UNIVERSITY OF TROMSØ UIT

FACULTY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF PHYSICS AND TECHNOLOGY

Non-Gaussian Statistical Analysis of Polarimetric Synthetic Aperture Radar Images

Anthony Paul Doulgeris

A dissertation for the degree of
Philosophiae Doctor

February 2011

Abstract

This thesis describes general methods to analyse polarimetric synthetic aperture radar images. The primary application is for unsupervised image segmentation, and fast, practical methods are sought.

The fundamental assumptions and statistical modelling are derived from the physics of electromagnetic scattering from distributed targets. The physical basis directly leads to the image phenomenon called speckle, which is shown to be potentially non-Gaussian and several statistical distributions are investigated. Speckle non-Gaussianity and polarimetry both hold pertinent information about the target medium and methods that utilise both attributes are developed. Two distinct approaches are proposed: a local feature extraction method; and a model-based clustering algorithm.

The local feature extraction approach creates a new six-dimensional description of the image that may be used for subsequent image analysis or for physical parameter extraction (inversion). It essentially extends standard polarimetric features with the addition of a non-Gaussianity measure for texture. Importantly, the non-Gaussianity measure is model independent and therefore does not unduly constrain the analysis. Unsupervised image segmentation is demonstrated with good results.

The model-based approach describes a Bayesian clustering algorithm for the K-Wishart model, with fast moment based parameter estimation, and incorporates both non-Gaussianity and polarimetry. The initial implementation requires the number of classes, and initial segmentation, and the effective number of looks (an important model parameter) to be given in advance. When compared to the more common Wishart model (i.e. Gaussian-based), the K-Wishart gives similar results for Gaussian image regions, but performs better for non-Gaussian regions.

Further development of the model-based method resulted in a novel technique to automatically determine the number of distinct classes supported by the data, given the model choice and a statistical confidence level. All relevant parameters are subsequently estimated within the algorithm, using the most up-to-date methods of matrix log-cumulants, and no special initialisation is required. These are significant advances from most existing methods, where key parameters are given in advance and the number of classes are usually determined after many full clustering results are obtained.

All methods are quite general and can be applied to all coherent imaging systems that exhibit product model based statistics. The methods are demonstrated on several real radar images, from both airborne and space-borne sensors, at different operating frequencies, and with different numbers of polarimetric channels.

Acknowledgements

I would like to acknowledge the University of Tromsø and the department of Physics and Technology for the opportunity and scholarship to study a topic that I enjoy greatly.

My sincerest gratitude goes to my supervisor, Torbjørn Eltoft, for years of patience and encouragement, and for having an open collaborative style. I appreciate having the freedom to explore my own research directions and yet have your broad knowledge and guidance to call upon when things were hard. I look forward to more great collaboration in the future.

Special thanks to Stian Anfinsen, for sharing both an office and ideas at work, and also for sharing much more as a friend. My time and development as a student would not have been so easy nor fun without you. Thanks also to all the new members in our ever expanding group, Camilla, Vahid, Ane, Mari-Ann, Tao Ding and Faozi, for making it an exciting place to work and for the pleasure, and pressure, of wanting to use my algorithms for further studies.

Thanks also to Kirsty Langley and Jack Kohler from the Norwegian Polar Institute for fruitful collaboration and knowledge about Arctic glaciers, and a wonderful field trip to Kongsvegen glacier on Svalbard. Thanks to Kjell Arild Høgda, Inge Lauknes, Tom Rune Lauknes, Rune Storvold, Harald Johnsen and Yngvar Larsen of Norut for data and support over the years, and I hope that some of my research results may now be useful to Norut in return.

Warmest thanks to Hajo Eicken, Matthew Druckenmiller, and everyone in the sea ice and remote sensing groups at the Geophysical Institute at the University of Alaska, Fairbanks. Thanks for enthusiastically welcoming me into your groups and teaching me all about sea ice, Alaska and the more human side of scientific endeavours. Special thanks to Franz Meyer and Don Atwood from the Alaskan Satellite Facility for both sharing data and ideas and for fixing the data quality problem that arose. I look forward to continuing all these friendships and collaborations, and finally working on the sea ice study in the near future.

Finally, I would like to thank my parents for early scientific encouragement and educational support for all those early, influential, years, and Corine Davids for encouragement, patience, opportunities, distractions and life in general in the recent years.

I dedicate this thesis to you all.

Contents

Chapter 1

Introduction

This chapter is intended to give a brief overview of the whole thesis. It introduces the motivation for the research, summarises and discusses the three research articles that make up the main content, and lists associated conference works.

The layout for the remainder of the thesis is to go into a more detailed discussion of the motivation and background theory in Chapter 2, to include the three articles directly as Chapters 3 to 5, to discuss the research achievements in the context of the broader research field in Chapter 6, and to finish with some concluding remarks in Chapter 7.

1.1 Motivation Summary

Synthetic aperture radar (SAR) systems are active devices that transmit and receive in the microwave region of the electromagnetic spectrum and are thereby independent of the sun's illumination and insensitive to cloud cover. This makes them highly suitable and reliable for environmental monitoring of the Earth's surface and several new systems are now in full operation. Being active coherent imaging devices, however, means that they exhibit the interference effect known as *speckle*. Understanding the physical basis and significance of speckle leads to improved methods that use the statistical nature of speckle to gain additional knowledge from SAR images. A similar motivation holds for investigating polarimetric SAR (PolSAR) images in that the additional information contained in multiple polarimetric channels can be of benefit for image interpretation and classification.

The works presented here are restricted to studying incoherent scattering from distributed natural terrain for which the scalar product model is deemed suitable. Extensions from the scalar to vector (or matrix) texture in the product models may be addressed in the future, and likewise for extending to partially coherent scattering, i.e. including dominant scatterers, including additive system noise, or including contextual information.

The physical foundations of radar backscatter theory, the random walk interpretation, speckle and causes of radar texture are discussed with sufficient detail in Chapter

2 to understand the significance of accounting for non-Gaussian statistics in PolSAR modelling. Even so, the mathematically simpler approach of assuming multivariate Gaussian statistics is still in common practise and is briefly discussed as background theory.

1.2 Publication Summary

The main body of this thesis is presented as three journal publications which are included as Chapters 3, 4 and 5. The following summaries for each article describe the key concepts and highlight the original contributions of the authors. The papers are not presented in publication order, but reflect the progress of the research itself. Paper 1 introduces the fundamentals of non-Gaussian modelling with the Scale Mixture of Gaussian scheme, Paper 2 introduces the K-Wishart clustering algorithm for multilook covariance matrix data, and Paper 3 extends this with goodness-of-fit testing to automatically select the appropriate number of clusters, plus other improvements.

Paper 1

A. P. Doulgeris and T. Eltoft, "**Scale Mixture of Gaussian Modelling of Polarimetric SAR Data**," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 874592, p. 12, 2010.

This paper introduces the techniques used for statistical modelling of single look complex (SLC) polarimetric synthetic aperture radar data. Non-Gaussian statistics of real PolSAR data are confirmed for many terrain types for which the Scale Mixture of Gaussian (SMoG) scheme is shown to be a suitable statistical model. Four different SMoG models are then compared for several real PolSAR images and their goodness-of-fit to the data is evaluated. Gaussian statistics are clearly shown to be insufficient for a significant proportion of all images and the flexible models, with both scale and shape parameters, are necessary to capture the full range of real PolSAR data distributions. The multivariate normal inverse Gaussian distribution (MNIG) appeared as the best choice, of the four, except for the sea ice image, where the multivariate K-distribution (MK) was slightly better. Both the MNIG and MK models are flexible models with two parameters, describing both a mean scale and non-Gaussian shape, and were approximately equally good fits to the data overall.

General PolSAR features were then extracted from the images using the modelling and some were logarithmically transformed to aid subsequent data processing. Polarimetric theory suggests five parameters from the complex covariance matrix and certain ratios and angles were used. In addition, a non-Gaussianity measure is added to the five to produce a new six dimensional feature space to describe PolSAR images. The preferred non-Gaussianity measure was a normalised moment of the data (because the moment would be independent of any particular SMoG model), rather than a particular

model's shape parameter (since they are derived directly from the moments anyway). The six-dimensional feature space can then be used with simple clustering algorithms to segment the images and good results were obtained for a real PolSAR image.

The most important aspects of this work are the addition of non-Gaussianity to the standard five polarimetric features and that this can be done in a model independent way. The example given is for image segmentation, but the local statistical modelling method is also well suited to physical property extraction and interpretation.

Paper 2

A. Doulgeris, S. Anfinsen, and T. Eltoft,"**Classification with a non-Gaussian Model for PolSAR Data**," *IEEE Trans. Geoscience and Remote Sensing*, vol. 46, pp. 2999–3009, Oct. 2008.

This paper details how to take the non-Gaussian, scale mixture of Gaussian modelling from the single look complex (SLC) vector data to the multilook complex (MLC) covariance matrix data and results in a generalisation of the complex Wishart distribution. The general case is described as an integral equation and the specific case of the K-distribution is subsequently derived and named the K-Wishart model. The K-Wishart classifier is then described as an iterative unsupervised clustering algorithm, with fast moment based parameter update expressions. The transformation factor from single-look to multilook non-Gaussianity parameter is also derived.

The algorithm requires that some parameters are given in advance, as is common with many clustering algorithms. The number of classes to cluster had to be predetermined and was manually chosen after performing segmentations for several different number of classes. The algorithm needs an initial segmentation as a starting point and a k-means clustering of the marginal intensities was found to converge faster than a purely random initialisation. In addition, the equivalent number of looks (ENL) is a key parameter in all generalised Wishart models and had to be pre-determined by analysing homogeneous regions within the image.

The clustering is demonstrated on simulated non-Gaussian data and compared to the standard Wishart classifier with improved classification accuracy. The K-Wishart and standard Wishart classifiers are then compared in detail on a real PolSAR scene, at two different levels of multilook averaging. There were evident differences which were discussed in terms of polarimetry, non-Gaussian intensity variations and the number of clusters. The non-Gaussian K-Wishart model includes the standard Wishart model as an asymptotic case and the model segmentations showed similar results for Gaussian regions, but differed for the most non-Gaussian regions, as was expected.

Paper 3

A. P. Doulgeris, S. N. Anfinsen, and T. Eltoft, "**Automated non-Gaussian Clustering of Polarimetric Synthetic Aperture Radar Images**," *IEEE Trans. Geoscience and Remote Sensing*, submitted 2010.

The most recent paper describes improvements to the K-Wishart clustering algorithm that addresses the optimum number of clusters, removes any initialisation concerns and includes a texture corrected ENL estimation internally. General improvements were made by utilising methods of matrix log-cumulants for all parameter estimation as well as improving the numerical evaluation of the K-Wishart probability density function. The optimum number of classes is determined by adding a split and merge stage to the standard algorithm with log-cumulant based goodness-of-fit tests.

At regular intervals, the model fit of each cluster is tested and bad fits, usually multimodal mixed clusters, are split into two clusters. Additionally, each pair of good clusters is tested for whether the combination has a good single model fit and such cases are merged. Thus the goodness-of-fit testing dynamically adjusts the number of clusters until a statistically good fit to all the data clusters is obtained. The sensitivity of the goodness-of-fit tests depends upon the chosen confidence level, e.g. 95%, and the number of data samples given. This sample size dependency allows for deliberate subsampling to obtain faster, less detailed clustering to the major classes only. Both coarse and fine detailed results were shown.

The algorithm for automatic clustering can be readily applied to different models and only requires that the probability density function and the first few log-cumulant expressions are known. Real world examples for the K-Wishart, the standard Wishart and the Relaxed-Wishart model were shown. Highly non-Gaussian regions of the images resulted in many more classes for the more restrictive Relaxed-Wishart and standard Wishart models, as was expected. K-Wishart clustering results were shown for several PolSAR data-sets, depicting quite different terrain types, with visually good, fully automatic results.

The individual class goodness-of-fit testing is a novel approach that not only finds an optimum number of classes but also guarantees that each class is a good fit to the model. The resulting clustering depicts the number of classes that are statistically supported by the data, given the data size, the confidence level and the choice of model.

1.3 Research Discussion Summary

Confirmation of non-Gaussianity in real images was important for two reasons. Firstly, to show the importance of including non-Gaussianity during PolSAR analysis, and secondly, to validate that the scalar product model was suitable for these images. This was important because other researchers have found examples where the scalar product model breaks down, as discussed in Section 2.2.

A conclusion from the Paper 1, that a single flexible model (either the K-distribution or the normal inverse Gaussian) is suitable for all image analysis, is discussed, in the light of more recent observations, and found to be lacking. However, at the time it seemed suitable and the mathematical advantage of working with a single model justified the subsequent research.

The computational speed of method of moments is noted as justification for their use, even though they may not be optimal in the sense of the total log-likelihood function. More recently, an observed clustering problem has been attributed to the use of a single moment for parameter estimation, and using multiple moments is discussed as a suggested solution.

A detailed discussion of the extended polarimetric features, proposed in Paper 1, includes noting the importance of being model independent, that the level of detail in the non-Gaussianity feature histogram signifies the value of the non-Gaussianity parameter, and that the good results from a simple clustering technique demonstrates the potential of the new feature space. It is noted, however, that this non-Gaussian feature extraction method may be more significant for physical parameter retrieval, rather than clustering, because the local values are maintained at each pixel.

An alternative approach to clustering is discussed in Paper 2, which describes the K-Wishart clustering algorithm for unsupervised clustering of PolSAR data. The importance of demonstrating unsupervised clustering of non-Gaussian models is discussed because many new articles continue to use the Gaussian-based Wishart classifier. This paper contributes to the field by showing that unsupervised non-Gaussian clustering can be achieved in practise, by discussing the differences when compared to the more common Wishart classifier, and by pointing out the most significant difficulties and how they may be addressed.

These suggested improvements are taken up in Paper 3 and the discussion turns to each in turn. The method of log-cumulants is used for parameter estimation because its estimates are known to have reduced bias and variance, even though solutions must be found numerically. The effective number of looks is automatically determined and corrected for textural bias. Goodness-of-fit tests are included to solve the problems of poor cluster fits, algorithm initialisation and, most importantly, determining the optimum number of clusters. The discussion further explains the usefulness of sub-sampling to adjust the sensitivity of the goodness-of-fit testing and therefore the level of detail of the final segmentation.

1.4 Other Publications and Presentations

As first author:

1. A. P. Doulgeris and T. Eltoft, "Automated Non-Gaussian Clustering of Polarimetric SAR.," in *8th European Conference on Synthetic Aperture Radar (EUSAR2010)*, (Aachen, Germany), June 7-10, 2010.

2. A. P. Doulgeris, "Finding the appropriate number of classes in mixture modelling", presented in *NOBIM2010*, (Tromsø, Norway), June 21-22, 2010. No proceedings published.

3. A. P. Doulgeris and T. Eltoft, "General Statistical Methods for SAR Analysis of the Cryosphere", presented in *Arctic Frontiers*, (Tromsø, Norway), January 25-29, 2010. No proceedings published.

4. A. Doulgeris, S. Anfinsen, Y. Larsen, K. Langley, and T. Eltoft, "Evaluation of Polarimetric Configurations for Glacier Classification," in *International POLinSAR Workshop (POLinSAR2009)*, (Frascati, Italy), January 26-30, 2009.

5. A. P. Doulgeris, K. Langley, and T. Eltoft, "Analysis and Classification of High Arctic Glaciers with ASAR Data," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS2008)*, (Boston, Massachusetts, USA), July 6-11, 2008.

6. A. Doulgeris, S. N. Anfinsen, and T. Eltoft, "Analysis of non-Gaussian PolSAR Data," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS2007)*, (Barcelona, Spain), July 23-27, 2007.

7. A. Doulgeris and T. Eltoft, "Scale Mixture of Gaussians Modelling of Polarimetric SAR Data," in *International POLinSAR Workshop (POLinSAR2007)*, (Frascati, Italy), January 22-26, 2007.

8. A. Doulgeris, R. Hall, T. Eltoft, and S. Tronstad, "Statistical Modelling and Classification of ASAR Alternating Polarisation Sea Ice Data," in *ENVISAT Symposium*, (Montreux, Switzerland), April 23-27, 2007.

9. A. P. Doulgeris and T. Eltoft, "Scale Mixture of Gaussians Modelling of Polarimetric SAR Data," in *Proceedings of NORSIG*, (Reykjavik, Iceland), June 7-9, 2006.

As coauthor:

1. S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Goodness-of-Fit Tests for Multilook Polarimetric Radar Data Based on the Mellin Transform," *IEEE Trans. Geoscience and Remote Sensing*, **49**(8): 18 pp., in press, August 2011.

2. F. J. Meyer and A. P. Doulgeris, "Interference Signatures in ALOS PALSAR Imagery in the American Arctic," in *CEOS SAR Cal/Val Workshop*, (Zürich, Switzerland), August 25-27, 2010. No proceedings published.

3. V. Akbari, A. P. Doulgeris, and T. Eltoft, "Non-Gaussian Clustering of SAR Images for Glacier Change Detection.," in *ESA Living Planet Symposium 2010*, (Bergen, Norway), 28 June - 2 July, 2010.

4. S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the Equivalent Number of Looks in Polarimetric Synthetic Aperture Radar Imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, pp. 3795–3809, Nov. 2009.

5. T. Eltoft, A. P. Doulgeris and S. N. Anfinsen, "Model-based Statistical Analysis of PolSAR Data.," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2009)*, (Cape Town, South Africa), July 12-17, 2009.

6. S. Anfinsen, T. Eltoft, and A. Doulgeris, "A Relaxed Wishart Model for Polarimetric SAR Data," in *Proc. 4th Int. Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR2009)*, (Frascati, Italy), January 26-30, 2009, ESA SP-668, p. 8 pp., April 2009.

7. K. Langley, A. Doulgeris, and T. Eltoft, "Analysis and classification of glacier facies with SAR and GPR data," Presented at *International symposium on Radioglaciology and its applications*, (Madrid, Spain), June 9-13, 2008. No proceedings published.

8. C. Davids and A. P. Doulgeris, "Unsupervised Change Detection of Multitemporal Landsat Imagery to Identify Changes in Land Cover Following the Chernobyl Accident.", in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS2007)*, (Barcelona, Spain), July 23-27, 2007.

Chapter 2

Research Motivation and Background

This chapter covers the early motivation, the scope and limitations of the presented works, and then introduces the fundamental physics of scattering theory and polarimetry to set the basic framework.

2.1 Early Motivation

There is no doubt of the importance of synthetic aperture radar (SAR) systems for monitoring the Earth's surface. Satellite-borne systems achieve wide area coverage on the ground and regular temporal coverage of the globe, while the aperture synthesis technique, of coherently combining very many individual signals as the satellite moves over the target area, maintains high resolutions of the order of metres on the ground. Their all weather and day/night operating capability makes for reliable imaging independent of cloud cover or Sun angle illumination. SAR sensors achieve these goals because they are active systems, providing their own illumination source, and they operate in the microwave region of the electromagnetic spectrum at wavelengths that propagate virtually unimpeded through both the atmosphere and cloud cover. The microwave signal also penetrates some distance into the surface coverage, depending upon the operating wavelength and the surface material's dielectric properties, and can therefore measure properties of the near surface volume that may better characterise the target media. These features have made SAR imaging a practical choice for environmental monitoring and several new systems are now in full operation, e.g. the Japanese ALOS/PALSAR and the Canadian Radarsat-2, with huge amounts of image data now, and continually, available. However, while optical imaging systems are familiar to us because of their similarity with human vision, radar interpretation is not a straight forward task and no dominant analysis technique exists. There was clearly scope for research in this field to try to take SAR image analysis mainstream.

A major concern for SAR image interpretation is known as *speckle* and appears as a grainy, noise-like variation of intensity across the whole image. While noise in optical images tends to have a constant and relatively low level and is therefore less significant

for brighter targets, the speckle 'noise' in radar images is of the same order as, and proportional to, the local intensity. This means that no matter how bright the target area is in the image, it will contain pixel-to-pixel variation from near zero to several times the targets mean brightness level. Speckle is not truly a noise because it originates from an interference mechanism found in all coherent imaging systems; such as radar, laser and ultrasound imaging. However, it is often considered as unwanted 'noise' and great effort is made to reduce or eliminate it from the scene. The view of this research is that speckle may hold useful information about the target media, since it originates from the interaction, or scattering, of the illuminated microwave signal with the imaged surface, and understanding and modelling the speckle variation could achieve potential benefits.

The characteristics of speckle are generally modelled as a Gaussian interference term multiplied by the terrain mean backscatter value, as is well described in the general literature such as [Oliver and Quegan, 2004]. This interference can be expected to follow a Gaussian distribution under certain general assumptions regarding scale and resolutions. Thus, the *product model* for the statistical variation of speckle results in Gaussian statistics for uniform mean terrain backscatter regions. More details are given in Section 2.3. Separation of the two terms - the desired mean terrain backscatter image, and the Gaussian speckle 'noise' term - is achieved by making the general assumption that the mean terrain backscatter value varies on a larger scale than the pixel-to-pixel interference at the scale of the imaging resolution.

From the early days of radar imaging, e.g. [Jakeman and Pusey, 1976], and as radar systems achieved higher resolutions, it became clear that the speckle statistics were not always Gaussian and often portrayed much heavier-tailed distributions [Quegan and Rhodes, 1993]. The concept of *radar texture* [Oliver and Quegan, 2004] was developed to encompass additional variation, above the Gaussian interference variation, within a single thematic target media. The product model was modified to allow the mean backscatter value to also vary as a random variable, known as texture, with its own statistical distribution. A constant backscatter value (i.e. a Dirac delta function distribution) for uniform, homogenous regions would still result in the appropriate Gaussian model, but different models for the *textured* terrain backscatter variable would prove to be better fitted models for other real world image regions [Yueh et al., 1989]. The specific details of the modelling and the choice of particular models is currently the focus of much research, e.g. [Freitas et al., 2005, Frery et al., 2007, Bombrun and Beaulieu, 2008, Bombrun et al., 2010, Vasile et al., 2010], and this thesis contributes to this field.

So far the textured speckle discussion is valid for all generic radar backscatter measurements. That is, with coherent, singly polarised, electromagnetic signals. Polarimetric synthetic aperture radar (PolSAR) deals with measuring the scattered return from multiple polarisation channels, (virtually) simultaneously and coherently, to build up the vector scattering characteristics of the target media. Thus, fully polarised, or quad-pol, data-sets contain four complex scattering coefficients, representing the four combinations of transmit and receive, for two orthogonal polarisation bases. Polarisation is most commonly implemented in the vertical (V) and horizontal (H) linear basis pair and leads

to quad-pol imaging containing HH, HV, VH and VV polarisation channels. Measuring the full polarimetric signal is advantageous not only because it contains more complete information to characterise and distinguish different targets, but also because it allows for additional physical scattering interpretation of the scene via the many polarimetric decomposition schemes [Cloude and Pottier, 1996, Freeman and Durden, 1998, van Zyl et al., 2008, Yamaguchi et al., 2008, Yamaguchi et al., 2005], both for coherent and incoherent scattering. The primary interest for this research was that polarimetric SAR holds potentially more target media information and that rigorous multivariate methods were not yet fully developed. Throughout this work, methods are developed that utilise the full vector or matrix information of the data, without discarding any terms, rather than the more common approach of averaging the results from individual polarisation channels processed independently [Freitas et al., 2005, Frery et al., 2007].

In summary, the primary motivation for the basic research in this thesis, is the expectation to obtain more detailed information of the remotely sensed target by using polarimetric SAR images and including non-Gaussian speckle statistics in the image analysis process.

2.2 Scope

The scope of this research is aimed at earth observation of natural environments with polarimetric SAR data and, therefore, certain assumptions and limitations are relevant.

Firstly from scattering theory, the case considered is that of incoherent scattering from distributed targets. This case may later be extended to include potential dominant scatterers and partial coherent scattering components, but for the moment dominant scatterers are ignored. In this case, the measured complex scattering coefficients are expected to have a mean value of zero in all channels and this is simply assumed to be the case during processing. Dominant scatterers will appear as consistently bright targets and will most likely cluster as separate classes by the algorithm, which is probably desired, so ignoring dominant scatterers may not cause great problems.

The product model under investigation is the simplest case with a scalar product term, which means that each dimension is assumed to have the same textural distribution. This model was chosen as the starting point because it is the easiest case and appeared to be suitable for the few images that were available at the time. The literature includes some studies [Fukuda et al., 1999, Sery and Lopes, 1997, Lombardo et al., 2001, Quegan and Rhodes, 1995, Oliver and Quegan, 2004] indicating that this is not always the case and the extended case of a diagonal matrix product model is now being investigated. It is believed that many of the techniques and experience gained in the simple case will remain appropriate for the extended case.

It is blatantly assumed that all images are already fully calibrated, corrected for atmospheric effects and topography, and free from image artefacts such as ghostly echoes. It is simply assumed that the modern SAR processors are sufficiently calibrated, although this should not be a great problem for clustering methods that primarily look at

relative values. Atmospheric affects such as Faraday rotation are either ignored or corrected for using the standard tools, such as the Alaskan Satellite Facility's MapReady application [http://www.asf.alaska.edu/downloads/software_tools]. Topography has not been corrected for in any current work but examples have been carefully chosen to have reasonably flat terrain, for example agricultural fields and sea ice. Additionally, any small variation due to topography may be absorbed into the textural variation parameter in the modelling. It is expected that topography has a major influence on mean backscatter values and this will certainly want to be corrected for in future research. Artefacts are currently inspected for, avoided if possible, but otherwise simply ignored. If they are present, then they would most likely cluster as separate clusters because the pixel data would represent mixed signals under the ghosted regions. That is to say that the clustering algorithm would be expected to produce a result for whatever data is input, no matter what the quality, but the subsequent interpretation would then be a new problem.

Although polarimetry and multivariate methods are fundamental to this research, there is no requirement to have full, quad-pol data-sets. The basic assumptions are that the data is from a coherent imaging system, and therefore exhibits multiplicative 'noise', and that any multivariate channels are zero-mean and may contain correlation, and that is all. It is therefore equally valid for radar sensors of any frequency, for any sub-sets of the full polarimetric channels, such as dual-pol and even mono-pol data-sets, and should also represent the distributions found in ultrasound or sonar images. This generality may also be extended to combining multiple frequency images into a single, stacked, data-set and the multivariate methods detailed in this thesis may be applied directly.

The presented algorithms assume that all pixels are independent and the clustering is essentially done on a pixel by pixel basis, and make no use of potential contextual information. Contextual methods would make use of the fact that neighbouring pixels are more likely to be same class which arises when the correlation length of natural variation is larger than the imaging resolution. Contextual information should, in general, improve the clustering because of this extra information and this may be addressed as a future extension.

2.3 Physical Foundations

This section will cover some of the physical foundations of scattering theory and statistical modelling that are the basis for the thesis. The various concepts introduced here, in often extremely simplified form, are not a complete, detailed description of scattering theory, but are intended to be sufficient to portray the importance of statistical analysis and non-Gaussianity for PolSAR image analysis.

2.3.1 Satellite Backscattering Geometry

At its most fundamental, a radar system transmits an electromagnetic signal as a pulse from its antenna, the pulse travels a distance to a target known as the range, the signal is scattered by the target which transforms the signal, and some of the signal travels back to the antenna to be received by the sensor. Information about the scattering target can be inferred from how it transforms the transmitted signal into the received signal, when all other factors, such as travelling the range distance, are accounted for. That is, the complex reflectivity, also know as complex scattering coefficients, denoted S, can be determined by inverting the scattering equation

$$E_r = \frac{e^{-2\pi jr/\lambda}}{r} S E_t \tag{2.1}$$

where E_r is the received electric field, E_t is the transmitted electric field, r is the range distance of the target from the antenna, λ is the radar wavelength, and $j = \sqrt{-1}$ is the imaginary unit.

The value of S depends upon the geometric and dielectric properties of the target, the incidence angle or orientation of the target with respect to the radar antenna, and the frequency of the radar signal. A SAR image, is essentially an image of S and remote sensing of the target media is by inference through S. The details of SAR image formation and signal processing are outside of the scope of this work, but an interested reader could refer to many texts, for example [Oliver and Quegan, 2004] and [Cumming and Wong, 2005].

This *backscattering* scenario may be visualised, if somewhat simplified, as shown in Figure 2.1. The antenna is in the upper left hand corner and the signal waves are shown propagating to and from the scattering target in the lower right hand corner. The target is shown as a small surface element, or facet, and the symbol also depicts the incident and specular reflection angles, following Snell's law, and the reflectivity envelope. The reflectivity envelope represents the reflected signal strength at different angles relative to the target. The shape of the envelope will depend on the geometric, roughness and material (dielectric) properties of the target surface, the incidence angle or orientation of the target with respect to the radar antenna, and the wavelength of the radar signal. The length of the path within the envelope, at the backscattering (incidence) angle, depicts how much power is reflected back to the antenna.

2.3.2 Surface Reflectivity

Figure 2.2 shows different scenarios for the backscatter envelope depending on the small scale roughness properties of the surface. Roughness is a relative term and must be considered at scales relative to that of the wavelength of the signal. A surface that is smooth, on scales smaller than the wavelength, will produce near specular reflection only and have no significant backscattered signal (Fig. 2.2 a), while an extremely rough

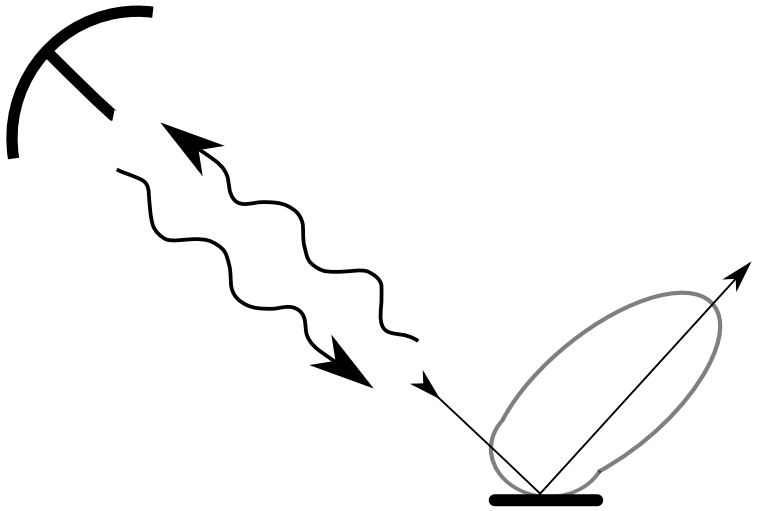


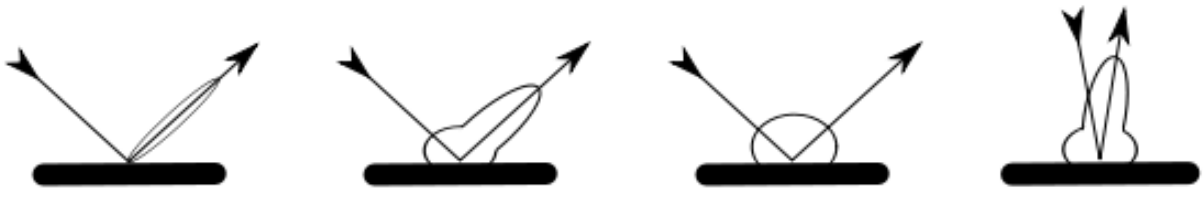
Figure 2.1: Symbolic representation of radar signal backscattering from a single target. The antenna is in the upper left and the electromagnetic signal is shown travelling to and from the target. The target is shown as a facet and includes the incident and specular reflection angles and a reflectivity envelope depicting the scattered power versus reflected angle. Most of the reflected energy goes in the direction of specular reflection, but some amount is backscattered depending on target properties.

surface will scatter in almost all directions equally and have a relatively large backscattered power (Fig. 2.2 c). For targets that are somewhere between the two extremes (Fig. 2.2 b), it is readily apparent that a steeper incidence angle will result in a higher backscattered power as the incidence angle approaches the diffused specular component (Fig. 2.2 d). Specular reflection has the maximum reflected power and in backscattering occurs when the local target incidence angle is zero (i.e. normal to the surface), however, even specular scattering from multiple surfaces so aligned as to reflect back to the target is generally brighter than direct diffuse surface scattering.

This is the case for a single isolated target, or a bright dominant point target within a weaker 'clutter', and results in a deterministic response, known as *coherent scattering*, dependent on the target properties and imaging geometry only. Its response would likely vary smoothly and continuously as the imaging geometry varies.

2.3.3 Distributed Target Scattering

The scale of the resolution of SAR images, in relation to the wavelength, means that, in general, the imaged pixel is an extended area of terrain on the surface and is referred to as *distributed scattering*. The measured electromagnetic response is then the coherent sum over the whole illuminated surface area. Although the response is actually the integrated coherent response over the entire surface area (or volume) it may be under-



(a) (b) (c) (d)

Figure 2.2: Backscatter envelope power at different surface roughness conditions: (a) smooth, (b) intermediate, (c) rough, and (d) intermediate at steeper incidence angle. The backscattered power level is shown to be increasing from (a) to (d) in this example.

stood and interpreted in a simple way by considering a discrete number of independent, locally coherent surface elements or facets. These are the most significant or dominant surface returns contributing to the overall response and the number of such scatterers within each cell is unknown and must be considered a random variable.

Natural terrain will consist of a random arrangement of scattering sites at the surface and therefore the number of significant sites from one resolution cell to the next will be random. Assuming that the surface structure is naturally random and that the significant, contributing scatterers can be assumed uniformly randomly distributed (i.e. equally likely anywhere) across the surface, then it is appropriate to use a Poisson distribution for the random variable N [Jakeman and Pusey, 1976]. Alternatively, under the assumption of small local correlation within a Poisson process framework, the random variable N could be modelled as a negative binomial distribution [Jakeman and Pusey, 1978]. In any case, since the scatterers are located randomly across the resolution cell, and since the resolution cell size in range is many time the wavelength in extent, then the distribution of local phases due to location, θ_i, can be considered uniformly distributed over 2π.

Each individual scatterer's complex reflectivity may be different and therefore the individual complex scattering coefficients, S_i, must also be considered as a random variables. Since S_i is complex, it will have a magnitude, $|S_i|$, and scattering phase, ϕ_i. However, any scattering phase information from the local complex reflectivity, ϕ_i, will be lost amongst the uniformly random phase due to location, θ_i. Additionally, the random phases, θ_i, ϕ_i and the reflectivity magnitude, $|S_i|$, are considered statistically independent.

This distributed scattering equation may be represented mathematically as

$$E_r = \frac{e^{-2\pi jr/\lambda}}{r} \left(\sum_{i=1}^{N} |S_i| \; e^{-j(\phi_i + \theta_i)} \right) E_t \tag{2.2}$$

where r now represents the mean range distance, and under the far-field conditions is considered equal for all scatterers within the cell, and the individual range variation,

relative to the mean range, has been taken into the individual location phase terms θ_i. Note that the local complex reflectivity, $|S_i|e^{-j\phi_i}$ and local phase, θ_i both depend upon the individual scatterer and its location within the cell. The total complex backscatter coefficient, S, now equates to the summation in parentheses.

The summation term in parentheses in (2.2) can be considered as a random walk model in the complex plane which leads to constructive and destructive interference amongst the scattered signals, dependent upon their random values and locations within each resolution cell, and introduces an essentially random process into the measured scattering response. This type of distributed, random scattering shall be referred to as *incoherent scattering*.

General discussion of the random walk model may be found in several sources, for example, [Oliver and Quegan, 2004, Jakeman and Pusey, 1976, Jakeman and Pusey, 1978, Goodman, 1976, Jakeman and Tough, 1987, Eltoft, 2005, Yueh et al., 1989, Barndorff-Nielsen, 1997].

2.3.4 Speckle

This random walk model and the interference effect is the physical basis of the phenomenon known as *speckle*. The complicated nature of interference from many distributed scatterers means that the total response depends on the relative geometric position of *all* contributing targets and now essentially varies randomly with the imaging geometry and from pixel to pixel. Neighbouring pixels of the same target medium will have different random walk summations because of the different random magnitudes and placement of individual scatterers within each pixel cell.

The results of the random walk sum may be summarised into two general cases depending, in general, on the imaging resolution and the variability of the scatterers.

The total scattering coefficient from the random walk summation will be zero-mean complex Gaussian distributed when the number of contributing scatterers is large. This result is largely a consequence of the Central Limit Theorem. These conditions will generally be the case for very low resolution (large pixel size) images or where the surface variability is approximately uniform on a fine scale relative to the cell's resolution. This is referred to as *fully developed* speckle.

The scattering coefficient will be zero-mean but non-Gaussian when there are very few contributing scatterers within the resolution cell. Few significant scatterers may result when the resolution is very high (small pixels) or because there is significant variability within the individual reflectivity magnitudes, resulting in very few of the most significant, and thus contributing, magnitudes to the total response. Furthermore, this type of heavy-tailed, non-Gaussian distribution can always be described by a product of a scaling term and a Gaussian random variable [Andrews and Mallows, 1974] [Oliver and Quegan, 2004, Appendix 11A]. This case of non-Gaussian interference is termed *under-developed* speckle.

Real SAR images have been measured to contain both Gaussian and non-Gaussian

local distributions which is indicative that the response is from both uniform and variable scattering on scales comparable to the resolution cell size. Therefore, incorporating non-Gaussianity into the image analysis methods should be beneficial.

2.3.5 Roughness Scales

One important cause of variability in the scattering sites is related to surface roughness at different scales [Ulaby and Elachi, 1990, Ulaby et al., 1982, Valenzuela, 1978]. The following is the author's interpretation of the basic concepts.

At the very smallest scales, less than the wavelength, the roughness contributes coherently to the facet's surface reflectivity, as discussed in Section 2.3.2. Small scale roughness may result in reflectivity scenarios ranging from specular to rough surface backscatter, as indicated by Figure 2.2. Variability of the small scale roughness, will appear as variability of the individual complex reflectivity values in the random walk summation.

The next scale, let's call it fine-medium scale, is larger then the wavelength, i.e. facet scale, but much smaller than the resolution cell size. In this case the fineness of the surface geometry will result in very many contributing scatterers spread uniformly across the illuminated resolution cell, resulting in fully developed speckle statistics. This may be seen in Figure 2.3, where there are very many of the most significant contributing scatterers.

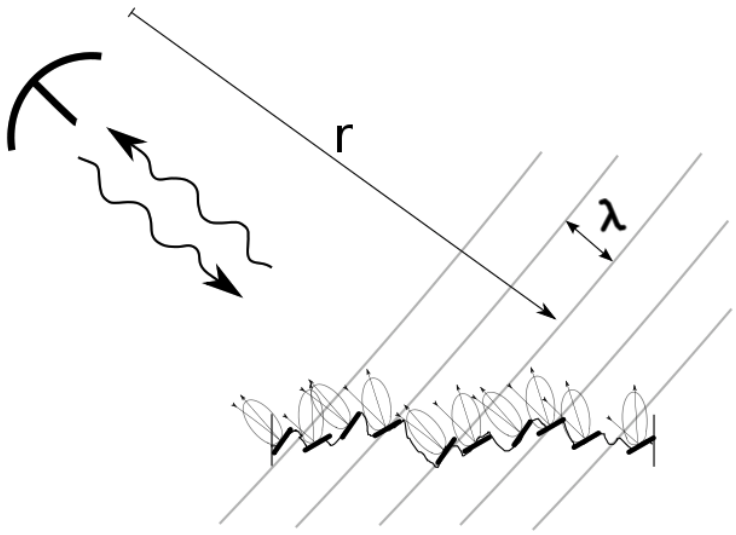


Figure 2.3: Distributed scattering from a fine-scale, uniformly rough surface with very many contributing scatterers in the resolution cell. The total response is the coherent sum of all scatterers within the area, with a uniform, random distribution of local phases and results in a zero-mean Gaussian random variable.

Next consider a coarse-medium scale that is approaching the resolution cell size, such that there is an uneven, though random, distribution of the few most significant scattering sites within the cell. This case is depicted in Figure 2.4, where there are only very few of the most specular aligned reflections across the resolution cell. The resulting response is non-Gaussian, under-developed speckle statistics.

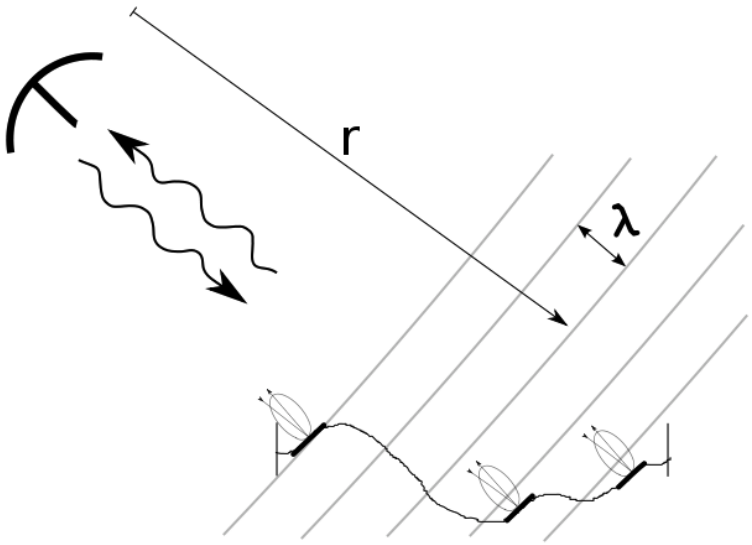


Figure 2.4: Distributed scattering from a coarse-scale rough surface with few contributing scatterers in the resolution cell. The total response is the coherent sum of all scatterers within the area, with a uniform, random distribution of local phases and results in a zero-mean but non-Gaussian random variable.

Note that if there was only one single contributing scattering site, then the result would be the case of dominant coherent scattering and will not exhibit speckle. This case may be evident in real images for point sources that are very bright, usually man-made, metallic, aligned reflectors (or corner reflectors), but is not the main focus of the presented works as noted in Section 2.2.

Lastly, for completeness, consider variation on the very large scale, greater than the resolution cell size, as shown in Figure 2.5. Such large scale variation is effectively smooth at the resolution cell scale, but will nevertheless introduce some random variation from pixel to pixel due to the mean local angle of incidence differences. The variation on this scale can (and should) be corrected for with digital elevation models and is called *radiometric terrain correction*.

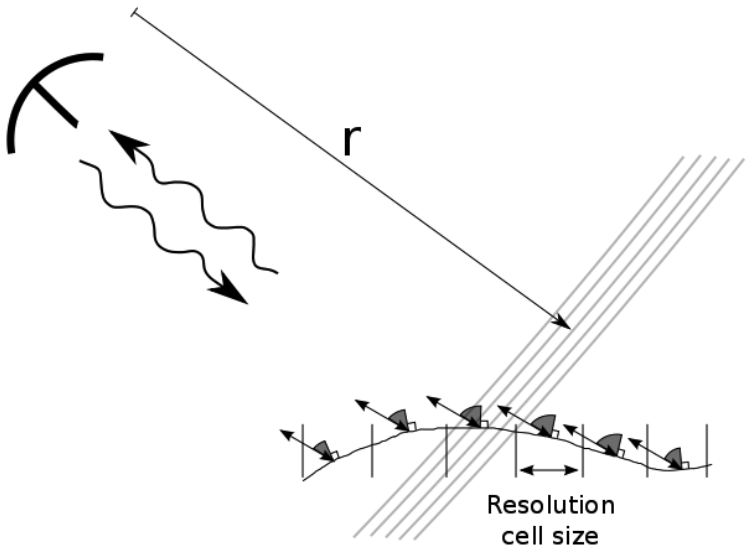


Figure 2.5: Large-scale surface variation over many resolution cells. The mean incidence angle variation can lead to mean backscatter variation for otherwise equally rough surface scattering.

2.3.6 Texture

The interference sum that includes variation in scattering magnitudes must now be considered a doubly stochastic mechanism, with a texture variation term and an independent Gaussian speckle term. This is referred to as the *product model* and is described in the general literature, for example in [Oliver and Quegan, 2004]. The distribution of $|S_i|$, however, is likely to be very difficult to define physically, as it depends upon the angular reflectivity envelope through the small scale surface roughness properties and the distribution of local angles at all other roughness scales.

There are yet more causes of variation from the mean backscatter value that are desirable to be considered as the same target media or thematic class for classification purposes. Besides medium scale geometric texture described previously, there may be minor natural variation in the small-scale roughness affecting backscattered power, natural variation of the local material properties such as moisture or density, or small variations in the mixing proportions of fine resolution composite media, i.e. mixtures on a finer scale than the resolution cell size, or often several causes combined. Examples of variable media include both forest and urban classes: forests have natural internal variation on the order of several metres due to the random placement of the trees, within the forest, yet may be desirable to simply class as forest; and the class known as urban only has meaning at scales where it depicts a mixture of buildings, roads, trees, and cars etc.

For example if the backscatter from vegetation is greater than that of smooth soil, as

is a likely case in many situations, then the significant contributing scattering sites will be governed by the distribution of vegetation in the resolution cell. This is depicted in Figure 2.6 for both the uniform, fully developed, case and the textured, non-Gaussian case.

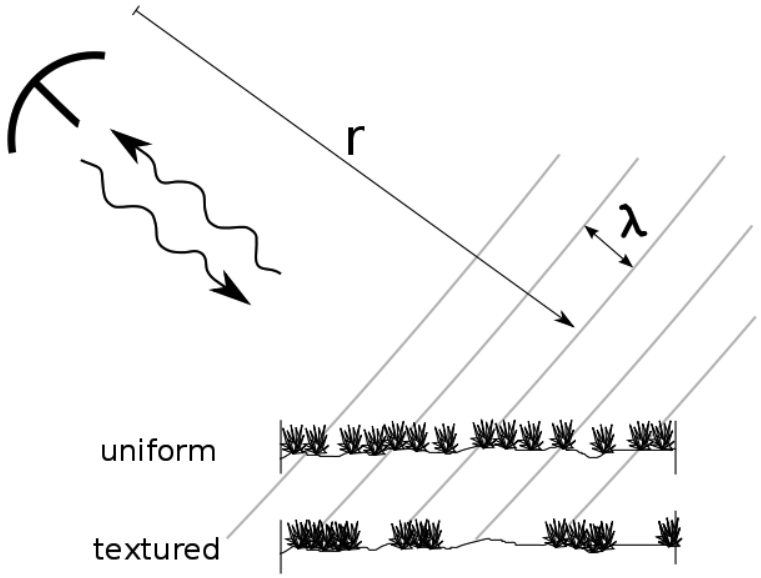


Figure 2.6: Vegetation 'mixture' variation over the resolution cells. The uniform case will lead to Gaussian speckle, and the textured case to non-Gaussian speckle statistics, and will be random from cell to cell due to the random distribution of the individual scattering sites.

All causes of additional variation in radar backscatter are encompassed in the term *radar texture* and their potentially non-Gaussian speckle statistics may be considered texture modulated interference. The different possibilities for textural variation may be indistinguishable in the final imaging and means that there is probably no single, simple mathematical model to suit all situations. Flexible empirical models are likely to be more useful in analysing images than more complicated physical modelling which would need several texture-like terms.

2.3.7 Polarimetric SAR

Polarimetry involves measuring the full vector nature of electromagnetic waves and describes the orientation and motion of the electric field oscillation with respect to the direction of wave propagation. Polarimetry is generally carried out in the vertical-horizontal linear polarimetric basis, which is sufficient since any other polarimetric basis can be synthesised from this. Understanding polarimetry and polarimetric SAR synthesis is outside the focus of this thesis, apart from recognising its importance as

a source of extra information about the target, and the interested reader is referred to review articles such as [Touzi et al., 2004] or the PolSAR Pro tutorial on the European Space Agency web-site [http://earth.esa.int/polsarpro/tutorial.html] as a reference.

Polarimetric SAR systems must therefore measure the four combinations of horizontal and vertical polarisation for both the transmitted and received signals. These four channels, HH, HV, VH and VV, are together referred to as quad-pol data, and contain the complete vector information of the scattered signal with respect to the satellite antenna system. SAR processing essentially produces four complex scattering coefficient images, one for each channel, and polarimetry is the study of inter-channel characteristics to infer more information about the target. Each scattering coefficient measurement is consistent with the previous discussion on speckle and texture and the vector nature of polarimetric SAR may be written as

$$\left[\begin{array}{c} E_h \\ E_v \end{array} \right]_r = \frac{e^{-2\pi jr/\lambda}}{r} \left[\begin{array}{cc} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{array} \right] \left[\begin{array}{c} E_h \\ E_v \end{array} \right]_t \tag{2.3}$$

The four measured coefficients are usually vectorised as the raw complex SAR image known as single-look complex (SLC) data.

$$\mathbf{s} = \left[\begin{array}{c} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{array} \right] \tag{2.4}$$

From the previous discussion for incoherent scattering from natural terrain, it is expected that the vector $\mathbf{s}$ will be zero-mean and therefore useful information must be obtained from the higher order statistical measures such as from the covariance or coherency matrices. Multi-look complex (MLC) data is either averaged in the frequency domain during image formation, or by spatial averaging of the SLC image pixels. In the latter case, the multi-look averaging is done in the (second order) power domain, since the mean is zero, and may be written as

$$\mathbf{C} = \frac{1}{L} \sum_{i=1}^{L} \mathbf{s}\,\mathbf{s}^H \tag{2.5}$$

$$= \left[\begin{array}{cccc} < |S_{hh}|^2 > & < S_{hh}S_{hv}^* > & < S_{hh}S_{vh}^* > & < S_{hh}S_{vv}^* > \\ < S_{hv}S_{hh}^* > & < |S_{hv}|^2 > & < S_{hv}S_{vh}^* > & < S_{hv}S_{vv}^* > \\ < S_{vh}S_{hh}^* > & < S_{vh}S_{hv}^* > & < |S_{vh}|^2 > & < S_{vh}S_{vv}^* > \\ < S_{vv}S_{hh}^* > & < S_{vv}S_{hv}^* > & < S_{vv}S_{vh}^* > & < |S_{vv}|^2 > \end{array} \right] \tag{2.6}$$

where L is the number of looks to be averaged, the superscript ^{H} denotes the Hemitian or complex transpose operator, the superscript * denotes complex conjugation, $< \cdot >$ refers to ensemble averaging, and here $| \cdot |$ refers to absolute magnitude. Multi-look

averaging also reduces the degree of speckle variation at the expense of resolution. Note that the diagonal elements are the mean intensity for each polarisation channel.

It is generally agreed that there are five free parameters observable from an incoherently scattered covariance matrix from natural surfaces with reflection symmetry. They are the three real intensities and the complex correlation magnitude and phase. Ratios of covariance matrix elements are often taken as features because they should be independent of any textural scaling since the ratio divides away the scale from the product model. A common set of five parameters are:

- An absolute backscatter value, often σ_{hh}
- A cross-polarisation fraction or ratio, often σ_{hv}/σ_{hh}
- A co-polarisation ratio, usually σ_{vv}/σ_{hh}
- The co-polarisation correlation magnitude, $|\rho| = |C_{hhvv}|/\sqrt{(|C_{hhhh}|\ |C_{vvvv}|)}$
- The co-polarisation correlation angle, $\angle\rho =< \phi_{hh} - \phi_{vv} >$

which are discussed in for example [Quegan et al., 2003], [Dierking et al., 2003] or [Thomsen et al., 1998].

Subsets of the full quad-pol data, such as dual-pol and mono-pol data, also have reduced features available as appropriate for the specific sub-sets.

2.3.8 Gaussian-based Statistics

The usual choice for PolSAR analysis is to assume Gaussian-based statistics because they are mathematically easier to deal with, but which effectively ignores any textural variation. The term Gaussian-based is chosen because the complex scattering coefficients are assumed to be Gaussian distributed, but the data analysis is likely to be carried out with some other dependent domain such as with the amplitude, intensity or covariance matrices which will obviously not be Gaussian data models. Given that the complex scattering coefficients are multivariate, zero-mean complex Gaussian variables, then it follows that the marginal amplitude distributions will be Rayleigh, the marginal intensities have negative exponential distributions and that the covariance matrix distribution will be complex Wishart distributed.

The complex Wishart distribution [Goodman, 1963] for the scaled covariance matrix, $\mathbf{C}_L$, with degrees of freedom, L, mean covariance matrix, $\mathbf{\Sigma}$, and dimension d, is defined for $L > d$ as

$$f_{\mathbf{C}_L}(\mathbf{C}_L; L, \mathbf{\Sigma}) = \frac{L^{Ld}|\mathbf{C}_L|^{L-d}\exp(-L\ \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}_L))}{|\mathbf{\Sigma}|^L I(L,d)}\ , \tag{2.7}$$

where $\mathrm{tr}(\cdot)$ denotes the trace operation, $|\cdot|$ the determinant,

$$I(L,d) = \pi^{\frac{d(d-1)}{2}} \prod_{i=1}^{d} \Gamma(L-i+1) \tag{2.8}$$

is a normalisation constant and $\Gamma(\cdot)$ is the standard Gamma function.

Wishart based clustering or classification is still common practise [Lumsdon et al., 2005,Lee et al., 2004] and often after dynamic speckle filtering [Cao et al., 2007,Singh and Venkataraman, 2008] which invalidates certain assumptions for the Wishart model. This work hopes to address and demystify some of the difficulties of using non-Gaussianity for general PolSAR analysis.

Chapter 6

Research Discussion

This chapter places particular research achievements of this thesis in context within the broader field of SAR image analysis. The concepts are introduced approximately chronologically and, thus, the evolution of the research is also apparent.

6.1 Suitability

To begin any analysis of non-Gaussian statistics it seemed appropriate to first investigate the need for non-Gaussianity and the suitability of the statistical scheme. Therefore, several available polarimetric SAR images were analysed and local histograms indicated areas of both Gaussian and non-Gaussian speckle statistics, thus confirming the need for non-Gaussianity for PolSAR analysis. This was reported in the result sections of both Paper 1 and Paper 2 to justify the modelling.

The local histograms of the real and imaginary (i.e. the in-phase and quadrature signals) showed that the speckle statistics appeared to be: (1) zero-mean, (2) symmetric, (3) with different widths for different polarimetric channels, (4) but with some pair-wise constraints, and (5) with different non-Gaussian shape at each location, but (6) with a similar shape for each dimension within each location. These properties must therefore be represented in the choice of statistical model, for which the scale mixture of (complex) Gaussian models are shown to be suitable. The scale mixture models derive from a scalar product model and can be easily made to be zero-mean, semi-symmetric, and use flexible non-Gaussian distributions. Furthermore, the pair-wise and real/imaginary constraints can be easily incorporated into the covariance structure matrix.

The images investigated showed no significant signs of violating the scalar texture, or 'global shape' assumption that some authors have noted in some situations (see Section 2.2). The small variation in shape parameter values in Paper 1, Figure 3, may not be significant with respect to the statistical variation of the parameter estimator, given the sample sizes involved, yet clearly show significant variation of the shape parameter at different locations in the image.

Therefore, these scale mixture of Gaussian models appear well suited for the task of

modelling PolSAR data-sets. Virtually all models for PolSAR being considered in the literature derive from the scalar product model, which is the same as the Scale Mixture of Gaussian scheme.

6.2 Flexible Parametric Models

Scale mixture of Gaussian models covers many different families of distributions depending upon the distribution for the scale parameter. In Paper 1, four different parametric models were investigated. Two with fixed shapes, the multivariate Gaussian and the multivariate Laplace distributions, and two with flexible shapes, i.e. with a range of non-Gaussianity measure, the multivariate K-distribution and the multivariate Normal Inverse Gaussian distribution. This is only a very limited choice of models, nevertheless, several observations were clear from the study. Firstly, the degree of non-Gaussianity varies throughout the images such that no single fixed-shape model would be suitable everywhere. This is also a confirmation that the Gaussian model is not sufficient for all areas. Secondly, the two flexible models studied did show a sufficient degree of flexibility to cover most natural image areas, with usually over 90% good-fitting coverage of the images. Thirdly, the Normal Inverse Gaussian model appeared to be a better general model for most cases. Lastly, terrain class boundaries and mixtures, showing extreme heterogeneity, do seem to be a problem for all models studied.

The conclusion, at that time, was that both the K-distribution and the Normal Inverse Gaussian distribution would be suitable single models because they were flexible enough to cover most degrees of non-Gaussianity found in natural images. Having a single parametric model was desirable because it would simplify the mathematics and the analysis. An alternative, would be a multi-model fitting scheme (called a dictionary-based scheme in [Moser et al., 2006]), but that would involve an extra level of testing, and probably time consuming parameter estimation for all models, that is not desirable. Subsequent research focussed on the K-distribution because it was well known in the literature, and the question of why the Normal Inverse Gaussian was superior was not addressed.

Recent research has brought new visualisation tools and greatly improved the statistical confidence of working with particular models such that even the single, flexible models are showing their limitations. The question of the best model amongst all proposed models, including those recently published by other authors such as the $\mathcal{G}^0$ and the Kummer-U distributions, shall be addressed in future research.

6.3 Moment Based Methods

One reason that non-Gaussian methods have become practical for image analysis and clustering purposes is because of fast estimation of parameters through method of moments techniques. In earlier works [Eltoft et al., 2006], the shape parameters were es-

timated through iterative techniques to maximise the log-likelihood function given the model and the data. This proved to be computationally quite slow, particularly for near Gaussian data-sets [Doulgeris, 2006]. One solution is to estimate one or more empirical moments from the data and equate them to the expected analytical moments given the chosen parametric model. This is called *method of moments*, and essentially needs one moment expression for each unknown in order to find a solution for each parameter.

This technique proved fast enough for full image analysis and reduced computational time manyfold, from days to minutes in some cases [Doulgeris, 2006]. However, it has become apparent in recent work that the method of moments, in particular using a single moment for estimates, is actually the cause of one of the clustering problems – of stable, mixed mode clusters. The stability is understood to be because a moment is a statistical average of some measure and is not unique in the sense that the same average may occur for different distributions of higher and lower values of the measure that just happen to balance each other to obtain the same mean value. Thus, the expected numerical average value was often obtained for a bi-modal mixture of two separate clusters, and since it was in balance there was no impetus to change from that stable point. Additionally, within iterative clustering algorithms, the situation is often very likely to re-occur because of the smoothly varying adjustments from one iteration to the next giving a high likelihood of passing through the stable point. This can be remedied, to some extent, by utilising more than one moment because they are less likely to all be in balance simultaneously, and this has been taken up in the more recent research.

Method of moment clustering, i.e. with parameter update algorithms utilising the method of moments, may also not be optimising the log-likelihood function and may therefore not be the absolute best method for parameter optimisation. However, they are often mathematically simple to implement, and usually very fast, which proved sufficient for furthering the techniques. It is the method of choice for several authors, for example in [Kuruoglu and Zerubia, 2004, Freitas et al., 2005, Frery et al., 2010], however, it has recently been superseded by the related method of log-cumulants (discussed in Section 6.8).

6.4 Extended Polarimetric Features

The idea of adding non-Gaussianity as extra information for PolSAR data analysis is central to this thesis. The suggestion to add it as an extra real valued feature to existing PolSAR features, as in Paper 1, was thought to be a novel addition to the field, however, a recently discovered conference proceeding paper [Quegan et al., 1994] suggested virtually the same idea. There are, however, significant new advances presented in Paper 1, because of the choice of features and their independence from explicit parametric models, and also because of putting the theory into practise by using it for image segmentation.

Polarimetry is usually interpreted through the covariance matrix, or a linear transformation of it as the coherency matrix. Since these are complex matrices in three or four

dimensions with internal correlation, it is reasonable to extract real valued features from the matrix for analysis, and five commonly used features are noted in Section 2.3.7. The ratio features are often chosen relative to the HH channel intensity, but it was thought to be more universal to make them relative to the total backscattered power or multivariate radar cross-section, defined as the d^{th} root of the determinant of the covariance matrix. Additionally, many of the features are logarithmically transformed to improve their numerical data spread for visualisation and clustering.

The non-Gaussianity measure used in [Quegan et al., 1994] was the order parameter from the K-distribution, however, Paper 1 argues for using a sample moment relation (the relative kurtosis) because it is an empirical statistical measure of the data and completely independent of any parametric model choice, apart from assuming the general scalar product model. This fact is quite an important result for general PolSAR analysis, because the exact parametric model, if one even exists, is never explicitly known.

Paper 1, Figure 8, clearly shows that a significant amount of detail is visible in the features from the example image. The overall appearance of the histograms encouraged the application of discrete mixture of Gaussian clustering to segment the 6-dimensional feature space. There is no specific reasons that the six features should have Gaussian distributions for natural data classes, but it seemed worthwhile to try it, given the softly rounded humps in the histograms. Besides being very fast, sub-sampling the data-set appeared to give the best results and is probably due to the reduced sensitivity of the mixture components to actually being Gaussian that comes with reduced sample sizes. The clustering results appeared to be visually quite good, as seen in Paper 1, Figure 13, although a rigorous testing against ground truth data has not been performed.

Although demonstrated for clustering, a more interesting aspect of this approach is that it maintains the local values of the features at each image location. Clustering, on the other hand, groups image pixels together and only obtains the groups mean features as its parameters, admittedly with more accurate parameter estimates because of increased sample size. The local values may be of interest for physical parameter retrieval, for example, estimating soil moisture or forest biomass. Physical parameter retrieval may benefit from first clustering to find out what type of land cover is present and then extracting physical parameters from terrain specific empirical inversion expressions. Therefore, this non-Gaussianity feature may be valuable for physical properties such as biomass because it reflects the geometric variation and spacing on scales smaller than the image resolution. Interpreting non-Gaussianity in this manner is yet another possibility for the future.

Recent work with log-cumulants suggests that substituting one, or more, sample log-cumulants (which would also be model independent) for the non-Gaussianity measure may be beneficial and will be investigated in the future.

6.5 K-Wishart Clustering

The second paper takes the specific non-Gaussian modelling approach and investigates statistical modelling of multilook complex (MLC) matrix data for image segmentation or clustering. Paper 1 argued for a flexible, two parameter model and the K-distribution was chosen because it was well known in the literature and generally a good fit in many real images. The statistics is then modelled through to the MLC covariance matrix data format and implemented in a clustering algorithm. The resulting MLC model was called the K-Wishart in Paper 2, but was originally presented 15 years earlier in [Lee et al., 1994b]. The original [Lee et al., 1994b] introduces the theoretical distribution, and related ratio distributions, and compares various marginal histograms to real data for validation. Very few works have actually used the non-Gaussian K-Wishart model for actual image analysis, and most intervening work still reverts to using the complex Wishart model for clustering. The exception being [Freitas et al., 2001], and other works by the same authors, where they demonstrate *supervised* classification using two non-Gaussian models, the K-Wishart and the G^0. Paper 2 tackles the more complicated case of unsupervised clustering.

Unsupervised clustering, via the expectation maximisation algorithm [Dempster et al., 1977] with method of moment parameter updates, was shown to be practical and successful and it produces similar clustering to the Wishart model for homogenous regions and arguably better results for non-Gaussian regions. Showing that this can be achieved in practise is a significant contribution on its own, however, the main contribution of the article was to point out several important influences or observations that make non-Gaussian modelling a difficult task and, in addition, to make some suggestions as to how to remedy these problems.

The most significant are that the number of classes still needs to be given *a priori*, or chosen after looking at many full clustering results, that the effective or equivalent number of looks is hugely influential and must be estimated manually, and that the initialisation conditions affect the resulting clusters and convergence speed. Many of these problems are very general problems with virtually all clustering algorithms, but here they were clearly observed to be limiting the effectiveness of the clustering or its interpretation. These ideas are subsequently addressed in Paper 3 and are discussed below.

6.6 Single to Multi-Look Texture Factor

One other novel factor discussed in Paper 2, is the change of non-Gaussianity degree that results from multilook averaging. It is well understood that multilook averaging will undoubtedly make the distributions more Gaussian, because of the central limit theorem, however the degree of the changes in relation to the number of looks is a new result. Whether this result proves to be particularly useful is yet to be seen, but it may potentially be used to determine, in advance, the likely importance of non-Gaussianity

after multilooking, help to choose how many looks to average, and be used to decide whether the simpler Wishart model may be appropriate. In Paper 2, it was simply used to convert between the estimated single look texture parameter and the multilook one.

6.7 Automated Analysis

Paper 3, also takes the specific model approach and picks up all the loose ends from Paper 2. The title calls it automated because it no longer requires prior choice of the number of classes, the ENL nor choice of initialisation, as these are all automatically determined internally. In addition, there are several other improvements. All parameter estimation has been converted to the more accurate method of log-cumulants (see Section 6.8), and in fact utilising multiple log-cumulants for better stability. The ENL estimate is handled internally and includes a correction for texture, which has never been seen elsewhere. By adding goodness-of-fit testing, the algorithm automatically determines the appropriate number of clusters based strongly on statistical significance through the tests, and in addition requires no special initialisation. Lastly, the evaluation of the Bessel K function has been significantly improved to avoid many numerical infinite results and to extend the calculable range.

All of these improvements are clearly described in the paper, and some are expanded upon below.

6.8 Methods of Log-Cumulants

The method of log-cumulants is similar to the method of moments (discussed in Section 6.3) and superior to them for product based distributions. First applied to SAR in [Nicolas, 2002] (translated in [Anfinsen, 2010]), the method of log-cumulants has become popular in recent works [Tison et al., 2004, Moser et al., 2006, Bombrun and Beaulieu, 2008], and has been rigorously expanded to matrixvariate forms in [Anfinsen, 2010]. Using log-cumulant methods for parameter estimation is considered superior because of reduced bias and variance of the estimators [Anfinsen and Eltoft, 2011], and because the mathematics becomes relatively simple, primarily due to the logarithmic transformation separating the product model origins of the parameters.

However, some drawbacks are that the log-cumulants often involve several polygamma functions that must be solved numerically and likewise for minimum distance optimisation as suggested in [Anfinsen and Eltoft, 2011]. Optimising using multiple log-cumulants, simultaneously, involves numerical gradient approach methods in multiple dimensions, but, fortunately, the expressions are not too complicated. Iterative optimisation to solve for a single log-cumulant vector for one or two parameters using the log-cumulant expressions should not be likened to the slow iterative clustering which involves evaluation (and re-evaluation) over every pixel with the PDF expressions.

The benefits of using methods of log-cumulants are well justified and allows for a consistent framework for many different models, and has been incorporated into Paper 3.

6.9 Goodness-of-fit Testing

The goodness-of-fit testing is a novel addition and solves three problems at once. Firstly, it breaks the stable, mixed points that had been observed in fixed number of classes clustering, and guarantees that all clusters are good fits to the models. Secondly, it requires no special initialisation, because the entire image can start as one (obviously mixed) cluster and it will automatically adapt from there. Thirdly, and most importantly, it solves the choice for the optimum number of clusters by dynamically splitting and merging clusters based upon the goodness-of-fit tests.

Finding reliable goodness-of-fit tests was the most difficult development, as they had to correctly account for different sample sizes and potentially extremely non-Gaussian distributions. Initial methods (presented in [Doulgeris and Eltoft, 2010a]) used a Pearson's Chi-squared test on a one-dimensional compacted distribution for trace$(\boldsymbol{\Sigma}^{-1}\mathbf{C})$. However, that involved knowing the compacted distribution, numerically integrating to obtain the cumulative distribution, broke down for small sample sizes because of various multinomial distribution assumptions, and introduced an additional undesirable tuning parameter - the number of bins. The final robust method was only possible once truly matrix-variate goodness-of-fit tests were developed in [Anfinsen et al., 2011]. These act directly on the matrix samples and are fast method of log-cumulant techniques, at least for large sample sizes. Monte Carlo methods are suggested for smaller sample sizes, that are accurate for any distributions but obviously computationally slower.

A clear advantage with using goodness-of-fit tests is that they have a statistically meaningful interpretation and only one tuning parameter - the confidence level, for example a 95% confidence level is commonly used. Another novelty is that the goodness-of-fit testing is applied to each cluster individually, in a sense locally tested, rather than globally on the total log-likelihood function. It is easy to understand that if each cluster is a good fit, then clearly the total must also be a good fit to the data, whereas the converse may not always be true. A third aspect is that the testing is dynamic within one iterative algorithm.

6.10 Number of Clusters

Finding a suitable number of clusters for clustering algorithms is a very important achievement, and to do so with a novel and general approach, with clear, intuitive statistical interpretation, is a strong contribution to the research field. The proposed approach to goodness-of-fit testing to determine the number of clusters is applicable to

any finite mixture modelling situation and may, in fact, be generally applied to many fields of study.

The optimum number of clusters is found dynamically within one iterative clustering algorithm, instead of the more usual hierarchical based, L-method [Salvador and Chan, 2004], that first obtains many full clustering results before evaluating each to find the optimum. Other authors are aware of the importance of finding the number of classes with several recent articles using the L-method, e.g. [Cao et al., 2007, Bombrun et al., 2010].

The proposed method includes a regular goodness-of-fit testing stage within the iterative clustering algorithm that tests each individual model fit to the data cluster, and each pairwise combination. Poorly fitting individual clusters are *split* into two, simply by partitioning the trace$(\mathbf{\Sigma}^{-1}\mathbf{C})$ at its expected mean value of d. This is because the poor fits were generally observed to be the result of multimodal mixed clusters of two, or more, distinct sub-groups of data, and is a clear sign of having too few classes in the clustering. The converse, of having too many classes in the clustering, usually results in some clusters converging to the same data group and having virtually identical parameters and simply sharing the total prior probability. This case is tested by considering each pair of good fitting clusters to see whether the combined group can be described by a single fitted cluster, and therefore involves estimating model parameters for the combined group before applying the goodness-of-fit test. Combinations that pass the test are *merged* together into one cluster. In this way, the optimum number of well fitting clusters dynamically evolves during the iterations.

The goodness-of-fit testing is performed only every 10, or so, iterations to give the standard expectation maximisation algorithm some chance to adapt in the intervening iterations. This means that the overall algorithm only ever partly converges for unsuitable number of clusters and is therefore faster than the hierarchical methods. The early stages of the algorithm very quickly splits towards the optimum number with potential doubling every test stage and then stabilises towards the optimum.

Some situations were observed that resulted in a cyclic splitting and re-merging of the same clusters, due to slight drifting of the cluster parameters in the intervening iterations. This was avoided by including a ramping confidence level that progressively reduces the sensitivity of the goodness-of-fit tests to be less sensitive to splitting and merging conditions. The ramping was only initiated after a suitably long constant period to allow the main splitting to occur, and then the sensitivity was ramped from, for example, 95% to 99.999% for splitting and from 95% to 85% for merging, before remaining fixed for another suitable length to allow stable convergence. Having to include this ramping does not appear to affect the majority of perfectly stable clustering scenarios, but does break the occasional cyclic behaviour.

6.11 Statistical Confidence

Another potentially useful outcome from the algorithm results from understanding that statistical confidence depends strongly on sample size and that this may be used to adjust the level of detail attained by simply sub-sampling the data-set.

The variance of estimators, like sample moments and log-cumulants, generally decreases inversely as the sample size increases. This means that statistical significance also varies with sample size, and hence the results of the goodness-of-fit testing, i.e. the confidence that the data variation falls under the null hypothesis, must reflect this variation with sample size. To turn this around and actively sub-sample the data to produce less sensitive goodness-of-fit tests is a potentially useful feature with several advantages. The clustering will be much faster if sub-sampled, because less pixels are involved in every calculation. The sensitivity reduction means that small sub-groups that are statistically near more major clusters will be less distinguishable and only the major cluster will be found to be significant for small sample sizes. This may be desirable for many segmentation applications where the whole point is to simplify the image into only a few thematic groups. The reduced sensitivity may also allow acceptable clustering even if the parametric model does not exactly match the physical phenomenon that it attempts to describe.

The last point is of particular interest because recent observations for large sample sizes may begin to indicate limitations with the choice of the K-Wishart model for some data areas. The strength of the goodness-of-fit testing together with improved visualisation through class histograms [Doulgeris et al., submitted 2010] or matrix log-cumulant diagrams [Anfinsen and Eltoft, 2011] are suggesting new avenues to investigate in the future.

Chapter 7

Conclusions

This chapter contains some final concluding remarks and discusses possible future research.

7.1 Concluding Remarks

The physics of radar backscatter from natural terrain can certainly lead to non-Gaussian statistical properties of radar images. This work demonstrates the need for flexible non-Gaussian statistical modelling, validates the suitability of a particular product model scheme, and describes two separate approaches for non-Gaussian polarimetric analysis.

The first approach involves local estimation of non-Gaussian features which can be used for subsequent image analysis or for physical parameter inversion. A significant property is that the resulting features are model independent and so makes very few prior assumptions on the data. The image clustering example exhibited quite smooth results and image speckle was clearly reduced in the new feature space.

The second approach assumes a particular parametric model and then takes statistical inference to a new level in fully unsupervised, non-Gaussian image segmentation. By including goodness-of-fit testing, the clustering algorithm automatically determines the number of clusters that are statistically supported by the data, guarantees good clustering results, and requires no special initialisation. In addition, all model parameters are determined automatically within the algorithm, using recent advances in methods of matrix log-cumulants, and the only required input is the PolSAR image data itself, the choice of model, and the statistical confidence level for the tests.

The improved level of statistical confidence achieved with these methods now raises the questions of the physical interpretation of these statistically distinct clusters and the suitability of particular models, which only future research can answer.

7.2 Future Applications

This section lists several important questions that this work has raised that still need to be addressed. Note that some ideas have already been mentioned where relevant in Chapter 6.

The automatic clustering, or segmentation, still needs to be interpreted. The statistical analysis can only determine the statistically distinct clusters supported by the data, but does not actually classify the terrain land cover type for each group. This may be investigated by obtaining suitably detailed ground truth data to determine each class, or by investigating polarimetric decompositions [Cloude and Pottier, 1996,Freeman and Durden, 1998, van Zyl et al., 2008, Yamaguchi et al., 2008, Yamaguchi et al., 2005] to determine basic physical scattering mechanisms to aid class assignment.

In addition to polarimetric decompositions, the class texture parameter may be worth including in any class labelling strategy. This may involve building a library of polarimetric and textural feature signatures for potential land cover types to choose from. This will be complicated because texture is scale dependent and some classes, like urban (as noted in 2.3.6), are only relevant on certain scales. The multilook texture adjustment factor, from Paper 2, may prove useful in converting all textures to a common level.

Substitution of one, or more, matrix log-cumulants as the texture feature(s) should also be investigated, as these are also model independent, are asymptotically Gaussian, and detailed enough to separate different model families. This may be important given the recent observations from detailed goodness-of-fit testing.

In fact the whole model independent, feature space approach should be investigated further. Firstly, to see whether the features are normalised in any sense such their values are consistent from image to image, or for different sensors, because this would be advantageous for identifying land cover signatures. Secondly, to investigate whether the automatic number of clusters method can be applied to the feature space approach. This may require finding a suitable parametric model for the feature space vector, or sub-sampling sufficiently such that the multivariate Gaussian assumption would suffice. The latter did give quite good results for the (fixed number) clustering presented in Paper 1. Results from the two different approaches to clustering, feature space and model-based, could also be directly compared.

Since the K-Wishart model may not always be appropriate, it would be worthwhile, and quite easy, to adapt the mixture model algorithm to include a range of several models, i.e. the dictionary-based approach. Since submitting Paper 3, the G^0 model has already been included as a choice of model. The evidence is, however, that more flexibility than either the K-Wishart or G^0 is required. Therefore, either a more complicated, mutli-parameter model is required, such as the Kummer U-distribution [Bombrun et al., 2010], along with more complicated parameter estimation, or a cluster by cluster choice of the best fitting distribution may attain similar results.

It would be worthwhile to investigate alternative, probably two texture parameter models that cover more of the space of matrix log-cumulants observed in real images. The Kummer U-distribution [Bombrun et al., 2010] is one such model, but there may

be others. The Kummer U-distribution covers the manifold between the K-distribution and the G^0 distribution, but not below either of them. The tools are now at hand to really investigate the whole range of potential models in matrix log-cumulant space to choose amongst them. The experience with numerical evaluation of the Bessel K function, should also be taken into consideration for a suitable model.

In Section 2.2, several extensions to the current framework were mentioned. It would be interesting to see whether extending the model, perhaps to a multivariate texture in the first instance, can still be incorporated into the current statistical framework and thereby keep the good features like automatically determining the number of clusters. Similarly, would extending to contextual methods agree with or contradict the automatic number of classes technique?

Also mentioned in Section 2.2, was that topography has not yet been implemented. Full DEM radiometric terrain correction should be included and the suitability of the modelling would need to be re-confirmed.

Finally, these techniques should be applied to specific application tasks to see if they are beneficial to general research. The author did spend ten months at the University of Alaska, Fairbanks to study sea ice for this purpose, but found that all of the PolSAR images acquired were severely corrupted with radio frequency interference and were unsuitable for classification. This has recently been corrected [Meyer and Doulgeris, 2010], and a future project will certainly look at sea ice classification with the automatic algorithm.

Bibliography

[Andrews and Mallows, 1974] A. F. Andrews and C. L. Mallows. *Scale mixtures of normal distributions*. *Journal of the Royal Statistical Society. Series B*, **36**: 99–102, no. 1 1974.

[Anfinsen et al., 2009] S. Anfinsen, T. Eltoft and A. Doulgeris. *A relaxed Wishart model for polarimetric SAR data*. In *Proc. 4th Int. Workshop on Science and Applications of SAR Polarimetry and Polarmietric Interferomentery (POLinSAR2009), Frascati, Italy, 26-30 January*, ESA SP-668, p. 8 pp. April 2009.

[Anfinsen, 2010] S. N. Anfinsen. *Statistical Analysis of Multilook Polarimetric Radar Images with the Mellin Transform*. Ph.D. thesis, University of Tromsø, Tromsø, Norway, May 2010.

[Anfinsen and Eltoft, 2011] S. N. Anfinsen and T. Eltoft. *Application of the Matrix-Variate Mellin Tranform to Analysis of Polarimetric Radar Images*. *IEEE Trans. Geoscience and Remote Sensing*, **49**(7): 15 pp., in press, July 2011.

[Anfinsen et al., 2011] S. N. Anfinsen, A. P. Doulgeris and T. Eltoft. *Goodness-of-Fit Tests for Multilook Polarimetric Radar Data Based on the Mellin Transform*. *IEEE Trans. Geoscience and Remote Sensing*, **49**(8): 18 pp., in press, August 2011.

[Barndorff-Nielsen et al., 1982] O. Barndorff-Nielsen, J. Kent and M. Sorensen. *Normal variance-mean mixtures and z distributions*. *International Statistical Review / Revue Internationale de Statistique*, **50**(2): 145–159, 1982.

[Barndorff-Nielsen, 1997] O. E. Barndorff-Nielsen. *Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling*. *Scand. J. Statist.*, **24**: 1–13, 1997.

[Bombrun and Beaulieu, 2008] L. Bombrun and J.-M. Beaulieu. *Fisher distribution for texture modeling of polarimetric sar data*. *Geoscience and Remote Sensing Letters, IEEE*, **5**(3): 512 –516, July 2008.

[Bombrun et al., 2010] L. Bombrun, G. Vasile, M. Gay and F. Totir. *Hierarchical segmentation of polarimetric sar images using heterogeneous clutter models*. *Geoscience and Remote Sensing, IEEE Transactions on*, pp. 1 –12, 2010.

[Cao et al., 2007] F. Cao, W. Hong, Y. Wu and E. Pottier. *An unsupervised segmentation with an adaptive number of clusters using the SPAN/H/α/A space and the complex wishart clustering for fully polarimetric SAR data analysis*. *IEEE Trans. Geoscience and Remote Sensing*, **45**(11): 3454–3467, November 2007.

[Cloude and Pottier, 1996] S. R. Cloude and E. Pottier. *A Review of Target Decomposition Theorems in Radar Polarimetry*. *IEEE Transactions on Geoscience and Remote Sensing*, **34, 2**: 498–518, March 1996.

[Cumming and Wong, 2005] I. G. Cumming and F. H. Wong. *Digital Processing of Synthetic Aperture Radar Data*. Artech House, 2005.

[Dempster et al., 1977] A. P. Dempster, N. M. Laird and D. B. Rubin. *Maximum likelihood from incomplete data via the EM algorithm*. *Journal of the Royal Statistical Society. Series B*, **39**(1): 1–38, 1977.

[Dierking et al., 2003] W. Dierking, H. Skriver and P. Gudmandsen. *SAR Polarimetry for Sea Ice Classification*. In *ESA SP-529: Applications of SAR Polarimetry and Polarimetric Interferometry*. April 2003.

[Doulgeris and Eltoft, 2007] A. Doulgeris and T. Eltoft. *Scale mixture of Gaussians modelling of polarimetric SAR data*. In *International POLinSAR Workshop (POLinSAR2007)*. Frascati, Italy, January 22-26 2007.

[Doulgeris et al., 2008] A. Doulgeris, S. Anfinsen and T. Eltoft. *Classification with a non-gaussian model for polsar data*. *IEEE Trans. Geoscience and Remote Sensing*, **46**(10): 2999–3009, Oct. 2008.

[Doulgeris et al., 2009] A. Doulgeris, S. Anfinsen, Y. Larsen, K. Langley and T. Eltoft. *Evaluation of polarimetric configurations for glacier classification*. In *International POLinSAR Workshop (POLinSAR2009)*. Frascati, Italy, January 26-30 2009.

[Doulgeris, 2006] A. P. Doulgeris. *Statistical Modelling of Polarimetric SAR Data*. Master's thesis, University of Tromsø, May 2006.

[Doulgeris and Eltoft, 2010a] A. P. Doulgeris and T. Eltoft. *Automated Non-Gaussian clustering of polarimetric SAR.*. In *8th European Conference on Synthetic Aperture Radar (EUSAR2010)*. Aachen, Germany, June 7-10 2010a.

[Doulgeris and Eltoft, 2010b] A. P. Doulgeris and T. Eltoft. *Scale mixture of gaussian modelling of polarimetric sar data*. *EURASIP Journal on Advances in Signal Processing*, **2010**(874592): 12, 2010b.

[Doulgeris et al., submitted 2010] A. P. Doulgeris, S. N. Anfinsen and T. Eltoft. *Automated non-Gaussian clustering of polarimetric synthetic aperture radar images*. *IEEE Trans. Geoscience and Remote Sensing*, submitted 2010.

[Eltoft, 2005] T. Eltoft. *The Rician inverse Gaussian distribution: A new model for non-Rayleigh signal amplitude statistics*. *IEEE Trans. Image Process.*, **14**: 1722–1735, November 2005.

[Eltoft et al., 2006] T. Eltoft, T. Kim and T.-W. Lee. *Multivariate scale mixture of gaussians modeling*. In *Independent Component Analysis and Blind Signal Separation*, edited by J. Rosca, D. Erdogmus, J. Príncipe and S. Haykin, volume 3889 of *Lecture Notes in Computer Science*, pp. 799–806. Springer Berlin / Heidelberg, 2006. 10.1007/11679363-99.

[Eltoft et al., 2006] T. Eltoft, T. Kim and T.-W. Lee. *Multivariate scale mixture of Gaussians models*. In *Proceedings of ICA 2006, Charleston, SC, USA*. March 2006.

[Folks and Chhikara, 1978] J. Folks and R. Chhikara. *The inverse Gaussian distribution and its application – A review*. *J. Roy. Statist. Soc. B*, **40**: 263–289, 1978.

[Freeman and Durden, 1998] A. Freeman and S. L. Durden. *A Three-Component Scattering model for Polarimetric SAR Data*. *IEEE Transactions on Geoscience and Remote Sensing*, **36, 3**: 963–973, May 1998.

[Freitas et al., 2001] C. Freitas, S. Sant'Anna, L. S. Soler, J. R. Santos, L. V. Dutra, L. S. de Araujo, J. C. Mura and P. Hernandez Filho. *The use of airborne p-band radar data for land use and land cover mapping in brazilian amazonia*. In *IEEE International Geoscience and Remote Sensing Symposium, 2001 (IGARSS '01)*, volume 4, pp. 1889 –1891. 2001.

[Freitas et al., 2005] C. C. Freitas, A. C. Frery and A. H. Correia. *The polarimetric G distribution for SAR data analysis*. *Environmetrics*, **16**: 13–31, 2005.

[Frery et al., 2007] A. C. Frery, A. H. Correia and C. d. C. Freitas. *Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information*. *IEEE Trans. Geoscience and Remote Sensing*, **45**(10): 3098–3109, October 2007.

[Frery et al., 2010] A. C. Frery, J. Jacobo-Berlles, J. Gambini and M. E. Mejail. *Polarimetric sar image segmentation with b-splines and a new statistical model*. *Multidimensional Syst. Signal Process.*, **21**: 319–342, December 2010.

[Fukuda et al., 1999] S. Fukuda, K. Suwa and H. Hirosawa. *Texture and statistical distribution in high resolution polarimetric sar images*. In *Geoscience and Remote Sensing Symposium, 1999. IGARSS '99 Proceedings. IEEE 1999 International*, volume 2, pp. 1268 –1270. 1999.

[Goodman, 1976] J. W. Goodman. *Some fundamental properties of speckle*. *J. Opt. Soc. Am.*, **66, 11**: 1145–1150, November 1976.

[Goodman, 1963] N. Goodman. *Statistical analysis based on certain multivariate complex gaussian distribution*. In *Ann. Math. Statist.*, volume 34, pp. 152–177. 1963.

[Jakeman and Pusey, 1976] E. Jakeman and P. N. Pusey. *A model for non-Rayleigh sea echo*. *IEEE Trans. Antennas Propagat.*, **24, 6**: 806–814, November 1976.

[Jakeman and Pusey, 1978] E. Jakeman and P. N. Pusey. *Significance of k distributions in scattering experiments*. *Phys. Rev. Lett.*, **40**(9): 546–550, Feb 1978.

[Jakeman and Tough, 1987] E. Jakeman and R. J. A. Tough. *Generalized K distribution: a statistical model for weak scattering*. *J. Opt. Soc. Am. A*, **4, 9**: 1764–1772, September 1987.

[Kuruoglu and Zerubia, 2004] E. E. Kuruoglu and J. Zerubia. *Modeling sar images with a generalization of the rayleigh distribution*. *Image Processing, IEEE Transactions on*, **13**(4): 527 – 533, April 2004.

[Lee et al., 1994a] J. Lee, M. Grunes and R. Kwok. *Classification of multi-look polarimetric SAR imagery based on the complex Wishart distribution*. In *Int. J. Remote Sensing*, volume 15. 11 1994a.

[Lee et al., 2004] J. Lee, M. Grunes, E. Pottier and L. Ferro-Famil. *Unsupervised terrain classification preserving polarimetric scattering characteristics*. *Transactions on Geoscience and Remote Sensing*, **42**(4): 722–731, April 2004.

[Lee et al., 1994b] J. S. Lee, D. L. Schuler, R. H. Lang and K. J. Ranson. *K-distribution for multi-look processed polarimetric sar imagery*. In *IEEE Int. Geosci. Remote Sensing Symp.*, pp. 2179–2181. 1994b.

[Lombardo et al., 2001] P. Lombardo, D. Pastina and T. Bucciarelli. *Adaptive polarimetric target detection with coherent radar. ii. detection against non-gaussian background*. *Aerospace and Electronic Systems, IEEE Transactions on*, **37**(4): 1207 –1220, October 2001.

[Lopes and Sery, 1997] A. Lopes and F. Sery. *Optimal speckle reduction for the product model in multilook polarimetric sar imagery and the wishart distribution*. *Geoscience and Remote Sensing, IEEE Transactions on*, **35**(3): 632–647, May 1997.

[Lumsdon et al., 2005] P. Lumsdon, S. Cloude and G. Wright. *Polarimetric classification of land cover for glen affric radar project*. *Radar, Sonar and Navigation, IEE Proceedings -*, **152**(6): 404 – 412, December 2005.

[Mardia, 1970] K. V. Mardia. *Measure of Multivariate Skewness and Kurtosis with Applications*. *Biometrica*, **57, 3**: 519–530, December 1970.

[Meyer and Doulgeris, 2010] F. J. Meyer and A. P. Doulgeris. *Interference Signatures in ALOS PALSAR Imagery in the American Arctic*. In *CEOS SAR Cal/Val Workshop, Zürich, Switzerland, August 25-27*. 2010.

[Moser et al., 2006] G. Moser, J. Zerubia and S. B. Serpico. *Dictionary based stochastic expectation-maximization for sar amplitude probability density function estimation*. *IEEE TGRS*, **44**(1): 88–200, January 2006.

[Nicolas, 2002] J.-M. Nicolas. *Introduction aux statistique de deuxième espèce: Application des logs-moments et des logs-cumulants à l'analyse des lois d'images radar*. *Traitement du Signal*, **19**(3): 139–167, 2002. In French.

[Oliver and Quegan, 2004] C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar Images*. SciTech Publishing, Raleigh, USA, second edition, 2004.

[Quegan and Rhodes, 1993] S. Quegan and I. Rhodes. *Statistical models for polarimetric sar data*. In *Texture analysis in radar and sonar, IEE Seminar on*, pp. 8/1 –8/8. 1993.

[Quegan and Rhodes, 1995] S. Quegan and I. Rhodes. *Statistical models for polarimetric data: consequences, testing and validity*. *International Journal of Remote Sensing*, **16**: 1183–1210, May 1995.

[Quegan et al., 1994] S. Quegan, I. Rhodes and R. Caves. *Statistical models for polarimetric sar data*. In *Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, volume 3, pp. 1371 –1373 vol.3. August 1994.

[Quegan et al., 2003] S. Quegan, T. Le Toan, H. Skriver, J. Gomez-Dans, M. C. Gonzalez-Sampedro and D. H. Hoekman. *Crop Classification with Multitemporal Polarimetric SAR Data*. In *Applications of SAR Polarimetry and Polarimetric Interferometry*, volume 529 of *ESA Special Publication*. April 2003.

[Salvador and Chan, 2004] S. Salvador and P. Chan. *Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms*. In *Tools with Artificial Intelligence, 2004. ICTAI 2004. 16th IEEE International Conference on*, pp. 576 – 584. nov. 2004.

[Sery and Lopes, 1997] F. Sery and A. Lopes. *Statistical properties of speckle and full polarimetric filters in sar*. In *Geoscience and Remote Sensing, 1997. IGARSS '97. Remote Sensing - A Scientific Vision for Sustainable Development., 1997 IEEE International*, volume 2, pp. 761 –763 vol.2. August 1997.

[Singh and Venkataraman, 2008] G. Singh and G. Venkataraman. *Los palsar data analysis of snow cover area in himalayan region using four component scattering decomposition technique*. In *Recent Advances in Microwave Theory and Applications, 2008. MICROWAVE 2008. International Conference on*, pp. 772 –774. November 2008.

[Skriver et al., 2005] H. Skriver, J. Dall, L. Ferro-Famil, T. Le Toan, O. Lumsdon, R. Moshammer, E. Pottier and S. Quegan. *Agriculture classification using POLSAR data*. In *International POLinSAR Workshop*. Frascati, Italy, January 2005.

[Thomsen et al., 1998] B. Thomsen, S. Nghiem and R. Kwok. *Polarimetric c-band sar observations of sea ice in the greenland sea*. In *Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98. 1998 IEEE International*, volume 5, pp. 2502 –2504 vol.5. July 1998.

[Tison et al., 2004] C. Tison, J.-M. Nicolas, F. Tupin and H. Maître. *A new statistical model for Markovian classification of urban areas in high-resolution SAR images*. *IEEE Trans. Geosci. Remote Sens.*, **42**(10): 2046–2057, October 2004.

[Tough et al., 1995] R. J. A. Tough, D. Blacknell and S. Quegan. *A statistical description of polarimetric and interferometric synthetic aperture radar data*. *Proc. Mathematical and Physical Sciences*, **449**(1937): 567–589, June 1995.

[Touzi et al., 2004] R. Touzi, W. Boerner, J. Lee and E. Lueneburg. *A review of polarimetry in the context of synthetic aperture radar: concepts and information extraction*. *Can. J. Remote Sensing*, **30**(3): 380–407, 2004.

[Ulaby and Elachi, 1990] F. T. Ulaby and C. Elachi. *Radar Polarimetry for Geoscience Applications*. Artech House, 1990.

[Ulaby et al., 1982] F. T. Ulaby, R. K. Moore and A. K. Fung. *Mircrowave Remote Sensign: Active and Passive, Vol. II – Radar Remote Sensing and Surface Scattering and Emission Theory*. Addison-Wesley, Advanced Book Program, Reading, Massachusetts, 1982.

[Valenzuela, 1978] G. R. Valenzuela. *Theories for the interaction of electromagnetic and oceanic waves — a review*. *Boundary-Layer Meteorology*, **13**: 61–85, 1978. 10.1007/BF00913863.

[Vasile et al., 2010] G. Vasile, J.-P. Ovarlez, F. Pascal and C. Tison. *Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric sar images*. *Geoscience and Remote Sensing, IEEE Transactions on*, **48**(4): 1809 –1826, April 2010.

[Yamaguchi et al., 2005] Y. Yamaguchi, T. Moriyama, M. Ishido and H. Yamada. *Four-component scattering model for polarimetric sar image decomposition*. *Geoscience and Remote Sensing, IEEE Transactions on*, **43**(8): 1699 – 1706, August 2005.

[Yamaguchi et al., 2008] Y. Yamaguchi, J. Nakamura, K. Aoyama and H. Yamada. *Coherent decomposition of fully polarimetric radar data*. In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, volume 4, pp. IV –161 –IV –164. July 2008.

[Yueh et al., 1989] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin and L. M. Novak. *K-Distribution and polarimetric terrain radar clutter*. *J. Electro. Waves Applic.*, **3**: 747–768, 1989.

[van Zyl et al., 2008] J. van Zyl, Y. Kim and M. Arii. *Requirements for model-based polarimetric decompositions*. In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, volume 5, pp. V –417 –V –420. July 2008.

ISBN 978-82-8236-022-7