbf{v}_s\cdot\mathbf{h}_s' & \mathbf{v}_s\cdot\mathbf{v}_s' \end{pmatrix} \mathbf{S}(\theta_{ic},\phi_{ic},\theta_{sc},\phi_{sc}) \begin{pmatrix} \mathbf{h}_i\cdot\mathbf{h}_i' & \mathbf{v}_i\cdot\mathbf{h}_i' \\ \mathbf{h}_i\cdot\mathbf{v}_i' & \mathbf{v}_i\cdot\mathbf{v}_i' \end{pmatrix} \tag{5.7-5}$$

After being scattered, the wave again propagates through the vegetation layer to the top before returning to the radar while being attenuated as shown in Eq. (5.7-3). Adding all these terms together, we find the scattering matrix of the scatterer, as seen from outside the layer, as

$$\mathbf{S}_{veg}\left(\theta_i,\phi_i,\theta_s,\phi_s,\theta_c,\phi_c,z\right)=\left[\boldsymbol{\alpha}\left(b-z\right)\right]\mathbf{S}\left(\theta_i,\phi_i,\theta_s,\phi_s,\theta_c,\phi_c\right)\left[\boldsymbol{\alpha}\left(b-z\right)\right] \tag{5.7-6}$$

The radar cross-section is derived from the elements of the covariance matrix, which for the backscatter case, are defined by

$$\boldsymbol{\Sigma}_{veg}\left(\theta_i,\phi_i,\theta_s=\theta_i,\phi_s=\phi_i,\theta_c,\phi_c,z\right)=\begin{pmatrix} S_{hh}S_{hh}^{*} & \sqrt{2}S_{hh}S_{hv}^{*} & S_{hh}S_{vv}^{*} \\ \sqrt{2}S_{hv}S_{hh}^{*} & 2S_{hv}S_{hv}^{*} & \sqrt{2}S_{hv}S_{vv}^{*} \\ S_{vv}S_{hh}^{*} & \sqrt{2}S_{vv}S_{hv}^{*} & S_{vv}S_{vv}^{*} \end{pmatrix}_{veg} \tag{5.7-7}$$

We note that if we expand the expression on the right in Eq. (5.7-6), we find that we can write the covariance matrix as

$$\boldsymbol{\Sigma}_{veg}\left(\theta_i,\phi_i,\theta_s=\theta_i,\phi_s=\phi_i,\theta_c,\phi_c,z\right)=\mathbf{K}\left(b-z\right)\boldsymbol{\Sigma}_{veg}\left(\theta_i,\phi_i,\theta_s,\phi_s,\theta_c,\phi_c\right)\mathbf{K}\left(b-z\right) \tag{5.7-8}$$

where

$$\mathbf{K}\left(z\right)=\begin{pmatrix} e^{-2\tau_h(b-z)} & 0 & 0 \\ 0 & e^{-\left(\tau_h(b-z)+\tau_v(b-z)\right)} & 0 \\ 0 & 0 & e^{-2\tau_v(b-z)} \end{pmatrix}=\begin{pmatrix} K_{hh} & 0 & 0 \\ 0 & K_{hv} & 0 \\ 0 & 0 & K_{vv} \end{pmatrix} \tag{5.7-9}$$

The average backscatter covariance matrix is found by calculating the ensemble average for the scatterer over all possible orientations. Note that only the inner matrices on the right-hand side of Eq. (5.7-6) and Eq. (5.7-9) are functions of the orientation of the scatterer. The average covariance matrix is then

$$\left\langle\boldsymbol{\Sigma}_{veg}\left(\theta_i,\phi_i,z\right)\right\rangle=\int_0^{2\pi}\int_0^{\pi}\boldsymbol{\Sigma}_{veg}\left(\theta_i,\phi_i,\theta_s=\theta_i,\phi_s=\phi_i,\theta_c,\phi_c,z\right)p\left(\theta_c,\phi_c\right)\sin\theta_c\,d\theta_c\,d\phi_c \tag{5.7-10}$$

Similarly, the average optical depth is found from

$$\tau_p\left(b-z\right)=\int_z^b\frac{\rho_s\left(\zeta\right)}{\cos\theta_i}\int_0^{2\pi}\int_0^{\pi}\frac{2\pi}{k_0}\mathrm{Im}\left[S_{pp}\left(\theta_i,\phi_i,\pi-\theta_i,\phi_i+\pi,\theta_c,\phi_c\right)\right]p\left(\theta_c,\phi_c\right)\sin\theta_c\,d\theta_c\,d\phi_c d\zeta \tag{5.7-11}$$

where $\rho_s(z)$ is the density of the scatterers in scatterers$/m^3$ at a height z above the ground. The final step is now to integrate the contribution from all the scatterers in the layer. This is gives

$$\left\langle \boldsymbol{\Sigma}_{veg}(\theta_i,\phi_i)\right\rangle = \int_0^b \left\langle \boldsymbol{\Sigma}_{veg}(\theta_i,\phi_i,z)\right\rangle \rho_s(z)\,dz \tag{5.7-12}$$

From Eq. (5.7-8) we can show that the average covariance matrix has the form

$$\left\langle \boldsymbol{\Sigma}_{veg}(\theta_i,\phi_i,z)\right\rangle = \begin{pmatrix} \left\langle S_{hh}S_{hh}^*\right\rangle K_{hh}K_{hh} & \left\langle \sqrt{2}S_{hh}S_{hv}^*\right\rangle K_{hh}K_{hv} & \left\langle S_{hh}S_{vv}^*\right\rangle K_{hh}K_{vv} \\ \left\langle \sqrt{2}S_{hv}S_{hh}^*\right\rangle K_{hh}K_{hv} & \left\langle 2S_{hv}S_{hv}^*\right\rangle K_{hv}K_{hv} & \left\langle \sqrt{2}S_{hv}S_{vv}^*\right\rangle K_{hv}K_{vv} \\ \left\langle S_{vv}S_{hh}^*\right\rangle K_{hh}K_{vv} & \left\langle \sqrt{2}S_{vv}S_{hv}^*\right\rangle K_{hv}K_{vv} & \left\langle S_{vv}S_{vv}^*\right\rangle K_{vv}K_{vv} \end{pmatrix} \tag{5.7-13}$$

If we assume the density of the scatterers to be uniform in elevation above the ground surface, then only the K terms are functions of z, and we find

$$\begin{aligned}
&\int_0^b \rho_s K_{hh}K_{hh}\,dz = \int_0^b \rho_s e^{-4\tau_h(b-z)}dz = \frac{b\rho_s\cos\theta_i}{4\tau_{hm}}\left\{1-e^{-4\tau_{hm}/\cos\theta_i}\right\} \\
&\int_0^b \rho_s K_{vv}K_{vv}\,dz = \int_0^b \rho_s e^{-4\tau_v(b-z)}dz = \frac{b\rho_s\cos\theta_i}{4\tau_{vm}}\left\{1-e^{-4\tau_{vm}/\cos\theta_i}\right\} \\
&\int_0^b \rho_s K_{hv}K_{hv}\,dz = \int_0^b \rho_s e^{-2\left[\tau_h(b-z)+\tau_v(b-z)\right]}dz = \frac{b\rho_s\cos\theta_i}{(2\tau_{hm}+2\tau_{vm})}\left\{1-e^{-(2\tau_{hm}+2\tau_{vm})/\cos\theta_i}\right\} \\
&\int_0^b \rho_s K_{hh}K_{hv}\,dz = \int_0^b \rho_s e^{-\left[3\tau_h(b-z)+\tau_v(b-z)\right]}dz = \frac{b\rho_s\cos\theta_i}{(3\tau_{hm}+\tau_{vm})}\left\{1-e^{-(3\tau_{hm}+\tau_{vm})/\cos\theta_i}\right\} \\
&\int_0^b \rho_s K_{hv}K_{vv}\,dz = \int_0^b \rho_s e^{-\left[\tau_h(b-z)+3\tau_v(b-z)\right]}dz = \frac{b\rho_s\cos\theta_i}{(\tau_{hm}+3\tau_{vm})}\left\{1-e^{-(\tau_{hm}+3\tau_{vm})/\cos\theta_i}\right\}
\end{aligned} \tag{5.7-14}$$

Here we used the short-hand notations

$$\tau_{hm} = \kappa_h^e b;\ \tau_{vm} = \kappa_v^e b \tag{5.7-15}$$

which represent the total vertical optical path depths of the layer for the two polarizations. The radar cross-section is defined as

$$\boldsymbol{\sigma}_{veg} = 4\pi\left\langle \boldsymbol{\Sigma}_{veg}(\theta_i,\phi_i)\right\rangle \tag{5.7-16}$$

At this point it is instructive to look at the behavior of the vegetation layer in two extreme cases. First, when the optical depth is large, the second terms in Eq. (5.7-14) vanish, and the total scattering (apart from the strength of the covariance matrix terms) approach a constant value. For example, in the case of HH scattering, this constant value can be written as

$$\sigma_{hh} = 4\pi \frac{b\rho_s \cos\theta_i}{4\tau_{hm}} \left\langle S_{hh} S_{hh}^* \right\rangle \tag{5.7-17}$$

In the other extreme, the layer optical depth is small enough that we can replace the exponential by a Taylor series expansion. In that case, the radar cross-section becomes

$$\sigma_{hh} = 4\pi \frac{b\rho_s \cos\theta_i}{4\tau_{hm}} \times \frac{4\tau_{hm}}{\cos\theta_i} \left\langle S_{hh} S_{hh}^* \right\rangle = 4\pi b \rho_s \left\langle S_{hh} S_{hh}^* \right\rangle \tag{5.7-18}$$

To first order, the backscatter from the layer is directly proportional to the density of the scatterers and the thickness of the layer.

All these expressions involving scattering from cylinders are functions of the dielectric constant of the cylinder. We used the expressions reported by Ulaby and El-Rayes (1987) [35] for the vegetation dielectric constant, which is given by

$$\varepsilon_v = \varepsilon_r + v_{fw}\varepsilon_f + v_b \varepsilon_b \tag{5.7-19}$$

In this expression, ε_r is the non-dispersive residual part of the dielectric constant. They report that in terms of the soil volumetric water content M_v, this quantity is given by

$$\varepsilon_r = 1.7 + 3.2 M_v + 6.5 M_v^2 \tag{5.7-20}$$

The dielectric constant of free water is

$$\varepsilon_f = 4.9 + \frac{75}{1 + j\, f/18} - j\frac{18\sigma_{sal}}{f} \tag{5.7-21}$$

Where f is the frequency in Gigahertz, and σ_{sal} is related to the salinity of the water S (measured in parts per thousand on a weight basis) by

$$\sigma_{sal} = 0.16S - 0.0013S^2 \tag{5.7-22}$$

The dielectric constant of bound water is

$$\varepsilon_b = 2.9 + \frac{55}{1+\sqrt{j\,f/0.18}} \tag{5.7-23}$$

The volume fraction of free water is

$$v_{fw} = M_v\left(0.82M_v + 0.166\right) \tag{5.7-24}$$

And the volume fraction of vegetation-bound water is

$$v_b = \frac{31.4M_v^2}{1+59.5M_v^2} \tag{5.7-25}$$

We shall express our simulations in terms of the vegetation volumetric moisture M_v and use the expressions above to calculate the dielectric constant.

5.7.2 Backscatter from the Underlying Ground Surface (Scattering Path 4)

The electromagnetic wave incident upon the ground surface travels through the entire vegetation layer before reaching the ground surface. As it does so, it is attenuated according to

$$\begin{pmatrix} E_h \\ E_v \end{pmatrix}^{igr} = \begin{pmatrix} e^{-\tau_{hm}/\cos\theta_i} & 0 \\ 0 & e^{-\tau_{vm}/\cos\theta_i} \end{pmatrix} \begin{pmatrix} E_h \\ E_v \end{pmatrix}^{inc} \tag{5.7-26}$$

This wave is incident upon the ground element. To calculate the scattering from this element, we shall assume that the scattering from the ground surface can be described by the small perturbation model, in which the scattering matrix for an untilted ground surface is given by

$$\begin{pmatrix} E_h \\ E_v \end{pmatrix}^{sc} = \begin{pmatrix} S_{hh'g} & 0 \\ 0 & S_{vv'g} \end{pmatrix} \begin{pmatrix} E_{h'} \\ E_{v'} \end{pmatrix}^{igr} \tag{5.7-27}$$

where

$$S_{hh'g} = k^2 h\cos^2\theta_i \alpha_{hh}\left(\varepsilon_g, \theta_i\right)\sqrt{W\left(2k\sin\theta_i\right)}\,, \tag{5.7-28}$$

$$S_{vv'g} = k^2 h\cos^2\theta_i \alpha_{vv}\left(\varepsilon_g, \theta_i\right)\sqrt{W\left(2k\sin\theta_i\right)}\,, \tag{5.7-29}$$

and

$$\alpha_{hh}\left(\varepsilon_g,\theta_i\right)=\frac{(\varepsilon_g-1)}{(\cos\theta_i+\sqrt{\varepsilon_g-\sin^2\theta_i})^2}, \tag{5.7-30}$$

$$\alpha_{vv}\left(\varepsilon_g,\theta_i\right)=\frac{(\varepsilon_g-1)\left[\left(\varepsilon_g-1\right)\sin^2\theta_i+\varepsilon_g\right]}{(\varepsilon_g\cos\theta_i+\sqrt{\varepsilon_g-\sin^2\theta_i})^2}, \tag{5.7-31}$$

$$W\left(2k\sin\theta_i\right)=\frac{2l_g^2}{\pi\left[1+4^2k^2l_g^2\sin^2\theta_i\right]^{3/2}}. \tag{5.7-32}$$

The roughness spectrum W given in Eq. (5.7-32) is that of a surface with an exponential correlation function with a correlation length l_g.

After scattering from the ground surface the wave travels again through the vegetation layer before returning to the radar, suffering attenuation given by Eq. (5.7-26) in the process. The covariance matrix of the ground layer is then

$$\left[\boldsymbol{\sigma}_g\right]=4\pi\begin{pmatrix} S_{hh'g}S_{hh'g}^{*}\,e^{-4\tau_{hm}/\cos\theta_i} & 0 & S_{hh'g}S_{vv'g}^{*}\,e^{-2\left(\tau_{hm}+\tau_{vm}\right)/\cos\theta_i} \\ 0 & 0 & 0 \\ S_{vv'g}S_{hh'g}^{*}\,e^{-2\left(\tau_{hm}+\tau_{vm}\right)/\cos\theta_i} & 0 & S_{vv'g}S_{vv'g}^{*}\,e^{-4\tau_{vm}/\cos\theta_i} \end{pmatrix} \tag{5.7-33}$$

This expression is true if we use the first order small perturbation model to represent the ground scattering. If higher order approximations are included, additional terms appear in this matrix, particularly in the cross-polarized element.

5.7.3 Double Reflection Scattering (Scattering Paths 2 and 3)

The next term to consider is the double reflection from a cylinder to the ground surface and back to the radar. This term involves calculating the product of the bistatic scattering matrix of the vegetation layer and the specular scattering matrix of the ground layer. Here we shall only consider the case of a flat ground surface. For a tilted surface, additional terms need to be considered; please see van Zyl (1993) [36] for details.

Consider a scatterer at a height $z<b$ inside the vegetation layer as shown in Figure 5-24. The incident wave is scattered as shown towards the ground surface where it is reflected in the specular direction after which it propagates through the entire vegetation layer back to the radar. The opposite path also

needs to be considered. In this case, the incident wave first travels through the entire vegetation layer to reach the ground surface where it is reflected in the specular direction. The reflected wave then propagates through part of the layer before it is scattered by the scatterer back in the direction of the radar. As the scattered wave propagates back to the radar, it again travels through part of the vegetation layer before reaching the radar.

Let us first consider the path that interacts with the ground surface first. It is well known that the specular ray for a facet characterized by a surface normal $\mathbf{n}$ and light incident along a unit vector $\mathbf{k}_i$ is given by

$$\mathbf{r} = \mathbf{k}_i - 2\left(\mathbf{k}_i \cdot \mathbf{n}\right)\mathbf{n} \tag{5.7-34}$$

Note that $\mathbf{r}$ as given in Eq. (5.7-34) must be expressed in the forward scattering alignment coordinate system. The scattered wave propagation vector in the backscatter alignment coordinate system that we are using is the negative of $\mathbf{r}$. At the ground surface, the incident and scattered propagation vectors are therefore

$$\mathbf{k}_{igs1} = \mathbf{k}_i; \quad \mathbf{k}_{sgs1} = -\mathbf{r} = 2\left(\mathbf{n} \cdot \mathbf{k}_i\right)\mathbf{n} - \mathbf{k}_i \tag{5.7-35}$$

The wave that is reflected in the specular direction from the ground is now incident upon the cylinder representing the vegetation, from which the scattered wave has to propagate back to the radar. This bistatic scattering at the cylinder is therefore characterized by incident and scattered wave propagation vectors

$$\mathbf{k}_{icb1} = \mathbf{r} = \mathbf{k}_i - 2\left(\mathbf{n} \cdot \mathbf{k}_i\right)\mathbf{n}; \quad \mathbf{k}_{scb1} = \mathbf{k}_i \tag{5.7-36}$$

where the subscript c is added to indicate that these refer to the cylinder. Note that Fig. 5-24 shows specular reflection at the layer inside the vegetation. This shows the propagation paths, and does not mean there is specular reflection at the cylinder itself. Since the cylinder may have any orientation, in general we have bistatic scattering at the cylinder.

Finally, for the case of bistatic scattering at the cylinder, followed by specular reflection at the ground, these are

$$\begin{aligned} &\mathbf{k}_{icb2} = \mathbf{k}_i; \quad \mathbf{k}_{scb2} = \mathbf{r} = \mathbf{k}_i - 2\left(\mathbf{n} \cdot \mathbf{k}_i\right)\mathbf{n} \\ &\mathbf{k}_{igs2} = -\mathbf{r} = 2\left(\mathbf{n} \cdot \mathbf{k}_i\right)\mathbf{n} - \mathbf{k}_i; \quad \mathbf{k}_{sgs2} = \mathbf{k}_i \end{aligned} \tag{5.7-37}$$

Figure 5-25 shows the scattering paths for specular scattering at the ground followed by bistatic scattering at the cylinder and vice versa.

For the ground surface, the bistatic scattering matrix describing the specular reflection can be written as

$$\mathbf{R}_g\left(\theta_i,\phi_i,\theta_s,\phi_s\right)=\begin{pmatrix} R_h\left(\varepsilon_g,\theta_i\right) & 0 \\ 0 & R_v\left(\varepsilon_g,\theta_i\right)\end{pmatrix} e^{-2k^2h^2/\cos^2\theta_i} \tag{5.7-38}$$

The two Fresnel reflections coefficients are

$$R_h\left(\varepsilon_g,\theta_i\right)=\frac{\cos\theta_i-\sqrt{\varepsilon_g-\sin^2\theta_i}}{\cos\theta_i+\sqrt{\varepsilon_g-\sin^2\theta_i}} \tag{5.7-39}$$

and

$$R_v\left(\varepsilon_g,\theta_i\right)=-\frac{\varepsilon_g\cos\theta_i-\sqrt{\varepsilon_g-\sin^2\theta_i}}{\varepsilon_g\cos\theta_i+\sqrt{\varepsilon_g-\sin^2\theta_i}} \tag{5.7-40}$$

The scattering matrix for the case that involves scattering at the ground first, can be written as

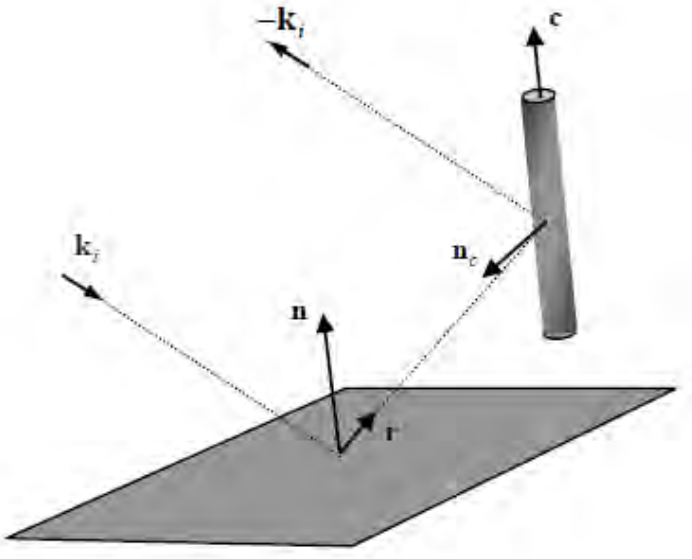


Fig. 5-25. Scattering paths for the double reflection cases.

$$\mathbf{S}_{gc2} = \begin{bmatrix} \quad (b-z) \end{bmatrix} \begin{pmatrix} S_{hhc}^{b} & S_{hvc}^{b} \\ S_{vhc}^{b} & S_{vvc}^{b} \end{pmatrix} \begin{bmatrix} \quad (z) \end{bmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{R} \begin{bmatrix} \boldsymbol{\alpha}_{\quad g} \end{bmatrix} \begin{bmatrix} \quad (b) \end{bmatrix} \tag{5.7-41}$$

The number 2 in the subscript refers to the scattering path in Fig. 5-24. The matrix in the middle of the right-hand side of these expressions is needed to take into account that each of the other matrices is described in the backscatter alignment system. The term involving scattering at the cylinder first is

$$\mathbf{S}_{g3} = \begin{bmatrix} \mathbf{R}(b) \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{\quad g} \end{bmatrix} \begin{bmatrix} \quad (z) \end{bmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S_{hhc}^{b'} & S_{hvc}^{b'} \\ S_{vhc}^{b'} & S_{vvc}^{b'} \end{pmatrix} \begin{bmatrix} \quad (b-z) \end{bmatrix} \tag{5.7-42}$$

Each of the matrices describing the scattering at the cylinder in Eq. (5.7-41) and Eq. (5.7-42) is of the form (see Appendix B)

$$\begin{pmatrix} S_{hhc}^{y} & S_{hvc}^{y} \\ S_{vhc}^{y} & S_{vvc}^{y} \end{pmatrix} = \begin{pmatrix} \mathbf{h}_s \cdot \mathbf{h}_s' & -\mathbf{v}_s \cdot \mathbf{h}_s' \\ \mathbf{v}_s \cdot \mathbf{h}_s' & \mathbf{h}_s \cdot \mathbf{h}_s' \end{pmatrix} \begin{pmatrix} S_{hhc} & S_{hvc} \\ S_{hvc} & S_{vvc} \end{pmatrix} \begin{pmatrix} \mathbf{h}_i \cdot \mathbf{h}_i' & \mathbf{v}_i \cdot \mathbf{h}_i' \\ -\mathbf{v}_i \cdot \mathbf{h}_i' & \mathbf{h}_i \cdot \mathbf{h}_i' \end{pmatrix} \tag{5.7-43}$$

where $y = b$ or b'.

For all these cases, we now need to find the elements of the transformation matrices shown in Eq. (5.7-43). We do this by defining local and global coordinate systems as described before using the propagation vectors derived previously for the interaction mechanisms. We shall describe each case separately below.

For the case where the wave first interacts with the ground surface, Eq. (5.7-34) describes the direction in which the specularly reflected energy will travel after interacting with the surface. The elements of the transformation matrices for this case are therefore (see Appendix B for the methodology of how to derive these expressions)

$$\mathbf{h}_{igs2} \cdot \mathbf{h}_{igs2}' = \frac{\mathbf{c} \cdot \mathbf{z} + (\mathbf{k}_i \cdot \mathbf{z})\left[\mathbf{k}_i \cdot \mathbf{c} - 2(\mathbf{k}_i \cdot \mathbf{z})(\mathbf{c} \cdot \mathbf{z})\right]}{\sqrt{1-(\mathbf{k}_i \cdot \mathbf{z})^2}\sqrt{1-(\mathbf{k}_i \cdot \mathbf{c} - 2(\mathbf{k}_i \cdot \mathbf{z})(\mathbf{c} \cdot \mathbf{z}))^2}} \quad , \tag{5.7-44}$$

$$\mathbf{v}_{igs2} \cdot \mathbf{h}_{igs2}' = \frac{\mathbf{k}_i \cdot (\mathbf{c} \times \mathbf{z})}{\sqrt{1-(\mathbf{k}_i \cdot \mathbf{z})^2}\sqrt{1-(\mathbf{k}_i \cdot \mathbf{c} - 2(\mathbf{k}_i \cdot \mathbf{z})(\mathbf{c} \cdot \mathbf{z}))^2}} \quad , \tag{5.7-45}$$

$$\mathbf{h}_{sgs2}\cdot\mathbf{h}'_{sgs2}=\frac{\mathbf{c}\cdot\mathbf{z}-(\mathbf{k}_i\cdot\mathbf{z})(\mathbf{k}_i\cdot\mathbf{c})}{\sqrt{1-(\mathbf{k}_i\cdot\mathbf{z})^2}\sqrt{1-(\mathbf{k}_i\cdot\mathbf{c})^2}}\quad , \tag{5.7-46}$$

and

$$\mathbf{v}_{sgs2}\cdot\mathbf{h}'_{sgs2}=\frac{\mathbf{k}_i\cdot(\mathbf{c}\times\mathbf{z})}{\sqrt{1-(\mathbf{k}_i\cdot\mathbf{z})^2}\sqrt{1-(\mathbf{k}_i\cdot\mathbf{c})^2}}\ . \tag{5.7-47}$$

where **c** is defined in Eq. (B.1).

For the case where the wave first scattered off the cylinder before reflecting in the specular direction from the ground, the propagation vectors are given by Eq. (5.7-37). The elements of the transformation matrices in this case are

$$\mathbf{h}_{igs3}\cdot\mathbf{h}'_{igs3}=\mathbf{h}_{sgs2}\cdot\mathbf{h}'_{sgs2}\,, \tag{5.7-48}$$

$$\mathbf{v}_{igs3}\cdot\mathbf{h}'_{igs3}=\mathbf{v}_{sgs2}\cdot\mathbf{h}'_{sgs2}\,, \tag{5.7-49}$$

$$\mathbf{h}_{sgs3}\cdot\mathbf{h}'_{sgs3}=\mathbf{h}_{igs2}\cdot\mathbf{h}'_{igs2}\,, \tag{5.7-50}$$

and

$$\mathbf{v}_{sgs3}\cdot\mathbf{h}'_{sgs3}=\mathbf{v}_{igs2}\cdot\mathbf{h}'_{igs2}\,. \tag{5.7-51}$$

The two signals propagating along the inverse paths add coherently. The total scattering matrix for this case is therefore

$$\mathbf{S}=\mathbf{S}_{gc2}+\mathbf{S}_{cg3}\,. \tag{5.7-52}$$

Performing the matrix multiplications shown in Eqs. 5.7-41 and 5.7-42, we find the elements of the scattering matrix to be

$$\begin{aligned}
S_{hh}&=-R'_h e^{-2\tau_h(b)}\left(S^b_{hhc}+S^{b'}_{hhc}\right)\\
S_{hv}&=e^{-(\tau_h(b)+\tau_v(b))}\left\{R'_v e^{-(\tau_v(z)-\tau_h(z))}S^b_{hvc}-R'_h e^{+(\tau_v(z)-\tau_h(z))}S^{b'}_{hvc}\right\}\\
S_{vh}&=e^{-(\tau_h(b)+\tau_v(b))}\left\{R'_v e^{-(\tau_v(z)-\tau_h(z))}S^{b'}_{hvc}-R'_h e^{+(\tau_v(z)-\tau_h(z))}S^{b}_{hvc}\right\}\\
S_{vv}&=R'_v e^{-2\tau_v(b)}\left(S^b_{vvc}+S^{b'}_{vvc}\right)
\end{aligned}\ . \tag{5.7-53}$$

Here we used the short-hand notation

$$\begin{aligned} R_h' &= R_h\left(\varepsilon_g, \theta_i\right) e^{-2k^2h^2/\cos^2\theta_i} \\ R_v' &= R_v\left(\varepsilon_g, \theta_i\right) e^{-2k^2h^2/\cos^2\theta_i} \end{aligned} . \tag{5.7-54}$$

Reciprocity dictates that the primed and unprimed scattering matrices are the same. Therefore,

$$\begin{aligned} S_{hh} &= -2R_h' e^{-2\tau_h(b)} S_{hhc}^b \\ S_{hv} &= e^{-\left(\tau_h(b)+\tau_v(b)\right)} \left\{ R_v' e^{-\left(\tau_v(z)-\tau_h(z)\right)} - R_h' e^{+\left(\tau_v(z)-\tau_h(z)\right)} \right\} S_{hvc}^b \\ S_{vh} &= S_{hv} \\ S_{vv} &= 2R_v' e^{-2\tau_v(b)} S_{vvc}^b \end{aligned} . \tag{5.7-55}$$

Note that for both S_{hh} and S_{vv} the terms describing the attenuation is simply twice the attenuation through the entire vegetation layer, and does not depend on where inside the layer the scatterer resides. For the cross-polarized term, the attenuation terms does depend on where inside the layer the scatterer is. This is a consequence of the fact that for the cross-polarized terms, the incident and scattered waves have orthogonal polarizations. For part of the path, therefore, the one polarization suffers attenuation, while for the rest of the path, the other polarization is attenuated. The terms that are dependent on z therefore is a function of the difference in attenuation between the two polarizations as shown in Eq. (5.7-55). We now have to form the covariance matrix corresponding to this scattering matrix. This is

$$\left[\boldsymbol{\Sigma}_{db}\left(\theta_i, \phi_i, \theta_s = \theta_i, \phi_s = \phi_i, \theta_c, \phi_c, z\right)\right] = \begin{pmatrix} S_{hh}S_{hh}^* & \sqrt{2}S_{hh}S_{hv}^* & S_{hh}S_{vv}^* \\ \sqrt{2}S_{hv}S_{hh}^* & 2S_{hv}S_{hv}^* & \sqrt{2}S_{hv}S_{vv}^* \\ S_{vv}S_{hh}^* & \sqrt{2}S_{vv}S_{hv}^* & S_{vv}S_{vv}^* \end{pmatrix} \tag{5.7-56}$$

The subscript *db* refers to the double reflection suffered by the signals. Next, we average this matrix over all scatterer orientations, which gives

$$\left\langle\left[\boldsymbol{\Sigma}_{db}\left(\theta_i, \phi_i, z\right)\right]\right\rangle = \int_0^{2\pi}\int_0^{\pi} \left[\boldsymbol{\Sigma}_{db}\left(\theta_i, \phi_i, \theta_s = \theta_i, \phi_s = \phi_i, \theta_c, \phi_c, z\right)\right] p\left(\theta_c, \phi_c\right) \sin\theta_c \, d\theta_c \, d\phi_c \tag{5.7-57}$$

Finally, we have to add the contributions from all the layers within the vegetation volume:

$$\left\langle \left[\boldsymbol{\Sigma}_{db}\left(\theta_i,\phi_i\right)\right]\right\rangle = \int_0^b \left\langle \left[\boldsymbol{\Sigma}_{db}\left(\theta_i,\phi_i,z\right)\right]\right\rangle \rho_s\left(z\right)dz \tag{5.7-58}$$

The radar cross section for this term is then

$$\left[\boldsymbol{\sigma}_{db}\right] = 4\pi \left\langle \left[\boldsymbol{\Sigma}_{db}\left(\theta_i,\phi_i\right)\right]\right\rangle \tag{5.7-59}$$

From Eq. (5.7-55) – (5.7-58), we find that if the density of scatterers is constant throughout the vegetation layer, the results are:

$$\begin{aligned}
\left\langle S_{hh}S_{hh}^*\right\rangle_{db} &= 4\rho b\left\langle S_{hhc}S_{hhc}^*\right\rangle \left|R_h'\right|^2 e^{-4\tau_{hm}/\cos\theta_i} \\
\left\langle S_{vv}S_{vv}^*\right\rangle_{db} &= 4\rho b\left\langle S_{vvc}S_{vvc}^*\right\rangle \left|R_v'\right|^2 e^{-4\tau_{vm}/\cos\theta_i} \\
\left\langle S_{hh}S_{vv}^*\right\rangle_{db} &= -4\rho b\left\langle S_{hhc}S_{vvc}^*\right\rangle R_h'R_v'^* e^{-2\left(\tau_{hm}+\tau_{vm}\right)/\cos\theta_i} \\
\sqrt{2}\left\langle S_{hh}S_{hv}^*\right\rangle_{db} &= 2\sqrt{2}\rho b\left\langle S_{hhc}S_{hvc}^*\right\rangle e^{-\left(3\tau_{hm}+\tau_{vm}\right)/\cos\theta_i} \\
&\quad \frac{\cos\theta_i}{\tau_{vm}-\tau_{hm}}\left\{-R_h'R_v'^*\left[1-e^{-\left(\tau_{vm}-\tau_{hm}\right)/\cos\theta_i}\right]-R_h'R_h'^*\left[1-e^{+\left(\tau_{vm}-\tau_{hm}\right)/\cos\theta_i}\right]\right\} \\
2\left\langle S_{hv}S_{hv}^*\right\rangle_{db} &= 2\rho b\left\langle S_{hvc}S_{hvc}^*\right\rangle e^{-2\left(\tau_{hm}+\tau_{vm}\right)/\cos\theta_i}\left(-R_h'R_v'^*-R_v'R_h'^*\right. \\
&\quad \left.+\frac{\cos\theta_i}{2\left(\tau_{vm}-\tau_{hm}\right)}\left\{R_v'R_v'^*\left[1-e^{-2\left(\tau_{vm}-\tau_{hm}\right)/\cos\theta_i}\right]-R_h'R_h'^*\left[1-e^{+2\left(\tau_{vm}-\tau_{hm}\right)/\cos\theta_i}\right]\right\}\right) \\
\sqrt{2}\left\langle S_{hv}S_{vv}^*\right\rangle_{db} &= 2\sqrt{2}\rho b\left\langle S_{hvc}S_{vvc}^*\right\rangle e^{-\left(\tau_{hm}+3\tau_{vm}\right)/\cos\theta_i} \\
&\quad \frac{\cos\theta_i}{\tau_{vm}-\tau_{hm}}\left\{R_v'R_v'^*\left[1-e^{-\left(\tau_{vm}-\tau_{hm}\right)/\cos\theta_i}\right]+R_h'R_v'^*\left[1-e^{+\left(\tau_{vm}-\tau_{hm}\right)/\cos\theta_i}\right]\right\}
\end{aligned} \tag{5.7-60}$$

Note that when the attenuation coefficients for the two polarizations are small, the expressions become

$$\left\langle S_{hh}S_{hh}^{*}\right\rangle_{db} \approx 4\rho b\left\langle S_{hhc}S_{hhc}^{*}\right\rangle\left|R_h'\right|^2\left[1-\frac{4\tau_{hm}}{\cos\theta_i}\right]$$

$$\left\langle S_{vv}S_{vv}^{*}\right\rangle_{db} = 4\rho b\left\langle S_{vvc}S_{vvc}^{*}\right\rangle\left|R_v'\right|^2\left[1-\frac{4\tau_{vm}}{\cos\theta_i}\right]$$

$$\left\langle S_{hh}S_{vv}^{*}\right\rangle_{db} = -4\rho b\left\langle S_{hhc}S_{vvc}^{*}\right\rangle R_h'R_v'^{*}\left[1-\frac{2\left(\tau_{hm}+\tau_{vm}\right)}{\cos\theta_i}\right]$$

$$\sqrt{2}\left\langle S_{hh}S_{hv}^{*}\right\rangle_{db} \approx 2\sqrt{2}\rho b\left\langle S_{hhc}S_{hvc}^{*}\right\rangle\left\{R_h'R_h'^{*}-R_h'R_v'^{*}\right\}\left[1-\frac{\left(3\tau_{hm}+\tau_{vm}\right)}{\cos\theta_i}\right]$$

$$2\left\langle S_{hv}S_{hv}^{*}\right\rangle_{db} \approx 2\rho b\left\langle S_{hvc}S_{hvc}^{*}\right\rangle\left|R_v'-R_h'\right|^2\left[1-\frac{2\left(\tau_{hm}+\tau_{vm}\right)}{\cos\theta_i}\right]$$

$$\sqrt{2}\left\langle S_{hv}S_{vv}^{*}\right\rangle_{db} \approx 2\sqrt{2}\rho b\left\langle S_{hvc}S_{vvc}^{*}\right\rangle\left\{R_v'R_v'^{*}-R_v'R_h'^{*}\right\}\left[1-\frac{\left(3\tau_{hm}+\tau_{vm}\right)}{\cos\theta_i}\right] \tag{5.7-61}$$

5.8 Simulation Results

In this section we present some results of a single layer vegetation model. To keep things simple, we will assume that the vegetation is grass-like, so that we can represent the vegetation elements with thin dielectric cylinders. We shall first discuss some general results, and then concentrate on the effects of soil moisture on the overall observed radar cross-section.

Putting the results of the previous Section together, we write the total covariance matrix of the vegetated layer as:

$$\left[\boldsymbol{\sigma}_{tot}\right] = \left[\boldsymbol{\sigma}_{veg}\right] + \left[\boldsymbol{\sigma}_{db}\right] + \left[\boldsymbol{\sigma}_{g}\right] \tag{5.8-1}$$

To illustrate the model, we assume the parameters for the vegetation layer and soil surface as shown in Table 5-1. We shall now discuss various different simulations to illustrate the effect of different vegetation parameters on the total scattering.

Table 5-1. Parameters used in simulating the backscatter from a vegetation layer.

Parameter	Value
Cylinder radius	2 mm
Cylinder length	50 cm
Cylinder density	900 cylinders/cubic meter
Vegetation layer height	50 cm
Orientation pdf	Cosine squared around vertical
Surface rms height	1 cm
Surface correlation length	15.2 cm
Radar wavelength	24 cm

5.8.1 Effect of the Angle of Incidence

Figure 5-26 shows the HH scattering cross-sections for the three scattering mechanisms, plus the total scattering cross-section, which is the sum of the three individual scattering mechanisms, as a function of the angle of incidence. We notice that for angles of incidence less than 30 deg, the HH scattering is dominated by direct backscattering from the ground. For angles between 30 and about 55, double reflections dominate the HH return, while for angles larger than about 55 deg, the scattering is dominated by the vegetation layer.

Fig. 5-27 shows the same plots, this time for the HV terms. Direct scattering from the ground is absent because we chose to model the surface scattering with a first-order small perturbation model, which predicts no cross-polarized return. In the cross-polarized return, we note that double reflections dominate for angles <30 deg, while vegetation scattering dominates for larger angles.

The situation is quite different for the VV case shown in Fig. 5-28. For angles less than 35 deg, the scattering is dominated by the underlying ground surface scattering. For angles larger than 35 deg, the vegetation scattering dominates. In the case of VV, double reflection scattering is typically much less than the other two, and does not really contribute much to the overall scattering.

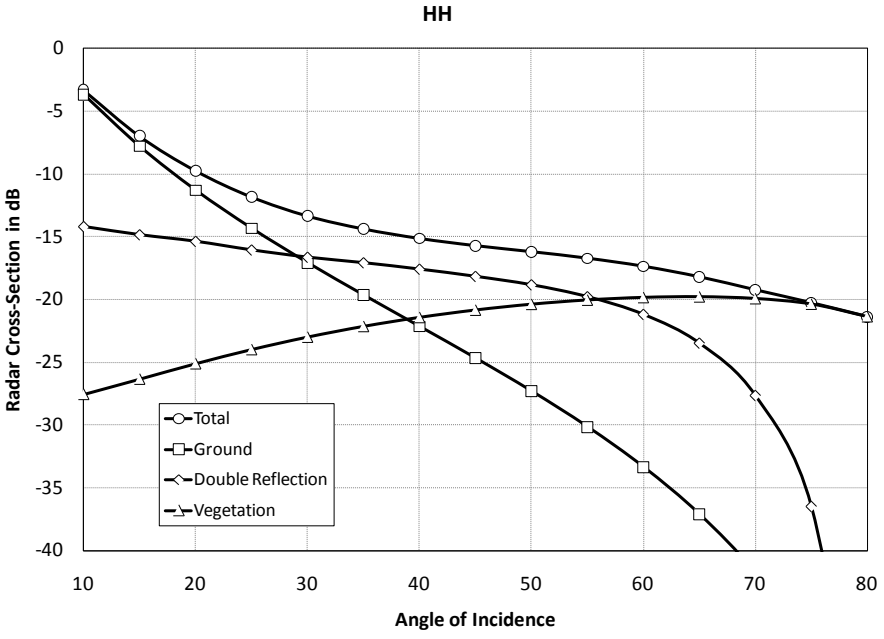


Fig. 5-26. Backscatter cross-sections for the three different scattering mechanisms as a function of angle of incidence at HH polarization.

The specifics of which mechanism dominates is obviously a function of the parameters used to characterize the vegetation and the soil. These results do point out an important point to keep in mind when analyzing scattering from vegetated surfaces, however. The dominant scattering mechanism may be different for the different polarizations. Focusing on an angle of incidence of 40 deg in Fig. 5-26 – Fig. 5-28, we note that the dominant scattering mechanism at HH is double bounce reflections, while at VV and HV the scattering from the vegetation canopy dominates.

5.8.2 Effect of Cylinder Radius

We shall evaluate the effect of the cylinder radius by fixing all other parameters at the values given in Table 5-1, and also fix the angle of incidence at 40 deg. We then vary the cylinder radius in our simulations. The results are shown for the different polarization combinations in Fig. 5-29 for the vegetation layer scattering only. From this calculation it is clear that larger cylinder sizes scatter much more efficiently and as a result the vegetation scattering contribution increases rapidly with increasing cylinder size. For larger cylinders, the increase in scattering is slower, so the rate of increase of the vegetation scattering also decreases.

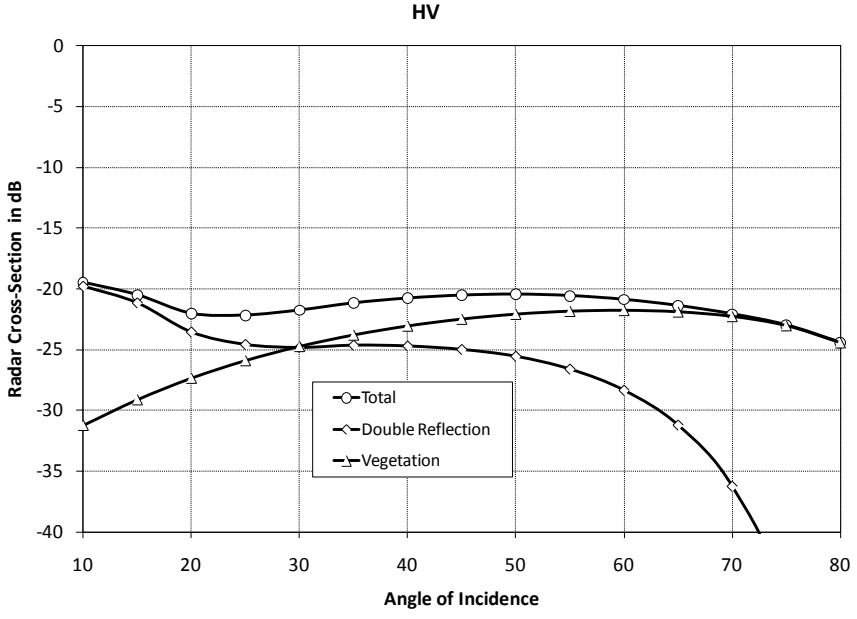


Fig. 5-27. Backscatter cross-sections for the three different scattering mechanisms as a function of angle of incidence at HV polarization.

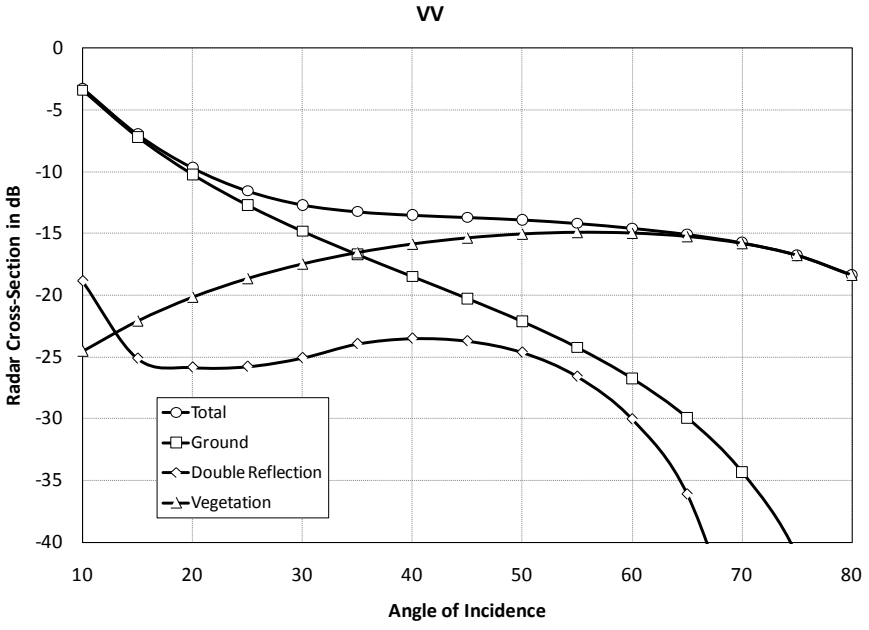


Fig. 5-28. Backscatter cross-sections for the three different scattering mechanisms as a function of angle of incidence at VV polarization.

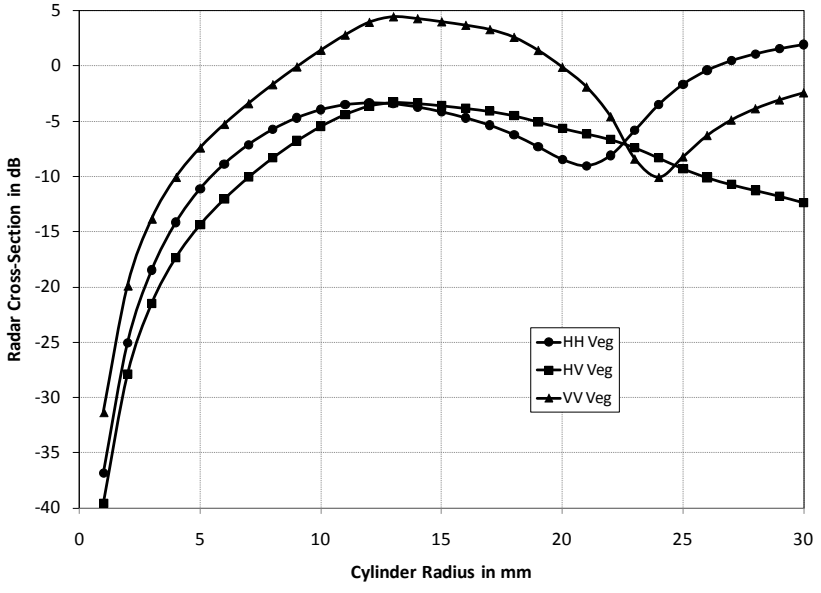


Fig. 5-29. Scattering from the vegetation layer as a function of the cylinder radius. For thin cylinders, there is a dramatic increase in scattering as the cylinder radius increases.

Figure 5-30 shows the scattering at HH polarization for all the scattering mechanisms. For very thin cylinders, the surface scattering dominates. As the cylinders become thicker, the double reflection term starts to dominate as the ground term is attenuated. As the cylinder radius increases further, the double reflection term also decreases, and the vegetation scattering becomes dominant. Not only is the vegetation scattering itself increasing, but the extinction through the vegetation layer is also increasing, as shown in Fig. 5-31.

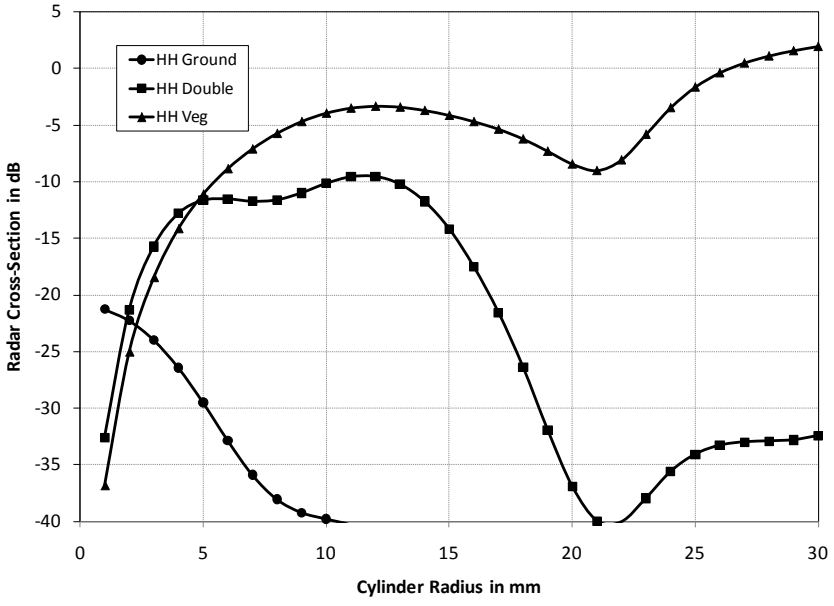


Fig. 5-30. The scattering contributions of the different scattering mechanisms as a function of cylinder radius. Only the HH terms are shown. See the text for discussion.

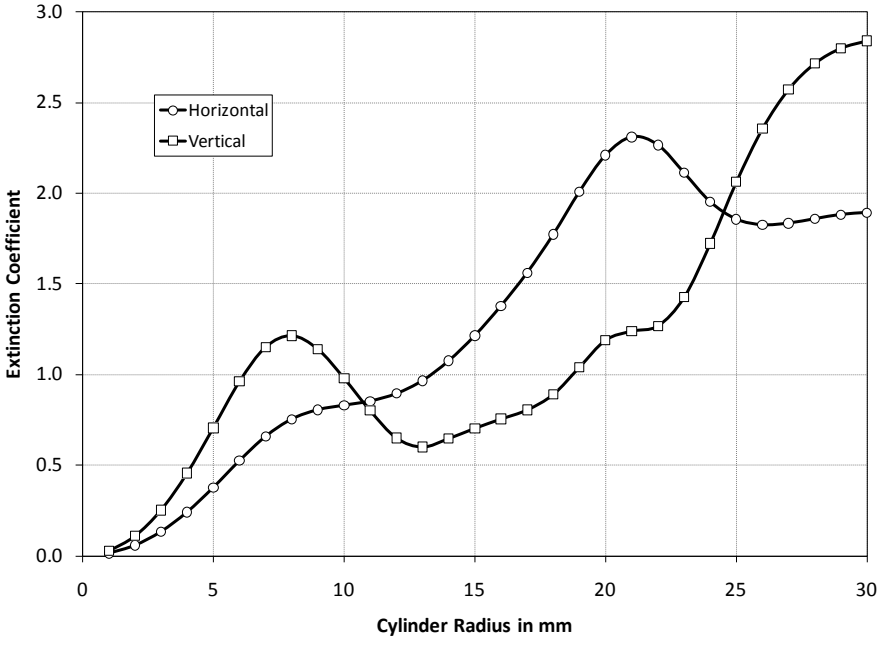


Fig. 5-31. Extinction as a function of cylinder radius. The extinction generally increases as the cylinders get thicker.

5.8.3 Effect of Cylinder Moisture

The cylinder moisture determines the cylinder dielectric constant as described in Section 5.7.1. We now fix all the parameters to their values in Table 5-1 in Section 5.8, and also fix the angle of incidence at 40 deg.

Panel (a) in Fig. 5-32 shows the vegetation scattering as a function of the cylinder water content. As expected, the radar backscatter increases as the cylinder water content (and hence the dielectric constant) increases. The higher dielectric constant means a larger dielectric contrast between the cylinder and the air, resulting in more efficient scattering. The stronger scattering at larger dielectric constant values (or equivalently higher cylinder moisture values) also leads to higher extinction as shown in the panel (b) in Fig. 5-32.

5.8.4 Radar Vegetation Index

We pointed out in Chapter 3 that the radar vegetation index is an indicator of the randomness observed in the scattering, similar to that measured by the entropy. We shall discuss this aspect next. The amount of vegetation is often characterized by the so-called vegetation water content. This quantity is the amount of water, typically expressed in kilograms per square meter, contained

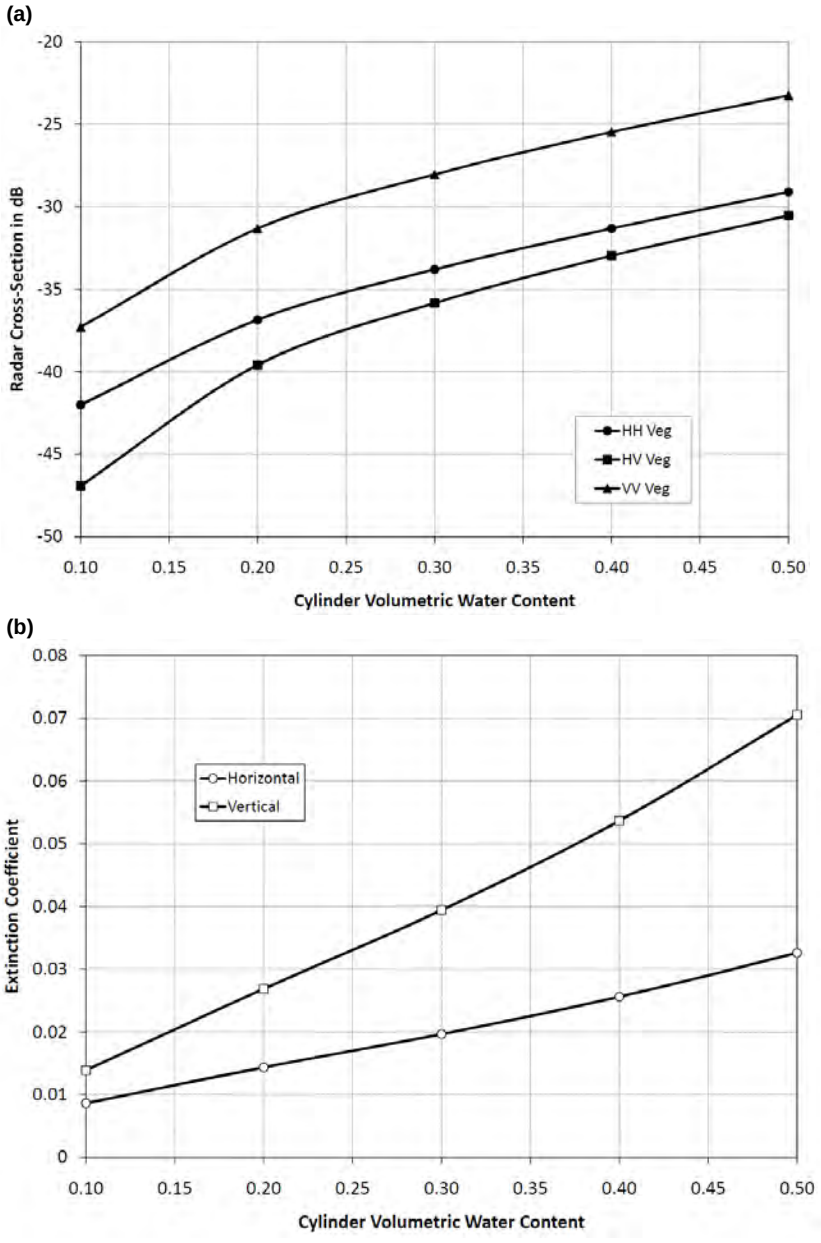


Fig. 5-32. Vegetation scattering (a) and extinction coefficient (b) as a function of the cylinder moisture content. Both quantities increase rapidly with increasing cylinder moisture.

in a column of vegetation with a surface area of 1 square meter. The example we analyzed above would translate to about 1.24 kg/m^2. Unfortunately, the

vegetation water content is a function of many vegetation parameters, namely the radius, length, and density of cylinders, the depth of the vegetation layer, and volumetric moisture of the cylinders.

Figure 5-33 shows the calculated radar vegetation index as a function of the vegetated water content for a number of different cases. The parameters for the different cases are shown in Table 5-2.

First, we notice that for all cases the RVI increases as the vegetation water content increases. This is the result of the fact that at higher vegetation water contents, the scattering from the vegetation layer itself dominates, resulting in a higher RVI. The details differ between the cases however.

Cases 1 and 2 are identical except for the vegetation moisture content. This quantity is used to calculate the dielectric constant of the vegetation as mentioned earlier. In each of these two cases, we keep the vegetation volumetric moisture constant, (and as a consequence the dielectric constant of the vegetation will be constant) and vary the number of cylinders per unit volume to increase the vegetation water content. Here vegetation water content refers to the water content of the layer, not to the moisture in each individual cylinder. These two cases would represent increasing amounts of vegetation

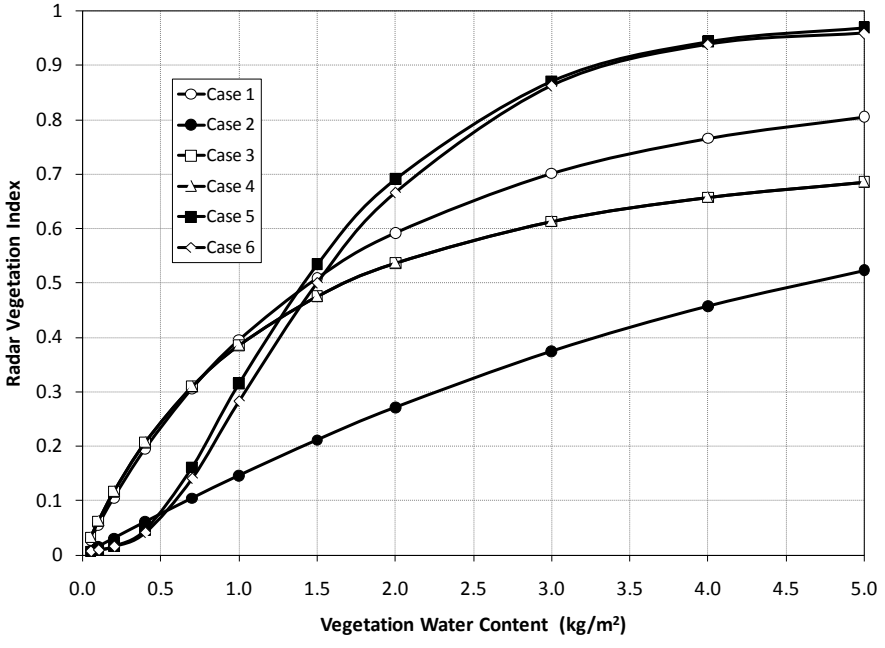


Fig. 5-33. Radar vegetation index as a function of the vegetation water content for various scenarios. See the text for discussion.

Table 5-2. Parameters for several cases of vegetated water content.

Case	Radius	Layer Height	Density	Vegetation Moisture	Soil Moisture
1	1 mm	1 m	Varies	0.5	0.2
2	1 mm	1 m	Varies	0.2	0.2
3	2 mm	1 m	Varies	0.2	0.2
4	2 mm	0.5 m	Varies	0.2	0.2
5	2 mm	1 m	1000	Varies	0.2
6	2 mm	1 m	1000	Varies	0.3

elements such as one would encounter during a growing season, although the constant vegetation water content and layer height would be artificial. In the case of the lower vegetation moisture (Case 2) the RVI increases slower than the higher vegetation water content case. This is a consequence of two factors. First, the higher moisture case means the dielectric constant of the cylinders is higher, leading to stronger scattering by the vegetation, and hence larger RVI values, as discussed earlier. At the same time, the higher dielectric constant also means higher attenuation through the vegetation because of the larger extinction, which decreases the surface related terms, and leads to higher RVI values.

Next, we compare cases 2 and 3. The only difference in this case is that for case 3 the individual cylinders are larger than for case 2. As shown before, larger cylinders scatter more efficiently, and hence the vegetation term is relatively speaking larger than the others, leading to a larger RVI. At the same time, the more efficient scattering means a larger extinction through the vegetation layer, decreasing the contribution from the ground related terms (see Fig. 5-31). This also increases the RVI, so that case 3 generally has larger RVI than case 2.

Case 4 is the same as case 3, except that the vegetation layer is half the height. This means that for the same vegetation water content, case 4 has double the cylinders that case 3 would have. The two RVI curves are identical. The reason for this is clear when the equations from the previous Section are considered. In all cases, we find the product of the density and the layer height. Therefore, if all other parameters are held constant, as long as this product is the same, we would expect the same result.

The cases discussed so far all have constant vegetation water content. For a given vegetation canopy over short time scales, the vegetation elements do not change their size and number, and the layer height can be considered constant. However, due to changing soil moisture and other environmental conditions, it is possible that the cylinder moisture content can change, thereby changing the vegetation water content. This is studied with cases 5 and 6. Both cases are identical except that the soil moisture of case 6 is higher than that of case 5. Here we keep the cylinder physical dimensions and number (density) fixed, and

vary the cylinder moisture content. Compared to case 3, this shows a different behavior. At first, when the vegetation water content is very low, (now vegetation water content is directly proportional to cylinder moisture), the dielectric constant of the cylinders is small, leading to inefficient scattering from the vegetation layer (see Figure 5-32), and a low RVI. As the moisture increases in the vegetation, the scattering from the vegetation quickly becomes dominant, leading to a rapid increase in the RVI. Once the vegetation scattering dominates, there is little additional change in the RVI. Case 6 shows indeed that the underlying soil moisture plays a relatively small part. When the surface is wetter (case 6), the soil term stays dominant a while longer, meaning that the rapid increase of the RVI happens for a slightly higher value of vegetation water content. This is a small effect, however.

5.8.5 Effect of Soil Moisture

The important question is whether we can observe changes in soil moisture by observing the overall scattering from the vegetation. As shown in the previous examples, under certain conditions, either the direct backscatter from the soil surface itself, or the double bounce scattering may dominate even though both these terms are attenuated. In these cases, we would expect to measure a change in the overall backscatter from the vegetation layer as the soil moisture changes.

Figure 5-34 shows the results of a simulation where we have assumed that the vegetation moisture does not change, even though the surface soil moisture changes. All other parameters, such as cylinder length, diameter, etc, as well as surface roughness also remain constant. The values of the parameters used in the simulation are as given in Table 5-1, except we used a cylinder density of 1600 cylinders per cubic meter.

The total scattering terms for both HH and VV show an approximately linear relationship with the soil moisture when plotted in dB. In fact, when fitting a linear function to these graphs, we find

$$\begin{aligned} \sigma_{hh}(dB) &= 9.8 m_{soil} - 21.2; \quad R^2 = 0.95 \\ \sigma_{vv}(dB) &= 13.7 m_{soil} - 18.7; \quad R^2 = 0.96 \end{aligned} \tag{5.8-2}$$

Both these functions fit the simulated data very well. From this one result it appears that the VV polarization shows more sensitivity to the underlying ground surface moisture. A look at the individual scattering mechanisms in Fig. 5-34 shows that the scattering from the ground surface actually dominates in this case. As the previous examples showed, when the vegetation water content increases, the scattering from the vegetation becomes more dominant.

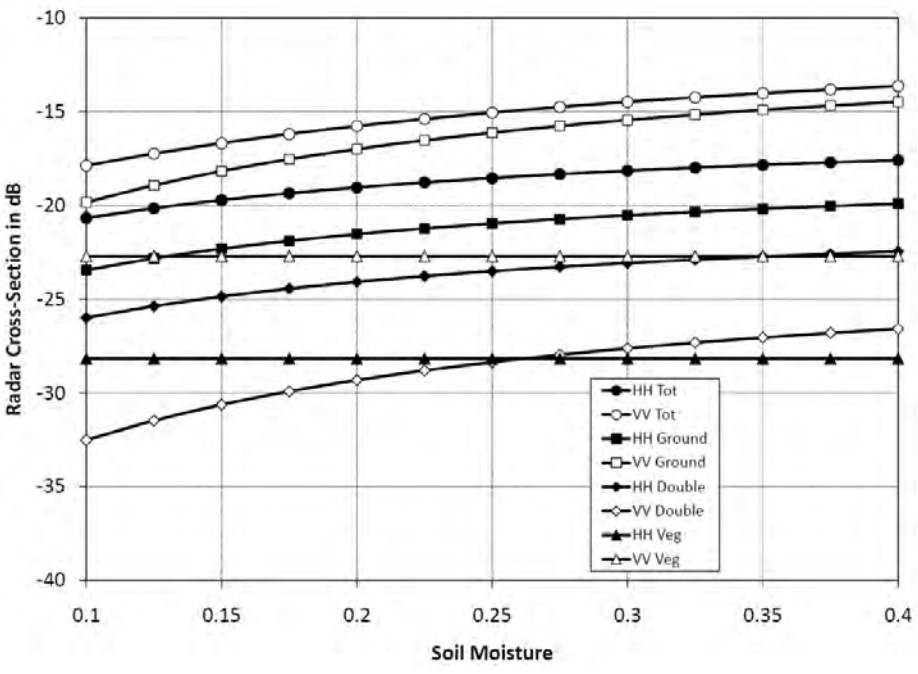


Fig. 5-34. These graphs show the sensitivity of the various scattering terms to changes in the surface soil moisture, assuming that all other parameters are constant. The vegetation water content is 0.8 kg/m^2. The incidence angle is 40 deg.

Figure 5-35 shows the slope of the linear fits to the simulated data for different values of vegetation water content. These values were calculated using different values of the cylinder moisture to vary the vegetation water content. The cylinder size and density were kept fixed.

The data shows, as expected, a decrease in sensitivity to surface soil moisture for increasing vegetation water content. The sensitivity at VV decreases faster, because the VV scattering from the vegetation and particularly the extinction for vertical polarization increases faster with increasing vegetation water content. Both these factors combine to decrease the sensitivity to scattering from the underlying soil surface. HH polarization retains a reasonable sensitivity to surface soil moisture even for the larger vegetation water content values.

Different graphs are shown for different cylinder densities. The cylinder density has a small effect on the overall sensitivity to soil moisture, with smaller densities generally resulting in smaller sensitivities. At first glance this seems counter intuitive as one would expect fewer cylinders to have less extinction, and therefore should show more sensitivity to the underlying soil moisture. The

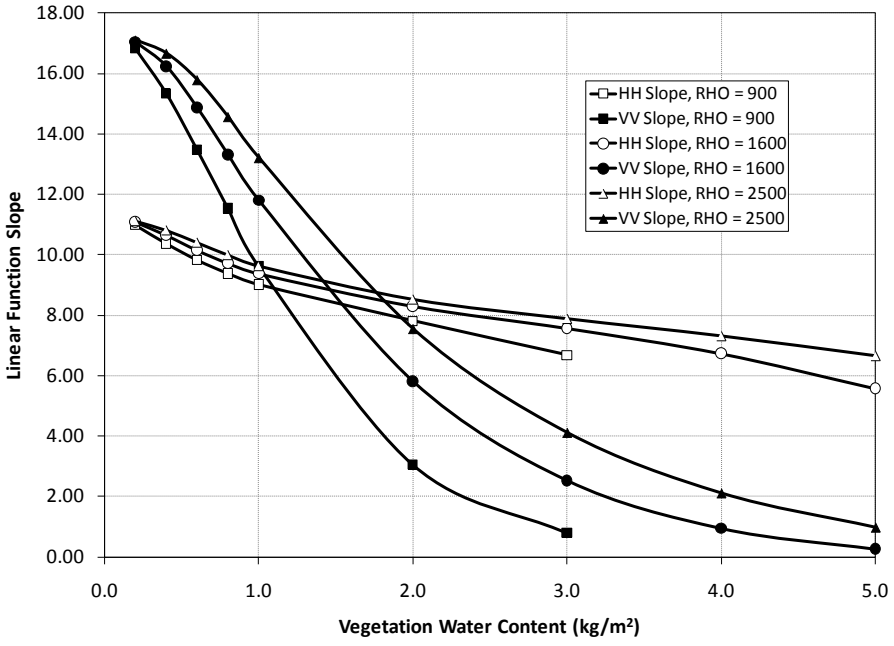


Fig. 5-35. This data shows a decrease in sensitivity to surface soil moisture as the vegetation water content increases. The VV sensitivity decreases faster than the HH sensitivity.

explanation for the graphs lies in the fact that the sensitivity is displayed as a function of the water content of the vegetation layer; not the moisture content of the cylinders. The layer water content is, among other factors, the product of the cylinder moisture and the cylinder density. Therefore, for a particular vegetation layer water content, the lower density layer has a higher cylinder moisture content, and hence a larger dielectric constant. The higher dielectric constant gives a bigger increase in scattering efficiency (and a larger associated extinction) than what is lost because of the lower density, resulting in a lower sensitivity to soil moisture. Note that in the lowest density case the graph stops at a vegetation water content of about 3 kg/m^2. At this value, the cylinders have to be 100% moisture in order for the vegetation layer to have this much water per unit area.

In the simulations above, we assumed that the vegetation moisture stays constant, regardless of what the underlying soil moisture might be. In practice one might expect that as the soil gets wetter, the vegetation moisture might increase, even if there might be a delay between the two values. Unfortunately, there seems to be little quantitative experimental data on this topic. Figure 5-36 shows the case where the vegetation moisture is 100% correlated with the soil

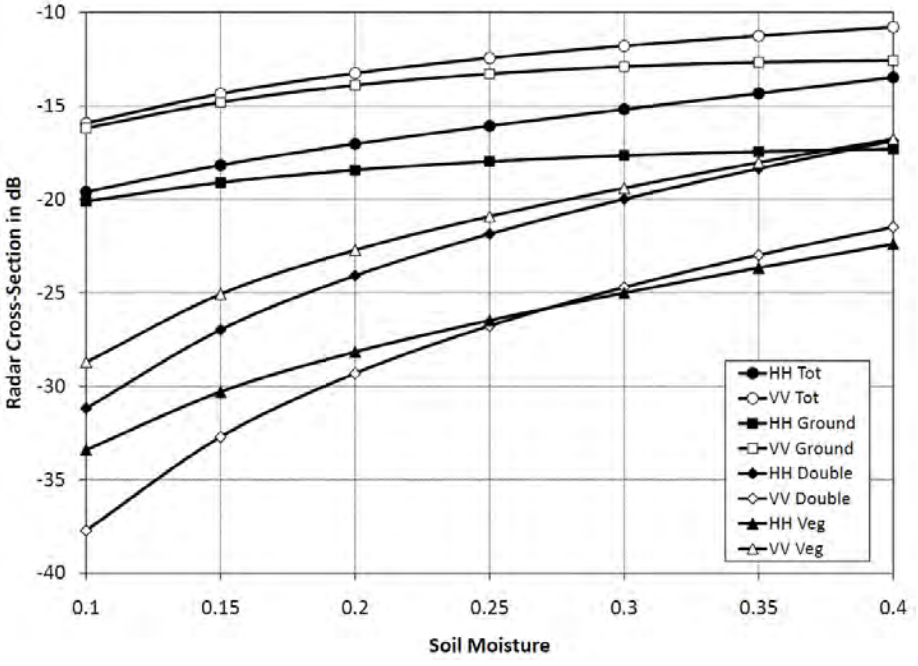


Fig. 5-36. These graphs show the strength of the various scattering components assuming that the soil moisture and the vegetation moisture are the same value. There is significant more sensitivity to moisture if these two quantities are the same.

moisture. This simulation shows that there is a significant increase in sensitivity to soil moisture if the vegetation and soil moistures are correlated. In fact, the two functions for this case are

$$\begin{aligned} \sigma_{hh}(dB) &= 19.9 m_{soil} - 21.2; \quad R^2 = 0.99 \\ \sigma_{vv}(dB) &= 16.5 m_{soil} - 16.9; \quad R^2 = 0.95 \end{aligned} \tag{5.8-3}$$

In fact, the HH sensitivity is nearly twice the maximum value we found before.

This last example is perhaps unrealistic, but can be considered a limiting case. In practice we can expect some correlation between the vegetation moisture and the soil moisture, even if it is not exactly 100% as assumed above. Any correlation will increase the sensitivity to soil moisture over that calculated assuming no correlation.

5.8.6 Inverting for Soil Moisture: The Data Cube

The aim of remote sensing is to be able to infer some of the geophysical properties of the terrain from measurements. With a single frequency

polarimetric radar, we have a limited number of measurements. In the case of terrain with reflection symmetry, we have at most 5 independent measurements, the cross-sections at HH, VV and HV, and the magnitude and phase of the correlation between HH and VV. Our model, on the other hand, has many more input parameters, even for the simple single layer model described here. When this model is extended to multiple layers, the number of input parameters increases dramatically. It is therefore impractical to solve uniquely for all of these input parameters with only a limited number of observables. It is therefore clear that we will have to combine some of the physical characteristics of the vegetation into a single descriptor in order to have any chance of inverting the observations.

In addition to the difficulty of having a limited number of observations, the modeling requires a large number of calculations for each case. This makes inversion of the data computationally quite challenging. One way to get around this problem is to construct "data cubes" for each model. This approach would calculate the polarimetric parameters for a range of input parameters. For example, one choice would be to fix many of the model parameters such as cylinder radius, length, and density, and then calculate the polarimetric parameters as a function of surface roughness, soil moisture, vegetation water content and angle of incidence. These values are then stored in files for later use. Each observation can then be "inverted" by searching these data cubes for the best simultaneous match to all the observables.

Many different ways could be used to define the best match between the observed and simulated data. As an example, for a multi-polarized radar that measures HH, VV and HV cross-sections, one could choose to define the error between the observations and the simulations as

$$\mathrm{E} = \sqrt{\left(\sigma_{hh\,obs} - \sigma_{hh\,sim}\right)^2 + \left(\sigma_{vv\,obs} - \sigma_{vv\,sim}\right)^2 + \left(\sigma_{hv\,obs} - \sigma_{hv\,sim}\right)^2} \ . \quad (5.8\text{-}4)$$

We can calculate this quantity for all values in the data cubes, and choose that set of parameters that minimizes this error to be representative of the vegetated terrain.

The cube inversion method has been studied extensively by Arii (2009) [37] in his thesis, where he compares different combinations of polarimetric parameters and the accuracy with which the vegetation parameters can be inferred. He not only looked at the multi-polarization case described above, but he also looked at including all the polarimetric parameters in the inversion. Since some of the elements of the covariance matrix are complex numbers, he proposes using the natural logarithm of the radar cross-sections in the error

calculations. This will allow the incorporation of the measured polarimetric phases in addition to the magnitudes.

Arii [37] found that for vegetation water content values less than about 0.5 kg/m^2 adding the other polarimetric parameters made little difference in the accuracy of the inversion for soil moisture. However, for vegetation water contents larger than this value, adding the cross-correlation between HH and VV made a substantial difference, improving the accuracy of the inversion substantially.

Arii also points out that the inversion accuracy is very sensitive to the exact values of the cylinder parameters. In his investigation, he simulated the data cubes with a fixed radius of 2 mm, and then generated test data with radii ranging between 2 and 3 mm. He then used the data cubes to invert the data. The resulting errors in soil moisture estimates were largest for larger vegetation moisture contents, with the error exceeding 10% at a vegetation water content of 2.5 kg/m^2 when the radius of the test cylinder was larger than 2.3 mm. Based on these results, Arii concluded that for such an inversion scheme to be successful, one would need use have a family of data cubes generated for different vegetation classes.

5.9 Time Series Estimation of Soil Moisture

So far we have discussed estimations of soil moisture that happens at a particular instant of time. We will call these "snap-shot" algorithms, since they would provide us with an estimate of the moisture as if we took a picture at a specific time. In doing so, we do not rely on any knowledge about the past values of the soil moisture at that specific location. In this Section we shall discuss a different approach in which we will track changes in the soil moisture over time.

To derive such a time series algorithm, we note that the polarimetric backscattering cross-section can be written as

$$\sigma_{pq} = f_{pq}(P_1, P_2, ..., P_N), \tag{5.9-1}$$

where P_i (i = 1, 2, …, N) represents a remote sensing variable such as soil moisture, surface roughness, and parameters characterizing vegetation. The subscripts p and q can be h or v depending upon the polarization configuration. The time variation of the backscattering cross-section is calculated as

$$\frac{\partial \sigma_{pq}}{\partial t} = \sum_{i=1}^{N} \frac{\partial f_{pq}}{\partial P_i} \frac{\partial P_i}{\partial t} . \tag{5.9-2}$$

Notice that $\partial f_{pq} / \partial P_i$ represents the backscattering cross section sensitivity to physical quantities such as soil moisture. This is the term that is used in the "snap-shot" algorithms to estimate soil moisture from polarimetric radar data. However, the time-series formulation shown in Eq. (5.9-2) has an additional term, $\partial P_i / \partial t$, that can be thought of as a temporal filter, which provides a weighting factor for each sensitivity. That is, if the time variation of a parameter such as soil texture and the vegetation structure is not significant enough to change the model function appreciably over a specified time scale, we can ignore the effect of that parameter on the temporal variation of the radar cross-section.

To best define a time-series algorithm, we have to consider the time scale over which the radar cross-section varies. There are three fundamental time scales of importance: 1) diurnal time scale, 2) soil moisture response time after a precipitation event and 3) the time scale over which a model function is invariant; this is the time scale that soil moisture will be retrieved. In order to minimize the diurnal effect on the soil moisture retrieval accuracy, time-series data must be collected at the same solar time (elapsed time after sun rise). In addition, from the time scale of the soil moisture response due to a precipitation event, we can define the sampling requirement of the time-series approach.

For the time-series approach, it is preferred that the radar data are collected using the same imaging geometry for each measurement. Especially the incidence angle should be constant for all time-series measurements unless the incidence angle effect on the backscattering cross-section can be properly compensated for. This can be accomplished with proper mission design. For instance, the Hydros radar configuration discussed by Entekhabi et al. (2004) [38] minimizes the imaging geometry change of a spaceborne radar. Even in this configuration, however, the azimuth angle of these observations can change; the impact of such azimuth angle variations was observed and characterized by Baup et al. (2007) [39] for surfaces in Mali. Except for one area that showed some azimuthal variation, they found little evidence, even in areas covered with sand dunes, for azimuth angle variations creating significant variations in radar cross-section. Nevertheless, it is obvious that a constant illumination geometry would eliminate any variations other than geophysical changes.

For bare surfaces, there are three parameters to be considered: soil moisture (P_1), surface roughness (P_2), and soil texture (P_3). If P_2 and P_3 vary in time much slower than P_1 (soil moisture), Eq. (5.9-2) can be approximated by

$$\frac{\partial \sigma_{pq}}{\partial t} \approx \frac{\partial f_{pq}}{\partial P_1}\frac{\partial P_1}{\partial t} \quad \text{if} \quad \frac{\partial P_2}{\partial t}, \frac{\partial P_3}{\partial t} << \frac{\partial P_1}{\partial t}. \tag{5.9-3}$$

Under this assumption, the temporal variation of a backscattering cross section depends only on the soil moisture variation ($\partial P_1/\partial t$). Therefore,

$$\Delta \sigma_{pq} \approx \frac{\partial f_{pq}}{\partial P_1}\Delta P_1. \tag{5.9-4}$$

It should be pointed out that if the surface roughness is influenced by soil moisture, this formulation can include this effect. For example, if the surface roughness changes when soil becomes dry after a precipitation event $(P_2 = g(P_1))$, then

$$\Delta P_2 = \frac{\partial g}{\partial P_1}\Delta P_1. \tag{5.9-5}$$

In this case the backscattering cross section can be written as

$$\Delta \sigma_{pq} \approx \left[\frac{\partial f_{pq}}{\partial P_1} + \frac{\partial f_{pq}}{\partial P_2}\frac{\partial g}{\partial P_1}\right]\Delta P_1. \tag{5.9-6}$$

The time scale over which this approach can be used to track soil moisture changes depends on the time scale over which the quantity inside the square brackets change. As long as the time is short enough that we can consider this quantity to be constant, we can attribute changes in the observed radar cross-section to changes in soil moisture.

For vegetated surfaces, at least two additional parameters must be added to include the vegetation scattering: vegetation structure (P_4) and vegetation dielectric constant or water content (P_5). If we ignore the vegetation structure (P_4) change and assume that the vegetation dielectric constant is affected by soil moisture ($P_5 = h(P_1)$), the backscattering cross section change can be written as

$$\Delta \sigma_{pq} \approx \left[\frac{\partial f_{pq}}{\partial P_1} + \frac{\partial f_{pq}}{\partial P_5}\frac{\partial h}{\partial P_1}\right]\Delta P_1. \tag{5.9-7}$$

Here, we have also assumed that the roughness change due to the soil moisture variation is much less significant than the dielectric constant change of vegetation. After a precipitation event, the dielectric constant of vegetation can change significantly as observed by McDonald et al. (2002) [40]. Both the vegetation dielectric constant and soil moisture will therefore be influenced by a precipitation event. This effect of the dielectric constant change of vegetation is represented by the second term in Eq. (5.9-7).

The next step is to derive model functions using experimental data: those shown in Eq. (5.9-6) for bare surfaces and in Eq. (5.9-7) for vegetated surfaces. We note that if these model functions are approximately linear, ΔP_1 does not have to be small since higher order terms of ΔP_1 can be ignored.

Model functions for bare surfaces can be derived using theoretical solutions such as SPM and IEM as shown in the previous Section. Kim and van Zyl (2009) [11] linked the radar cross-section expressed in dB to soil moisture directly. A similar time-series approach was successfully applied to European Remote Sensing Satellite (ERS) data by Wagner and Scipal [41]. For vegetated surfaces, Kim and van Zyl (2009) [11] used L-band radar data collected using a truck-mounted radar system from May to October 2002 at USDA-ARS (U.S. Dept. of Agriculture-Agricultural Research Service) OPE3 (Optimizing Production Inputs for Economic and Environmental Enhancement) test site in Beltsville, MD [42]. The results are shown in Fig. 5-37.

From the experimental data Kim and van Zyl (2009) [11] found that both H- and V-polarization backscattering cross sections at L-band can be modeled better using a linear relationship between the radar cross-section expressed in dB and volumetric soil moisture (m_v) as

$$10\log_{10}(\sigma_{pp}) = Cm_v + D \tag{5.9-8}$$

A similar result was reported previously at C-band by Ulaby et al. (1986) [17]. Therefore, using this linear relationship, one can write the time series algorithm for the soil moisture as

$$m_v = A\left(10\log_{10}\sigma_{pp}\right) + B \tag{5.9-9}$$

Since the two unknowns, A and B, may vary from one pixel to the next, one must develop a method to reliably estimate A and B for each pixel. Kim and van Zyl (2009) [11] point out that in order to determine the two unknowns (A and B) in Eq. (5.9-9), it is necessary to estimate at least two soil moisture values corresponding to two backscattering cross sections. As an example, if

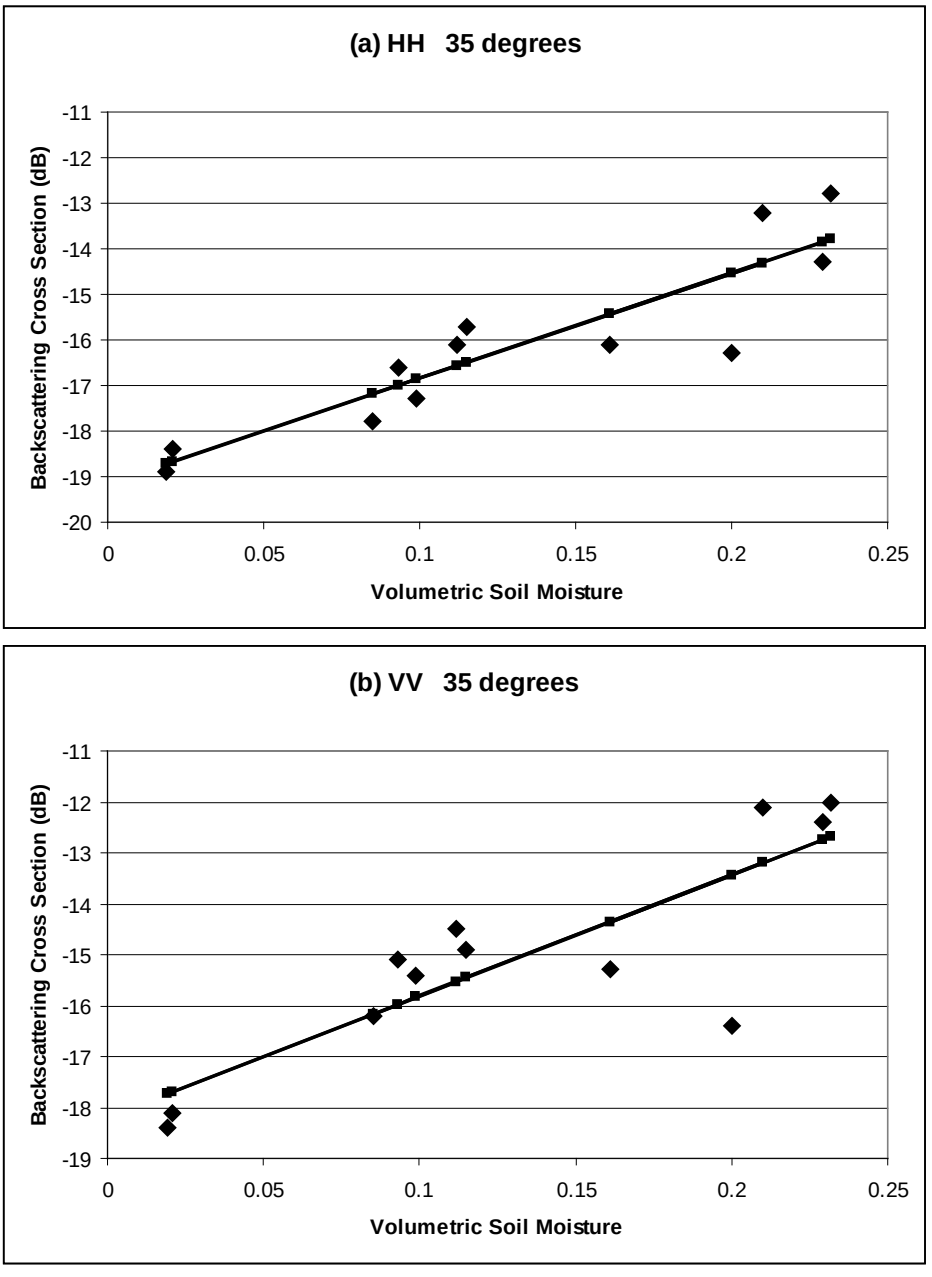


Fig. 5-37. Linear fit of L-band HH and VV backscattering cross sections (dB) for various soil moisture (unitless) collected during the USDA-ARS OPE3 (2002) [42] field experiment for a corn field (biomass > 2.5 kg/m2). In order to remove the diurnal effect, the radar data collected near 8:00 AM were used in this figure. The incidence angle is 35 deg.

the extreme soil moisture values (as an example, 35% for the maximum value and 5% for the minimum value) of time-series data are known, A and B can be estimated using the corresponding maximum and minimum backscattering cross sections. This might not be very practical since these extreme values may not be known for each pixel. On the other hand, if the time period includes completely dry and completely wet (i.e., run-off) conditions, the estimation of A and B are relatively straightforward.

Using experimental data, Kim and van Zyl (2009) [11] showed that for bare surfaces, it is possible to estimate volumetric soil moisture better than 2% if the extreme soil moisture values are known exactly. The accuracy slightly degrades for vegetated surfaces. If incorrect extreme values are used, the retrieval accuracy becomes worse. In order to evaluate the effect of the error in extreme soil moisture values statistically, Kim and van Zyl (2009) [11] performed a simple simulation by adding a random error to the exact soil moisture value of two extreme conditions (wet and dry). The random error is simulated using a uniform random variable. The simulation results are shown in Fig. 5-38. The abscissa value (x) in Fig. 5-38 represents the uniform random error over [–x, +x]. The retrieval error is calculated by averaging the magnitude of estimation errors when these random errors are introduced to the inversion process. For both bare and vegetated surfaces, the extreme soil moisture values (wet and dry) must be known better than 6% in order to reduce the retrieval error to be less than 4%.

Our previous simulations showed that as the vegetation water content increase, the slope of the linear function relating radar cross-section to soil moisture changes (see Figure 5-35). Therefore, over the full growing season, one may have to segment the radar data in order to apply the correct model function for a vegetated surface. The Radar Vegetation Index (RVI) has been proposed to identify vegetated surfaces [11]. The RVI values for the time-series data from the USDA-ARS OPE3 are shown in Fig. 5-39. When RVI > 0.35, the biomass level of the corn field was higher than 2.5 kg/m^2 except one data point. The peak biomass was about 7 kg/m^2. From this analysis, a vegetation model function must be used when RVI > 0.35. When we consider the attenuation due to vegetation (biomass as high as 7 kg/m^2), it is not obvious that the radar sensitivity to the soil moisture variation is due to the direct scattering from a soil surface. It is possible that the dielectric constant change of vegetation enhances the sensitivity.

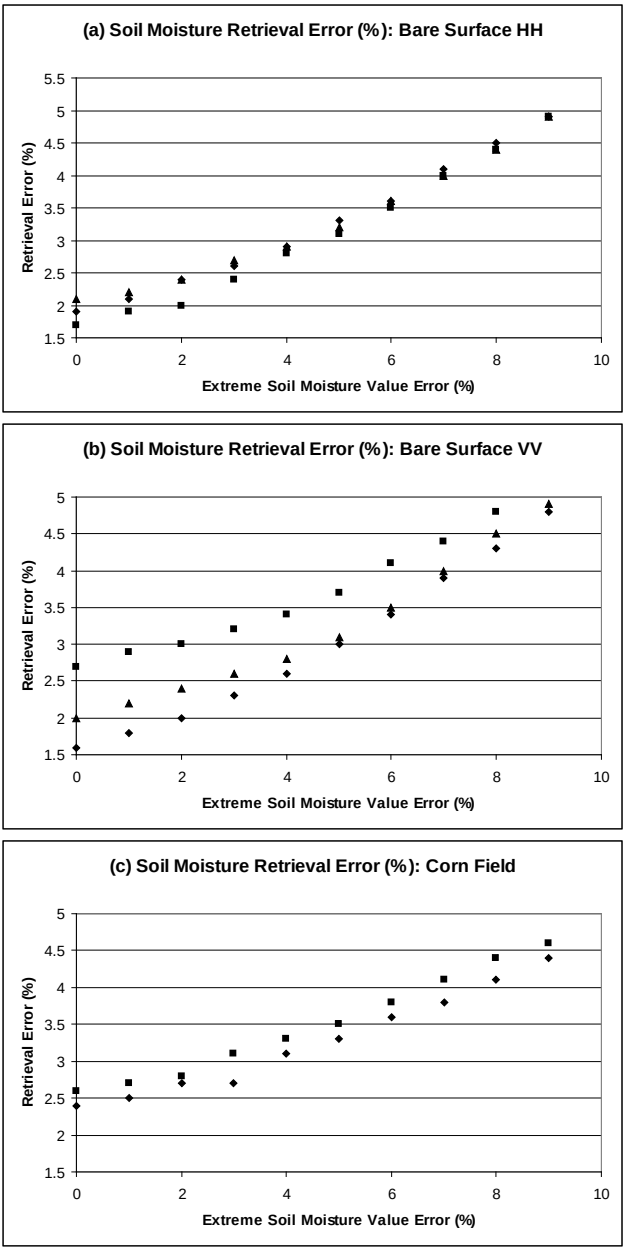


Fig. 5-38. Average retrieval errors of volumetric soil moisture (%): (a) HH for a bare surface (diamond: 30 deg, square: 40 deg. and triangle: 50 deg.), (b) VV for a bare surface (diamond: 30 deg, square: 40 deg. and triangle: 50 deg.) and (c) HH and VV for a corn field (diamond: HH and square: VV) (Kim and van Zyl, 2009) [11].

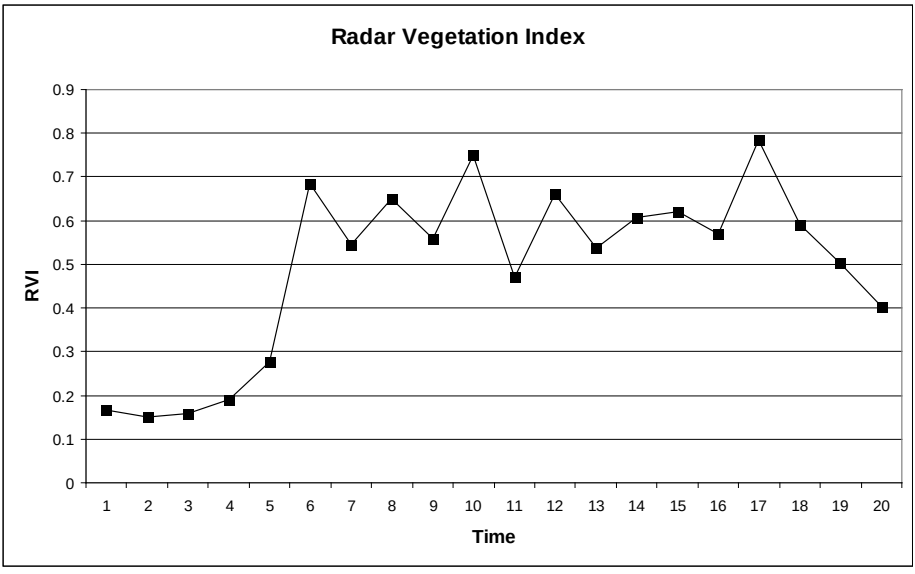


Fig. 5-39. RVI (Radar Vegetation Index) estimated from polarimetric radar data collected during the USDA-ARS OPE3 (2002) field experiment. The abscissa represents time in weeks. When RVI is compared with biomass measurements, the condition that RVI > 0.35 defines all data with biomass > 2.5kg/m2 except one point (6th week).

5.10 Summary

In this Chapter, we examined methods for using radar data to estimate surface soil moisture and roughness. Many different models exist to calculate the radar cross-section as a function of surface roughness and surface dielectric constant. Herein, we looked only at a small subset of these models and showed that they all predict sensitivity to soil moisture.

We also looked at a few of the more commonly used inversion algorithms and compared the results. Most of these models perform reasonably well for bare surfaces. For vegetated surfaces, however, the models become very complicated and the soil moisture inversion becomes less accurate.

Most of the models we discussed here fall into the class of “snapshot” algorithms; that is, they try to estimate the soil moisture using data only from one instant in time, effectively ignoring all prior information that might exist. We also discussed a different approach that holds the promise of being applicable to many different types of surface covers. This time series approach fits a linear function through the radar cross section as a function of soil moisture using the extreme values of the radar cross-sections as indicators of saturated wet and completely dry conditions. As long as these extreme values are accurate to within 6 percent, soil moisture can be estimated to an accuracy

of better than 4 percent. The obvious disadvantage of the time series approach is that inversions can only be performed once the extreme values have been observed.

The science of inverting radar data for soil moisture will receive a significant increase in attention in the near future. As of the writing of this book, the Japanese Advanced Land Observing Satellite Phased Array type L-band Synthetic Aperture Radar (ALOS PALSAR) system is acquiring polarimetric SAR data that could be used to check many of these algorithms. The National Aeronautics and Space Administration (NASA) is also planning the Soil Moisture Active Passive (SMAP) mission for launch in 2015. This system will carry both a radar and a radiometer to routinely cover the globe specifically for the purpose of estimating surface soil moisture. Undoubtedly, this area of research will expand rapidly to meet this challenge, and the climate research community can look forward to excellent results from these missions.

References

[1] J. Wang, E. Engman, J. Shiue, M. Ruzek and C. Steinmeier, "The SIR-B Observations of Microwave Backscatter Dependence on Soil Moisture, Surface Roughness, and Vegetation Covers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 24, pp. 510–516, 1986.

[2] M. C. Dobson and F.T. Ulaby, "Preliminary Evaluation of the SIR-B Response to Soil Moisture, Surface Roughness, and Crop Canopy Cover," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 24, pp. 517-526, 1986.

[3] T. J. Jackson and F. R. Shiebe, *Washita '92 Data Report*, NAWQL Report 101, USDA National Agricultural Water Quality Lab, Durant, Oklahoma, 1993.

[4] K. Schneider and N. Oppelt, "The Determination of Mesoscale Soil Moisture Patterns with ERS Data," *1998 IEEE International Geoscience and Remote Sensing Symposium Proceedings (IGARSS1998)*, Vol. 4, pp. 1831–1833, 1998.

[5] A. S. Quesney, S. Le Hegarat-Mascle, O. Taconet, D. Vidal-Madjar, J. P. Wigneron, C. Loumagne, and M. Normand. "Estimation of Watershed Soil Moisture Index from ERS/SAR Data," *Remote Sensing of the Environment*, vol. 72, pp. 290–303, 2000.

[6] P. C. Dubois, J. J. van Zyl, and E. T. Engman, "Measuring Soil Moisture with Imaging Radars," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, issue 4, pp. 915–926; plus corrections to, vol. 33, issue 7, p. 1340, 1995.

[7] J. Shi, J. Wang, A. Hsu, P. O'Neill, and E. T. Engman, "Estimation of Bare Surface Soil Moisture and Surface Roughness Parameters Using L-Band SAR Image Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 1254-1266, 1997.

[8] Y. Oh, K. Sarabandi, and F. T. Ulaby, "An Empirical Model and Inversion Technique for Radar Scattering from Bare Soil Surface," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 370–381, 1992.

[9] U. Wegmuller, *Active and Passive Microwave Signature Catalogue on Bare Soil (2-12 GHz)*, Institute of Applied Physics, University of Berne, Switzerland, 1993.

[10] A. K. Fung, Z. Li, and K. S. Chen, "Backscattering from a Randomly Rough Dielectric Surface," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 356–369, 1992.

[11] Y. Kim and J. J. van Zyl, "A Time-Series Approach to Estimate Soil Moisture Using Polarimetric Radar Data: Bare Surfaces," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 2519–2527, August 2009.

[12] T.G. Farr, "Microtopographic Evolution of Lava Flows at Cima Volcanic Field, Mojave Desert, California" *Journal of Geophysical Research*, vol. 97, pp. 15,171–15,179, 1992.

[13] J. R. Wang and T. J. Schmugge, "An Empirical Model for the Complex Dielectric Permittivity of Soils as a Function of Water Content," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 18, pp. 288–295, 1980.

[14] M. T. Hallikainen, F. T. Ulaby, M. C. Dobson, M. A. El-Rayes, and L. Wu, "Microwave Dielectric Behavior of Wet Soil – Part I: Empirical Models and Experimental Observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 23, pp. 25–34, 1985.

[15] B. Brisco, T. J. Pultz, R. J. Brown, G. C. Topp, M. A. Hares, and W. D. Zebchuck, "Soil Moisture Measurement Using Potable Dielectric Probes and Time Domain Reflectometry," *Water Resources Research*, vol. 28, pp. 1339–1346, 1992.

[16] M. C. Dobson, F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes, "Microwave Dielectric Behavior of Wet Soil – Part II: Dielectric Mixing Models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 23, pp. 35-46, 1985.

[17] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, vol. II*, Artech House, Norwood, Massachusetts, pp. 860-862, 1986.

[18] L. Bruckler, H. Witono, and P. Stengel, "Near Surface Soil Moisture Estimation from Microwave Measurements," *Remote Sensing of Environment*, vol. 26, pp. 101-121, 1988.

[19] J. P., Walker, P. A. Troch, M. Mancini, G. R. Willgoose, and J. D. Kalma, "Profile Soil Moisture Estimation Using the Modified IEM," *1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings (IGARSS'97)*, Vol. 3, pp. 1263-1265, 1997.

[20] A. Le Morvan, M. Zribi, N. Baghdadi, and A. Chanzy, "Soil Moisture Profile Effect on Radar Signal Measurement," *Sensors*, vol. 8, pp. 256-270, 2008.

[21] D. Entekhabi, H. Nakamura, and E. G. Njoku, "Solving the Inverse Problem for Soil Moisture and Temperature Profiles by Sequential Assimilation of Multifrequency Remotely Sensed Observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 438–448, March 1994.

[22] J. F. Galantowitz, D. Entekhabi, and E. G. Njoku, "Tests of Sequential Data Assimilation for Retrieving Profile Soil Moisture and Temperature from Observed L-band Radiobrightness," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 1860–1870, July 1999.

[23] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, Vol. I. Addison-Wesley, Reading, Massachusetts, pp. 82–87, 1981.

[24] A. Tabatabaeenejad and M. Moghaddam, "Bistatic Scattering from Three-Dimensional Layered Rough Surfaces," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 2102–2114, August 2006.

[25] C. Kuo and M. Moghaddam, "Electromagnetic Scattering From Maltilayer Rough Surfaces With Arbitrary Dielectric Profiles for Remote Sensing of Subsurface Soil Moisture," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 349–366, February 2007.

[26] A. Tabatabaeenejad and M. Moghaddam, "Inversion of Subsurface Properties of Layered Dielectric Structures With Random Slightly Rough Interfaces using the Method of Simulated Annealing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 2035–2046, July 2009.

[27] S. H. Lou, L. Tsang, C. H. Chan, and A. Ishimaru, "Application of the Finite Element Method to Monte Carlo Simulations of Scattering of Waves by Random Rough Surfaces With the Periodic Boundary Condition," *Journal of Electromagnetic Waves and Applications*, vol. 5, pp. 835–855, 1991.

[28] K. Pak, L. Tsang, L. Li, and C. H. Chan, "Combined Random Rough Surface and Volume Scattering Based on Monte Carlo Simulations of Solutions of Maxwell's Equations," *Radio Science*, vol. 28, pp. 331–338, 1993.

[29] W. H. Peake and T. L. Oliver, *The Response of Terrestrial Surfaces at Microwave Frequencies*, Technical Report AFAL-TR-70301, Electroscience Laboratory, Ohio State University, Columbus, Ohio, 1971.

[30] T. D. Wu, K. S. Chen, J. Shi, and A. K. Fung, "A Transition Model for the Reflection Coefficient in Surface Scattering," *IEEE Transactions on Geoscience and Remote Sensing.*, vol. 39, pp. 2040–2050, 2001.

[31] S. O. Rice, ''Reflection of Electromagnetic Waves from Slightly Rough Surfaces,'' *Communications on Pure and Applied Mathematics*, vol. 4, pp. 361–378, 1951.

[32] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Artech House, Norwood, Massachusetts, 1987.

[33] J. R. Wang, P. E. O'Neill, E. T. Engman, R. Pardipuram, J. C. Shi, and A. Y. Hsu, "Estimating Surface Soil Moisture from SIR-C Measurements over the Little Washita Watershed," *1995 International Geoscience and Remote Sensing Symposium (IGARSS '95)*, Vol. III, IEEE, pp. 1982–1984, 1995.

[34] S. L. Durden, J. J. van Zyl, and H. A. Zebker, "Modeling and Observations of the Radar Polarization Signatures of Forested Areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 27, pp. 290–301, 1989.

[35] F. T. Ulaby and M. A. El-Rayes, "Microwave Dielectric Spectrum of Vegetation Part II: Dual-Dispersion Model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 25, pp. 550–557, 1987.

[36] J. J.van Zyl, "The Effects of Topography on the Radar Scattering from Vegetated Areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, pp. 153–160, 1993.

[37] M. Arii, "Soil Moisture Retrieval under Vegetation Using Polarimetric Radar," Ph.D. dissertation, California Institute of Technology, Pasadena, California, 2009.

[38] D. Entekhabi, E. G. Njoku, P. Houser, M. Spencer, T. Doiron, Y. Kim, J. Smith, R. Girard, S. Belair, W. Crow, T. J. Jackson, Y. H. Kerr, J. S. Kimball, R. Koster, K. C. McDonald, P. E. O'Neill, T. Pultz, S. W. Running, J. Shi, E. Wood, and J. van Zyl, "The Hydrosphere State (Hydros) Satellite Mission: an Earth System Pathfinder for Global Mapping of Soil Moisture and Land Freeze/Thaw," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 2184–2195, 2004.

[39] F. Baup, , E. Mougin, P. Hiernaux, A. Lopez, P. D. Rosnay, and I. Chenerie, "Radar Signatures of Sahelian Surfaces in Mali Using ENVISAT-ASAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 2354–2363, 2007.

[40] K C. McDonald, R. Zimmermann, and J. S. Kimball, "Diurnal and Spatial Variation of Xylem Dielectric Constant in Norway Spruce (Picea abies [L.] Karst.) as Related to Microclimate, Xylem Sap Flow, and Xylem Chemistry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, pp. 2063–2082, September 2002.

[41] W. Wagner and K. Scipal, "Large-Scale Soil Moisture Mapping in Western Africa Using the ERS Scatterometer," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 1777–1782, 2000.

[42] P. E. O'Neill, A. Joseph, G. De Lannoy, R. Lang, C. Utku, E. Kim, P. Houser, and T. Gish, "Soil Moisture Retrieval Through Changing Corn Using Active/Passive Microwave Remote Sensing," *2003 IEE International Geoscience and Remote Sensing Symposium (IGARSS 2003: Learning from Earth's Shapes and Sizes*, Toulouse, France, pp. 407–409, July 2003.

Appendix A
Tilted Small Perturbation Model Details

Consider a tilted surface element as shown in Fig. A-1. The plane of incidence is the (x,z) plane. The coordinate system in which the radar measures the scattering coefficients is the (h,v,k) system and is described by

$$\begin{aligned}\hat{\mathbf{k}} &= -\sin\theta\,\hat{\mathbf{x}} - \cos\theta\,\hat{\mathbf{z}} \\ \hat{\mathbf{h}} &= \frac{-\hat{\mathbf{k}}\times\hat{\mathbf{z}}}{\left|\hat{\mathbf{k}}\times\hat{\mathbf{z}}\right|} = -\hat{\mathbf{y}} \\ \hat{\mathbf{v}} &= \frac{-\hat{\mathbf{h}}\times\hat{\mathbf{k}}}{\left|\hat{\mathbf{h}}\times\hat{\mathbf{k}}\right|} = -\cos\theta\,\hat{\mathbf{x}} + \sin\theta\,\hat{\mathbf{z}}\end{aligned} \quad . \tag{A.1}$$

We shall refer to the radar coordinate system as the global coordinate system and the tilted coordinate system as the local coordinate system. The local and global incidence angles are defined as

$$\cos\theta_l = -\hat{\mathbf{k}}\cdot\hat{\mathbf{z}}_l \,. \tag{A.2}$$

and

$$\cos\theta = -\hat{\mathbf{k}}\cdot\hat{\mathbf{z}} \,. \tag{A.3}$$

We define the local coordinate system such that the local z-axis coincides with the local normal vector of the tilted surface. Let α denote the tilt angle in the range direction; that is, in the $x-z$ plane. We define positive values of α as those that tilt the surface towards the radar. Similarly, let β denote the tilt in

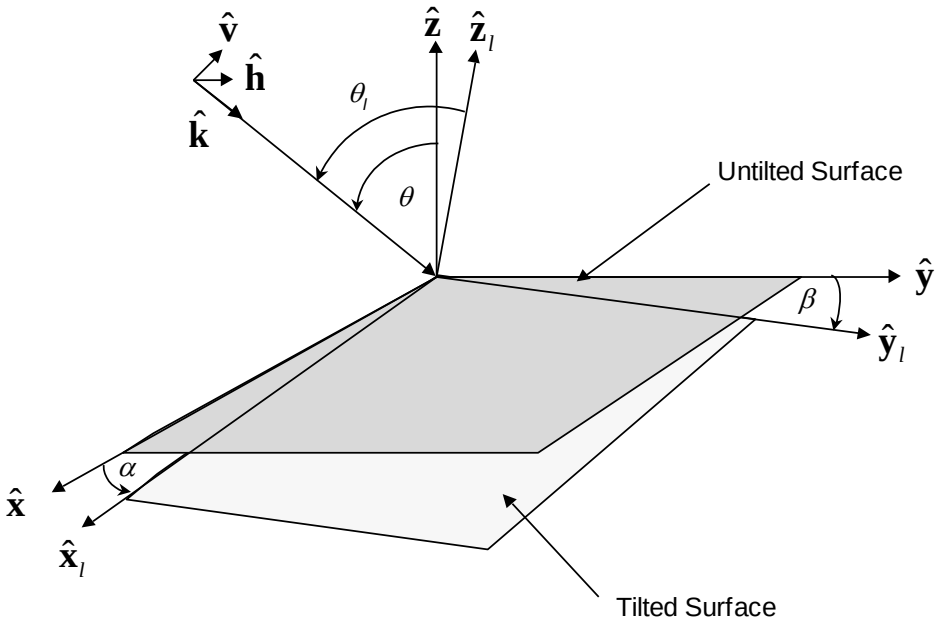


Fig. A-1. Scattering geometry for a tilted surface. The surface is tilted by an angle α in the range (or cross-track) direction and by an angle β in the azimuth(or along-track) direction. The radar measures the scattering coefficients in the global coordinate system, whereas scattering models predict the scattering coefficients in the local coordinate system.

the azimuth direction; that is, in the $y-z$ plane. We define positive values of β as those that tilt the surface in a clockwise direction. The local normal to the surface can then be written as

$$\hat{\mathbf{n}}_l = \frac{\tan\alpha\hat{\mathbf{x}} + \tan\beta\hat{\mathbf{y}} + \hat{\mathbf{z}}}{\sqrt{\tan^2\alpha + \tan^2\beta + 1}} = \frac{h_x\hat{\mathbf{x}} + h_y\hat{\mathbf{y}} + \hat{\mathbf{z}}}{\sqrt{1 + h_x^2 + h_y^2}}. \tag{A.4}$$

with h_x and h_y the slopes in the *x* and *y* directions, respectively. Therefore, the z-axis unit vector in the local coordinate system is

$$\hat{\mathbf{z}}_l = \frac{h_x\hat{\mathbf{x}} + h_y\hat{\mathbf{y}} + \hat{\mathbf{z}}}{\sqrt{1 + h_x^2 + h_y^2}}. \tag{A.5}$$

Using Eq. (A.5) in Eq. (A.2), we find the local incidence angle as

$$\cos\theta_l = \frac{h_x\sin\theta + \cos\theta}{\sqrt{1 + h_x^2 + h_y^2}}. \tag{A.6}$$

The radar coordinates expressed in the local coordinate system are needed to calculate the scattering cross-section in the global coordinate system. The scattering matrix in the local coordinate system is calculated assuming

$$\begin{aligned}\hat{\mathbf{h}}_l &= -\hat{\mathbf{k}} \times \hat{\mathbf{z}}_l \Big/ \left|\hat{\mathbf{k}} \times \hat{\mathbf{z}}_l\right| \\ \hat{\mathbf{v}}_l &= -\hat{\mathbf{h}}_l \times \hat{\mathbf{k}} \Big/ \left|\hat{\mathbf{h}}_l \times \hat{\mathbf{k}}\right|\end{aligned}. \tag{A.7}$$

From Eq. (A.1) and Eq. (A.5) we find that

$$\hat{\mathbf{h}}_l = \frac{-h_y \cos\theta \hat{\mathbf{x}} - \left(\sin\theta - h_x \cos\theta\right)\hat{\mathbf{y}} + h_y \sin\theta \hat{\mathbf{z}}}{\sqrt{h_y^2 + \left(\sin\theta - h_x \cos\theta\right)^2}}. \tag{A.8}$$

Using Eq. (A.8) and Eq. (A.1) in Eq. (A.7), we find

$$\hat{\mathbf{v}}_l = \frac{\cos\theta\left(h_x \cos\theta - \sin\theta\right)\hat{\mathbf{x}} + h_y \hat{\mathbf{y}} - \sin\theta\left(h_x \cos\theta - \sin\theta\right)\hat{\mathbf{z}}}{\sqrt{h_y^2 + \left(\sin\theta - h_x \cos\theta\right)^2}}. \tag{A.9}$$

Note that in the absence of tilts, Eq. (A.9) and Eq. (A.8) reduce to Eq. (A.1).

The scattering matrix in the local coordinate system is defined in these coordinates. We can now calculate the relationship between the scattering matrix measured in the local coordinates, which is usually what is predicted by models, and the one measured in the global coordinates as seen by the radar.

Note that for the purpose of calculating the scattering cross-sections in the global coordinate system, we do not need to know the exact expressions for the local $\left(x_l, y_l, z_l\right)$ coordinates; they were already taken into account during the modeling process when the scattering model was originally formulated. All we need to do here is to transform from the local $\left(h_l, v_l, k\right)$ coordinate system to the global $\left(h, v, k\right)$ system. Let $\vec{E}^t$ be the electric field transmitted by the radar. Then

$$\vec{E}^t = E_h^t \hat{\mathbf{h}} + E_v^t \hat{\mathbf{v}} = E_{hl}^t \hat{\mathbf{h}}_l + E_{vl}^t \hat{\mathbf{v}}_l. \tag{A.10}$$

From this, it follows that

$$E_{hl}^t = \vec{E}^t \cdot \hat{\mathbf{h}}_l = E_h^t\left(\hat{\mathbf{h}} \cdot \hat{\mathbf{h}}_l\right) + E_v^t\left(\hat{\mathbf{v}} \cdot \hat{\mathbf{h}}_l\right) \tag{A.11}$$

and

$$E_{vl}^{t} = \vec{E}^{t} \cdot \hat{\mathbf{v}}_l = E_h^t \left(\hat{\mathbf{h}} \cdot \hat{\mathbf{v}}_l \right) + E_v^t \left(\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}_l \right). \tag{A.12}$$

Thus, we can write

$$\vec{E}_l^t = \left[\mathbf{T} \right] \vec{E}^t, \tag{A.13}$$

where the transformation matrix is given by

$$\left[\mathbf{T} \right] = \begin{pmatrix} \hat{\mathbf{h}} \cdot \hat{\mathbf{h}}_l & \hat{\mathbf{v}} \cdot \hat{\mathbf{h}}_l \\ \hat{\mathbf{h}} \cdot \hat{\mathbf{v}}_l & \hat{\mathbf{v}} \cdot \hat{\mathbf{v}}_l \end{pmatrix}. \tag{A.14}$$

Similarly,

$$E_h^t = \vec{E}_l^t \cdot \hat{\mathbf{h}} \;\; = E_{hl}^t \left(\hat{\mathbf{h}} \cdot \hat{\mathbf{h}}_l \right) + E_v^t \left(\hat{\mathbf{h}} \cdot \hat{\mathbf{v}}_l \right) \tag{A.15}$$

and

$$E_v^t = \vec{E}_l^t \cdot \hat{\mathbf{v}} \;\; = E_{hl}^t \left(\hat{\mathbf{v}} \cdot \hat{\mathbf{h}}_l \right) + E_v^t \left(\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}_l \right). \tag{A.16}$$

Thus, we can write the transformation from the local to the global coordinates as

$$\vec{E}^t = \left[\tilde{\mathbf{T}} \right] \vec{E}_l^t, \tag{A.17}$$

where the ~ character means transpose.

Now we are ready to derive an expression for the scattering matrix in the global coordinate system. First, we transform the electric field transmitted by the radar into the local coordinate system. This is defined by Eq. (A.13) and Eq. (A.14) above. Next, the scattered electric field in the local coordinate system is given by

$$\vec{E}_l^{sc} = \left[\mathbf{S}_l \left(\theta_l \right) \right] \vec{E}_l^t = \left[\mathbf{S}_l \left(\theta_l \right) \right] \left[\mathbf{T} \right] \vec{E}^t. \tag{A.18}$$

The addition of the local angle of incidence to the local scattering matrix is a reminder that the appropriate angle of incidence should be calculated using Eq. (A.6). After transforming this scattered electric field to the global coordinate system, we find

$$\vec{E}^{sc} = \left[\tilde{\mathbf{T}} \right] \vec{E}_l^{sc} = \left[\tilde{\mathbf{T}} \right] \left[\mathbf{S}_l \left(\theta_l \right) \right] \left[\mathbf{T} \right] \vec{E}^t. \tag{A.19}$$

Therefore, the scattering matrices in the global and local coordinate systems are related by

$$\left[\mathbf{S}(\theta)\right] = \left[\tilde{\mathbf{T}}\right]\left[\mathbf{S}_l(\theta_l)\right][\mathbf{T}]. \tag{A.20}$$

To understand the effect of this transformation, we now need explicit expressions for the elements of the transformation matrix. These are

$$[\mathbf{T}] = \frac{1}{\sqrt{h_y^2 + u^2}}\begin{pmatrix} u & h_y \\ -h_y & u \end{pmatrix}; \quad u = \sin\theta - h_x\cos\theta\,. \tag{A.21}$$

These expressions show that the scattering matrix in the global coordinate system is related to the one in the local coordinate system through a rotation angle that is a function of three parameters: the azimuth slope h_y of the pixel, the range slope h_x of the pixel, and the angle of incidence θ for the pixel. This rotation angle is

$$\tan\varphi = \frac{h_y}{u}\,. \tag{A.22}$$

We can write Eq. (A.21) as

$$[\mathbf{T}] = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}. \tag{A.23}$$

To be really useful in our studies, we need to know how the covariance matrix in the global coordinate system is related to the one in the local coordinate system. This can be derived as follows. First, we note that we can write

$$\vec{S} = \begin{pmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{pmatrix} = [\mathbf{Q}]\vec{S}_l \tag{A.24}$$

with

$$[\mathbf{Q}] = \frac{1}{2}\begin{pmatrix} 1+\cos 2\varphi & -\sqrt{2}\sin 2\varphi & 1-\cos 2\varphi \\ \sqrt{2}\sin 2\varphi & 2\cos 2\varphi & -\sqrt{2}\sin 2\varphi \\ 1-\cos 2\varphi & \sqrt{2}\sin 2\varphi & 1+\cos 2\varphi \end{pmatrix}. \tag{A.25}$$

The covariance matrix is defined as

$$[\Sigma] = \vec{S}\tilde{\vec{S}}^* = [\mathbf{Q}]\vec{S}_l\tilde{\vec{S}}_l^*\left[\tilde{\mathbf{Q}}\right]^* = [\mathbf{Q}][\Sigma_l]\left[\tilde{\mathbf{Q}}\right]. \tag{A.26}$$

This expression allows us to calculate the covariance matrix in the global coordinate system once we know the covariance matrix in the local coordinate system.

Appendix B
Bistatic Scattering Matrix of a Cylinder with Arbitrary Orientation

In this appendix we summarize the equations describing the bistatic scattering from a finite length dielectric cylinder with arbitrary orientation. To describe the scattering, we shall use a global coordinate system (see Fig. B-1) wherein the cylinder orientation is described using a unit vector aligned with the long axis of the cylinder, as follows:

$$\mathbf{c}\left(\theta_c, \phi_c\right) = \sin\theta_c \cos\phi_c \, \mathbf{x} + \sin\theta_c \sin\phi_c \, \mathbf{y} + \cos\theta_c \, \mathbf{z} \tag{B.1}$$

We shall assume that an electromagnetic wave is incident upon this cylinder in such a way that the propagation vector of the incident wave can be written as

$$\mathbf{k}_i = -\sin\theta_i \cos\phi_i \, \mathbf{x} - \sin\theta_i \sin\phi_i \, \mathbf{y} - \cos\theta_i \, \mathbf{z} \,. \tag{B.2}$$

We are interested in the scattered field that is represented by the propagation vector

$$\mathbf{k}_s = -\sin\theta_s \cos\phi_s \, \mathbf{x} - \sin\theta_s \sin\phi_s \, \mathbf{y} - \cos\theta_s \, \mathbf{z} \,. \tag{B.3}$$

Using the backscatter alignment coordinate system, we define two triplets of local coordinates to describe the transverse components of the incident and scattered fields. These coordinates are defined as

$$\mathbf{v}_i = -\cos\theta_i \cos\phi_i \, \mathbf{x} - \cos\theta_i \sin\phi_i \, \mathbf{y} + \sin\theta_i \, \mathbf{z} \,, \tag{B.4}$$

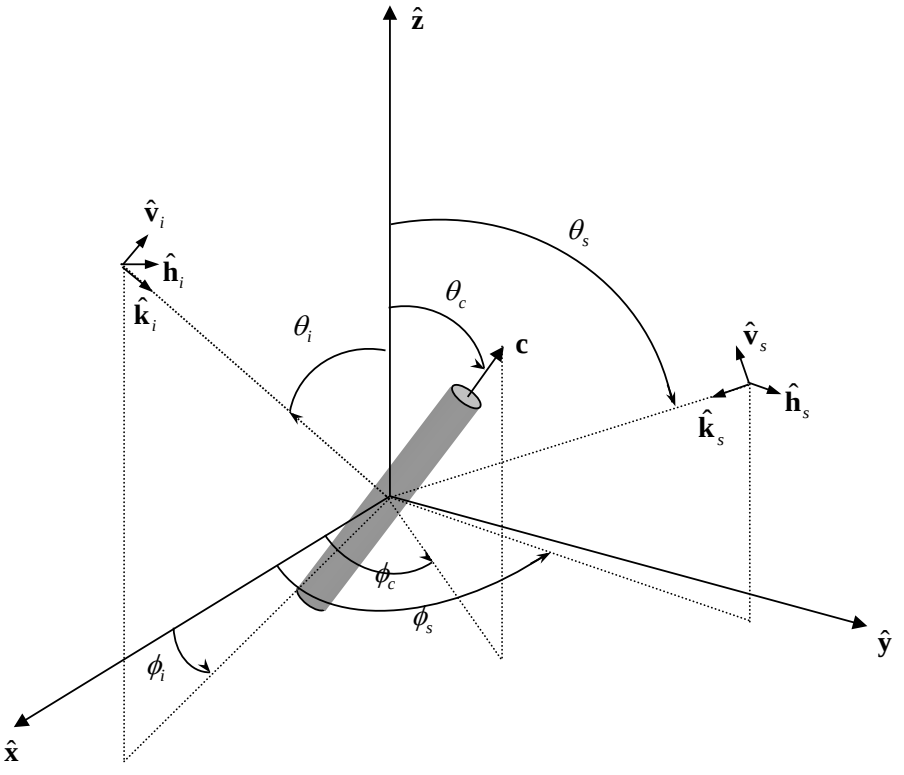


Fig. B-1. Global backscattering alignment coordinate system used in the calculations.

$$\mathbf{h}_i = \sin\phi_i\,\mathbf{x} - \cos\phi_i\,\mathbf{y}\,, \tag{B.5}$$

$$\mathbf{v}_s = -\cos\theta_s\cos\phi_s\,\mathbf{x} - \cos\theta_s\sin\phi_s\,\mathbf{y} + \sin\theta_s\,\mathbf{z}\,, \tag{B.6}$$

and

$$\mathbf{h}_s = \sin\phi_s\,\mathbf{x} - \cos\phi_s\,\mathbf{y}\,. \tag{B.7}$$

We shall assume that the incident wave is a plane electromagnetic wave. Because the cylinder is of finite size, the scattered wave is, in general, a spherical wave that propagates away from the cylinder. We shall assume that we know the expressions for the bistatic scattering matrix of a cylinder that is oriented vertically and denote this matrix by $\mathbf{S}\left(\theta_i,\phi_i,\theta_s,\phi_s\right)$. We shall use the same definition as proposed by van Zyl and Ulaby [1], which relates the incident and scattered waves as follows:

$$\mathbf{E}^{sc} = \left[\mathbf{S}\right]\mathbf{E}^{inc}\,\frac{e^{ikr}}{r}\,. \tag{B.8}$$

The term on the extreme right $\left(\frac{e^{ikr}}{kr}\right)$ explicitly shows the amplitude and phase of a spherical wave. Note that, by this definition, the scattering matrix is not dimensionless. Van de Hulst [2] defines the denominator of the spherical wave as *kr*; in Van de Hulst's definition the scattering matrix is dimensionless. In our case, the dimension of the scattering matrix elements is meters.

The elements of the bistatic scattering matrix for a vertically oriented cylinder are derived by Senior and Sarabandi [3] as

$$S_{hh}\left(\theta_i,\phi_i,\theta_s,\phi_s\right) = -\frac{1}{2}\frac{k_0 a l}{\sin\theta_i}\frac{\sin\left(k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2\right)}{k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2}D^h, \tag{B.9}$$

$$S_{hv}\left(\theta_i,\phi_i,\theta_s,\phi_s\right) = -\frac{1}{2}\frac{k_0 a l}{\sin\theta_i}\frac{\sin\left(k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2\right)}{k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2}i\bar{D}^h, \tag{B.10}$$

$$S_{vh}\left(\theta_i,\phi_i,\theta_s,\phi_s\right) = +\frac{1}{2}\frac{k_0 a l}{\sin\theta_i}\frac{\sin\left(k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2\right)}{k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2}i\bar{D}^e, \tag{B.11}$$

and

$$S_{vv}\left(\theta_i,\phi_i,\theta_s,\phi_s\right) = -\frac{1}{2}\frac{k_0 a l}{\sin\theta_i}\frac{\sin\left(k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2\right)}{k_0 l\left(\cos\theta_i+\cos\theta_s\right)/2}D^e. \tag{B.12}$$

In these equations, a is the radius of the cylinder, $k_0 = 2\pi/\lambda$ is the wave number of the incident wave, and l is the length of the cylinder. Also,

$$\begin{aligned} D^e = \sum_{m=-\infty}^{+\infty}(-1)^m \Bigg\{ & J_m'(x_0)J_m(y_0) - \frac{\sin\theta_i}{B}J_m(x_0)J_m'(y_0) \\ & + C_m^{TM}\left[H_m^{(1)\prime}(x_0)J_m(y_0) - \frac{\sin\theta_i}{B}H_m^{(1)}(x_0)J_m'(y_0)\right] \\ & + \frac{m\cos\theta_i}{x_0}\left(1 - \frac{\mathbf{k}_s\cdot\mathbf{c}(0,0)}{\cos\theta_i}\frac{x_0\sin\theta_i}{y_0 B}\right)\bar{C}_m H_m^{(1)}(x_0)J_m(y_0)\Bigg\} e^{im\tilde{\phi}} \end{aligned}, \tag{B.13}$$

$$D^h = \sum_{m=-\infty}^{+\infty} (-1)^m \left\{ J'_m(x_0) J_m(y_0) - \frac{\sin\theta_i}{B} J_m(x_0) J'_m(y_0) \right.$$
$$+ C_m^{TE} \left[H_m^{(1)\prime}(x_0) J_m(y_0) - \frac{\sin\theta_i}{B} H_m^{(1)}(x_0) J'_m(y_0) \right] \quad , \qquad \text{(B.14)}$$
$$\left. + \frac{m\cos\theta_i}{x_0} \left(1 - \frac{\mathbf{k}_s \cdot \mathbf{c}(0,0)}{\cos\theta_i} \frac{x_0 \sin\theta_i}{y_0 B} \right) \bar{C}_m H_m^{(1)}(x_0) J_m(y_0) \right\} e^{im\tilde{\phi}}$$

$$\bar{D}^e = \sum_{m=-\infty}^{+\infty} (-1)^m e^{im\tilde{\phi}} \left\{ \bar{C}_m \left[H_m^{(1)\prime}(x_0) J_m(y_0) - \frac{\sin\theta_i}{B} H_m^{(1)}(x_0) J'_m(y_0) \right] \right. \quad , \qquad \text{(B.15)}$$
$$\left. + \frac{m\cos\theta_i}{x_0} \left(1 - \frac{\mathbf{k}_s \cdot \mathbf{c}(0,0)}{\cos\theta_i} \frac{x_0 \sin\theta_i}{y_0 B} \right) \left[J_m(x_0) + C_m^{TE} H_m^{(1)}(x_0) \right] J_m(x_0) \right\}$$

$$\bar{D}^h = \sum_{m=-\infty}^{+\infty} (-1)^m e^{im\tilde{\phi}} \left\{ \bar{C}_m \left[H_m^{(1)\prime}(x_0) J_m(y_0) - \frac{\sin\theta_i}{B} H_m^{(1)}(x_0) J'_m(y_0) \right] \right. \quad , \qquad \text{(B.16)}$$
$$\left. + \frac{m\cos\theta_i}{x_0} \left(1 - \frac{\mathbf{k}_s \cdot \mathbf{c}(0,0)}{\cos\theta_i} \frac{x_0 \sin\theta_i}{y_0 B} \right) \left[J_m(x_0) + C_m^{TM} H_m^{(1)}(x_0) \right] J_m(x_0) \right\}$$

$$B = \frac{1}{2} \sqrt{\left(\sin\theta_i + \sin\theta_s \cos(\phi_s - \phi_i) \right)^2 + \sin^2\theta_s \cos^2(\phi_s - \phi_i)^2} \, , \qquad \text{(B.17)}$$

$$\cos\tilde{\phi} = \frac{1}{2B} \left(\sin\theta_s \cos(\phi_s - \phi_i) + \sin\theta_i \right); \quad \sin\tilde{\phi} = \frac{1}{2B} \sin\theta_s \sin(\phi_s - \phi_i) \, , \qquad \text{(B.18)}$$

$$y_0 = k_0 a B \, , \qquad \text{(B.19)}$$

$$x_0 = k_0 a \sin\theta_i \, , \qquad \text{(B.20)}$$

and

$$C_m^{TM} = -\frac{V_m P_m - q_m^2 J_m(x_0) H_m^{(1)}(x_0) J_m^2(x_1)}{P_m N_m - \left[q_m H_m^{(1)}(x_0) J_m(x_1) \right]^2} , \qquad \text{(B.21)}$$

$$C_m^{TE} = -\frac{M_m N_m - q_m^2 J_m(x_0) H_m^{(1)}(x_0) J_m^2(x_1)}{P_m N_m - \left[q_m H_m^{(1)}(x_0) J_m(x_1) \right]^2} , \qquad \text{(B.22)}$$

$$\bar{C}_m = i \frac{2}{\pi x_0 \sin\theta_i} \frac{q_m J_m^2(x_1)}{P_m N_m - \left[q_m H_m^{(1)}(x_0) J_m(x_1) \right]^2} , \qquad \text{(B.23)}$$

with

$$x_1 = k_0 a\sqrt{\varepsilon - \cos^2\theta_i}\ , \tag{B.24}$$

$$q_m = mk_0 a\cos\theta_i\left(\frac{1}{x_1^2} - \frac{1}{x_0^2}\right), \tag{B.25}$$

$$V_m = k_0 a\left\{\frac{\varepsilon}{x_1}J_m(x_0)J_m'(x_1) - \frac{1}{x_0}J_m'(x_0)J_m(x_1)\right\}, \tag{B.26}$$

$$P_m = k_0 a\left\{\frac{1}{x_1}H_m^{(1)}(x_0)J_m'(x_1) - \frac{1}{x_0}H_m^{(1)\prime}(x_0)J_m(x_1)\right\}, \tag{B.27}$$

$$N_m = k_0 a\left\{\frac{\varepsilon}{x_1}H_m^{(1)}(x_0)J_m'(x_1) - \frac{1}{x_0}H_m^{(1)\prime}(x_0)J_m(x_1)\right\}, \tag{B.28}$$

and

$$M_m = k_0 a\left\{\frac{1}{x_1}J_m(x_0)J_m'(x_1) - \frac{1}{x_0}J_m'(x_0)J_m(x_1)\right\}. \tag{B.29}$$

The cylinder complex dielectric constant is ε. Note the dimension of each element of the scattering matrix.

Equations (B.9) – (B.29) apply to the case of a vertically oriented cylinder. Returning to the case of a cylinder with arbitrary orientation, we will now define two local coordinate systems for the incident and scattered waves such that we can use these expressions to characterize the scattering in those two coordinate systems. We shall denote these coordinate systems by primed vectors. Starting with the incident electric field, we note that we can write this field as

$$\mathbf{E}^{inc} = E_h^{inc}\mathbf{h}_i + E_v^{inc}\mathbf{v}_i = E_{h'}^{inc}\mathbf{h}_i' + E_{v'}^{inc}\mathbf{v}_i', \tag{B.30}$$

from which it is easily shown that

$$\begin{pmatrix} E_{h'} \\ E_{v'} \end{pmatrix}^{inc} = \begin{pmatrix} \mathbf{h}_i\cdot\mathbf{h}_i' & \mathbf{v}_i\cdot\mathbf{h}_i' \\ \mathbf{h}_i\cdot\mathbf{v}_i' & \mathbf{v}_i\cdot\mathbf{v}_i' \end{pmatrix}\begin{pmatrix} E_h \\ E_v \end{pmatrix}^{inc}. \tag{B.31}$$

The bistatic scattering matrix links the incident and scattered waves in the local coordinate systems aligned with the cylinder axis, as follows:

$$\begin{pmatrix} E_{h'} \\ E_{v'} \end{pmatrix}^{sc} = \mathbf{S}\left(\theta_{ic}, \phi_{ic}, \theta_{sc}, \phi_{sc}\right) \begin{pmatrix} E_{h'} \\ E_{v'} \end{pmatrix}^{inc}, \tag{B.32}$$

where the subscripts ic and sc indicate that the angles are relative to the cylinder orientation, rather than the z-axis, as was the case in Eqs. (B.9) – (B.29).

The scattered wave can also be written as

$$\mathbf{E}^{sc} = E_s^{sc}\mathbf{h}_s + E_s^{sc}\mathbf{v}_s = E_{h'}^{sc}\mathbf{h}'_s + E_{v'}^{sc}\mathbf{v}'_s, \tag{B.33}$$

from which we can show that

$$\begin{pmatrix} E_h \\ E_v \end{pmatrix}^{sc} = \begin{pmatrix} \mathbf{h}_s \cdot \mathbf{h}'_s & \mathbf{h}_s \cdot \mathbf{v}'_s \\ \mathbf{v}_s \cdot \mathbf{h}'_s & \mathbf{v}_s \cdot \mathbf{v}'_s \end{pmatrix} \begin{pmatrix} E_{h'} \\ E_{v'} \end{pmatrix}^{sc}. \tag{B.34}$$

Combining Eqs. (B.31), (B.32), and (B.34), we find the bistatic scattering matrix of the cylinder as

$$\mathbf{S}\left(\theta_i, \phi_i, \theta_s, \phi_s, \theta_c, \phi_c\right) = \begin{pmatrix} \mathbf{h}_s \cdot \mathbf{h}'_s & \mathbf{h}_s \cdot \mathbf{v}'_s \\ \mathbf{v}_s \cdot \mathbf{h}'_s & \mathbf{v}_s \cdot \mathbf{v}'_s \end{pmatrix} \mathbf{S}\left(\theta_{ic}, \phi_i, \theta_{sc}, \phi_s\right) \begin{pmatrix} \mathbf{h}_i \cdot \mathbf{h}'_i & \mathbf{v}_i \cdot \mathbf{h}'_i \\ \mathbf{h}_i \cdot \mathbf{v}'_i & \mathbf{v}_i \cdot \mathbf{v}'_i \end{pmatrix}. \tag{B.35}$$

The local coordinate systems are defined as

$$\mathbf{h}'_i = \frac{\mathbf{k}_i \times \mathbf{c}\left(\theta_c, \phi_c\right)}{\left|\mathbf{k}_i \times \mathbf{c}\left(\theta_c, \phi_c\right)\right|} \tag{B.36}$$

and

$$\mathbf{v}'_i = \mathbf{h}'_i \times \mathbf{k}_i = \frac{1}{\left|\mathbf{k}_i \times \mathbf{c}\left(\theta_c, \phi_c\right)\right|} \left\{\mathbf{c}\left(\theta_c, \phi_c\right) - \mathbf{k}_i \left(\mathbf{k}_i \cdot \mathbf{c}\left(\theta_c, \phi_c\right)\right)\right\}. \tag{B.37}$$

The scattered wave coordinate axes are defined in the same way. We note that we can write

$$\left|\mathbf{k}_i \times \mathbf{c}\left(\theta_c, \phi_c\right)\right| = \sqrt{1 - \left(\mathbf{k}_i \cdot \mathbf{c}\left(\theta_c, \phi_c\right)\right)^2}. \tag{B.38}$$

Also, note that the unprimed coordinate systems can be written like Eq. (B.36) and Eq. (B.37) with $\mathbf{c}\left(\theta_c, \phi_c\right)$ replaced by $\mathbf{z}$. It then follows that

$$\mathbf{h}_i \cdot \mathbf{h}_i' = \frac{\mathbf{c}(\theta_c,\phi_c)\cdot\mathbf{z} - (\mathbf{k}_i\cdot\mathbf{z})(\mathbf{k}_i\cdot\mathbf{c}(\theta_c,\phi_c))}{\sqrt{1-(\mathbf{k}_i\cdot\mathbf{z})^2}\sqrt{1-(\mathbf{k}_i\cdot\mathbf{c}(\theta_c,\phi_c))^2}} \quad , \tag{B.39}$$

$$\mathbf{v}_i \cdot \mathbf{v}_i' = \mathbf{h}_i \cdot \mathbf{h}_i' , \tag{B.40}$$

and

$$\mathbf{h}_i \cdot \mathbf{v}_i' = -\mathbf{v}_i \cdot \mathbf{h}_i' = -\frac{\mathbf{k}_i\cdot(\mathbf{c}(\theta_c,\phi_c)\times\mathbf{z})}{\sqrt{1-(\mathbf{k}_i\cdot\mathbf{z})^2}\sqrt{1-(\mathbf{k}_i\cdot\mathbf{c}(\theta_c,\phi_c))^2}} \quad . \tag{B.41}$$

It can be shown that

$$\begin{aligned}&\mathbf{c}(\theta_c,\phi_c)\cdot\mathbf{z} - (\mathbf{k}_i\cdot\mathbf{z})(\mathbf{k}_i\cdot\mathbf{c}(\theta_c,\phi_c))\\ &= \sin\theta_i\left[\cos\theta_c\sin\theta_i - \sin\theta_c\cos\theta_i\cos(\phi_c-\phi_i)\right]\end{aligned} \tag{B.42}$$

and

$$\mathbf{k}_i\cdot(\mathbf{c}(\theta_c,\phi_c)\times\mathbf{z}) = -\sin\theta_i\sin\theta_c\sin(\phi_c-\phi_i) . \tag{B.43}$$

Using Eq. (B.43) and Eq. (B.42) in Eq. (B.41) and Eq. (B.39), respectively, we find that

$$\mathbf{h}_i \cdot \mathbf{h}_i' = \mathbf{v}_i \cdot \mathbf{v}_i' = \frac{1}{\sin\theta_{ic}}\left\{\cos\theta_c\sin\theta_i - \sin\theta_c\cos\theta_i\cos(\phi_c-\phi_i)\right\}, \tag{B.44}$$

and

$$\mathbf{h}_i \cdot \mathbf{v}_i' = -\mathbf{v}_i \cdot \mathbf{h}_i' = -\frac{\sin\theta_c\sin(\phi_c-\phi_i)}{\sin\theta_{ic}} . \tag{B.45}$$

For the scattered wave, we find that

$$\mathbf{h}_s \cdot \mathbf{h}_s' = \mathbf{v}_s \cdot \mathbf{v}_s' = \frac{\mathbf{c}(\theta_c,\phi_c)\cdot\mathbf{z} - (\mathbf{k}_s\cdot\mathbf{z})(\mathbf{k}_s\cdot\mathbf{c}(\theta_c,\phi_c))}{\sqrt{1-(\mathbf{k}_s\cdot\mathbf{z})^2}\sqrt{1-(\mathbf{k}_s\cdot\mathbf{c}(\theta_c,\phi_c))^2}} \tag{B.46}$$

and

$$\mathbf{h}_s \cdot \mathbf{v}_s' = -\mathbf{v}_s \cdot \mathbf{h}_s' = -\frac{\mathbf{k}_s\cdot(\mathbf{c}(\theta_c,\phi_c)\times\mathbf{z})}{\sqrt{1-(\mathbf{k}_s\cdot\mathbf{z})^2}\sqrt{1-(\mathbf{k}_s\cdot\mathbf{c}(\theta_c,\phi_c))^2}} \quad , \tag{B.47}$$

which can be written as

$$\mathbf{h}_s \cdot \mathbf{h}'_s = \mathbf{v}_s \cdot \mathbf{v}'_s = \frac{1}{\sin\theta_{sc}} \left\{ \cos\theta_c \sin\theta_s - \sin\theta_c \cos\theta_s \cos\left(\phi_c - \phi_s\right) \right\} \tag{B.48}$$

and

$$\mathbf{h}_s \cdot \mathbf{v}'_s = -\mathbf{v}_s \cdot \mathbf{h}'_s = -\frac{\sin\theta_c \sin\left(\phi_c - \phi_s\right)}{\sin\theta_{sc}}. \tag{B.49}$$

The angles θ_{ic} and θ_{sc} are defined by

$$\cos\theta_{ic} = -\left(\mathbf{k}_i \cdot \mathbf{c}\right) = \cos\theta_c \cos\theta_i + \sin\theta_c \sin\theta_i \cos\left(\phi_c - \phi_i\right) \tag{B.50}$$

and

$$\cos\theta_{sc} = -\left(\mathbf{k}_s \cdot \mathbf{c}\right) = \cos\theta_c \cos\theta_s + \sin\theta_c \sin\theta_s \cos\left(\phi_c - \phi_s\right). \tag{B.51}$$

It is also useful to look at simpler expressions reported in the literature. Barrick [4] used the expressions for the fields scattered by an infinitely long cylinder and then accounted for the finite length l of the cylinder by multiplying by a factor

$$l\sqrt{\frac{k_0}{\pi}} \sin\theta_{sc} \frac{\sin\left[k_0 l\left(\cos\theta_{ic} + \cos\theta_{sc}\right)/2\right]}{k_0 l\left(\cos\theta_{ic} + \cos\theta_{sc}\right)/2}. \tag{B.52}$$

Under this assumption, and evaluating Eqs. (B.9) – (B.12) on the cone $B = \sin\theta_i$, we find

$$S_{hh}\left(\theta_i, \phi_i, \theta_s, \phi_s\right) = -\frac{il\sin\theta_s}{\pi\sin\theta_i} \frac{\sin V}{V} \sum_{m=-\infty}^{+\infty} \left\{ \left(-1\right)^m C_m^{TM} e^{im\left(\phi_s - \phi_i\right)} \right\}, \tag{B.53}$$

$$S_{hv}\left(\theta_i, \phi_i, \theta_s, \phi_s\right) = -\frac{il\sin\theta_s}{\pi\sin\theta_i} \frac{\sin V}{V} \sum_{m=-\infty}^{+\infty} \left\{ \left(-1\right)^m \bar{C}_m e^{im\left(\phi_s - \phi_i\right)} \right\}, \tag{B.54}$$

$$S_{vh}\left(\theta_i, \phi_i, \theta_s, \phi_s\right) = +\frac{il\sin\theta_s}{\pi\sin\theta_i} \frac{\sin V}{V} \sum_{m=-\infty}^{+\infty} \left\{ \left(-1\right)^m \bar{C}_m e^{im\left(\phi_s - \phi_i\right)} \right\}, \tag{B.55}$$

and

$$S_{vv}\left(\theta_i,\phi_i,\theta_s,\phi_s\right)=-\frac{il\sin\theta_s}{\pi\sin\theta_i}\frac{\sin V}{V}\sum_{m=-\infty}^{+\infty}\left\{\left(-1\right)^m C_m^{TE}e^{im\left(\phi_s-\phi_i\right)}\right\}, \qquad \text{(B.56)}$$

where

$$V=\frac{1}{2}k_0 l\left(\cos\theta_i+\cos\theta_s\right). \qquad \text{(B.57)}$$

Note that the series coefficients given in Eqs. (B.21) – (B.29) have the following symmetry relations

$$C_m^{TM}=C_{-m}^{TM};\quad C_m^{TE}=C_{-m}^{TE};\quad \bar{C}_m=-\bar{C}_{-m};\quad \bar{C}_0=0. \qquad \text{(B.58)}$$

This means that the far-zone scattered field in the plane of incidence is not depolarized. This, however, is not the case of other azimuth angles. Also, in the forward scattering direction, required for the calculation of the extinction coefficient, these expressions are

$$S_{hh}\left(\theta_i,\phi_i,\pi-\theta_i,\phi_i+\pi\right)=-\frac{il\sin\theta_s}{\pi\sin\theta_i}\frac{\sin V}{V}\sum_{m=0}^{+\infty}C_m^{TM}, \qquad \text{(B.59)}$$

$$S_{hv}\left(\theta_i,\phi_i,\pi-\theta_i,\phi_i+\pi\right)=S_{vh}\left(\theta_i,\phi_i,\pi-\theta_s,\phi_i+\pi\right)=0, \qquad \text{(B.60)}$$

and

$$S_{vv}\left(\theta_i,\phi_i,\pi-\theta_i,\phi_i+\pi\right)=-\frac{il\sin\theta_s}{\pi\sin\theta_i}\frac{\sin V}{V}\sum_{m=0}^{+\infty}C_m^{TE}. \qquad \text{(B.61)}$$

Senior and Sarabandi [3] built on the work of Barrick [4] by integrating the current distribution over a cylinder of finite length. In contrast, Barrick [4] added a $\sin x/x$ term to the scattered field of an infinitely long cylinder to account for the finite length. As such, Senior and Sarabandi's equations are possibly more accurate than those of Barrick, but require significantly more calculations.

We shall use these expressions to calculate the composite scattering from a layer of vegetation. If we consider such a layer, there are three basic calculations for the layer as a whole that we need to perform. These include the backscatter from the layer, the bistatic forward scattering that would interact with the ground surface, and the attenuation through the layer. The latter is needed to calculate the attenuated backscatter from the underlying soil surface. The expressions listed so far can be used directly to calculate the first two

components. To calculate the attenuation through the layer, we shall make use of the optical theorem that states that the extinction cross section of a single particle is related to the forward scattering field through

$$\sigma_p^e = \frac{2\pi}{k_0}\mathrm{Im}\left[S_{pp}\left(\theta_{ic},\phi_i,\pi-\theta_{ic},\phi_i+\pi,\theta_c,\phi_c\right)\right], \tag{B.62}$$

where p denotes the polarization of the wave. The dimension of this is meters squared, since the scattering matrix has dimension meters. The total extinction coefficient of a medium containing a random distribution of N cylinders per unit volume is obtained by performing an ensemble average over the particles, as follows:

$$\kappa_p^e = N\left\langle \sigma_p^e \right\rangle. \tag{B.63}$$

The dimension of this quantity is m^{-1}. Using this definition, the strength of the incident wave after propagating through a layer of thickness d at an angle θ_i with respect to the vertical direction, is given by

$$\begin{pmatrix} E_h \\ E_v \end{pmatrix}^{tr} = \begin{pmatrix} e^{-\kappa_h^e d/\cos\theta_i} & 0 \\ 0 & e^{-\kappa_v^e d/\cos\theta_i} \end{pmatrix} \begin{pmatrix} E_h \\ E_v \end{pmatrix}^{inc}. \tag{B.64}$$

To find the total field propagating in this direction, we need to add the bistatic scattered field in this direction to Eq. (B.64).

With these expressions for the scattering from an arbitrarily oriented cylinder, one can define models to describe the scattering from vegetated terrain.

References

[1] J. J. van Zyl and F. T. Ulaby, "Scattering Matrix Representations (FSA – BSA) for Simple Target ," Chapter 2 in *Radar Polarimetry for Geoscience Applications*, F.T. Ulaby and C. Elachi (eds.), Artech House, Norwood, Massachusetts, 1990.

[2] H. C. van de Hulst, *Light Scattering by Small Particles*, Dover, New York, New York, 1981.

[3] T. B. A. Senior and K. Sarabandi, "Scattering Models for Point Targets," Chapter 3 in *Radar Polarimetry for Geoscience Applications*, F. T. Ulaby and C. Elachi (eds.), Artech House, Norwood, Massachusetts, 1990.

[4] D. E. Barrick, "Cylinders," Chapter 4 in *Radar Cross Section Handbook, Volume 1*, G. T. Ruck (ed.), Plenum Press, New York, 1970.

Appendix C
Nomenclature

C.1 Acronyms and Abbreviations

AIRSAR	Airborne Synthetic Aperture Radar
ALOS PALSAR	Advanced Land Observing Satellite Phased Array L-band Synthetic Aperture Radar
ASAR	Advanced Synthetic Aperture Radar (on (European Space Agency Environmental Satellite [Envisitat])
dB	power in decibels
deg	degree
DEM	digital elevation model
ed.	editor
Envisitat	European Space Agency Environmental Satellite
Eq.	Equation
ERS	European Remote Sensing Satellite
Fig.	Figure
GHz	gigahertz

HH	horizontal-horizontal
HV	horizontal-vertical
IEEE	Institute for Electrical and Electronics Engineers
IEM	Integral Equation Method
IGARSS	International Geoscience and Remote Sensing Symposium
JPL	Jet Propulsion Laboratory
NASA	National Aeronautics and Space Administration
NNED	non-negative Eigenvalue decomposition
PFR	pulse repetition frequency
RASAM	Radiometer-Scatterometer
rms	root mean square
RVI	radar thin vegetation index
SAR	synthetic aperture radar
SIR-B	Shuttle Imaging Radar-B (-C)
SIR-C/X	Synthetic Imaging Radar C-band and X-band
SLAR	side-looking aperture radars
SMAP	Soil Moisture Active Passive
SNR	signal-to-noise ratio
SPIE	Society of Photographic Instrumentation Engineers
SPM	small perturbation method
VH	vertical-horizontal

vol.	volume
VV	vertical-vertical
w.r.t	with respect to

C.2 Commonly Used Symbols

R : Range

c : Speed of light

θ or θ_i : Incidence angle

B : Radar bandwidth

τ_p : Physical pulse length

λ : Wavelength

f_0 : Center frequency

σ_N : Noise equivalent backscatter cross section

W : Antenna width

L : Antenna length

P_r : Received power by radar

P_t : Transmit power by radar

G_r : Receive antenna gain

G_t : Transmit antenna gain

σ_0 : Normalized backscattering cross section

v : Velocity of the radar platform

PRF : Pulse repetition frequency

$\hat{\mathbf{h}}$: Horizontal basis vector

$\hat{\mathbf{v}}$: Vertical basis vector

$S_0,\ S_1,\ S_2,\ S_3$: Stokes parameters

$[\mathbf{S}]$: 2x2 complex scattering matrix

$\mathbf{E}^{sc}$: Scattered electric field vector

$\mathbf{E}^{tr}$: Transmitted electric field vector

$\mathbf{p}^{tr}$: Normalized transmit antenna polarization vector

$\mathbf{p}^{rec}$: Normalized receive antenna polarization vector

V : Voltage measured by radar system

$\mathbf{A}$: Antenna polarization vector

$[\mathbf{C}]$: Scatter covariance matrix

$\mathbf{T}$: Scatter vector

$S_{hh},\ S_{vv},\ S_{hv},\ S_{vh}$: Elements of 2x2 complex scattering matrix

$[\mathbf{M}]$: Stokes scattering operator

$\lambda_1,\ \lambda_2,\ \lambda_3$: Eigenvalues of covariance matrix when $S_{hv} = S_{vh}$

$\hat{\mathbf{e}}_i$: Eigenvectors of covariance matrix

$\hat{\mathbf{p}}_i$: Eigenvectors of Pauli covariance matrix when $S_{hv} = S_{vh}$

H_T : Entropy

$\overline{\alpha}$: Average alpha angle for the alpha-entropy algorithm

RVI : Radar vegetation index

R_{co} : Co-polarization ratio (S_{hh} / S_{vv})

R_x : Cross-polarization ratio (S_{hv} / S_{vv})

$\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3$: Eigenvectors of covariance matrix when $S_{hv} = S_{vh}$

$[\mathbf{T}_R]$: Tilt transformation matrix from the local to the global coordinates

$[\mathbf{R}]$: Receive system distortion matrix for calibration

$[\mathbf{T}]$: Transmit system distortion matrix for calibration

$[\mathbf{D}]$: System distortion matrix for calibration

m_v : Volumetric soil moisture

ε : Dielectric constant

σ_{hh} : HH normalized backscattering cross section

σ_{hv} : HV normalized backscattering cross section

σ_{vh} : VH normalized backscattering cross section

σ_{vv} : VV normalized backscattering cross section