University College London

Coherent Radar Clutter Statistics

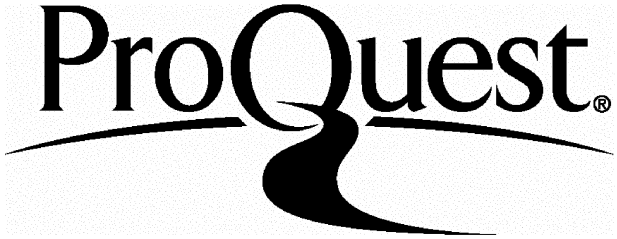
Ph.D. Thesis

Mohammed Jahangir

January 2000

This thesis is submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy



Abstract

Coherent radar clutter statistics

Mohammed Jahangir

Ground surveillance radars have been used for a number of years for target detection and terrain mapping. Neither of these tasks can be carried out without a proper understanding of the ground clutter statistics. Traditionally, radar signal processing has been optimised for Gaussian interference. However, ground clutter in high resolution systems is rarely Gaussian. Thus in this thesis the non-Gaussian statistics of a coherent radar sensor are investigated.

The clutter statistics are a function of both the terrain features and the imaging process. One of the major influencing factors is the radar resolution cell size. Thus as a prerequisite to studying the clutter characteristics the theory concerning radar resolution is discussed.

A review of the non-Guassian clutter models is carried out and the K-distribution is selected as a reasonable model for land clutter statistics based on empirical evidence. However, a major drawback with the K-distribution is a lack of efficient optimum estimators for its parameters.

Neural nets are devised to give an optimum estimator for the order parameter of the K-distribution. The minimum error is obtained with a net trained on the mean normalised log intensity and the amplitude contrast. A new estimator is proposed based on the multiple moments model. It is shown that nearly optimum performance is achieved with the new estimator.

The parameter estimators are also investigated for a more general form of the K-distribution where the data statistics are such that a simple K-distribution is no longer sufficient to describe the clutter statistics. Estimators are proposed for the parameter values of generalised-K and homodyned-K based on a large order parameter approximation. Error analysis is carried out and a comparison is made between theoretical and simulated results.

Contents

Acknowledgements

I am indebted to a number of people for their kind help and assistance without which it would not have been possible to complete this thesis. Most of all I would like to thank Dr. Chris Baker and Dr. Richard White for firstly offering me the research scientist post at the Defence Evaluation Research Agency and secondly encouraging me to undertake the external PhD at University College London. I am grateful to Prof. Hugh Griffiths for taking me as a research student. I thank him for all the guidance I received in my research work and the time and effort he devoted to reviewing my written work. I especially thank him for the suggestion to extend the parameter analysis work to weak scattering models. During the course of this research extensive use was made of the neural net code developed by Martin Warner and the image interpretation language, Imanip, developed by Alan Blake at DERA Malvern to whom I would like to extend my thanks. I am also especially grateful to Dr. David Blacknell for agreeing to be my second supervisor, he has been very generous with his time and offered tremendous guidance through out. I would also like to acknowledge the use of his Mathematica perturbation routine which was modified to carry out the small perturbation results presented in this thesis. Thanks are also extended to Prof. Chris Oliver of DERA Malvern for useful comments and discussions throughout the course of this study. I would also like to express by thanks to Prof. Pier Lombardo of Rome University, Prof. Maurizio Di Bisceglie of Naples University and Dr Robert Tough of TW Research whose advice was sought at various stages. Finally, I wish to thank my friends for general encouragement and my wife Sayqa for endeavouring with me through the seemingly endless task.

Chapter 1

1. Introduction

The Concise Oxford Dictionary gives the meaning of clutter as 'crowded confusion' or 'untidy state'. In a radar context clutter refers to unwanted signals - signals which confuse the detection of objects of interest. In ground surveillance radar the main objects of interest for a military application are the man-made vehicles. The radar is being used for target detection and in this situation the signals from sea, ground, buildings and vegetation are considered as clutter. However, increasingly radar is being used for terrain mapping and, in this remote sensing application, the signals from land and sea are the ones which are of interest. Nevertheless in this thesis the term clutter will refer to the radar returns from land and sea irrespective of the application. The question arises why there is a requirement for studying clutter characteristics. The answer to this very much depends upon the application of the radar. Thus there are essentially two reasons for studying clutter which follows from the two distinct applications of the ground surveillance radar, i.e. target detection and terrain mapping. A review of these two applications of ground surveillance radar would give an appreciation for studying clutter characteristics.

1.1 Importance of understanding clutter

The principle of a radar is fairly straightforward. Radar is an active sensor whereby a signal is transmitted and the energy backscattered from an object is received by the radar antenna. The nature of this signal provides information regarding the object. The delay between the transmitted pulse and the received echo from the object relates to its range. The strength of the received echo relates to the object size, orientation and electromagnetic properties and is an indicator for the object

type. The use of a directional antenna allows the radar to sense the angle of arrival of the received echo and this identifies the angular location of the target.

From a target detection consideration the simplest case is where the target, the object of interest, is in an isolated location, with no interfering clutter. This situation is typical of airborne targets. During the second world war this usage of the radar was first brought into service and deployed in air defence systems for detecting attacking aircraft [65]. These original radar systems were noncoherent, in the sense that they transmitted a burst of Radar Frequency (RF) energy and detected the amplitude of that proportion of the signal that came back from a target of interest. A receiver contains other spurious signals as well which are background electrical noise which invariably exists in the output of a receiver. The echo received from the target will thus only be detected once it can be discerned above this background electrical noise. On a radar video image such target returns show up as bright spots or 'blips' which are recognised by a trained human operator. In sophisticated radars, detection is performed automatically. The received signal is compared against a threshold and a target is registered only when it exceeds this threshold. The system will have some undetected targets because of the target signal being too low and it will also have some false alarms due to the noise signal exceeding the threshold level. The threshold is set according to the background signal level. The criteria for the signal processing are to minimise the false alarm rate and maximise the detection probability. When the target is just competing with the noise background it is easier to set this threshold.

The task becomes more difficult when the target also has to compete with simultaneous received echoes from interfering objects like ground. These clutter returns may in some instances be stronger than the noise signal and the detection performance drops. Since clutter is always present for ground targets and for low flying targets there could be clutter at the same range as the target, this would seriously limit the effectiveness of the radar system. Modern radars overcome this limitation by using a coherent system where the detection is based not only on the amplitude of the signal, but on its phase characteristics as well. The phase of the signal changes linearly with target distance, providing a means to measure the ve-

locity of the target. When the target is moving the received, echo signal is shifted in frequency due to the target's Doppler effect. This enables a radar to separate the signals from the desired moving target (such as vehicles) from stationary clutter even though the stationary echo signal may be many orders of magnitude greater than the moving target. Thus coherent pulse radars use Doppler filtering to detect moving targets; a process that is termed Moving Target Indication (MTI). Ground based systems use this method to detect both airborne and ground based moving targets. For fixed antenna all the returns from stationary clutter will be in zero Doppler and the Doppler returns from moving targets will be generally free of clutter interference. The detection threshold can, therefore, be set using simplified algorithms based on thermal noise limited data.

The clutter returns, however, do not always reside at zero Doppler. The clutter can have real Doppler or apparent Doppler. Real Doppler is when the clutter has intrinsic motions. The internal clutter motions are mainly wind induced and this produces only limited Doppler spread. Apparent Doppler is due to one of two reasons. A scanning antenna gives the clutter an apparent Doppler. The clutter Doppler varies linearly along the aperture resulting in a Doppler spread. The extent of the spread is dependent on the scan rate and the antenna beamwidth. Secondly, a moving platform, as in the case of an airborne or spaceborne radar system, introduces an apparent Doppler for the clutter. The apparent Doppler is due to the component of platform motion in the direction of the clutter. Since the radar beam has a finite width there is a spread in the Doppler frequencies of the clutter return. The extent of this spread depends up on the platform velocity and also on the antenna pointing angle in relationship to the platform velocity vector. The widening of the clutter band can be quite extensive due to the high speeds of moving platforms. Where the clutter Doppler spread is large, the target will only clear the clutter dominant region of the Doppler spectrum if it has a significant Doppler shift. In the case of the airborne targets where the velocities are relatively large the targets have a significant possibility of residing in the clutter free region of the Doppler spectrum. However, ground moving targets are much slower and are far more likely to fall within the Doppler spread of the clutter. Traditionally, in radar

image processing, regions which are clutter dominated are discarded [82]. This results in a failure to detect any of the slow moving targets. Therefore, the airborne surveillance of ground targets has had limited success to date despite the strategic importance of such systems. In order to ensure that slow moving targets can be detected, the target detection has to be carried out in a clutter dominated background. This is not possible without a full understanding of the statistical properties of clutter. An improved understanding of the clutter characteristics will allow us to set appropriate detection threshold in clutter dominated regions hence extending the range of velocities at which a target can be detected. Considerable work has been done on understanding the characteristics of land and sea clutter. A number of statistical models have been put forward for the clutter data [17, 41, 52, 74, 75, 82]. The bulk of this work has been performed by empirical fitting of distribution to measured data [13, 16, 27, 86, 88, 94]. However, this work has to be put on a firmer physically based model.

The understanding of clutter characteristics also plays a vital role in interpreting images of ground-mapping radar. These radars are generally side-looking. During flight the radar radiates a microwave beam typically at right angles to the aircraft, then detects and records the reflections received from the earth's surface. As it travels forwards, successive strips of terrain are exposed to the radar beam and are detected at the aircraft. The reflections are used to produce a map of the terrain covered. In these images the clutter is now the object of interest.

There is continuous push to produce maps with finer resolution. Improved range resolution is achieved using a narrower pulse. Improved azimuth resolution in conventional radar was achieved using a narrower beam either by employing a larger antenna or using shorter wavelength. In modern pulsed Doppler radar, azimuth resolution is achieved by deploying a technique called Synthetic Aperture Radar (SAR), whereby the aircraft forward motion is utilised to synthesize a very long antenna. This method produces near photographic quality images. High resolution radar mapping has proved particularly useful in areas that are unmapped, poorly known, sparsely populated, intensively cloud covered and in areas that otherwise have not lent themselves to traditional aerial mapping methods.

In order to extract useful information about the scene from such images post processing tasks such as image segmentation and classification are carried out. These techniques are based upon the statistics of the underlying cross section values of the clutter [60]. An added complication with the coherent imaging process is that there is speckle noise present. Speckle is random noise which results from the interference of the coherent radar echo with the surface scatterer. This speckle, which gives a granular appearance to a SAR image, seriously degrades the image quality. Therefore, both the statistical variation in the underlying cross section values of the background along with the statistical properties of the speckle has to be taken into account for the proper modelling of the observed data. This type of coherent clutter statistics analysis is not unique to the radar sensor but other coherent imaging processes like sonar and laser have similar statistical behaviour and an identical approach can be applied to understanding the clutter statistics from such a scenario.

We have therefore established that studying the clutter properties is a vital link in a) the detection of radar targets and b) the understanding and interpretation of images of coherent systems. In this thesis we address this task of characterising the clutter model. The discussion will be based on a coherent radar system but the modelling technique is a generic approach, applicable to other coherent imaging methods. One particular example of this is coherent sonar systems where clutter modelling techniques, developed for radar systems, are increasingly being applied.

The clutter characterisation problem can be broken down into a number of sub-questions:

1. What parameterised statistical model can be used to characterise the clutter?
2. What physical justification is there to use the statistical models?
3. What are the optimal parameter solutions to the statistical model?
4. What are the errors associated with estimating these parameter values?

These questions comprise the main aspects of the clutter characterisation problem addressed in the following chapters.

In the remainder of this introductory chapter an outline plan of the thesis is presented, the novel aspects of the research presented in this thesis are listed and a brief description is given of the radar data that will be used in later chapters.

1.2 Outline

Before we can address the task of understanding the clutter characteristics we need to establish a full understanding of the imaging process. We need to know what the parameters of the radar systems are and what information is represented by the data. This is a general prerequisite for interpreting any type of data. Hence in Chapter 2 the general theory of radar imaging is discussed. A pulsed Doppler system and a high resolution SAR system is described, discussing in particular the image resolution achieved with each system. The information concerning the resolution of the system will be used in later chapters to devise models for the clutter statistics.

After a brief explanation of the radar principles we go on to address the task of clutter characterisation covered by the questions highlighted above. The first two of these questions are addressed in Chapter 3. A general review is carried out of the clutter modelling techniques used by various researchers. A short background is given on direct modelling techniques based on the electromagnetic theory followed by a detailed discussion of statistical based models used to describe the clutter statistics. The search for a proper phenomenological based model leads us to the K-distribution as an adequate representation of the statistics of medium resolution land clutter data. Evidence from goodness-of-fit tests on SAR and MTI data is used to justify the K-distribution model. The discussion on the K-distribution is extended in Chapter 4 where we address the third and fourth question on clutter modelling regarding parameter estimation. Neural nets are used to investigate optimal parameter estimators for the K-distribution. A new multi-moment estimator is proposed for the estimation of the parameter of the K-distribution and its error analysis is carried out. It is shown that near optimal parameter estimates are obtained with this hybrid estimator.

In Chapter 5 we extend our discussion to high resolution clutter where the statistics deviate from a K-distribution. Once again we address the four questions as to what is the most suitable noise model and how to obtain the optimal parameter estimators. We introduce weak scattering models which are a more appropriate representation of the scattering process when the phase of the detected field is not uniformly distributed. Consideration of a random walk in a weak scattering regime leads us to the generalised-K and the homodyned-K distribution. Estimators are proposed for the parameters of the distributions and error analysis is carried out for the estimators.

Chapter 6 contains conclusions on the clutter statistics given the results in Chapter 3 to 5. It also suggests ideas for further research.

A number of appendices are included. Appendix A lists a complete evaluation of the predicted errors for the various estimators discussed in this thesis. Appendix B contains an extended proof of a result referred to in the main body of the thesis. The details of the chi-squared goodness-of-fit test and the maximum likelihood classification test used in the evaluation of real radar data are described in appendices C and D respectively. Appendix E lists the output distributions of basic binary operators in terms of the distributions of their inputs; these results are required at various stages in the main body of the thesis. Appendix F is a glossary which briefly defines various mathematical functions and relationships referred to in the thesis. For ease of reference topics listed in the glossary are printed in **bold** in the main body of the thesis wherever the reader may require further explication of the topic (e.g., where it is first introduced). Finally, appendix G gives a list of symbols used in the thesis.

1.3 Novel aspects of the research

The main achievement of the results presented in this thesis is the development of suitable statistical models for surface clutter. New estimators for the most viable clutter distribution models are proposed and their performance assessed using theoretical predictions and simulated results. This work, on the main, built on the many years of research on sea and land surface scattering carried out by the Sur-

face Surveillance group at DERA Malvern but also drew upon the work of the wider radar clutter modelling research community. There are a number of original results presented in this thesis and these are listed as follows

- Using new SAR data for land, the K-distribution is shown to be a good empirical model for medium resolution system which confirmed previously published results.
- New results are presented for the empirical fit to MTI data for land clutter which for the first time showed that the K-distribution is also a suitable non-Gaussian model for this case.
- Neural nets are used in a novel way to construct estimators for the t parameter of the K-distribution which outperformed existing texture measure estimators. A key element to the success of the neural net estimators is the introduction of the pre-processing stage. Thus instead of presenting directly the K-distributed intensity data, moments calculated from this data are presented to the input layer of the neural net which enables a much more successful convergence of the neural nets' weights.
- A new hybrid estimator is proposed which combines the normalised log and the amplitude contrast texture measures to give a near optimal estimator for the t parameter of the K-distribution. A new expression is derived for the predicted error performance of this estimator which is compared with simulated results.
- The homodyned-K and the generalised-K distribution, which are candidate distribution for weak scattering, lacked any optimal parameter estimators. New results are presented which propose some simple estimators for the parameters of these distributions. The corresponding expressions for the error predictions are derived and their performance assessed using simulations.

1.4 Real radar data

Results presented in this thesis are based on a number of data sets from DERA Malvern airborne radars:

Band	Allocation	
	Frequency GHz	Wavelength cm
C	5.25-5.35	5.71-5.62
X	9.50-9.80	3.16-3.06

Table 1.1 Frequency band allocations for radar remote sensing [89].

1. Single-look complex SAR data taken by the DERA Malvern Canberra X-band airborne system near Amesbury, Wiltshire, England.
2. Single polarisation MTI data taken by the DERA Malvern Canberra X-band airborne system of the Hull area, England.
3. Single-look complex SAR data taken by the DERA Malvern Andover C-band airborne system near Pershore, Worcestershire, England.

The frequency allocations corresponding to the radar frequency bands referred to above are listed in Table 1.1. Other relevant imaging parameters will be listed when the need arises.

Chapter 2

2. Radar resolution theory

Radar resolution has a significant bearing on the statistics of the clutter returns. Thus in this chapter, as a prerequisite to studying clutter characteristics, the theory of radar resolution for ground imaging surveillance systems is described. For modern surveillance systems there has been continuous pressure and desire to achieve finer resolution. There are a number of techniques available for improving radar resolution. An airborne system has certain limitations but there are also particular features which can be exploited to improve radar resolution. The high resolution techniques which are applicable to airborne surveillance systems are described in this chapter.

The resolution of a system can be defined as a length scale which characterises how far apart two point targets must be for them to be distinguished as such in an image. A common choice for the resolution is the width of the main lobe of the point spread function (PSF) [70]. However, this is only a nominal resolution as the actual ability to resolve two scatterers depends upon their relative cross-section and phase difference [69]. Nevertheless, for the purpose of the discussion here we will treat this nominal resolution as the system resolution.

Figure 2.1 shows a typical geometry for a sideways-looking radar mounted on an airborne platform. The aircraft flies along a straight line track. The direction of flight is known as the along-track or azimuth direction and the direction perpendicular to this is known as the cross-track or range direction. As the aircraft flies along its track, it emits a continuous train of radar pulses. Each radar pulse illumi-

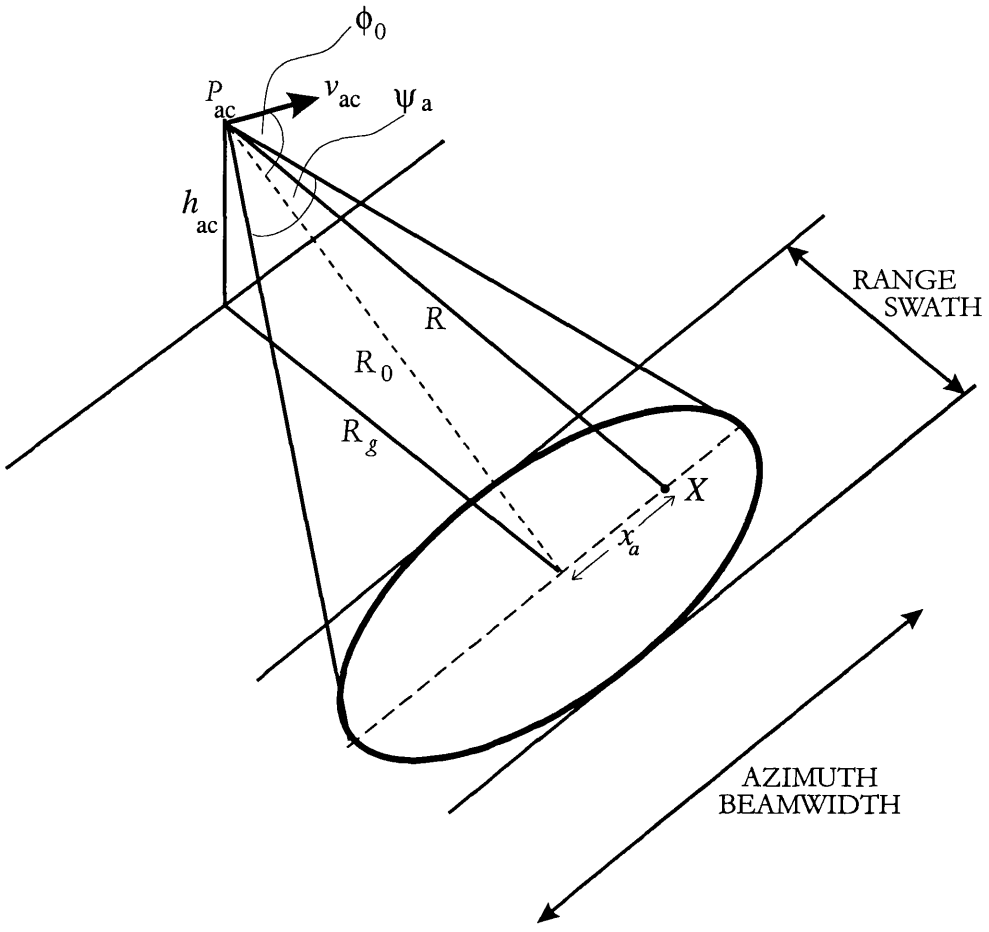


Figure 2.1: Geometry of an airborne sideways-looking radar

nates the ground over a region determined by the antenna beamwidth. As a rule of thumb, the beamwidth of an antenna of length D is [83]

$$\psi = \lambda / D \tag{2.1}$$

where λ is the radar wavelength. Typically the antenna is rectangular with dimensions $D_a \times D_e$, where a and e denote azimuth and elevation, respectively. Thus the antenna footprint has both a range and azimuth extent given by (2.1). These two spatial components make up the two-dimensional radar image. The range resolution of a sideways-looking radar is discussed in Section 2.1. The range resolution is inversely proportional to the bandwidth of the radar pulse. High range

resolution requires larger bandwidth which can be achieved by using frequency modulation. When linear modulation (i.e., a chirp pulse) is used the system response is described by a *sinc* function, provided no amplitude taper is used. The azimuth resolution for a real aperture antenna is described in Section 2.2. The azimuth resolution is inversely proportional to the aperture of the antenna expressed as number of wavelengths. Higher azimuth resolution is achieved using SAR techniques. The azimuth resolution for a fully focused synthetic aperture is equal to half the real aperture length, and the system response is described by a *sinc* function.

An alternative to the two-dimensional spatial image is to represent the received signal as a range Doppler map. The Doppler data is obtained by a frequency transformation of the azimuth data. The Doppler of a target is proportional to its radial velocity towards the radar. This feature is exploited in coherent MTI radars to separate the echoes of moving targets from stationary clutter. The clutter echoes also have a Doppler return determined by the platform speed, the antenna beamwidth and the antenna pointing direction. Section 2.3 discusses the Doppler resolution in an MTI system. The Doppler resolution is equal to the reciprocal of the coherent integration time. The total width of the Doppler spectrum is given by the Pulse Repetition Frequency (PRF).

The imaging geometry will assume a flat earth model. This is valid for airborne surveillance radar operating over a narrow swath width. However, surveillance systems which operate at a very long range, for example spaceborne surveillance radars, the imaging geometry must also take into account the Earth's curvature and rotation, and the satellite orbit [25, 66]. However, these factors are not considered further since they affect only the details of the range and azimuth processing rather than the general theory. The effect of range curvature [90] will also be ignored.

2.1 Range resolution

The range component of a radar image is given by the time it takes a radar pulse to return from the target when it is broadside. This is termed the slant range. Figure 2.1 shows the range swath of a sideways-looking radar.

For a pulsed radar the width of the radar pulse determines the ability of a radar to resolve closely spaced targets in range. Thus for a radar with a pulse width τ_p the slant range resolution is given as

$$d_{sr} = \frac{c\tau_p}{2} \tag{2.2}$$

where c is the speed of electromagnetic propagation through the atmosphere ($c \approx 3\times10^8$ m/s). For an unmodulated pulse, τ_p is approximately related to the bandwidth B of the radar as $\tau_p \approx 1/B$. Therefore, the slant range can be written as

$$d_{sr} = \frac{c}{2B} \tag{2.3}$$

For clutter modelling, the resolution that is of interest is the corresponding ground range resolution. For (2.3) this is given as

$$d_{gr} = \frac{d_{sr}}{\sin\vartheta} \tag{2.4}$$

where ϑ is the local incident angle of the beam with the surface. ϑ changes over the range swath of the antenna footprint. The maximum attainable swath in the range direction is given by the antenna elevation beamwidth ψ_e. For a side looking radar flying at a height h_{ac} and looking down at an angle ϕ_0 the maximum range swath is [25]

$$\begin{aligned} S_{sr} &= h_{ac}\left(\cot\left(\phi_0 - \frac{\psi_e}{2}\right) - \cot\left(\phi_0 + \frac{\psi_e}{2}\right)\right) \\ &= \frac{2h_{ac}\sin\psi_e}{\cos\psi_e - \cos 2\phi_0} = \frac{h_{ac}\psi_e}{\sin^2\phi_0} \qquad for \quad \psi_e << 1 \end{aligned} \tag{2.5}$$

The ground range resolution varies nonlinearly across this swath. This can have important consequences for image properties, particularly for spaceborne or short-range airborne systems. The actual slant range swath, bounded by the minimum and maximum range, is determined by the time between the pulse transmission and the start and end of the reception period. The swath is always selected to lie

between the main lobe of the elevation beam. Since transmission and reception cannot overlap, the slant swathwidth is restricted by the PRF, and the condition [60]

$$S_{sr} < \frac{c}{2 \times PRF} \tag{2.6}$$

must hold to avoid range ambiguities. For real aperture systems there is no lower limit on the PRF, however, for SAR systems the PRF must exceed a certain lower limit (see Section 2.2).

A simple sinusoidal pulse of length 1 μs will give a slant range resolution of the order of 150 m. To achieve higher range resolution would require a pulse with a much shorter length. This would require higher transmitter peak power to maintain the detection ranges. The bandwidth of the unmodulated 1 μs pulse is 1 MHz. The range resolution can be improved by using frequency modulation, which increases the bandwidth of the pulse without reducing the length of the transmitted pulse. This results in the system response in range being a *sinc* function [45, 58, 90] when no amplitude taper is used.

For a linear modulated pulse (i.e., a chirp) of bandwidth B and length τ_p', where the frequency $f(\tau)$ at time τ is given by

$$f(\tau) = \begin{cases} f_0 - \dfrac{B\tau}{\tau_p'} & |\tau| \le \dfrac{\tau_p'}{2} \\ 0 & |\tau| > \dfrac{\tau_p'}{2} \end{cases} \tag{2.7}$$

and where f_0 is the carrier frequency, the transmitted pulse is of the form[1]

$$p(\tau) = \exp\left\{ j2\pi \left(f_0 - \frac{B\tau}{2\tau_p'} \right) \tau \right\} \quad \text{for} \quad |\tau| \le \frac{\tau_p'}{2} \tag{2.8}$$

The received pulse from a point target at range R is a delayed and scaled version of the transmitted pulse. Processing the return signal involves stripping off the carrier

[1] $j = \sqrt{-1}$

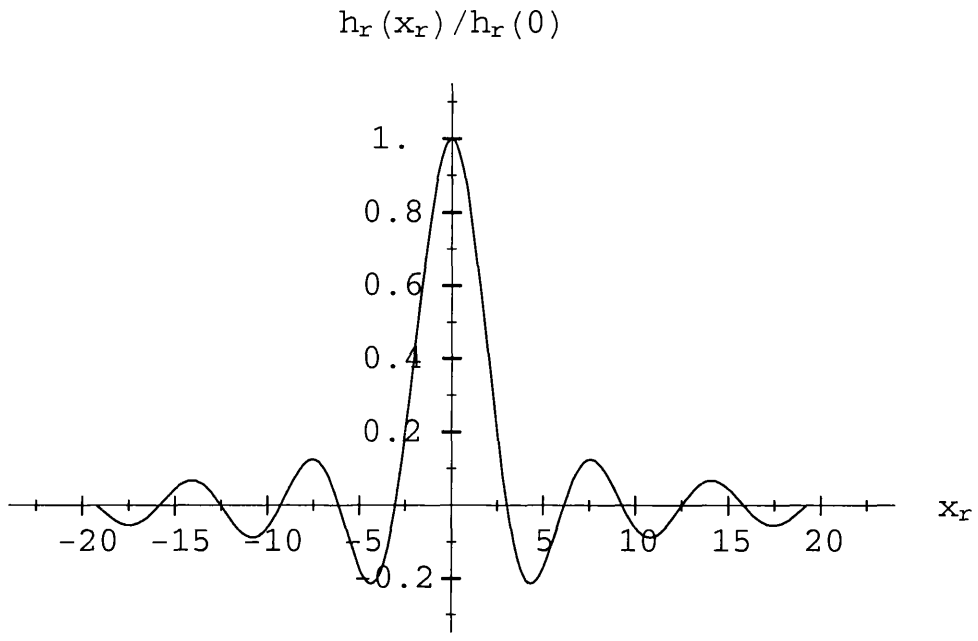


Figure 2.2: Compressed chirp pulse for B =100 MHz and $\tau_p' = 100$ sec

frequency and performing a correlation with a copy of the transmitted signal (this is known as matched filtering [63]). The output for a point target is given by [62]

$$\begin{aligned} h_r(t) &= \int_{-\tau_p'/2}^{\tau_p'/2} \exp\left(j\frac{\pi B}{\tau_p'} s^2 \right) \exp\left(-j\frac{\pi B}{\tau_p'} [s+\tau]^2 \right) \mathrm{rect}\left(\frac{s+\tau}{\tau_p'} \right) ds \\ &= \left(\tau_p' - |\tau|\right) \mathrm{sinc}\left(\frac{B}{\tau_p'} \tau \left[\tau_p' - |\tau|\right] \right) \mathrm{rect}\left(\frac{\tau}{\tau_p'} \right) \end{aligned} \tag{2.9}$$

where

$$\mathrm{rect}(\tau) = \begin{cases} 1 & |\tau| \le \frac{1}{2} \\ 0 & |\tau| > \frac{1}{2} \end{cases} \tag{2.10}$$

The time resolution r_τ (related to range resolution as $d_{sr} = cr_\tau/2$) is given by the first positive zero of the *sinc* function. This occurs when

$$\frac{B}{\tau_p'} \tau(\tau_p' - \tau) = 1 \tag{2.11}$$

which gives the solution for r_τ as

$$\tau = r_\tau = \frac{\tau_p'}{2}\left(1 - \sqrt{1 - \frac{4}{B\tau_p'}}\right) \tag{2.12}$$

For high resolution systems, the *time-bandwidth product* $B\tau_p'$ is large and (2.12) approximates to

$$r_\tau \approx \frac{1}{B} \tag{2.13}$$

r_τ is the compressed pulse length. Using (2.13) we get the expression for the *compression ratio* as

$$compression \quad ratio = \frac{\tau_p'}{r_\tau} = B\tau_p' \tag{2.14}$$

Thus a pulse of length $\tau_p' = 1$ μs using a chirp with a *compression ratio* of 100 will have a range resolution of 1.5 m. The corresponding bandwidth is B = 100 MHz. The plot of the compressed chirp pulse for this system is shown in Figure 2.2, where the x-axis has been converted to distance units by the substitution $x_r = c\tau$.

2.2 Azimuth resolution

The azimuth information makes up the second component in a radar image. The azimuth resolution of a real aperture radar is governed by its azimuth beamwidth. The azimuth geometry of a side looking radar imaging is illustrated by Figure 2.1.

The azimuth resolution at a slant range R_0 is given as

$$d_a = R_0\psi_a \tag{2.15}$$

Using (2.1) this gives

$$d_a = \frac{R_0\lambda}{D_a} \tag{2.16}$$

The returns from a point scatterer at range R_0 would be smeared out in azimuth by a distance given by (2.16). An X-band system using a 2-m antenna at a range of 10

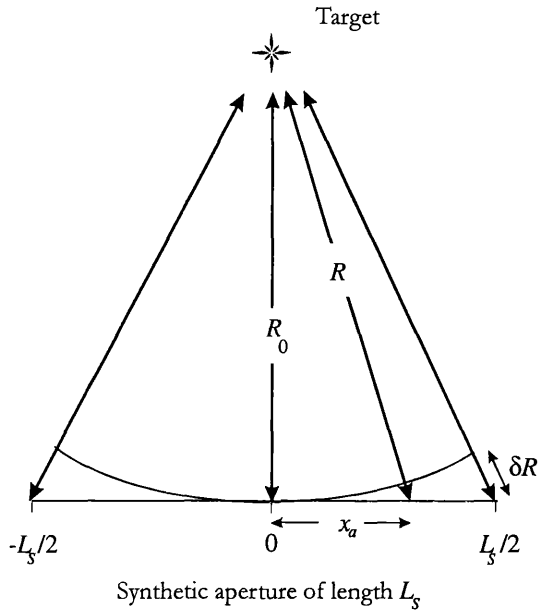


Figure 2.3: Geometry of the illumination of a point target across a synthetic aperture of length L_S.

km will have an azimuth resolution of 150 m. For most imaging purposes this resolution is very poor. The brute force method for increasing the azimuth resolution would be to a) reduce the wavelength of the transmitted signal, and b) increase the aperture of the antenna. Neither of these solutions are practical for long range surveillance radars, the former due to high propagation losses for frequencies higher than X-band and the latter due to the constraint on the antenna size due to platform dimensions.

A more practical method for achieving higher azimuth resolution for a sideways-looking radar on a moving platform is to coherently add the returns from a series of different positions along the line of flight. This effectively forms a synthetic aperture from which this method derives the name 'synthetic aperture radar' (SAR). The geometry of the illumination of a point target along a straight synthetic aperture of length L_S is shown in Figure 2.3.

As the radar beam passes over a point target, the range of the target to the platform changes. This produces a change in the phase of the detected field at each point along the synthetic aperture. The range R of a point target X can be ex-

pressed in terms of its range when it is broadside to the radar platform R_0, and its azimuth lag x_a, along the ground track of the radar beam away from being broadside

$$R = \sqrt{R_0^2 + x_a^2} \tag{2.17}$$

For a narrow beam, X is only illuminated when $x_a << R_0$ in which case

$$R = R_0 + \frac{x_a^2}{2R_0} \tag{2.18}$$

Thus the path difference δR between when a target is at an azimuth lag x_a compared to when it is broadside of the antenna is given by

$$\delta R = \frac{x_a^2}{2R_0} \tag{2.19}$$

where the corresponding two-way phase difference $\delta\phi$ as a function of azimuth lag is given by

$$\delta\phi(x) = \frac{2\pi x_a^2}{\lambda R_0} \tag{2.20}$$

If E_0 is the field detected when the point target is broadside of the antenna then the detected field $F(x_a)$ at point x_a along the synthetic aperture is given by

$$F(x_a) = \begin{cases} E_0 \exp\left(\dfrac{j2\pi x_a^2}{\lambda R_0}\right) & |x_a| \le \dfrac{L_S}{2} \\ 0 & |x_a| > \dfrac{L_S}{2} \end{cases} \tag{2.21}$$

where the variations in the magnitude of the detected field across the synthetic aperture are ignored. This is a valid assumption as the variation in the target's range across the synthetic aperture is small relative to the range itself. In the absence of any phase correction, only those returns for which $\delta\phi$ is less than a fraction of a cycle can be summed together. With proper phase correction, radar returns re-

ceived over the full length of the synthetic aperture can be summed together which results in very fine azimuth resolution.

Full phase correction involves convolving the detected field with the matched filter of the phase difference term. The azimuth response to a point target is then given as

$$\begin{aligned} h_a(x_a) &= F(x_a) * \exp(-j\delta\phi(-x_a)) \\ &= \int_{-L_S/2}^{L_S/2-|x_a|} F(|x_a| + y)\exp(-j\delta\phi(-y))dy \\ &= E_0 L_S\left(1-\frac{|x_a|}{L_S}\right)\operatorname{sinc}\left(\frac{2L_S x_a}{\lambda R_0}\left(1-\frac{|x_a|}{L_S}\right)\right) \end{aligned} \tag{2.22}$$

for $|x_a| \le L_S$, otherwise it is zero[2]. The output response is a distorted sinc function with the first null occurring at

$$\frac{2L_S x_a}{\lambda R_0}\left(1-\frac{|x_a|}{L_S}\right) = 1 \tag{2.23}$$

which gives the solution for the azimuth resolution as

$$x_a = d_a = \frac{\lambda R_0}{2L_S} \tag{2.24}$$

The maximum value for L_S is determined by the distance over which a target is within the main beam of the antenna. This is equal to the azimuth beamwidth. Thus substituting for maximum L_S from (2.16), the azimuth resolution of a fully focused SAR is given as

$$d_a = \frac{D_a}{2} \tag{2.25}$$

The azimuth resolution is independent of range and wavelength and is determined solely by the azimuthal length of the real aperture. Thus an X-band system with a 2-m antenna can achieve an azimuth resolution of 1 m. The associated cost of SAR

[2] When the symbol * is used as in x^*, it represents the complex conjugate of x, whilst when used as a binary operator as in $x * y$, it represents the **convolution** of x and y.

processing is high data storage since the Nyquist criterion demands that the sampling rate must exceed the resolution length [90]. This imposes a lower limit on the PRF which then also limits the maximum slant swathwidth (see (2.6)).

The azimuth resolution, however, cannot be increased indefinitely by reducing the real aperture size. Firstly, reducing the size will lower the gain of the antenna and thus the detection performance of the radar. Secondly, the required minimum PRF will increase, reducing the slant swathwidth [90]. Higher SAR resolution is, however, possible with *spotlight mode* SAR which uses beam steering to keep the target in the beam for a longer time and thus form a longer synthetic aperture. This gives considerable improvement on the azimuth resolution, however at the expense of spatial coverage [17].

2.3 Doppler resolution

A target can be defined in three state variables; range, azimuth and Doppler. The Doppler of a target is a function of its radial velocity toward the radar and is used in MTI radars to form range-Doppler maps which facilitate the detection of moving targets.

The Doppler frequency of a target is related to its radial velocity v_t, toward the radar, as [81]

$$f_t = \frac{2v_t f_0}{c} \tag{2.26}$$

The maximum unambiguous Doppler is given by the PRF. In a coherent MTI system the received pulses are processed along the azimuth direction using a one-dimensional Fast Fourier Transform (FFT) to generate a Doppler profile. The FFT is performed for each range gate. This produces a two dimensional range-Doppler map. The Doppler resolution is given as [83]

$$f_d = \frac{PRF}{N_p} \tag{2.27}$$

where N_p is the number of pulses that are coherently integrated. $N_p/PRF = \tau_{int}$, the coherent integration time. Therefore, the Doppler resolution, and hence the bandwidth, is given as

$$f_d = \frac{1}{\tau_{int}} \tag{2.28}$$

The total number of Doppler bins is equal to the number of pulses that are coherently integrated. The Doppler resolution is increased by extending the coherent integration interval. For a given PRF this is achieved by increasing the number of pulses that are processed together by the FFT operation. The upper bound on t_{int} is the dwell time on the target. The dwell time is a function of antenna beamwidth, platform velocity and antenna scan rate. Therefore, there is a trade-off between azimuth accuracy, swath coverage and Doppler resolution. Furthermore, t_{int} has to be kept as short as possible to avoid target migrating through the resolution cells. A typical value of t_{int} for ground surveillance radar is 0.1 sec.

For a moving platform, ground clutter also has a relative velocity toward the radar which results in the clutter echo being shifted in Doppler. The relative clutter velocity depends on the aircraft velocity and the direction of the clutter relative to the platform velocity vector. The mean value of the clutter Doppler shift is given by

$$f_c = \frac{2v_{ac}f_0}{c}\cos\phi_0\cos\theta_0 \tag{2.29}$$

where v_{ac} is the platform velocity, θ_0 is the antenna's azimuth angle, and ϕ_0 is the antenna's depression angle. Since the radar footprint has a spatial extent (see Figure 2.1) the clutter spectrum has a Doppler spread. For small depression angles, the Doppler spread from clutter returns is generally determined by the radar's azimuth beamwidth. The Doppler spectral width is found by subtracting the radial Doppler frequency components, which occur at the half-power edges of the azimuth beam as [75]

$$\begin{aligned}\Delta f_c &= \frac{2v_{ac}f_0}{c}\cos\phi_0\left[\cos\left(\theta_0 - \psi_a/2\right) - \cos\left(\theta_0 + \psi_a/2\right)\right] \\ &= \frac{2v_{ac}f_0}{c}\psi_a \cos\phi_0 \sin\theta_0\end{aligned} \tag{2.30}$$

for small azimuth beamwidths. An X-band airborne radar moving at 200 m/s with a 2-m antenna at 5^0 depression angle and pointing sideways will have a clutter spectral width due to platform motion of 199 Hz.

The clutter Doppler spectrum is divided into a number of Doppler bins determined by the Doppler resolution. This has the effect of dividing the azimuth beamwidth into subapertures whose beamwidth is determined by the Doppler resolution. The Doppler frequency corresponding to the kth subaperture is given as

$$f_{d_k} = \frac{2v_{ac}f_0}{c}\cos\phi_0 \cos\theta_k \tag{2.31}$$

where θ_k is the azimuth angle between the kth subaperture and the aircraft velocity vector. Since Doppler resolution is the frequency difference between adjacent Doppler bins, therefore

$$f_d = f_{d_{k+1}} - f_{d_k} = \frac{2v_{ac}f_0}{c}\cos\phi_0\left(\cos\theta_{k+1} - \cos\theta_k\right) \tag{2.32}$$

Letting $\Delta\theta$ be the differentiating angle represented by each subaperture such that $\theta_k = \theta_m - \Delta\theta/2$ and $\theta_{k+1} = \theta_m + \Delta\theta/2$, where θ_m is the medium angle to the centre of the clutter cell and substituting in (2.32) gives

$$f_d = \frac{2v_{ac}f_0}{c}\cos\phi_0\Delta\theta\sin\theta_m \tag{2.33}$$

and solving for the azimuth resolution angle gives

$$\Delta\theta = \frac{cf_d}{2v_{ac}f_0\cos\phi_0\sin\theta_m} \tag{2.34}$$

The Doppler processing has resulted in an azimuth resolution better than that achieved with just the real aperture. This technique is called Doppler beam sharp-

ening [75]. Although the azimuth resolution achieved is inferior to that of a fully focused SAR and it is not independent of range, it is used in MTI radars to produce simple clutter maps. For the example of the X-band system considered earlier, a PRF of 1 kHz and an FFT of 64 pulses will give a Doppler resolution of 16Hz. The corresponding azimuth angle resolution at 5^0 depression angle and sideways-looking antenna is 0.07^0. At a slant range of 10 km this equates to a spatial azimuth resolution of 12 m. Thus in MTI radars, Doppler processing also effects the size of the radar footprint on ground which in turn has a bearing on the statistics of the observed scattering. Clutter models that are able to characterise the observed scattering are discussed in the following chapter.

2.4 Summary

In this chapter

- The theory of how a side looking radar flown on a moving platform achieves high resolution in range, azimuth and Doppler domain has been described.
- It was shown that the range resolution is inversely proportional to radar bandwidth. The chirp pulse technique for achieving high range resolution was described and the system response to a point target was shown to be a *sinc* function.
- The azimuth resolution improvement achieved with a synthetic aperture compared to a real aperture was discussed. The system response to a point target was shown to be again a *sinc* function.
- The relationship between the coherent integration time and the Doppler resolution was described and the technique of Doppler beam sharpening for obtaining clutter maps was discussed.

Chapter 3

3. Clutter models

Clutter modelling is essential for optimising target detection and terrain classification. The clutter echoes result from a coherent electromagnetic scattering process. When an electromagnetic wave scatters from a position (x, y) on the Earth's surface, the physical properties of the terrain cause changes in both the phase $\Phi(x, y)$ and amplitude $A(x, y)$ of the wave. Therefore, the most direct method for modelling the clutter returns is to obtain a full wave solution using Maxwell's equations while applying the appropriate boundary conditions. This method, however, suffers from the difficulty of selecting and applying the correct boundary conditions to represent typical real surfaces. A more simplified approach is to represent the received complex signal as a sum of contributions from discrete elementary scatterers. The statistical distribution of the received signal can then be modelled in terms of the probability distribution of the amplitude, phase and number of the discrete scatterers. The choice of statistical models can be based on some physical justification for the scatterer distribution or on ad hoc models based on empirical evidence. The major part of this chapter is devoted to a discussion of the statistical models.

The *forward problem*, as the electromagnetic modelling of rough surface scattering is commonly known, is briefly reviewed in Section 3.1. An in-depth study of the forward problem is outside the scope of this thesis and the discussion is limited to a general overview of the technique. This method predicts the scattered field based on the properties of the scattering medium and the incoming wave. The first results in this area were obtained for scattering from periodic surfaces. Subse-

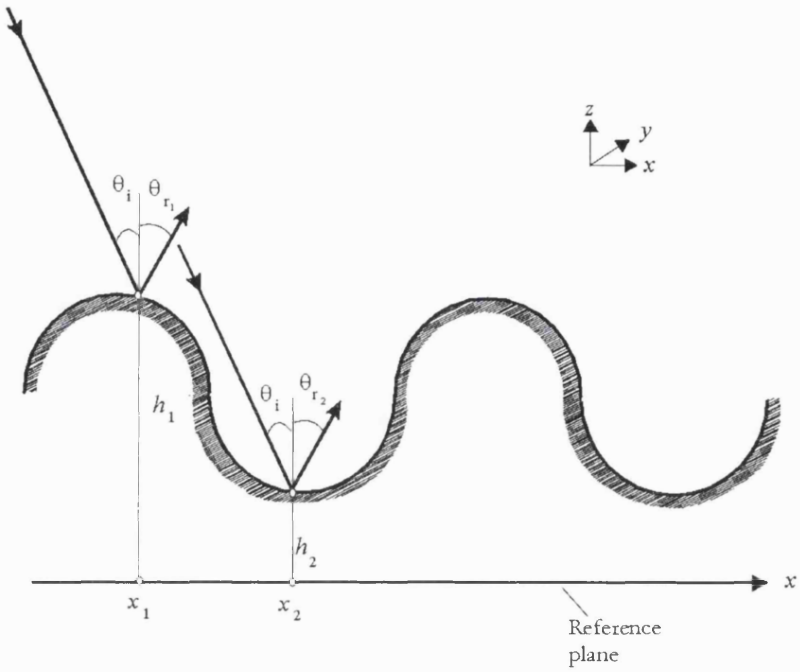


Figure 3.1: The Rayleigh criterion for smooth surfaces.

quent work addressed the problem of scattering from random rough surfaces. Section 3.1.1 to 3.1.5 summarises the various approaches taken to solve the forward problem.

Although the forward problem analysis has helped to identify the dominant scattering mechanisms, the full wave solutions in most cases are unavailable. Therefore, an analysis of the statistics of the observed field is a more realistic approach for characterising the clutter models. The statistics of the received signal are investigated in terms of a general model for the electromagnetic field detected by a radar system; this is described in Section 3.2. When the resolution cell is large compared to the wavelength the phase of the scatterers is uniformly distributed and contains no information regarding the target. Therefore the useful information regarding the scattering surface is contained only within the intensity of the detected field. For the completeness of the discussion we briefly describe the statistics of the phase and complex components of the detected field in Section 3.2. This is followed by a more extensive discussion of the statistics of the intensity of the detected fields in Section 3.2.1 onwards.

A review is carried out of the statistical models used to describe the intensity characteristics of ground clutter. Where applicable the physical models underlining the statistical models are described. For each model the **probability density function** (pdf), **cumulative distribution function** (cdf) and the **moments** are listed. The most promising of these models are selected for testing with real data. The results are quoted for goodness-of-fits using the chi-squared test.

3.1 Electromagnetic models

The problem of backscattering from a rough surface is of interest to radar imaging, sonar detection and optics. There has been considerable work published on this subject in a wide spectrum of journals. The very first work published in this area was by Rayleigh [67] which analysed the scattering from sinusoidal surfaces for normal incidence. Schouten and De Hoop [76] extended this for any analytical rough surface. Although periodic surfaces proved useful in indicating the general behaviour of rough surfaces, real surfaces are rarely periodic. A number of researchers have investigated the scattering from random rough surfaces. Beckmann and Spizzichino [7] considered scattering from normally distributed rough surfaces using the Kirchhoff solution which assumes a slowly varying height profile. However, many surfaces of practical interest are not normally distributed. The distribution of terrain with sharp ridges and round valleys and that of a rough sea is asymmetric and therefore not normal. A standard approach for modelling scattering from non-Gaussian surfaces is the Small Perturbation Method (SPM) which is suitable when the surface roughness is small [68]. The purely analytical approach, however, is limited in the type of surface that can be modelled, due to the lack of closed form solutions. Various numerical solution based approaches have been used to analyse the scattering from arbitrarily defined rough surfaces. Section 3.1.1 to 3.1.4 summarises each of these approaches. Besides these methods there have been a few other notable approaches, which are mentioned in Section 3.1.5.

3.1.1 Periodic surfaces

Historically, this class of rough surfaces was the first to be treated. A rough surface is represented with periodic irregularities such as sinusoidal undulations, saw-tooth

profiles, protrusions of equal shape spaced at regular intervals, etc. This approach enables the problem to be solved in a simplified non-statistical way.

Rayleigh [67] analysed the problem for sinusoidal surfaces and developed the well-known *Rayleigh criterion*, for qualifying the roughness of a surface. The process in which an incident wave is reflected depends upon the surface roughness. For smooth surfaces the reflection is specular. Specular reflection is directional and has a constant predictable phase with respect to the incident wave. In this case θ_{r_1} and θ_{r_2} in Figure 3.1 will be equal and the reflections will add coherently to give a strong return in a single direction. For rough surfaces, the reflection is diffused. Diffused reflection is much more dispersed and the phase of scattered energy varies over 0 - 2π. Rayleigh stated that a surface may be regarded as smooth if the phase difference between two reflected waves is less that $\pi/2$. Using the geometry of Figure 3.1 this is

$$\Delta h < \frac{\lambda}{8\cos\theta_i} \tag{3.1}$$

where Δh is the height difference, θ_i is the incident angle and λ is the wavelength of the incident wave. (3.1) states that roughness of any scattering surface is not an intrinsic property of that surface but depends on the frequency and angle of the incident wave. A surface appears rougher the smaller the incident wavelength or the closer the angle of incidence is to the surface normal. Although, the Rayleigh criterion was originally defined for periodic surfaces it has been universally adopted for all type of rough surfaces, where Δh is replaced by σ_h the **standard deviation** of surface heights.

Rayleigh's work was for the special case of normal incidence onto a sinusoidal surface which was extended by LaCasce and Tamarkin [47] to oblique angles. Deryugin [23] applied it to different profiles and Schouten and De Hoop [76] generalised it for any analytically given rough surface.

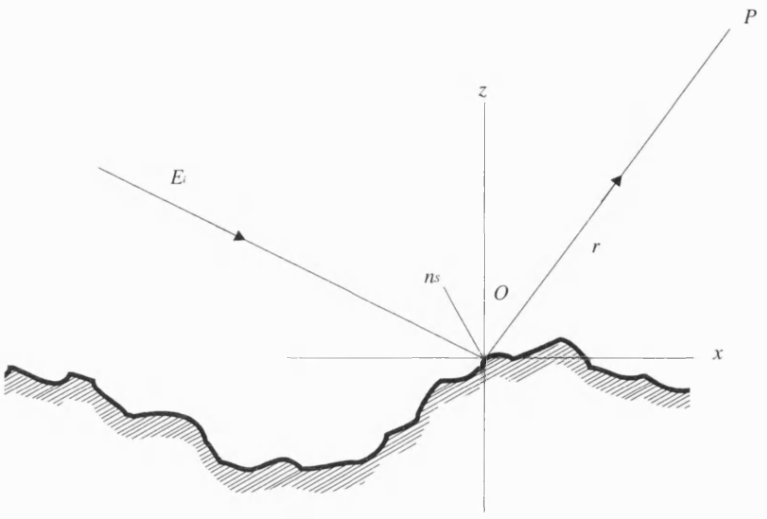


Figure 3.2: Scattering from random rough surfaces.

3.1.2 Normal distributed surfaces

The scattered field E_P observed at a point P at a distance r from a point O on a surface S (Figure 3.2) is given by the Helmholtz integral [7] as

$$E_P = \frac{1}{4\pi} \iint_S E_S \frac{\partial \psi_S}{\partial n_S} - \psi_S \frac{\partial E_S}{\partial n_S} \quad dS \tag{3.2}$$

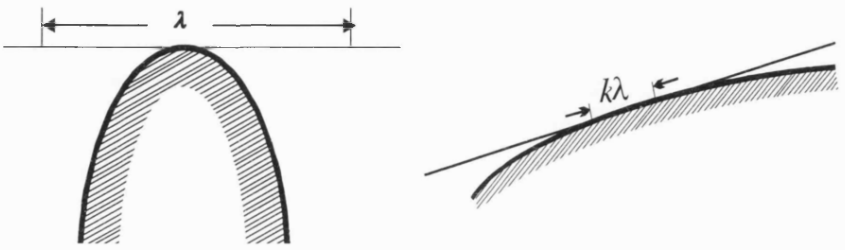
where E_S is the field on S, n_S is a normal to the surface and ψ_S is given by

$$\psi_S = \frac{\exp(\,/\frac{2\pi}{\lambda} r)}{r} \tag{3.3}$$

Beckmann and Spizzichino [7] derived an analytical solution to the scalar problem of (3.2) for a normally distributed surface using the Kirchhoff approximation whereby the field at any point on the surface is expressed as the sum of the incident and reflected field at the tangent at the considered point. Thus E_S in (3.2) is replaced by

$$E_S = \left(1 + R_{coeff}\right) E_i \tag{3.4}$$

where R_{coeff} is the reflectivity coefficient of a smooth plane and E_i is the incident field. This approximation is very good when the radius of curvature of the irregu-



(a) (b)

Figure 3.3: The tangent plane at a general point of the rough surface. The radius of curvature is (a) small, and (b) large in comparison with the wavelength.

larities is very large compared to the wavelength (Figure 3.3b) but breaks down completely for sharp edges (Figure 3.3a). Thus the Kirchhoff solution is suitable for smoothly varying normally distributed surfaces. The analyses produced by Beckmann and Spizzichino [7] assumed small slopes and small standard deviation of surface heights σ_h. Wu and Fung [98] extended this to surfaces with large σ_h using a vector formulation of the Kirchhoff method.

Many surfaces of practical importance are, however, not normally distributed. Beckmann [8] showed that the surface height probability distribution has a significant bearing on the characteristics of the scattered field where the surface roughness and the grazing angle of the incident wave are large. Introducing a correlation function into the Kirchhoff model can compensate for an incorrect surface model. Shaw and Dougan [80] proposed suitable correlations to model ocean-like surfaces. The assumption of gently undulating terrain, nevertheless severely limits the applicability of this model to land surface scattering and therefore there have been alternative techniques developed for predicting scattering from non-Gaussian surfaces. These are described in the following section.

3.1.3 Non-Gaussian surfaces

Rice [68] developed a scattering model suitable for non-Gaussian surfaces based

on SPM analysis. Unlike the Kirchhoff method, SPM does not impose a restriction on the surface curvature and permits the surface height to vary within the distance of a wavelength. The method, however, is applicable only to slightly rough surfaces where

$$k\sigma_h < 0.3 \tag{3.5}$$

It also requires that l_S, the average slope of the surface, should meet the criterion

$$\sqrt{2}\sigma_h / l_S < 0.3 \tag{3.6}$$

SPM calculates the field at any point in the scattering medium using the solutions for scattering from a smooth surface, together with 'perturbative' terms arising from the slight surface roughness. Under the assumption of slightly rough surfaces the quantities that are a function of the surface height may be expanded as a Taylor series about their value on the mean scattering surface

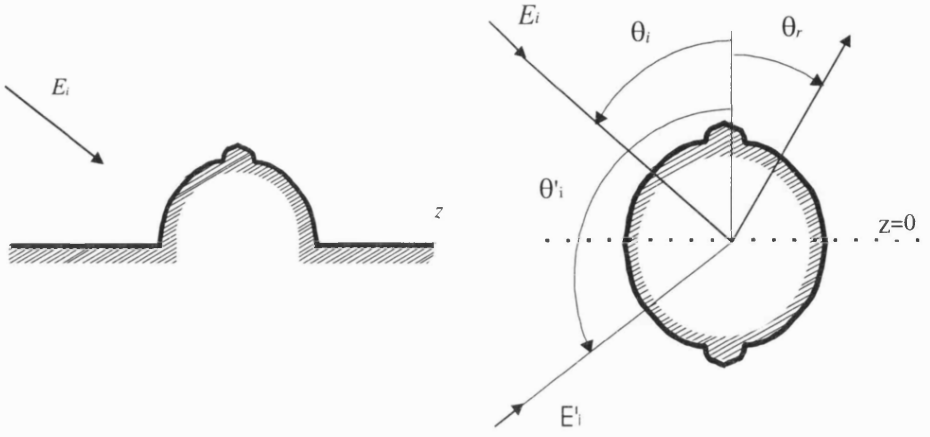
$$f(x,y,h) = f(x,y,0) + h\frac{\partial f}{\partial z}(x,y,0) + \frac{h^2}{2}\frac{\partial^2 f}{\partial z^2}(x,y,0) + \ldots \tag{3.7}$$

where f is a quantity defined in the Cartesian co-ordinates x, y and z, and the mean scattering plane is $z = 0$. The accuracy of the perturbation theory depends on the number of terms retained in the expansion of (3.7). Perturbation theory assumes that the total scattered field E_r may be written as a series

$$E_r = E_r^0 + E_r^1 + E_r^2 + \ldots \tag{3.8}$$

where E_r^k is the contribution from the h^k term in (3.7). The earlier analysis of the perturbation theory was carried out using just the first order terms in h [15]. More recent work by Bass and Fuks [5] and Watson and Keller [96] carried out the analysis for the second order perturbation theory.

Wright [97] tested the SPM model for sea clutter and found good agreement to vertically polarised data. The gap between the Kirchhoff method, suitable for high frequencies, and the SPM, suitable for slightly rough surfaces, was bridged by Bahar [2] who produced a unified full wave solution to evaluate the scattering from



(a) (b)

Figure 3.4: The image method. The surface field in (a) is equivalent to the surface field in (b) where the conducting plane has been replaced by a mirror image of the surface protrusion.

rough surface with arbitrary slopes. The full wave solution also extends the solution to take into account shadowing and multiple scattering.

3.1.4 Numerical solutions

This approach uses the numerical solutions of integral equations for the unknown surface fields to calculate the scattered field. The surface height function $h(x, y)$ is treated as a random correlated discrete variable. The numerical solutions are not restricted to an analytical expression for the scattering surface and thus the solution can be calculated for any realisation of surface height profile. A variety of methods have been adopted from spectral analysis to generate surface height profiles for example moving average methods and autoregressive processes [55]. Each set of generated data represents one surface realisation for which the scattered field may be calculated. The statistical parameters of the scattered field are then determined by generating large number of surface realisation and analysing the scattered fields from these surfaces. There is, however, a large price to pay in terms of increased computational effort with this method.

Numerical solutions have been used to study the accuracy of the SPM [18, 30] and the Kirchhoff theory [29, 48] approximate techniques. A further use of numerical simulation techniques is in the modelling of many aspects of realistic scattering geometries. Factors that complicate, and often preclude, analytical theory, such as non-planar incident waves with a spatially varying amplitude profile, incident pulses of arbitrary form and finite-sized surfaces of non-ideal shape, may all be included in a numerical study, however, at the expense of increased computational effort.

3.1.5 Other methods

The methods described in the previous sections are for the more general type of rough surfaces. There has been other less general method developed which over come some of the limitations of the more general models. Twersky [87] developed a scattering model based on mirror images. He considered a perfectly conducting surface with arbitrary shaped protrusions. Under this situation the field scattered by the protruding object under an incident plane wave E_i can be derived by evaluating the field scattered from the object in free space illuminated by two incident waves. This is so since the surface field in both instances is the same, as illustrated diagramatically in Figure 3.4. The scattered field is easily derivable using this method and it allows the modelling of multiple scattering. This approach is particularly useful when the surface can be represented by a collection of randomly located discrete scatterers. Fung and Ulaby [28] and others [19, 44] used the discrete scatter model to represent scattering from vegetation clutter.

Rather then attempting to calculate the field from a realistic model of a rough surface an alternate approach is to adopt a model of a rough surface that will scatter a realistic field. A simple model of rough surfaces that will have this effect is one which consists of plane facets of random slope where the random slopes are generated by a Markov chain (Figure 3.5). Beckmann [6] used this model to obtain an approximation to the scattered field. The analysis was simplified by assuming that the reflection are locally specular (this is true when the facets are large compared with the wavelength) and that the elementary waves reflected in the same

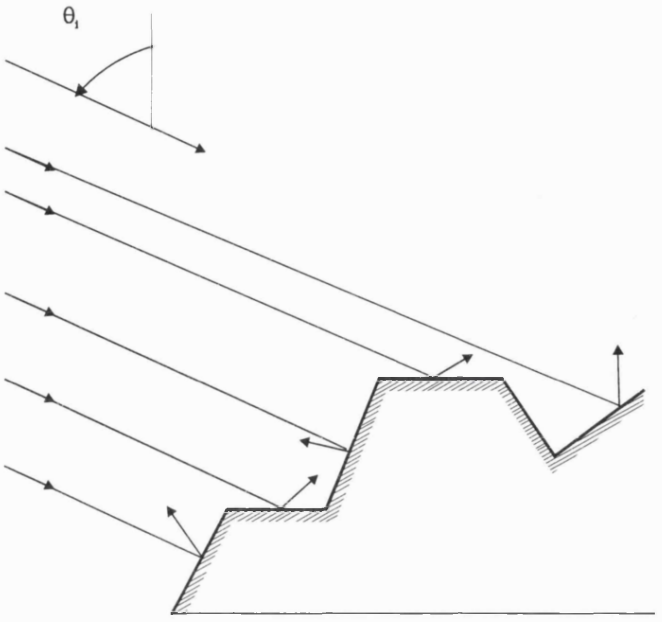


Figure 3.5: Scattering by a surface of large plane facets.

direction (by facets of the same slope) are summed together with respect to their phases to form a single resultant wave in that direction. The plane facet approach is, however, more suited to man-made objects. The facets can be generated deterministically using the object geometry. This method has been used to generate accurate scattering predictions for vehicles [34].

3.2 Statistical models

An image generated by a coherent imaging system is affected by coherent interference between scatterers [32] which causes the detected intensity to fluctuate from resolution cell to resolution cell. This fluctuation in pixel intensities is evident in both MTI and SAR images (see Figure 3.6 and Figure 3.7) of land. The observed intensity fluctuations can be characterised by statistical model without performing a rigorous solution of the Maxwell equations for the scattering surface.

The statistical model assumes that the received electric field at an observation point $\boldsymbol{r}$ can be represented as a sum of contributions from elementary scatterers [32, 37, 56, 60]

$$E(\mathbf{r}) = \sum_{k=1}^{N} a_k(\mathbf{r}) \exp\left(j\phi_k(\mathbf{r})\right) = A(\mathbf{r}) \exp\left(j\Phi(\mathbf{r})\right) \tag{3.9}$$

where a_k (**r**) and ϕ_k(**r**) are the amplitude and phase of the kth scatterer with respect to the observation point **r**. a_k(**r**) includes experimental parameters, e.g. range, transmitter power etc. a_k and ϕ_k are assumed to be statistically **independent** random variables. This discrete scatterer model representation is obviously very simple but is based on the belief that any model that gives the same value for the scattered field per pixel of the image may be regarded as an equivalent representation.

The detected field of (3.9) can be viewed as arising from an N-step random walk in the complex plane where a_k(**r**) and ϕ_k(**r**) now constitute the length and orientation of the random steps. The probability distribution of the detected field in a given pixel can be analysed in terms of the probability distribution of a_k(**r**), ϕ_k(**r**) and N.

As long as a target is rough relative to the illuminating wavelength and incident angle, or the depth of the resolution cell is much greater than the wavelength, the phase of the scatterers within the resolution cell will be randomly distributed over many cycles. The phase of each of the scatterers will then be uniformly distributed over the interval $-\pi$ and π and will be statistically independent of its amplitude [32]. This implies a uniform distribution of the step orientation in the random walk. The resultant phase of the detected field will have a phase that is uniformly distributed and independent of the detected field amplitude.

The scatterers will not have a uniformly distributed phase when the resolution cell is only a few wavelengths in depth. At X and C-band (3 and 6 cm wavelength) this would occur at very high resolutions. In other situations, for example sonar scattering from sea bed, this type of scattering may occur at much lower frequency (a few tens of kHz) owing to the slower velocity of propagation of the signal. The random walk model represents this type of scattering process as a bias in the random walk. Alternatively, if there is a smooth target, e.g. a corner reflector or a strong scatterer, immersed in a resolution cell otherwise containing scatterers with

uniformly distributed phase, the detected field again will have a non uniformly distributed phase. The analysis of the statistical models resulting from a biased random walk is a more complicated task and is dealt in Chapter 5. Before that, in this chapter, we consider the simple scenario of uniformly distributed phase for the scatterers and analyse the statistics of the detected field.

When analysing the data of a coherent imaging system we have the choice of working in the complex domain, phase domain or the intensity domain. However, when the scattering process results in the uniform distribution of the phase, the phase component contains no information. Similarly the complex components also have little information content. This is explained as follows:

When the phase of the detected field is uniformly distributed, the real and imaginary component of the detected field given by

$$E_{\mathcal{R}}(\mathbf{r}) = \sum_{k=1}^{N} a_k(\mathbf{r}) \cos \phi_k(\mathbf{r}) \tag{3.10}$$

$$E_{\mathcal{I}}(\mathbf{r}) = \sum_{k=1}^{N} a_k(\mathbf{r}) \sin \phi_k(\mathbf{r}) \tag{3.11}$$

will both have **mean** zero[3]

$$\left\langle E_{\mathcal{R}}(\mathbf{r}) \right\rangle = \sum_{k=1}^{N} \left\langle a_k(\mathbf{r}) \right\rangle \left\langle \cos \phi_k(\mathbf{r}) \right\rangle = 0 \tag{3.12}$$

$$\left\langle E_{\mathcal{I}}(\mathbf{r}) \right\rangle = \sum_{k=1}^{N} \left\langle a_k(\mathbf{r}) \right\rangle \left\langle \sin \phi_k(\mathbf{r}) \right\rangle = 0 \tag{3.13}$$

equal **variance**, which depends only on the step length

$$\left\langle E_{\mathcal{R}}(\mathbf{r})^2 \right\rangle = \sum_{i=1}^{N} \sum_{k=1}^{N} \left\langle a_i(\mathbf{r}) a_k(\mathbf{r}) \right\rangle \left\langle \cos \phi_i(\mathbf{r}) \cos \phi_k(\mathbf{r}) \right\rangle = \frac{1}{2} \sum_{k=1}^{N} \left\langle a_k(\mathbf{r})^2 \right\rangle \tag{3.14}$$

$$\left\langle E_{\mathcal{I}}(\mathbf{r})^2 \right\rangle = \sum_{i=1}^{N} \sum_{k=1}^{N} \left\langle a_i(\mathbf{r}) a_k(\mathbf{r}) \right\rangle \left\langle \sin \phi_i(\mathbf{r}) \sin \phi_k(\mathbf{r}) \right\rangle = \frac{1}{2} \sum_{k=1}^{N} \left\langle a_k(\mathbf{r})^2 \right\rangle \tag{3.15}$$

---

[3] The operator $\langle \cdot \rangle$ indicates **ensemble averaging**.

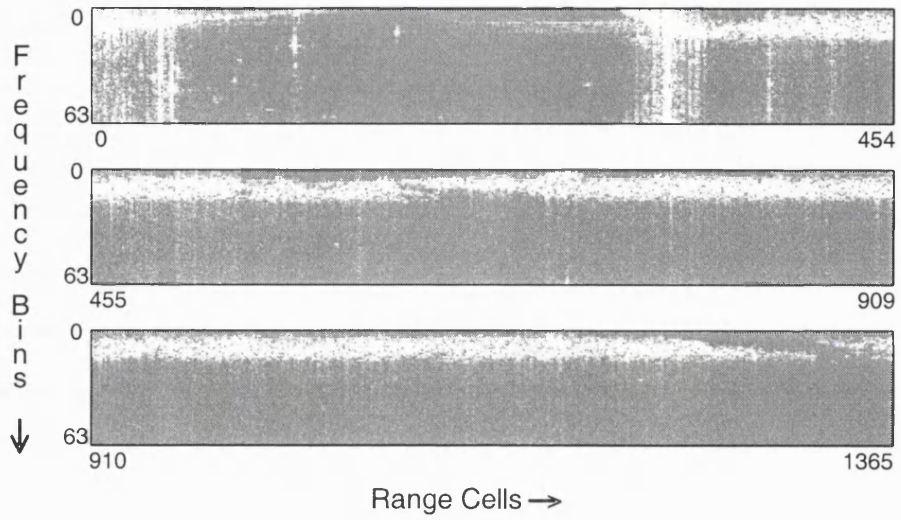
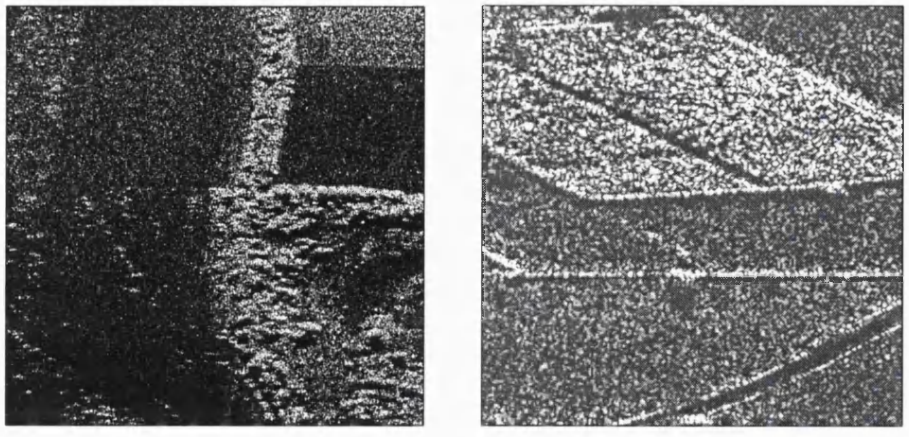
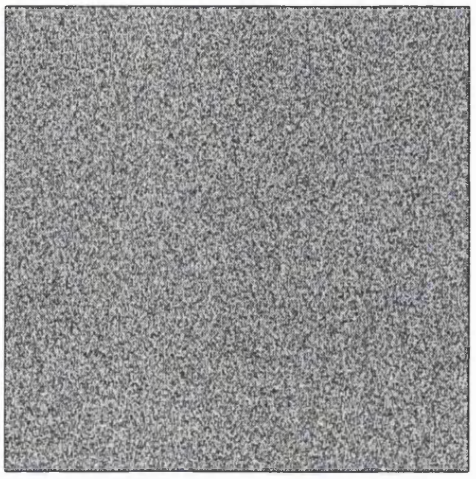


Figure 3.6: MTI intensity image from DERA Canberra X-band radar containing 64 frequency bins and 1365 range bins. The clutter dominant frequency bins appear bright. The clutter band is several frequency bins wide. The data is from an imaging pass over the Hull area and the strongest clutter return corresponds to returns from the Humber bridge.

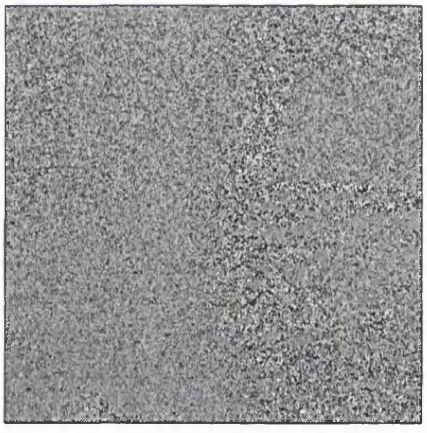


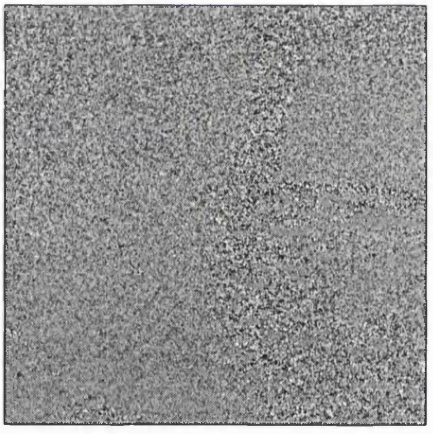
(a) (b)

Figure 3.7: Single look SAR intensity images of rural area from two DERA airborne platforms (a) Andover C-band SAR and (b) Canberra X-band SAR. Woodlands appear as bright regions, whereas grass and fields appear as dark regions.



(a)





(b) (c)

Figure 3.8: (a) the phase and, (b) and (c) the real and imaginary components of the SAR image whose intensity is shown in Figure 3.7a.

and they will be uncorrelated [32]. This is because

$$\left\langle E_{\mathcal{R}}(\mathbf{r})E_{\mathcal{I}}(\mathbf{r})\right\rangle = \sum_{i=1}^{N}\sum_{k=1}^{N}\left\langle a_i(\mathbf{r})a_k(\mathbf{r})\right\rangle\left\langle \cos\phi_i(\mathbf{r})\sin\phi_k(\mathbf{r})\right\rangle = 0 = \left\langle E_{\mathcal{R}}(\mathbf{r})\right\rangle\left\langle E_{\mathcal{I}}(\mathbf{r})\right\rangle \qquad (3.16)$$

Equations (3.12) to (3.16) rely on the fact that when $\phi_k(\mathbf{r})$ is an independent random variable uniformly distributed on the interval $(-\pi, \pi)$

$$\langle \cos \phi_k(\mathbf{r}) \rangle = \langle \sin \phi_k(\mathbf{r}) \rangle = 0 \tag{3.17}$$

$$\langle \cos \phi_i(\mathbf{r}) \cos \phi_k(\mathbf{r}) \rangle = \langle \sin \phi_i(\mathbf{r}) \sin \phi_k(\mathbf{r}) \rangle = \begin{cases} \frac{1}{2} & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} \tag{3.18}$$

and

$$\langle \cos \phi_k(\mathbf{r}) \sin \phi_k(\mathbf{r}) \rangle = 0 \tag{3.19}$$

for all i and k.

Figure 3.8 shows the phase and complex components of the C-band SAR image whose intensity is shown in Figure 3.7a. Structural features present in the intensity image are barely discernible in the complex components. No structure is apparent in the phase image. The results for the MTI data are similar as can be seen from Figure 3.9 which shows the phase and real component of the MTI intensity image shown in Figure 3.6. As most of the information regarding clutter is in the intensity component, the remainder of the chapter will discuss the intensity statistics of the detected field.

The initial statistical models used to described the radar clutter intensity were based on simple Gaussian scattering which gives rise to a negative exponential distribution for the intensity field; this is described in Section 3.2.1. As the radar systems became more sophisticated and the resolution achieved with these systems improved, the simple clutter models no longer represented the radar data adequately. Recently, in an attempt to remedy this situation several non-Gaussian clutter models have been introduced. Some of these models are purely empirical, whereas the K-distribution uses a plausible physical model based on the random walk in a complex plane to arrive at a pdf for the detected field intensity. In Section 3.2.2 and 3.2.3 we describe the log-normal and the Weibull distribution, two widely used empirical models prior to a detailed description of the K-distribution

model in Section 3.2.4. We present results from real data to show the suitability of K-distribution for land clutter.

3.2.1 Negative exponential distributed intensity

A simple model for the scatterer distribution is where the number of scatterers tends to infinity and the phase is uniformly distributed. The probability distribution for the detected field can be derived from the random walk model. Jakeman [38] used the characteristic function to derive the intensity pdf. The characteristics function for (3.9) is given as[4]

$$\begin{aligned} C(\mathbf{u}) &= \left\langle \exp\left(j\mathbf{u}.\mathbf{E}\right)\right\rangle \\ &= \left\langle \prod_{k=1}^{N} \exp\left(ja_k u\cos\left(\phi_k + \gamma\right)\right)\right\rangle \end{aligned} \tag{3.20}$$

where γ depends on the components of $\mathbf{u}$. Using the assumption that a_k and ϕ_k are statistically independent, the product sign can be taken outside the average, so

$$C(\mathbf{u}) = \prod_{k=1}^{N} \left\langle \exp\left(ja_k u\cos\left(\phi_k + \gamma\right)\right)\right\rangle \tag{3.21}$$

Integrating over ϕ then gives

$$C(\mathbf{u}) = \prod_{k=1}^{N} \left\langle \mathrm{J}_0\left(a_k u\right)\right\rangle \tag{3.22}$$

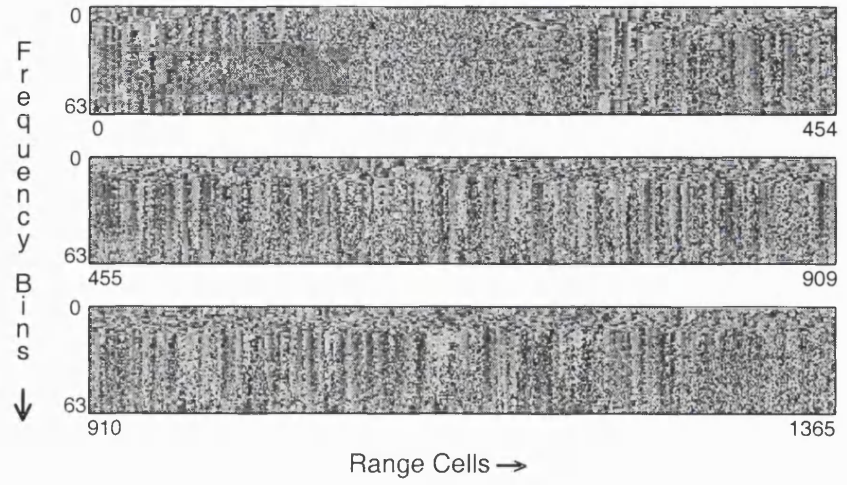
where J_0 is the zeroth order **Bessel function** of the first kind. Since a_k are assumed to be statistically independent, (3.22) becomes

$$C(\mathbf{u}) = \left\langle \mathrm{J}_0(au)\right\rangle^N \tag{3.23}$$

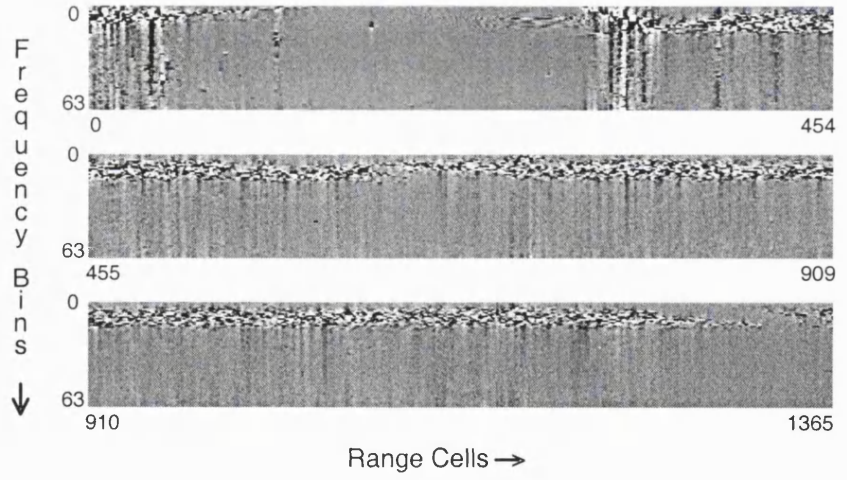
where $\langle a\rangle$ is the ensemble average of all possible realisations of the step length. By scaling the step length a through $a \to a/\sqrt{N}$ and allowing N to tend to infinity (3.23) becomes

---

[4] The notation has been simplified by removing the explicit reference to $\mathbf{r}$.



(a)



(b)

Figure 3.9: (a) The phase and (b) the real component of the MTI image whose intensity is shown in Figure 3.6.

$$C(\mathbf{u}) = \exp\left(-\frac{1}{4}u^2\langle a^2\rangle\right) \tag{3.24}$$

This corresponds to the Rayleigh distribution for the amplitude A

$$f_A(A) = 2\left(A/\langle A^2\rangle\right)\exp\left(-A^2/\langle A^2\rangle\right) \tag{3.25}$$

Defining the intensity by

$$I = A^2 \tag{3.26}$$

leads to the negative exponential distribution for the intensity fluctuation

$$f_I(I) = \left(1/\langle I\rangle\right)\exp\left(-I/\langle I\rangle\right) \tag{3.27}$$

Rewriting (3.27) in terms of a general random variable x with mean μ_x gives the pdf as

$$f_x(x) = \frac{1}{\mu_x}\exp\left(-\frac{x}{\mu_x}\right) \tag{3.28}$$

the cumulative distribution function as

$$F_x(x) = 1-\exp\left(-\frac{x}{\mu_x}\right) \tag{3.29}$$

and moments[5] as

$$\langle x^n\rangle = \Gamma(n+1)\mu_x^n \tag{3.30}$$

From (3.27) and (3.28) the mean of the negative exponential intensity is given as

$$\mu_I = \langle I\rangle \tag{3.31}$$

where the subscript I indicates an intensity pdf. The resultant radar cross section of the surface, ρ[6] is given by the mean of the intensity. Therefore,

$$\rho = \mu_I = \langle I\rangle \tag{3.32}$$

The negative exponential model is based on the phenomenon that the resolution cell contains a large number of scatterers with no single scatterer dominating the

---

[5] $\Gamma(z)$ is the **gamma function**.

[6] The symbol ρ is used for the radar cross-section of the surface rather than the more normal σ so as not to confuse σ's indicating the radar cross-section of the surface with those indicating variances.

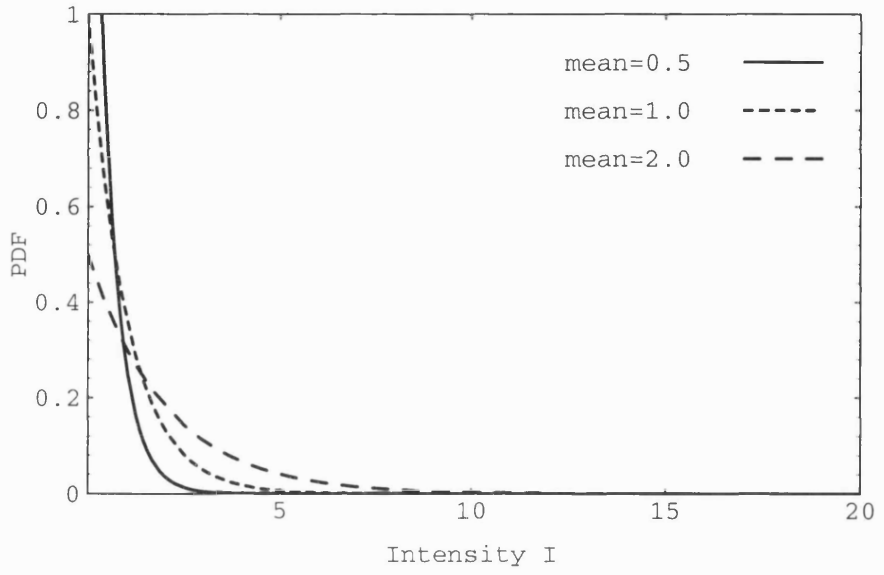


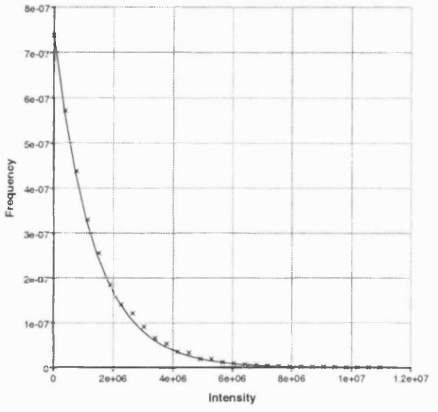
Figure 3.10: Negative exponential pdf for three different values of the mean.

radar return. Figure 3.10 gives the plot of the intensity pdf for various values of the mean. As the mean increases the tail of the distribution gets longer.

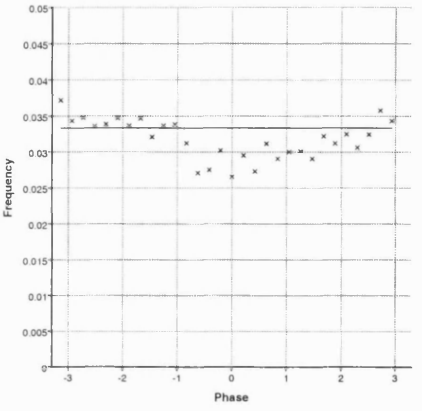
The result that the detected intensity is negative exponentially distributed when the number the of independent scatters tends to infinity can also be reached by applying the **central limit theorem** to the complex components of the detected field [32]. In this case the real and imaginary parts of the vector E will be independent **Gaussian distributions** with zero mean and the same variance. They will have a **joint** pdf given as

$$f_E\left(E_{\mathcal{R}}, E_{\mathcal{I}}\right) = \frac{1}{2\pi\sigma_E^2}\exp\left(-\frac{\left(E_{\mathcal{R}}^2 + E_{\mathcal{I}}^2\right)}{2\sigma_E^2}\right) \tag{3.33}$$

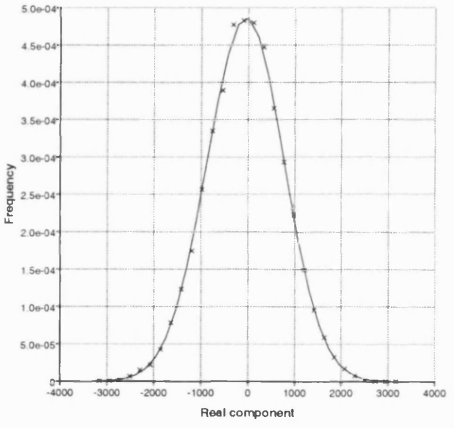
where σ_E^2 is the variance of both the real and imaginary components ((3.14), (3.15)). The joint pdf of the intensity and phase of the detected field is obtained from (3.33) by transforming to polar co-ordinates



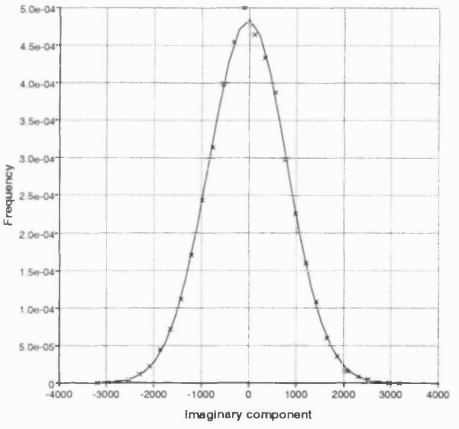
(a)



(b)



(c)



(d)

Figure 3.11: Histograms showing the distribution of the (a) intensity, (b) phase, (c) real component and (d) imaginary component, of pixel values from a homogeneous region (128×128 pixels) in the image shown in Figure 3.7a. The curves represent expected distributions fitted to the data, (a) negative exponential, (b) uniform, and (c) and (d) zero mean Gaussian.

$$f_E(I,\theta) = f_E\left(\sqrt{I}\cos\theta, \sqrt{I}\sin\theta\right)|J| = \frac{1}{4\pi\sigma_E^2}\exp\left(-\frac{I}{2\sigma_E^2}\right) \qquad (3.34)$$

where $J = 1/2$ is the **Jacobian** of the transform. The **marginal** distribution of the intensity is obtained by integrating over the phase which gives the negative exponential distribution as expected

$$f_I(I) = \int_{-\pi}^{\pi} f_E(I,\phi)\,\mathrm{d}\,\phi = \frac{1}{\mu_I}\exp\left(-\frac{I}{\mu_I}\right) \tag{3.35}$$

where $\mu_I = 2\sigma_E^2$. Similarly, integrating over intensity confirms that the marginal distribution of the phase is uniformly distributed

$$f_\phi(\phi) = \int_0^{\infty} f_E(I,\phi)\,\mathrm{d}\,I = \frac{1}{2\pi} \tag{3.36}$$

(3.34), (3.35) and (3.36) also confirm that the intensity and phase are independent and therefore **circularly symmetric** as the product of their marginal distribution gives their joint distribution

$$f_E(I,\theta) = f_I(I) f_\theta(\theta) \tag{3.37}$$

Radar data from homogeneous clutter at low to medium resolution exhibit this type of statistical behaviour. Figure 3.11 shows the histograms of the intensity, the phase and the complex component, of pixel values from a homogeneous region in a C-band SAR image (the grass field in the top left hand corner of the image in Figure 3.7a). The histogram values are given by cross marks and the solid lines represent the expected distributions. For this set of data there is a good fit between histogram and expected distribution.

In order to get a more overall assessment of the clutter statistics chi-squared goodness-of-fit test is performed over the whole of the SAR region in Figure 3.7a. Appendix C describes the procedure for performing the goodness-of-fit test. The test is performed using a stepping window of 5×9. Over each window a comparison is made between the observed histogram of the pixel intensities and the expected histogram from a negative exponential distribution using estimated parameters obtained from the data values. An indicator of the goodness-of-fit is the failure rate at the 5% level for the upper tail probabilities. In Figure 3.12, 12.9% of

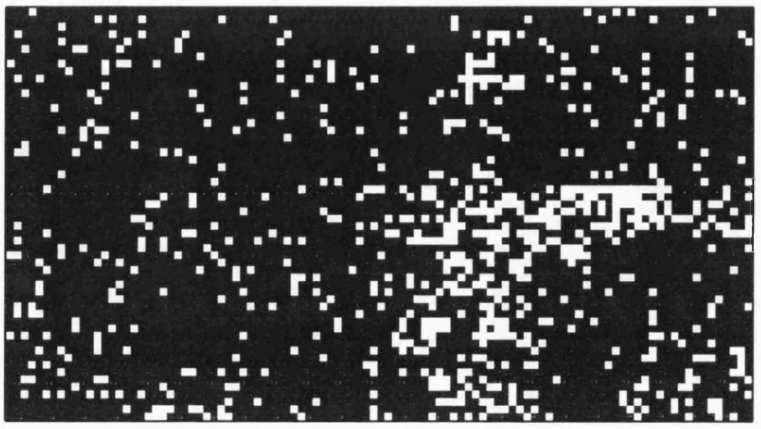


Figure 3.12: The chi-squared failure rate at the 5% level on the upper tail probabilities for the negative exponential distribution applied to the Andover C-band SAR image of Figure 3.7a. The test was performed using a stepping window of 5×9. This reduces the 512×512 image to 102×56 image. The white pixels mark a failure and the black pixels a success at the 5% level for the data fitting to the negative exponential distribution.

the image fail at the 5% level with the negative exponential model. A majority of the failures are in regions of the image where the clutter exhibits significant texture variation. This suggests that the negative exponential distribution is an insufficient representation of the clutter characteristics. Similar tests on the X-band SAR image of Figure 3.7b and the MTI image of Figure 3.6 gave a failure rate of 28.3% and 28.9% respectively. A possible explanation for the failure rate is that the resolution cell is too small for the assumption of a large number of scatterers to hold true. Therefore, the central limit theorem may not be invoked for the real and imaginary parts of the scattered field in (3.33), and non-Gaussian statistics may arise. A number of non-Gaussian models have been considered for high resolution radar data. A few of the more widely used non-Gaussian models are described in the following sections.

3.2.2 Log-normal distributed intensity

The log-normal distribution has been fitted to high resolution radar data from both land and sea clutter [26, 31, 62, 73, 86]. The log-normal is a two parameter distribution. The pdf of the intensity log-normal data is given as

$$f_I(I) = \frac{1}{\sqrt{2\pi}\sigma_{\ln I} I_m} \exp\left(-\frac{\ln^2(I/I_m)}{2\sigma_{\ln I}^2} \right) \tag{3.38}$$

where I_m is the median value and $\sigma_{\ln I}$ is the standard deviation of the underlying normal generating distribution (i.e., standard deviation of $\ln I$). The cumulative distribution function for this model is [7]

$$F_I(I) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\ln\left(\frac{I}{I_m}\right)}{\sqrt{2}\sigma_{\ln I}} \right) \right) \tag{3.39}$$

and the moments are given as

$$\langle I^n \rangle = \exp\left[n\left(\ln I_m + n\sigma_{\ln I}^2/2 \right) \right] \tag{3.40}$$

Since the log-normal distribution has two degrees of freedom it can fit the tail of the data better where it deviates from the negative exponential distribution. Figure 3.13 shows the plot of the log-normal pdf for several values of $\sigma_{\ln I}$ and a median value of 1. The mean-to-median ratio of the distribution is usually higher than those for members of the generalised chi-squared family.

The other attractive feature of the log-normal distribution is that, like the negative exponential distribution, it has simple closed form solutions for its parameter values. Thus when using m independent normalised log samples to **estimate** the parameters I_m and $\sigma_{\ln I}$, the minimum variance estimators are given as [84]

---

[7] erf(z) is the **error function**.

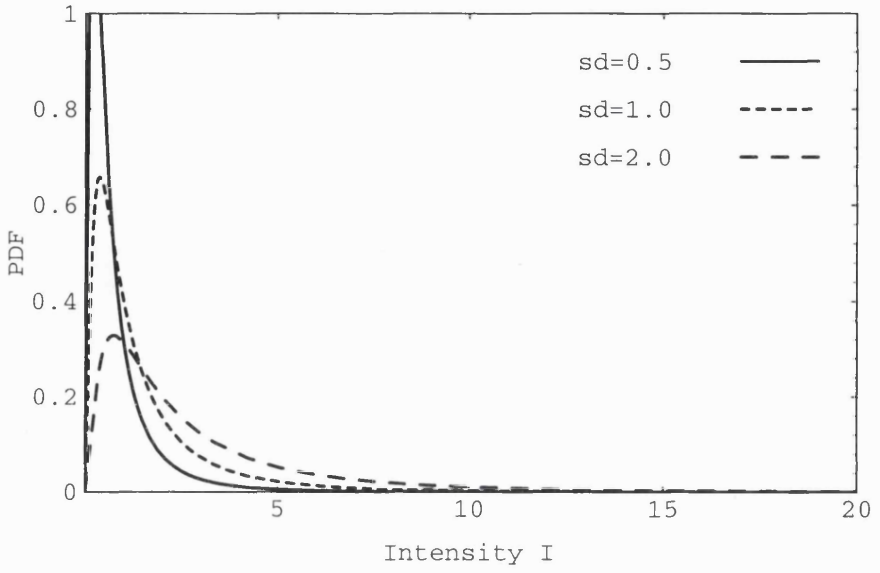


Figure 3.13: The log-normal pdf for median 1 and standard deviation $\sigma_{\ln I} = 0.5, 1, 2$.

$$\hat{I}_m = \left(\prod_{i=1}^{m} I_i \right)^{1/m} \tag{3.41}$$

and

$$\hat{\sigma}^2_{\ln I} = \frac{1}{m} \sum_{i=1}^{m} \ln^2 \left(\frac{I_i}{\hat{I}_m} \right) \tag{3.42}$$

where the 'hat' symbol signifies an estimated quantity.

The log-normal distribution has been fitted to sea clutter [73, 86]. It has been found to be appropriate for urban clutter [95] where the data is very spiky. Oliver and Quegan [62] investigated the log-normal model for land clutter. Using C-band SAR data similar to that shown in Figure 3.7a they found a reasonable fit to homogeneous clutter from fields, but woodlands gave a very poor fit to the log-normal distribution. The texture variation of clutter in a wooded area did not follow a log-normal distribution. The log-normal generally tends to overestimate the dynamic range of the real clutter distribution. Fay *et al.* [27] showed that the portion of the

curve where the fit is poorest is the tail; this is the region most critical for setting False Alarm Rate (FAR). The log-normal distribution has a further drawback, that it predicts a zero probability at zero intensity which is not the case for coherent clutter. Due to these limitations and because of the lack of a suitable physical model, the log-normal distribution is not considered as a strong candidate for the statistics of scattering from land.

3.2.3 Weibull distributed intensity

The Weibull family of distributions, like the log-normal, is a two parameter family. The pdf of Weibull distributed clutter intensity is of the form

$$f_I(I) = \left(\frac{I}{b_W}\right)^{c_W-1}\left(\frac{c_W}{b_W}\right)\exp\left[-\left(\frac{I}{b_W}\right)^{c_W}\right] \tag{3.43}$$

where b_W is the scale parameter relating to the median value of the intensity and c_W is a parameter relating to the skewness of the distribution. The cumulative distribution is given as

$$F_I(I) = 1-\exp\left[-\left(\frac{I}{b_W}\right)^{c_W}\right] \tag{3.44}$$

and the moments as

$$\langle I^n\rangle = {b_W}^n\Gamma\left(1+n/c_W\right) \tag{3.45}$$

Figure 3.14 shows the pdf plot of the Weibull distribution for b_W =10 and three values of c_W. The negative exponential distribution is a special case of the Weibull model for c_W =1.

In the early 70's there were many researchers investigating the non-Gaussian characteristics of high resolution clutter statistics. The log-normal distribution described in the previous section gave a better fit compared to the negative exponential model but still fell well short of describing adequately the single point statistics of coherent surface clutter. Goldstein [31] called attention to the Weibull distribution as a suitable candidate. Boothe [13] carried out an extensive study of land

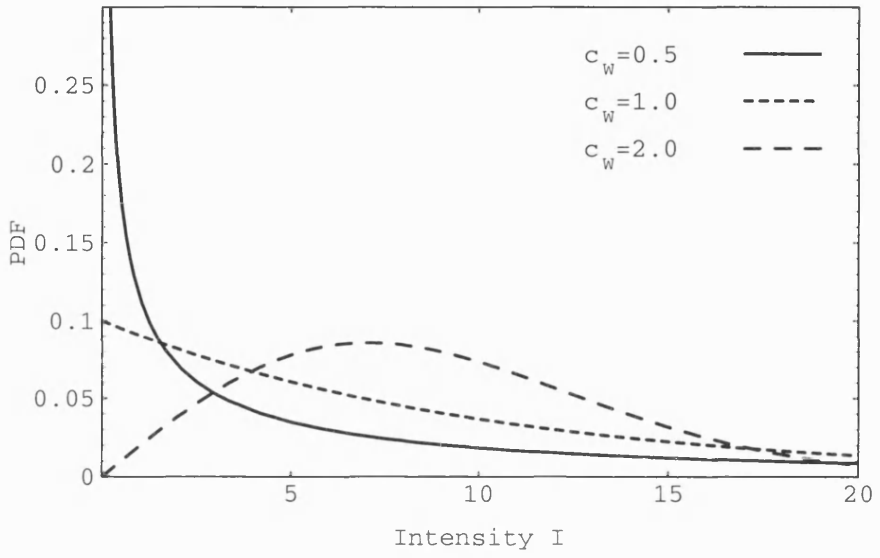


Figure 3.14: Weibull pdf for b_W =10 and c_W =0.5, 1, 2. For c_W =1 the Weibull distribution reduces to a negative exponential distribution.

clutter and showed a very good fit of the data to the Weibull distribution. This provided a strong stimulus for many other researchers to investigate the suitability of the Weibull distribution for describing clutter characteristics. Ekstrom [24] showed that the skewness of the Weibull distribution increases as the radar depression angle decreases. Schleher [74] investigated the Weibull distribution for both land and sea clutter and showed that generally the Weibull distribution represents the real clutter distribution more accurately then either the log-normal or the negative exponential model. Fay *et. al.* [27] using the DERA Canberra X-band radar data showed that the skewness of the distribution varies with the resolution cell size. Investigations carried out by Japanese researchers using L and X-band radar data showed that the Weibull distribution is appropriate for a wide range of clutter types ranging from land [79], sea [78], weather [77] and sea-ice [54].

The Weibull model gives a good empirical fit to high resolution radar clutter. The additional degree of freedom enables the model to get a much better fit to real clutter data compared to the single parameter negative exponential model. As the clutter data gets more spiky the skewness parameter of the Weibull distribution

gets smaller. However, with such low skewness parameters the Weibull yields an incorrectly large probability for zero intensity [62].

The Weibull distribution, unlike the negative exponential and the log-normal distribution, does not have closed form solution for the parameter estimators [35]. The optimum parameter estimators have to be calculated using an iterative process which is computationally expensive and is undesirable in real time systems.

Despite the empirical evidence supporting the Weibull model there is no theoretical justification which links the discrete scatterer model to this distribution. This means that there is no definite way of linking the distribution parameter values to a physical model of the scattering scene. The gap between an acceptable empirical model and a plausible physical justification was bridged by the introduction of the K-distribution model which is described in the next section.

3.2.4 K-distributed intensity

At low resolution many of the clutter features like trees and houses are much smaller than a resolution cell. The contributions from the radar cross section (RCS) fluctuations are averaged out so that no spatial variation is visible. However, at higher resolution the length scale of many of the clutter features are longer than the resolution and the clutter RCS fluctuates from resolution cell to resolution cell. This variation in the underlying RCS is visible as a change in image contrast. This can be seen within the wooded region of Figure 3.7a (the vertical narrow strip) and between the wooded and grass region.

A possible model to describe the scattered field is that it is the result of a negative exponential process due to the coherent interference of many scatterers, but the mean characterised by the underlying surface RCS is also random variable which depends on the physical properties (such as dielectric constant) of those elements. The pdf of the intensity is then a combination of the negative exponential pdf and the pdf of the clutter RCS which is given by

$$f_I(I) = \int_0^\infty f_{I|\rho}(I|\rho) f_\rho(\rho) d\rho \tag{3.46}$$

where $f_{I|\rho}$ is the negative exponential distribution, termed the speckle component, and f_ρ is the RCS component. This representation of the observed intensity statistics as a product model [91] was a major breakthrough in the understanding of radar image properties.

Empirical analysis of sea clutter [91, 92] and land clutter [72] showed that the underlying surface RCS, was usually consistent with a gamma pdf, given by

$$f_\rho(\rho) = \frac{1}{\Gamma(\nu)}\left(\frac{\nu}{\mu_\rho}\right)^\nu \rho^{\nu-1} \exp\left[-\frac{\nu\rho}{\mu_\rho}\right] \tag{3.47}$$

where μ_ρ is the mean RCS and ν is the order parameter. The cumulative distribution is given as

$$F_\rho(\rho) = 1 - \exp\left(-\frac{\nu\rho}{\mu_\rho}\right)\sum_{k=0}^{\nu-1}\frac{1}{k!}\left(\frac{\nu\rho}{\mu_\rho}\right)^k \tag{3.48}$$

and the moments as

$$\langle \rho^n \rangle = \left(\frac{\mu_\rho}{\nu}\right)^n \frac{\Gamma(n+\nu)}{\Gamma(\nu)} \tag{3.49}$$

Figure 3.15 shows how the shape of the gamma distribution changes with order parameter. When ν =1 the gamma distribution reduces to the negative exponential distribution. As the order parameter tends to infinity the gamma distribution tends to a Gaussian distribution with mean μ_ρ and zero variance.

For the gamma distributed RCS the integral in (3.46) has a closed form and is given by

$$f_I(I) = \frac{2}{I\Gamma(\nu)}\left(\sqrt{\frac{\nu I}{\mu_I}}\right)^{\nu+1} K_{\nu-1}\left(2\sqrt{\frac{\nu I}{\mu_I}}\right) \tag{3.50}$$

where μ_I is the mean, ν the order parameter and $K_{\nu-1}[\cdot]$ is the **modified Bessel function** of the second kind of order ν -1. This distribution has been termed the

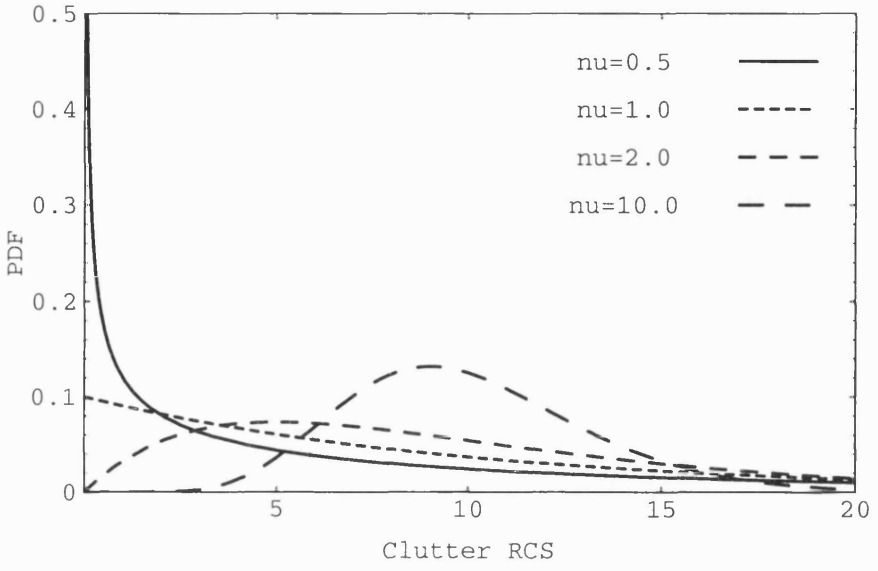


Figure 3.15: Gamma pdf of the underlying surface RCS for μ_ρ =10 and ν =0.5, 1, 2, 10. As ν gets large the gamma distribution tends to a Gaussian.

K-distribution and its mean and order parameters are same as those of the gamma distribution. The cumulative distribution is given as

$$F_I(I) = 1 - \frac{2}{\Gamma(\nu)}\left(\sqrt{\frac{\nu I}{\mu_I}}\right)^{\nu} K_\nu\left(2\sqrt{\frac{\nu I}{\mu_I}}\right) \tag{3.51}$$

and the moments as

$$\langle I^n \rangle = n!\frac{\Gamma(n+\nu)}{\Gamma(\nu)}\left(\frac{\mu_I}{\nu}\right) \tag{3.52}$$

Figure 3.16 shows the K-distribution for the intensity corresponding to the gamma distributed RCS shown in Figure 3.15. For large order parameter value the K-distribution is equivalent to a negative exponential model. This is so since the underlying gamma distributed surface RCS is constant for large order parameter values. For small values of ν the tail of the distribution gets longer. Longer tails are a feature of spiky clutter which have a much higher contrast. This can be seen from the **coefficient of variation**, given as

$$v_I^2 = \frac{\sigma_I^2}{\mu_I^2} = 1 + \frac{2}{\nu} \tag{3.53}$$

The square root of the coefficient of variation gives the image contrast. For $\nu \to \infty$ the right hand side of (3.53) is equal to 1 and the standard deviation of the intensity equals to its mean value. This means that the intensity will on average fluctuate away from its mean value by an amount equal to its mean value. This is often interpreted as a signal to noise ratio of one and is the case for negative exponential distributed intensity. For smaller values of ν, it is evident from (3.53) that the intensity will fluctuate away from its mean value by an amount significantly greater than its mean value.

One of the first evidence that the K-distribution is a reasonable model for the scattered field was obtained for non-coherent sea clutter data [91]. Since then it has been fitted to coherent sea clutter [3, 94], land clutter [41, 56, 57] and weather clutter [53]. The essential features of the K-distribution results from the radar observing a multiscale process [93]. The scales smaller than the resolution cell size cause the negative exponential intensity 'speckle', whilst the larger scales are resolved by the radar and produce the modulation. Jakeman and Pusey [37] proposed a theoretical model that provided important insight into the basis for this empirical model.

The simple scattering model considered by (3.9) assumed a fixed number of scatterers and in the limit of $N \to \infty$ the detected field intensity is negative exponentially distributed. However, at high resolution the effective number of scatterers within a resolution cell will undoubtedly fluctuate and therefore N would be a random variable itself. For the random walk model described in Section 3.2, Jakeman [38] suggested that the number of steps could be modelled by a Markov process controlled by the birth-death-immigration equation

$$\frac{df_N}{dt} = \mathcal{D}(N+1)f_{N+1} - \left((\mathcal{B}+\mathcal{D})N + \mathcal{J}\right)f_N + \left(\mathcal{B}(N-1) + \mathcal{J}\right)f_{N-1} \tag{3.54}$$

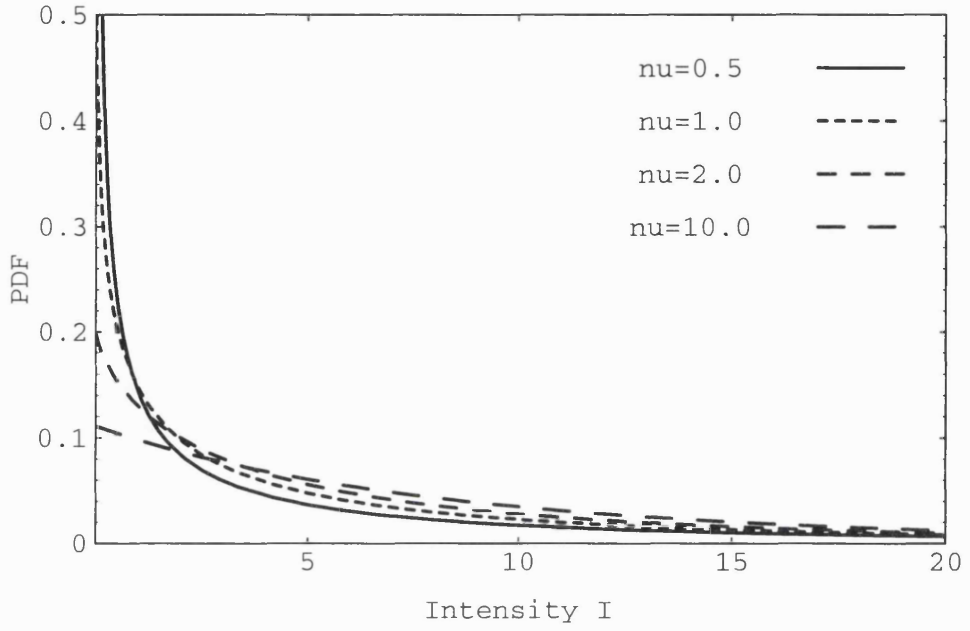


Figure 3.16: Pdf's of K-distribution for μ_I =10 and ν =0.5, 1, 2, 10. At this linear scale a K-distribution with order parameter 10 or higher is virtually indistinguishable from a negative exponential distribution with the same mean value.

where $\mathcal{B}$ is the birth rate, $\mathcal{D}$ is the death rate and $\mathcal{J}$ is the spontaneous immigration rate. An equilibrium solution to (3.54) exists when the death rate is greater than the birth rate, in which case f_N tends to a negative binomial distribution given as

$$f_N = \binom{N+\nu+1}{N} \frac{\left(\mu_N/\nu\right)^N}{\left(1+\mu_N/\nu\right)^{N+\nu}} \tag{3.55}$$

where the mean $\mu_N = \mathcal{J}\left(\mathcal{D}-\mathcal{B}\right)$ and the order parameter $\nu = \mathcal{J}/\mathcal{B}$ as $t \to \infty$. The order parameter characterises the clustering or bunching of scatterers in a resolution cell and is related to the variance of the number of steps as

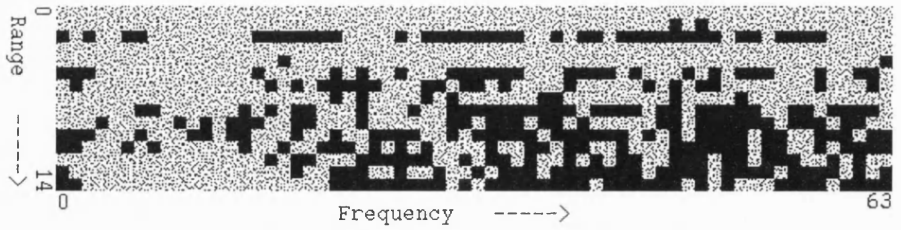


Figure 3.17: The ML decision between the K-distribution and the negative exponential distribution for the Canberra X-band MTI image shown in Figure 3.6. For the original data, blocks of 91 range gates in each of the 64 frequency bins are used for the ML decision. This gives an ML classified image which has 64 frequency bins (horizontal axis) and 15 blocks of data in range (vertical axis).

$$\sigma_N^2 = \mu_N^2\left(\frac{1}{\mu_N}+\frac{1}{\nu}\right) \tag{3.56}$$

Considering (3.23) for a negative binomial distributed N and averaging over the fluctuations in N gives

$$C_{\mu_N}(u) = \left[1+\frac{\mu_N}{\nu}\left(1-\langle J_0(ua)\rangle\right)\right]^{-\nu} \tag{3.57}$$

Once again scaling the step length a through $a \to a/\sqrt{\mu_N}$ then in the limit $\mu_N \to \infty$ the **characteristic function** of the output is given as

$$\lim_{\mu_N\to\infty} C_{\mu_N}(u) = \left[1+\frac{u^2\langle a^2\rangle}{4\nu}\right]^{-\nu} \tag{3.58}$$

The Fourier inversion of (3.58) corresponds to the K-distribution amplitude pdf

$$f_A(A) = \frac{4}{\Gamma(\nu)}\left(\frac{\nu}{\langle A^2\rangle}\right)^{(\nu+1)/2} A^{\nu} K_{\nu-1}\left(2\sqrt{\frac{\nu}{\langle A^2\rangle}}A\right) \tag{3.59}$$

for which the intensity pdf is that of (3.50).

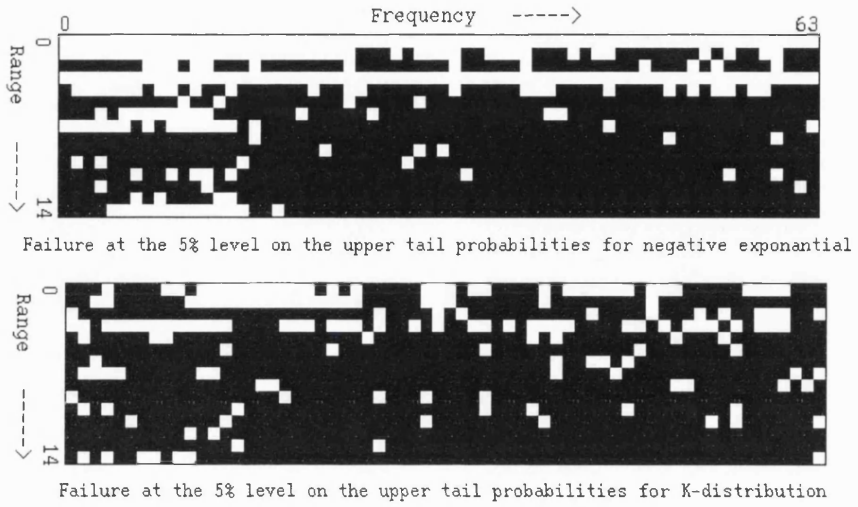


Figure 3.18: The chi-squared test performed on the MTI image shown in Figure 3.6. The white pixels mark a failure and the black pixels a success at the 5% level for the data fitting to the distribution under test.

Thus when the distribution of the number of steps is negative binomial the resultant intensity is K-distributed [37, 38]. The negative binomial distribution is in fact the discrete analogue of the gamma distribution, and it has been shown that the output of a continuous birth-death-immigration process, with a death rate greater than the birth rate, tends to a gamma distribution as $t \rightarrow \infty$ [39]. The mean and the order parameter of the gamma distribution (and therefore of the K-distribution) are then defined in the same way as the mean and order parameter of the negative binomial distribution in the discrete case.

A comparison was made between the suitability of the K-distribution and the negative exponential distribution to model the intensity statistics of the MTI data shown in Figure 3.6. A maximum likelihood (ML) classifier (see Appendix D) was used to select the model which best describes the data. This test calculates the total probability of occurrence P_T for m random samples assuming a given probability distribution for the data samples and chooses the distribution for which P_T is maximum. The ML test was carried for blocks of 91 range gates in each of the 64 frequency bins. Figure 3.17 shows that the whole of the clutter band and the side-lobe dominated region is classified as K-distribution. The area beyond the main

clutter band comprises a mixture of distributions. Altogether, 66% of the image is classified as K-distributed data.

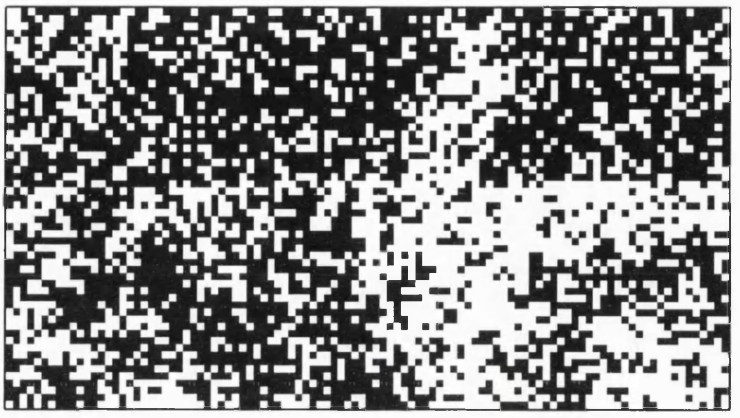
In order to get an idea about how well each of the models fitted to the data the chi-squared test was carried out on the MTI data. The size of the data blocks selected for the chi-squared test were chosen to be same as those used for the ML classification test. The top image in Figure 3.18 shows 28.9% of the image failed at the 5% level with the negative exponential model confirming the inadequacy of this type of model to describe land clutter. The same test with the K-distribution gives a failure rate of 17.8% (bottom image in Figure 3.18). This shows that K-distribution is a better fit than the negative exponential model, although this distribution alone is not able to describe all the features in land clutter.

The ML classification results for the SAR intensity images shown in Figure 3.7 are given in Figure 3.19. The ML results were obtained using a stepping window of 5×9, same as for the chi-squared test on the SAR image detailed in Section 3.2.1. 43.3% of the C-band SAR (Figure 3.19a) and 30% of the X-band SAR image (Figure 3.19b) are classified as K-distributed. Area of woodland and edges have been classified as K-distributed whereas low contrast areas of grass and fields as negative exponential distributed. Where the image is classified as negative exponential the values of the ML ratio are close to 1 indicating that both distributions are equally suitable for the data. Thus whereas the negative exponential model fits the low contrast homogeneous regions, the K-distribution is a reasonable fit for the whole scene.

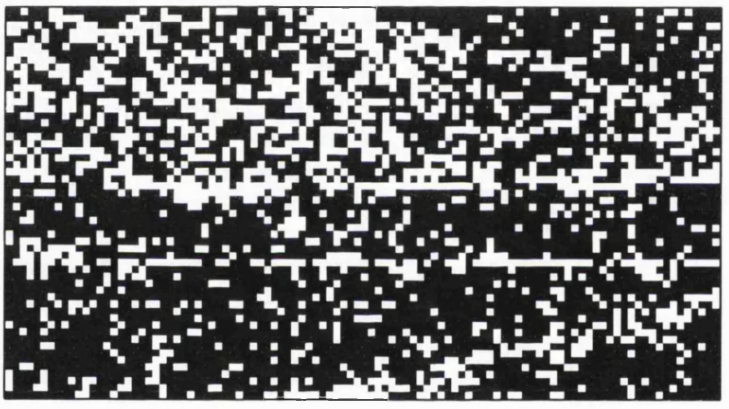
The chi-squared test and the ML classification were carried out by estimating parameters of the distribution from the data samples. The negative exponential is a single parameter distribution and the optimum estimate is just the average of the data samples i.e.,

$$\hat{\mu}_I = \hat{I} = \frac{1}{m}\sum_{i=1}^{m} I_i \tag{3.60}$$

where m is the number of independent samples.



(a)



(b)

Figure 3.19: The ML decision between the K-distribution (marked as white pixels) and the negative exponential distribution (marked as black pixels) for (a) the Andover C-band and (b) the Canberra X-band SAR images shown in Figure 3.7. The ML test performed using a stepping window of 5×9 on a 512×512 intensity image.

The K-distribution, however, does not have a simple closed form solution for the optimum parameter estimates [36]. Numerous approximate solutions have been considered which aim to minimise the error on the estimates of the parameter values for the K-distribution. The next chapter addresses the problem of optimum

parameter estimates for the K-distribution and derives a new estimator for the order parameter.

3.3 Summary

In this chapter:

- An overview was given for the forward problem which solves the Maxwell equations to predict the scattered field from a rough surface.
- It was shown that solutions are available for only very simple scattering scenarios and the basic assumptions were described for the Kirchhoff and the SPM approximate methods. The scattered field from a surface with an arbitrarily defined surface height profile can be calculated using numerical solutions. However, it was shown that for a realistic surface the computational complexity can be prohibitively large.
- It has been described how the field detected from a resolution cell may be represented as a sum of contributions from discrete elementary scatterers.
- The conditions required for strong scattering arising from uniformly distributed phase of the scatterers were described.
- It was shown that, for strong scattering, the phase and the complex components of the detected field contained little information.
- It was shown how in the limit of large number of scatterers the observed field will have a negative exponential intensity distribution which is a reasonable model for low resolution radar data. However, using chi-square goodness-of-fit test on real examples of MTI and SAR data, it was shown that at high resolution the assumption of negative exponential distributed intensity is no longer valid.
- The shapes of the log-normal and Weibull pdf were described and their suitability for modelling non-Gaussian clutter was discussed.

- It was shown how the intensity of the detected field may be modelled as the product of a negative exponential speckle process and a gamma distributed surface RCS which gives rise to the K-distribution.
- The K-distribution was shown to be the result of a negative binomial distribution in the number of effective scatterers in a resolution cell.
- Considering the results in open literature and the analysis carried out on SAR and MTI data it was concluded that the K-distribution is the preferred model for coherent land clutter.

Chapter 4

4. K-distribution parameter estimators

Statistical models which can accurately describe coherent land clutter are used to define image analysis algorithms like detection, segmentation, despeckling and clutter classification. The statistical models are characterised by a finite number of parameters. These parameters for the statistical models are not known *a priori* and have to be estimated from the radar data using local statistics. Therefore, a major portion of radar image analysis involves the estimation of the parameters of the clutter models.

In the previous chapter the K-distribution was shown to be a good model for coherent land clutter. The parameters that are required to be estimated from the data for this model are the mean intensity and the order parameter.

The optimum parameter estimate is obtained using the maximum likelihood (ML) solution. The K-distribution does not have a closed form for the ML estimates. Therefore, sub-optimal estimators have to be devised for the K-distribution parameters. Before discussing the sub-optimal estimators, the full expression for the ML solution of the K-distribution parameters is described in Section 4.1 and it is shown that it cannot be solved analytically to obtain the parameter estimates. A suitable sub-optimal estimator for mean intensity is simply the arithmetic average of the pixel intensities. This is in fact an optimum estimator for the mean intensity when there is no texture variation and the clutter is negative exponentially distributed. Section 4.2 derives the error performance for the mean intensity estimator.

Estimating the second parameter of the K-distribution is addressed in Section 4.3. Current texture measure estimators for the order parameter are shown to be suboptimal. A neural net based estimator is devised which improves upon the existing texture measure estimators. Prompted by the neural net results, a new hybrid texture measure estimator is proposed which almost matches the ML performance.

4.1 ML estimation of the mean and order parameter

The use of statistical models to describe the detected field requires us to have suitable parameter values that can be assigned to the model. These parameters have to be estimated from the data in the absence of prior knowledge. The ML estimation gives the optimal solution for the parameter estimators when the form of the distribution to be estimated is known. This is because the limiting distribution of the ML estimate is normal around the true value as mean and with a variance achieving the Cramer-Rao lower bound [20].

If m independent random samples, $\{x_1, x_2, \ldots, x_m\}$ are drawn from a distribution with l parameters, $\alpha_1, \alpha_2, \ldots, \alpha_l$, then the total probability of occurrence is given as[8]

$$P_T\left(\{x_i\}|\alpha_1, \alpha_2, \ldots, \alpha_l\right)d\mathbf{x} = \prod_{i=1}^{m} f_x\left(x_i|\alpha_1, \alpha_2, \ldots, \alpha_l\right)dx \tag{4.1}$$

where $f_x\left(x_i|\alpha_1, \alpha_2, \ldots, \alpha_l\right)$ is the pdf for the given statistical model. The ML solution is then a set of values $\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_l$ that maximises the likelihood function $L\left(\alpha_1, \alpha_2, \ldots, \alpha_l\right) = P_T\left(\{x_i\}|\alpha_1, \alpha_2, \ldots, \alpha_l\right)$. Assuming that the joint pdf is unimodal, this is found by solving the partial differential equation

$$\frac{\partial L}{\partial \alpha_\kappa}\left(\alpha_1, \alpha_2, \ldots, \alpha_l\right) = 0 \tag{4.2}$$

where α_k is the kth parameter. Since the logarithm function for a positive real variable is a strictly monotonic increasing function, the ML estimate will also maximise

[8] $\{x_i\}$ denotes a vector containing all x_i for $i = 1, \ldots, m$.

the log-likelihood function, ln L. The ML estimate is often more easily found by differentiating the log-likelihood with respect to the parameter values. The ML estimates for the parameters are then given as

$$\begin{aligned}\frac{\partial \ln L(\alpha_1,\alpha_2,\ldots,\alpha_l)}{\partial \alpha_k} &= \frac{\partial \ln \prod_{i=1}^{m} f_x(x_i|\alpha_1,\alpha_2,\ldots,\alpha_l)}{\partial \alpha_k} \\ &= \frac{\partial}{\partial \alpha_k}\sum_{i=1}^{m} \ln f_x(x_i|\alpha_1,\alpha_2,\ldots,\alpha_l) = 0\end{aligned} \tag{4.3}$$

Given m statistically identical independent K-distributed intensity values $I_1, I_2, \ldots, I_m$ with unknown mean and order parameter, the log-likelihood that they have mean μ_I and order parameter ν is given by

$$\begin{aligned}\ln L(\mu_I,\nu) = m\left(\ln 2 + \frac{\nu+1}{2}(\ln \nu - \ln \mu_I) - \ln \Gamma(\nu)\right) \\ + \frac{\nu-1}{2}\sum_{i=1}^{m} \ln I_i + \sum_{i=1}^{m} K_{\nu-1}\left(2\sqrt{\frac{\nu I_i}{\mu_I}}\right)\end{aligned} \tag{4.4}$$

The ML estimates of the mean and order parameter of the K-distributed intensity are given by the values of μ_I and ν which maximises the log-likelihood. Solving for where the derivatives of the log-likelihood with respect to the mean and the order parameter both equal zero, and using (F.14) gives [61]

$$\sqrt{\frac{\mu_I}{\nu}}(\nu+1) = \frac{1}{m}\sum_{i=1}^{m}\sqrt{I_i}\,\frac{K_\nu\left(2\sqrt{\frac{\nu I_i}{\mu_I}}\right) + K_{\nu-2}\left(2\sqrt{\frac{\nu I_i}{\mu_I}}\right)}{K_{\nu-1}\left(2\sqrt{\frac{\nu I_i}{\mu_I}}\right)} \tag{4.5}$$

and [9]

$$\ln \psi(\nu) + \ln\sqrt{\frac{\mu_I}{\nu}} = \frac{1}{m}\sum_{i=1}^{m}\sqrt{I_i} + \frac{1}{m}\frac{\frac{\partial}{\partial \nu}K_{\nu-1}\left(2\sqrt{\frac{\nu I_i}{\mu_I}}\right)}{K_{\nu-1}\left(2\sqrt{\frac{\nu I_i}{\mu_I}}\right)} \tag{4.6}$$

[9] $\psi(z)$ is the **Digamma function**.

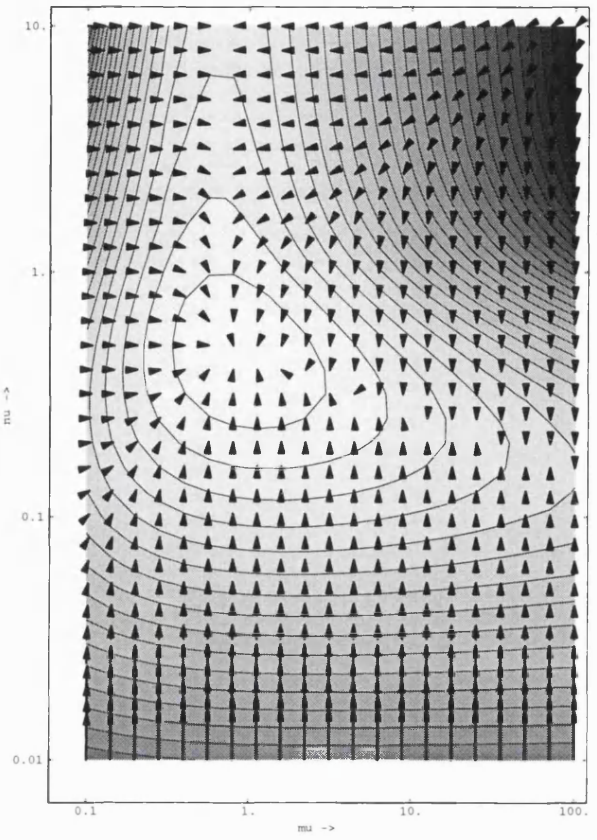


Figure 4.1: The likelihood surface for 256 samples of K-distributed intensity data with μ_I =1 and ν =0.5. The likelihood values are increasing as the colour changes from dark to light. The arrow heads indicate the direction of the slope at any point on the surface.

(4.5) and (4.6) do not have a closed form solution for μ_I and ν. The optimum solution, therefore, requires a two dimensional numerical search over μ_I and ν. This is computationally very expensive. Figure 4.1 shows the likelihood surface for 256 K-distributed intensity samples with mean 1 and order parameter 0.5. Superimposed on the plot of the likelihood surface are arrow marks showing the direction in which the numerical search for optimum parameters of the K-distribution will progress at any point on the surface. The arrows indicate the convergence of the numerical solution to the peak on the surface. However, the convergence rate is very slow and this method for estimating the parameters is not very efficient. Considerable attention has been given to various simpler estimators, suitable for real

time evaluation, which approximate the ML solution for the K-distribution parameters. The following sections address the issue of non-optimal parameter estimators for the K-distribution. The mean parameter has a reasonable estimator in terms of the average intensity and Section 4.2 analyses the performance of this estimator. A suitable estimator for the order parameter is less forthcoming and Section 4.3 describes in detail the development of appropriate estimators.

4.2 Non-optimal estimation of the mean

A simple estimator for the mean value is the arithmetic average of the pixel intensities. This gives an unbiased estimate of the mean if the samples are from a **stationary** random process.

For a homogeneous region where the surface RCS is constant, which implies a negative exponentially distributed intensity, the ensemble average intensity gives the ML estimate of the mean intensity. From (4.1) the log-likelihood that the intensity values have mean μ_I when the pixel intensities are negative exponentially distributed is given as

$$\ln\left(L(\mu_I)\right) = \ln \prod_{i=1}^{m} \frac{1}{\mu_I} \exp\left(-\frac{I_i}{\mu_I}\right) = -m \ln \mu_I - \frac{m}{\mu_I} \hat{I} \tag{4.7}$$

where $\hat{I}$ is the average intensity defined by

$$\hat{I} = \frac{1}{m} \sum_{i=1}^{m} I_i \tag{4.8}$$

Using (4.3) it can be shown by solving

$$\frac{d(\ln L)}{d\mu_I} = \frac{m}{\mu_I^2} \hat{I} - \frac{m}{\mu_I} = 0 \tag{4.9}$$

that the ML estimate of the mean is given by the average intensity when the intensities are negative exponentially distributed.

The performance of an estimator can be judged by the bias and the variance of the estimated quantity. The bias and variance of a function of estimated parameters may be approximated by expanding the function around the expected

values of the estimate [64]. For the negative exponential distribution, the bias and the variance in the average intensity estimator of the mean can therefore be approximate by expanding $\hat{\mu}_I$ around the true mean value (see Appendix A.1) and are given by

$$\Delta\mu_I = \langle \hat{\mu}_I \rangle - \mu_I = 0 \tag{4.10}$$

$$\sigma^2_{\hat{\mu}_I} = \left\langle \hat{\mu}_I{}^2 \right\rangle - \langle \hat{\mu}_I \rangle^2 = \frac{1}{m}\mu_I^2 \tag{4.11}$$

where m is the number of samples used to estimate the mean and the error predictions are derived to first order in $1/m$. The average intensity is an unbiased estimate for the mean of the negative exponential distribution and the variance of the estimate decreases as the number of samples increases.

For the K-distribution, the first moment of the intensity also completely defines the mean μ_I (3.52). Therefore, the average of m statistically independent K-distributed intensity pixels can also be used to estimate the unknown mean μ_I. This average intensity is also an unbiased estimate for the K-distribution mean but the variance of the estimate is a function of the order parameter and is given by (see Appendix A.1)

$$\sigma^2_{\hat{\mu}_I} = \frac{1}{m}\mu_I^2\left(1+\frac{2}{\nu}\right) \tag{4.12}$$

Thus the variance in the estimate of the mean obtained from the average intensity is higher over a homogeneous region if the intensity values are K-distributed than if they were negative exponentially distributed with the same mean value. Figure 4.2 compares the variance in the estimated intensity mean for K-distributed and negative exponential data. The variance of the mean estimator for the K-distribution increases as ν decreases, i.e. the estimate is poorer for spikier data. From (4.12) it can also be seen that the variance of the mean estimator is inversely proportional to the sample size. Therefore, this estimator will not be suitable for very small sample sizes.

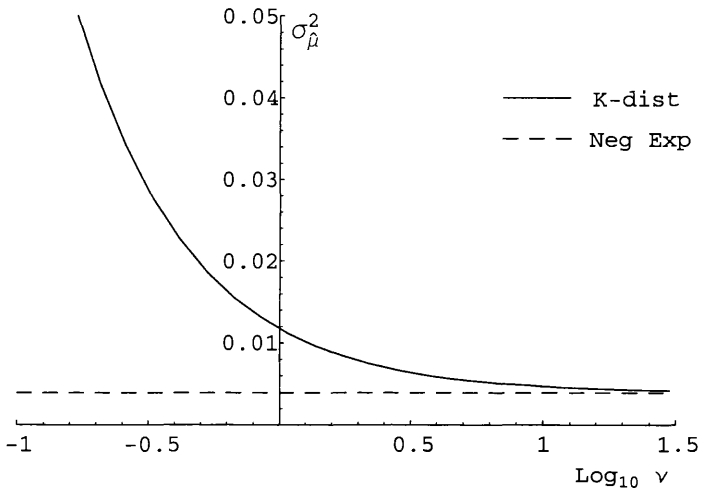


Figure 4.2: Comparison of the variance of the average intensity estimator for the K-distribution and negative exponential distribution intensity for μ_I =1 and m =256 samples.

Although the local mean is not an optimum estimator for the mean of the K-distribution, it greatly reduces the complexity of the two dimensional numerical search required to solve (4.5) and (4.6) for the optimum parameters of the K-distribution. The one dimensional numerical search for the optimum value of ν would still be computationally prohibitively expensive. In the following section various sub-optimum estimators for the ν parameter of the K-distribution intensity are discussed which attempt to approximate the ML solution for ν.

4.3 Non-optimal estimation of the order parameter

The intensity contrast gives a simple estimate for ν but has large errors for small ν or small sample numbers [57]. Oliver [61] proposed an estimator based on the normalised log texture measure, which has a much better error performance.

Texture estimators consist of two operations. A texture measure is derived from the data followed by an inversion operation to obtain the order parameter.

Analysis of the estimator then involves evaluating the error in the texture measure and converting it into an error in the estimate of ν.

Using first order approximation in the expansion of the estimate about its true value the bias and variance in the estimate of ν from the normalised log measure can be derived [61]. Whilst this estimator is well behaved for small values of ν, the estimates become increasingly suboptimal as ν gets larger. Lombardo and Oliver [50] demonstrated that the problem is associated with the highly non-linear inversion process from the texture measure to the order parameter of the (assumed) K-distributed texture. Introducing a reciprocal order parameter $t \equiv 1/\nu$ results in a more stable inversion since the texture measure now tends to a linear dependence on this parameter for large t. This estimator has a more direct physical relevance to the texture inhomogeneity since it is related to the variance of the underlying gamma distribution.

Section 4.3.1 describes the intensity contrast estimator and drives the theoretical errors in estimating the order parameter. Section 4.3.2 then goes on to describe the normalised log estimator and derives the estimates in terms of the t parameter. It is shown that this alternative estimator has better variance results compared to the ν estimator. Although Lombardo and Oliver [50] proved that the normalised log provides the best performance of the measures they considered, it is still suboptimal. Joughin *et al.* [42] demonstrated that the numerical ML approach does indeed give improved estimates.

Since radar detection and estimation performance in K-distributed clutter are limited by the errors in the estimate of t (and therefore ν) parameter, there is a requirement for estimators that can approach the optimum solution. In an attempt to obtain a better estimator for the t parameter of the K-distribution a neural net based solution is considered. Many authors have proposed the use of neural nets for clutter classification [10, 22, 43, 59]. Neural nets provide a means for developing an estimator which carriers out the non-linear transformation of the K-distributed data to the t parameter value. One particular drawback of texture based estimators is that they assume that independent samples are available for parameter

estimation which is often not the case with high resolution images. Neural nets have the potential to perform parameter estimation with correlated data since they infer subtle unknown relationships from the data. However, since in this thesis we merely aim to demonstrate the viability of neural nets as a parameter estimator, we restrict the investigation to uncorrelated K-distributed data.

In Section 4.3.3 we discuss the development of the neural net based estimators. In using neural nets as functional estimators, the data have to have some appropriate pre-processing before they can be presented to the net. When the preprocessing involved just log scaling the intensity data the net was unable to produce a good solution. However, training the net on various texture measures of the data resulted in better estimates for the t parameter. Experimentation with the neural nets showed that two texture measures, the mean normalised log and the contrast of the amplitude, when used as a pair of inputs to train the net resulted in an estimator which outperforms the normalised log estimator. So with the help of the neural nets it was possible to identify a solution which closed the gap between the performance of the texture measure estimators and the ML solution.

Prompted by the results of the neural net research a new estimator is proposed which combines the normalised log and the amplitude contrast texture measure. The performance of this hybrid estimator depends upon the relative weighting of the two moments. A constant weighting provides a good estimate only over a certain range of the parameter values [36]. In order to obtain the best estimates over the whole of range of t values of interest optimum weight values are required which would minimise the variance on the values. These optimum weight values cannot be defined without knowing the parameter value. To overcome this problem an iterative procedure is adopted. An initial estimate of t is made using an arbitrary weight value which is used to make a more informed guess of the weight value. This weight value gives a better estimate on t which in turns allows for a better choice for the weight value. This way an estimate of t can be made using the most appropriate weight value to be used in the texture measure. In Section 4.3.4 we analyse the predicted errors for this new estimator and discuss the implications

for optimal parameter estimation. It is shown that near optimal estimates are obtained using an adapted weighting scheme for the hybrid normalised log estimator.

4.3.1 Intensity contrast estimator

The intensity contrast can be used to encode the image texture information and is given by

$$V \equiv \frac{\langle I^2 \rangle}{\langle I \rangle^2} - 1 \tag{4.13}$$

Using (3.52) this is related to the order parameter as

$$V \equiv 1 + \frac{2}{\nu} \tag{4.14}$$

The expectation values are true theoretical values of the moments which cannot be obtained by working within the constraint of a finite sample size. For m statistically identical independent K-distributed intensity values with unknown parameter ν, the first moment can be estimated by (4.8) and the second moment by

$$\hat{I^2} = \frac{1}{m} \sum_{i=1}^{m} I_i^2 \tag{4.15}$$

therefore the estimated values of the texture measure, represented by $\hat{V}$, can be expressed in terms of these quantities. The estimator error is the difference between the expectation value and the estimated quantity and establishes the usefulness of the estimator. As the order parameter is obtained from the texture measure using the inversion relationship of (4.14), the errors in the estimate of the texture measure can be translated into errors in the estimate of the order parameter values.

In Appendix A.2 the error predictions for the intensity contrast estimator are derived to first order in $1/m$ [50]. The fractional bias and the variance of the estimated texture measure $\hat{V}$ are given as

$$\frac{\Delta \hat{V}}{V} = -2 \frac{1}{m} \left(1 + \frac{1}{\nu}\right) \left(1 + \frac{6}{\nu}\right) \left(1 + \frac{2}{\nu}\right)^{-1} \tag{4.16}$$

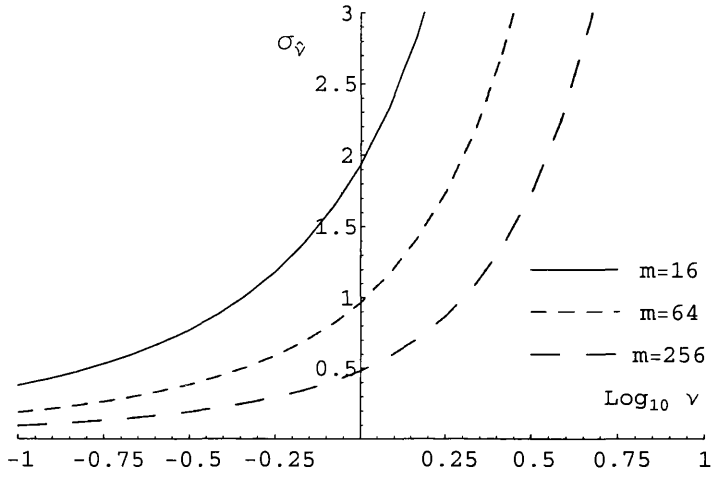


Figure 4.3: The predicted standard deviation of $\hat{v}$ obtained using the intensity contrast estimator for m =16, 64, 256.

$$\sigma_{\hat{v}}^2 = 4\frac{1}{m}\left(1+\frac{1}{v}\right)\left(1+\frac{4}{v}\right)\left(1+\frac{5}{v}\right) \tag{4.17}$$

(4.17) gives the theoretical error in obtaining the intensity contrast measure, V, from estimated quantities. This can then be translated into errors in v using the relationship (see Appendix B)

$$\sigma_{\hat{v}}^2 = \frac{\sigma_{\hat{V}}^2}{|dV\,/\,dv|^2} \tag{4.18}$$

which gives the predicted variance of $\hat{v}$ as

$$\sigma_{\hat{v}}^2 = \frac{v^4}{m}\left(1+\frac{1}{v}\right)\left(1+\frac{4}{v}\right)\left(1+\frac{5}{v}\right) \tag{4.19}$$

the square root of which gives the standard deviation of the estimate of the order parameter obtained using the intensity contrast estimator, which is a measure of the error in the estimate. Figure 4.3 shows the plot of the error in the intensity

contrast for various sample sizes as a function of the order parameter. For small sample sizes the error is very large. Although the error falls as the sample size increases it is still large even for m =256. The intensity contrast estimator has a simple form and may be attractive in some circumstances where there are large homogeneous regions within the clutter, however, for most high resolution images the number of samples which can be used to estimate the parameter values are small and the contrast estimator would give a very sub-optimal performance.

4.3.2 Normalised log estimator

The expectation value for the normalised log texture measure (U) proposed by Oliver [61] is related to the order parameter through[10]

$$U = \langle \ln I \rangle - \ln \langle I \rangle = \psi(\nu) - \ln \nu - \gamma_E \tag{4.20}$$

The fraction bias and the variance of the normalised texture measure are derived in Appendix A.3 and are given as[11]

$$\frac{\Delta \hat{U}}{U} = \frac{1}{m}\left(1+\frac{2}{\nu}\right)\left(\psi(\nu) - \ln \nu - \gamma_E\right)^{-1} \tag{4.21}$$

$$\sigma_{\hat{U}}^2 = \frac{1}{m}\left(\psi'(\nu) + \frac{\pi^2}{6} - 1\right) \tag{4.22}$$

Using the relationship (4.18), the errors in the estimate of the texture measure can be translated into errors in ν estimate. This gives the predicted variance of $\hat{\nu}$ as

$$\sigma_{\hat{\nu}}^2 = \frac{\nu^2\left(\psi'(\nu) + \frac{\pi^2}{6} - 1\right)}{m\left(\nu\psi'(\nu) - 1\right)^2} \tag{4.23}$$

---

[10] γ_E is the **Euler's constant**.

[11] $\psi'(z)$ is the **Trigamma function**.

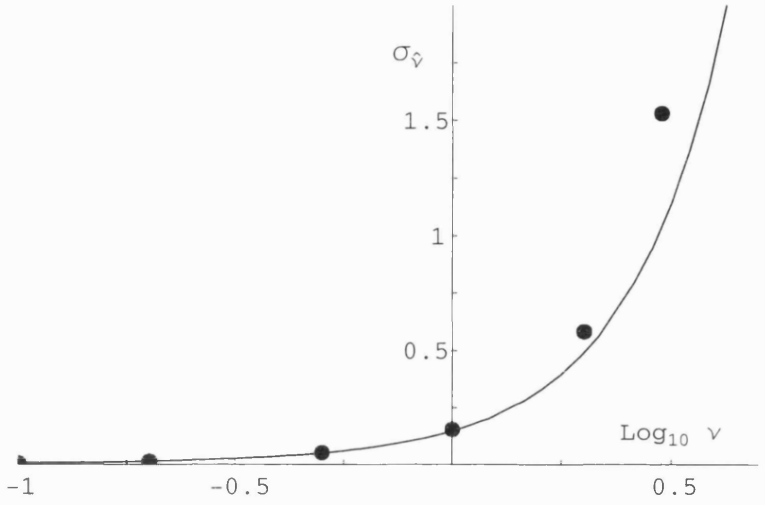


Figure 4.4: Standard deviation in $\hat{\nu}$ for the normalised log estimator with m =256 samples; (—) prediction; (•) simulations, 10^4 trials.

which can be compared with simulation results using uncorrelated K-distributed noise. The comparison is performed for a K-distribution mean value of unity as variation of the mean value simply scales the data and so does not affect the relative shape of the error curves. The simulated data samples were obtained using the product model (3.46) whereby uncorrelated gamma distributed noise of mean one and defined order parameter is generated and multiplied with negative exponential generated noise of unity mean to give uncorrelated K-distribution samples with unity mean and order parameter ν. As an example, the predicted errors in $\hat{\nu}$ for the normalised log texture measure are compared with simulations for a sample size of m =256. The simulation results were obtained by generating 256 uncorrelated K-distributed data samples of unity mean and given order parameter value for which ν is estimated using the normalised log estimator. A total of 10^4 trials are carried out and the standard deviation in $\hat{\nu}$ measured. This is repeated for each

point plotted along the ν axis in Figure 4.4 and the error results are determined over the whole range of ν values of interest. For $\nu > 1$ there is a discrepancy between theory and simulation. Lombardo and Oliver [50] attributed this discrepancy to the pronounced nonlinearity of the inversion process of obtaining ν from the texture measure. They proposed a reciprocal estimator $t \equiv 1/\nu$, which tends to a linear dependence and therefore has a more stable inversion. This reciprocal estimator is directly related to the variance of the texture, which is a measure of the spikiness of the data. Both the error and mean of this improved estimator, therefore, tend to zero as the texture contrast is reduced (i.e. as ν tends to infinity).

Substituting for $t \equiv 1/\nu$ in (4.23) gives the variance in $\hat{t}$ for the normalised log estimator as

$$\sigma_{\hat{t}}^2 = \frac{t^2\left(\psi'(1/t) + \frac{\pi^2}{6} - 1\right)}{m\left(\psi'(1/t)/t - 1\right)^2} \tag{4.24}$$

Figure 4.5 plots the predicted standard deviation in $\hat{t}$ values together with the errors obtained in estimating t from simulated data. The close match between simulated and predicted results shows that the derivation of (4.24) to first order in $1/m$ is valid. The errors are also compared to the numerical ML solution for the t parameter in order to asses the optimality of the estimator. The ML solution was simplified by using the local mean (4.8) for μ_I. Despite this simplification that reduces the ML solution to a 1 dimensional search in $\hat{t}$, the computational load is still 500 times greater than the normalised log estimator. The crosses in Figure 4.5 mark the ML error values and it is evident that the normalised log estimates for t are suboptimal.

These increased errors in the estimates of t translate into poorer performance in radar detection and classification tasks that make use of the estimated parameter values. We need to improve upon the texture measure models to get closer to the ML solution without necessarily increasing the computational complexity of the estimator.

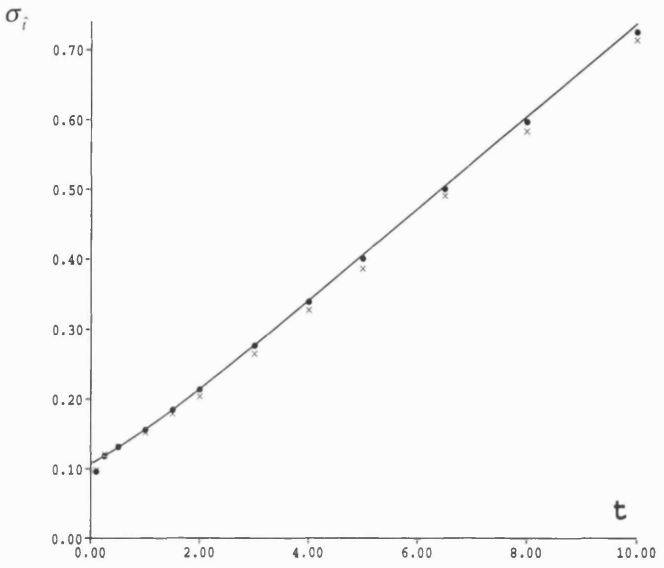


Figure 4.5: Comparison of standard deviation of $\hat{t}$ obtained with parameter estimators over m =256 samples; (—) normalised log prediction; (•) normalised log simulation, 10^4 trials; (×) ML simulation, 10^4 trials.

In the absence of a clear direction for modifying the texture model we use a neural net based technique to search for an improved parameter estimator. The next section details the work carried out using a neural net to obtain an estimator for the t parameter of the K-distribution.

4.3.3 Neural net based estimator

The texture measure discussed in Section 4.3.2 provides for an approximate ML solution of the t parameter. The resultant errors in t depend upon the errors in the estimate of the texture measure which themselves depend upon the moments of the data which are used to obtain the texture measure. Thus the performance of the estimator is dictated by the moments used for the texture measure. There are various moments which can be used to write the texture measure and the error performance varies from texture measure to texture measure. However, it is not obvious which moments of the data most accurately encode the texture information of the imaged data. This has hindered the development of improved models

for the parameter estimation because an exhaustive search through possible combinations of moments would be required to find the best texture measure. To overcome this problem a non-committal neural network was used whereby the properties of the data in relation to the order parameter are introduced through training.

Parameter estimation requires a non-linear mapping of the K-distributed values to the t parameter. For the particular task considered in this paper labelled data is available with the labelling being defined by the simulation process. For this reason, neural nets which adapt themselves to match a desired mapping have been used. The particular network on which this work is based is the multi-layer-perceptron (MLP) which has been described by, amongst others, Rumelhart and McClelland [71] and Lippmann [49]. A MLP consists of a set of simple processing elements, or nodes, interconnected via a set of weights. The MLP used in the course of the work of this thesis consists of an input layer, a hidden layer and an output layer. Data are presented to the input layer which has a number of nodes equivalent to the sample size being used for the texture measure. Non-linear weights combine together the input data and the outcome is translated to the hidden layer nodes. The hidden layer node values are combined together using linear weights to give a single output which is the parameter value corresponding to the data presented to the net.

The net performance is influenced by a number of parameters, for example the size and number of hidden layers, the type of weights used, the amount of training data etc. The choice for these parameters is made through experimentation. The operation of the neural net is carried out as a two stage process; a training stage and a test stage.

In the training stage the net starts with random weights. Simulated K-distributed data of a fixed mean and a range of t values are presented to the input nodes of the neural net. The weights of the net are adjusted through back propagation [49]. The network passes each input pattern through the hidden layers to generate a result at the output node. It then subtracts the actual result from the target parameter value to find the output layer errors. These errors are then used to

adjust the weights. This is called one epoch of training. A large number of epochs are required for training and the nets are regularly tested on unseen tuning data to note the output errors. When there is no more improvement on the error on the estimates for the tuning data the training is stopped. Further training will result in the net learning the training data rather then generalising for the t parameter.

The testing stage involves the net being presented with a large amount of unseen data and measuring the standard deviation in the network estimate of the K-distribution t parameter value. The performance of the neural net estimator is then assessed by comparing its standard deviation in $\hat{t}$ with that obtained with a texture measure estimator.

Neural nets have the potential to infer subtle, unknown relationships from the data. Unlike the texture measure which has an exact analytical form, the neural net acquires the relationship between texture characterisation and the underlying parameter value through training on data. The training data has to include examples of data that covers parameter values over the whole range for which the net is to obtain estimates. There is a trade-off to be made between how representative the training data is to the parameter space, and the memory and processing capacity of the computer system used to implement the neural net. If not enough examples of data with different t values are included in the training set then the neural net will not generalise over the whole range of t parameter values. However, the number of images cannot be increased indefinitely due to consideration of computational load. After some trial and error, eleven images were selected for the training data set. The images were simulated with a constant mean of unity, and t parameter equal to 0.1, 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.5, 8.0, 10.0.

The size of each of the images was 800 by 640 pixels. A 16 by 16 stepping window was used to select individual patterns to be presented to the neural net. The window size matched the input layer size of the neural net. The input layer was chosen to be of this size so that a direct comparison can be made with the performance of the texture measure estimators for m =256 samples. A total of 2000

independent realisations for the t value were obtained from each image using the stepping window.

In addition to the training data set, two further data sets were also generated. The first of these was a tuning data set of eleven images containing the same range of t values as the training data set but only 80 by 640 pixels in size. For a 16 by 16 stepping window this yielded 200 independent realisations for the t value per tuning image. The tuning data set was used to check the neural net error for estimating t at every 50 epoch interval and to halt the training when the performance tailed-off on the tuning data set (i.e. the standard deviation in $\hat{t}$ shows no further reduction). The second data set consisted of 33 test images. The size of each of the images was kept same as those in the training data set. The test data set, however, incorporated 3 times as many examples of the t parameter values in the range 0.1 to 10.0. This data set was used to measure the performance of the neural net estimators. A complete list of the size and the t parameter values of the images used in each of the three data sets is given in Table 4.1.

Data type	**Image Size**	**No. of 16x16 subimages**	**No. of images**	***t* parameter for each image**
Train	800x640	2000	11	0.1 0.5 1.0 1.5 2.0 3.0 4.0 5.0 6.5 8.0 10.0
Tune	80x640	200	11	0.1 0.5 1.0 1.5 2.0 3.0 4.0 5.0 6.5 8.0 10.0
Test	800x640	2000	33	0.1 0.25 0.5 0.75 1.0 1.25 1.5 1.75 2.0 2.25 2.5 2.75 3.0 3.25 3.5 3.75 4.0 4.25 4.5 4.75 5.0 5.25 5.5 5.75 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10.0

Table 4.1: Description of the neural net data sets.

Initially the net was trained on just the intensity K-distributed data. However, the net failed to converge. The intensity data were log-scaled to limit the dynamic

range of the data. With this simple pre-processing the neural net was able to provide estimates of t when tested on unseen data but the standard deviation on $\hat{t}$ did not compare well with the normalised log estimator. Thus the neural net estimator, with this simple pre-processing, proved incapable of automatically extracting the relevant texture measure information from the data.

From previous work on texture estimators [11, 50, 61] we know that the moments of the data are capable of coding the information regarding the underlying order parameter value of the imaged data. It was therefore decided to train the neural net on specific moments of the data. The procedure then involves calculating the texture measure for the data samples over the 16 by 16 window and presenting the net with the moment value. The input layer then has just a single node. The remainder of the structure of the MLP remains unchanged. Figure 4.6 shows the structure of the neural net estimator when using texture measures as input to the net. The training procedure for this type of neural net is outlined as follows:

1. Select 11 equally sized simulated K-distributed images which have the same mean but different t values spanning the range from 0.1 to 10.0.
2. Using a 16×16 stepping window, calculate moments of the data.
3. Present the moment images to the net and train the weights to give the true parameter value corresponding to each input image.
4. Test the performance of the net on a small set of tuning data and note the overall standard deviation. Stop training when the performance has tailed off on this tuning data.
5. Test the neural net on unseen data and note the standard deviation on $\hat{t}$.

The neural net shown in Figure 4.6 has 16 nodes in the hidden layer. Various other configurations for the neural net were also tried with different number of nodes in the hidden layer but it was found that a net with 16 hidden nodes is complex enough to generalise for the t parameter value but still small enough to avoid learning the training data exactly.

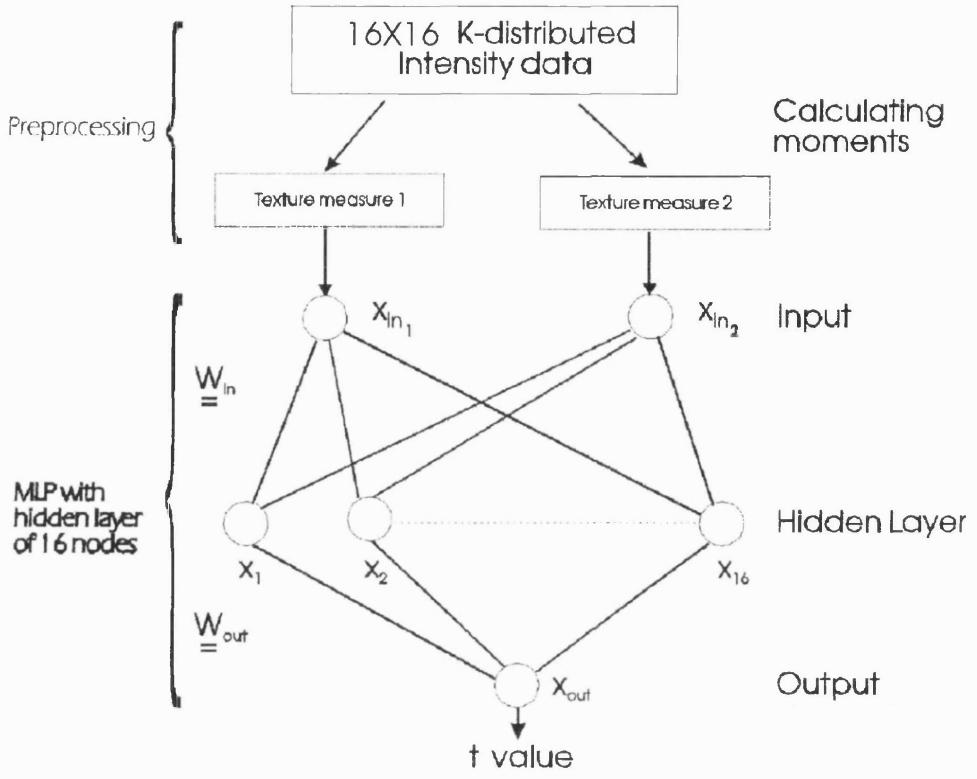


Figure 4.6: Neural net estimator using texture measures to estimate t parameter using 256 data samples. For the hidden layer, $x_k = f\left(\sum_{j=1}^{2} W_{in_j,k} x_{in_j}\right)$ where $f(y) = \{1 + \exp(y)\}^{-1}$. For the output layer, $x_{out} = \sum_{k=1}^{16} w_{k,out} x_k$.

The types of moments that can be calculated from the intensity data are infinite. A key question in the design of the neural net estimator is what moment will result in the best estimator performance? A whole series of neural nets were trained using different moments of the K-distributed data as the input to the net. Most of the choices for the moments used to train the neural net came from previous estimators for the K-distribution parameters, for example the mean of the intensity,

$$x_{in} = \langle I \rangle \tag{4.25}$$

the contrast of the intensity,

$$x_{in} = \frac{\langle I^2 \rangle}{\langle I \rangle^2} - 1 \tag{4.26}$$

the normalised log intensity [61],

$$x_{in} = \langle \ln I \rangle - \ln \langle I \rangle \tag{4.27}$$

the contrast of amplitude [51],

$$x_{in} = \frac{\langle I \rangle}{\langle \sqrt{I} \rangle^2} - 1 \tag{4.28}$$

and variance of log of intensity [46]

$$x_{in} = \langle \ln^2 I \rangle - \langle \ln I \rangle^2 \tag{4.29}$$

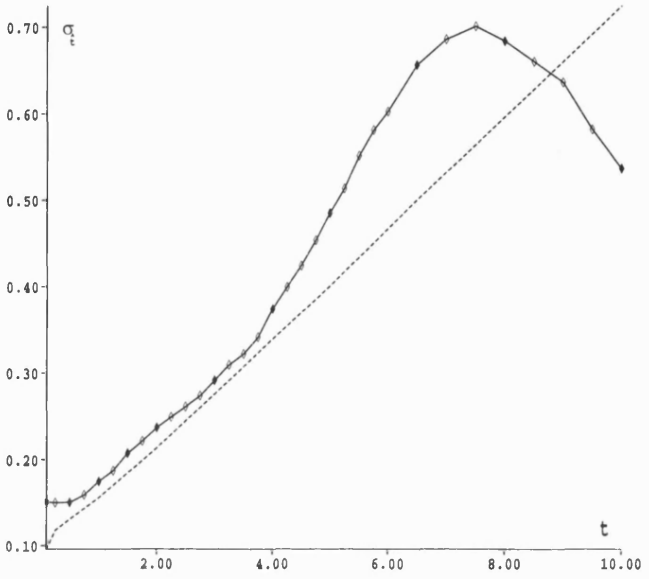
It was found that the net trained on some moments had standard deviation results for $\hat{t}$ which were comparable to the normalised log estimator. These nets have acquired, through training, the appropriate texture measure inversion process to obtain the t parameter value. It was then decided to train the net using pairs of various moments as inputs to the net. The input layer then has two nodes corresponding to the two measured moments. With this arrangement the nets gave some very good estimates for the t parameter. Figure 4.7 shows the error on the estimates for t obtained with three of these nets. The data statistics on which the three nets were trained are as follows:

1. variance of log of intensity ($\sigma^2_{\ln I}$) and contrast of amplitude (CoA).

2. $\sigma^2_{\ln I}$ and normalised log.

3. CoA and normalised log.

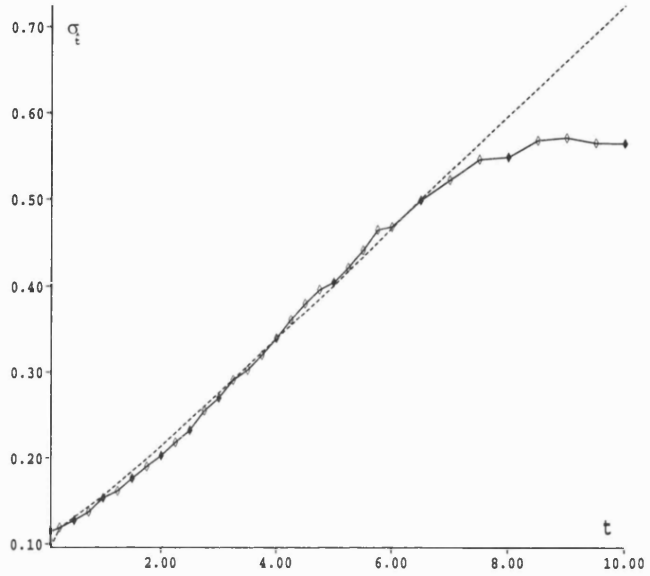
In Figure 4.7 the error performance of each net has been compared to that of the normalised log estimator. The values of t for which the net estimates were obtained are marked with the diamond symbols. The nets were trained on only a few examples of the t parameter (see Table 4.1). Points in Figure 4.7 for which there was a corresponding t parameter image in the training data set are marked by shading the diamond symbols black. It can be seen by the plots in Figure 4.7 that



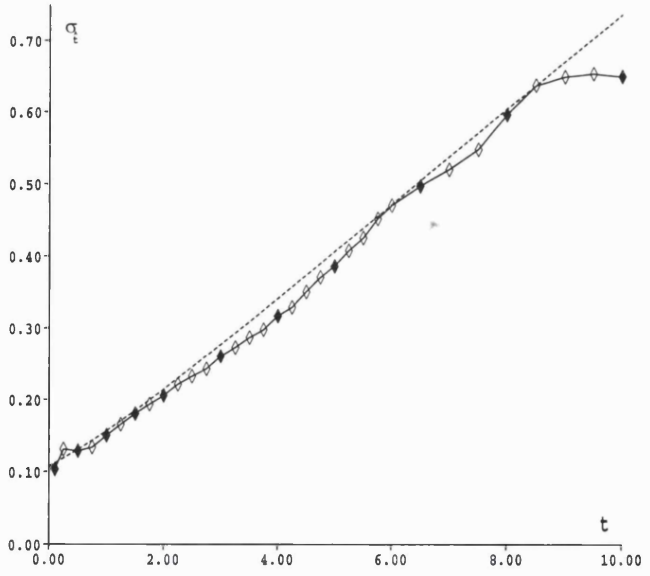
(a)

Figure 4.7: Comparison of the error in $\hat{t}$ obtain using the normalised log estimator (- - -) with the neural net estimator (—). The neural nets have 2 input nodes, 16 hidden nodes and a single output node. Results from three networks are shown. The texture measure used to train each of the nets were (a) $\sigma^2_{\ln I}$ and CoA, (b) $\sigma^2_{\ln I}$ and normalised log and (c) CoA and normalised log. The t values of the test images used to obtain the neural net error results are marked as diamonds. The t values of those test images for which there was a corresponding training image with the same t value, are marked as black diamonds (continued on next page).

the nets have been able to interpolate between t parameter values which were not contained within the training data set. Thus a net trained on just a limited example of t parameter values is able to produce estimates over a continuous range of t values.



(b)



(c)

Figure 1.7: (continued).

The comparison of the three nets shows that the Figure 4.7a net falls well short of the normalised log performance, the Figure 4.7b net outperforms it for

only a few values of t whereas the Figure 4.7c net has been successful in achieving better errors for the majority of t values. From the neural net research it has emerged that the net trained on the normalised log and the CoA produces the best estimate for t. This net improved upon the normalised log estimator for a wide range of t values. The errors for the neural net for Figure 4.7c are much closer to the ML performance. However, comparing the results with Figure 4.5 it appears as if the net is producing lower errors for $\hat{t}$, for values close to 10. This apparent improvement on the ML estimator is due to the fact that the net is only able to give estimates for t values which lie within the range of parameter values included in the training set. This restriction in the range of t values which the net is able to estimate, lowers the standard deviation near the top of the range of the t values. Thus in actual fact the neural nets never exceed the performance of the ML estimator.

The test data used to produce the plots of Figure 4.7 had the same mean intensity value as the training data. The nets were also tested on data which had different μ_I values and error results identical to those of Figure 4.7 were obtained. Thus the performance of the net is invariant to μ_I changes. So the fact that the same μ_I value has been used for all the training data is not a limitation of the neural net estimator.

Previous parameter estimators used just single moments of the data to obtain the texture measure [46, 51, 61]. With those types of estimators the normalised log proved to give the lowest error in regions of large t. The normalised log measure was also found to be the most accurate estimator with the neural nets when using just a single moment to calculate the texture measure . However, the net revealed that a better estimator is achieved if the amplitude contrast is also used to obtain the texture measure. Amplitude contrast is a reasonable estimator for the t parameter but is outperformed by the normalised log estimator for uncorrelated K-distributed data. However, for noisy K-distributed clutter data the amplitude contrast does rather better [51] and is the preferred estimator for the parameter. This is because the normalised log is sensitive to lower values of the K-distribution whereas the CoA is sensitive to higher values of the distribution which are not so

affected by the additive noise. This sensitivity to different regions of the distribution makes it possible for the two moments to contribute independent information to the texture measure. The variance of the log, on the other hand, is sensitive to the same pdf region as the normalised log and hence a net trained on these two moments (see Figure 4.7b) shows no improvement on the normalised log measure. The net has now confirmed that there is independent information in the amplitude contrast which hence results in better estimates for t when combined with the information in the normalised log texture measure.

Neural net research has shown that multiple moments of the data can be used to obtain improved texture measures since there is independent information to be had from the different texture measures. These improvements upon the current model for the approximate ML solution lead us to propose a new model which can be taken as a better approximation to the ML solution. The next section discusses the hybrid estimator for the t parameter of uncorrelated K-distributed data based on the pair of moments: the normalised log intensity and the contrast of the amplitude.

4.3.4 Hybrid normalised log estimator

A new model is proposed for the t parameter estimate of K-distributed clutter data which uses multiple moments to obtain more accurate texture measures. Following on from the neural net research it was found that a useful combination of moments is the normalised log intensity and the contrast of the amplitude. The model can be written as

$$\begin{aligned} W &\equiv \alpha\left(\langle \ln I\rangle - \ln\langle I\rangle\right) + (1-\alpha)\left(\frac{\langle I\rangle}{\langle\sqrt{I}\rangle^2} - 1\right) \\ &= \alpha\left(\psi\left(t^{-1}\right) + \ln t - \gamma_E\right) + (1-\alpha)\left(\frac{4\Gamma^2\left(t^{-1}\right)}{\pi t\Gamma^2\left(t^{-1}+\frac{1}{2}\right)} - 1\right) \end{aligned} \tag{4.30}$$

where α is the weight ratio according to which the two moments are combined.

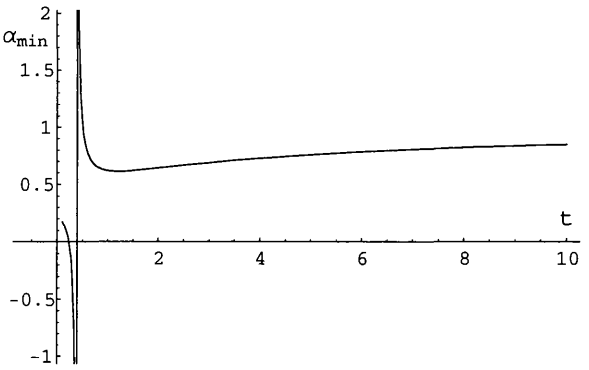


Figure 4.8: Optimum α values that minimise the error in $\hat{t}$ for hybrid normalised log estimator.

Error predictions were derived to first order in $1/m$ using the same approach as detailed in Section 4.3.2. The theoretical error in obtaining the hybrid normalised log measure, W, from the estimated quantities is (see Appendix A.4)

$$\begin{aligned}\sigma_{\hat{W}}^2 = &\frac{\alpha^2}{m}\left(\psi'(t^{-1})+\frac{\pi^2}{6}-1\right)\\ &+16\frac{(1-\alpha)^2}{m}\frac{\Gamma^4(t^{-1})}{\pi^3 t^3\Gamma^6\left(\frac{1}{2}+t^{-1}\right)}\left(16\Gamma^2(t^{-1})-\pi t^2\Gamma^2\left(\tfrac{1}{2}+t^{-1}\right)-5\pi t\Gamma^2\left(\tfrac{1}{2}+t^{-1}\right)\right)\\ &+4\frac{\alpha(1-\alpha)}{m}\frac{\Gamma^2(t^{-1})}{\pi t\Gamma^6\left(\frac{1}{2}+t^{-1}\right)}\left(-6+8\ln 2+t+4\psi(t^{-1})-4\psi\left(\tfrac{1}{2}+t^{-1}\right)\right)\end{aligned} \quad (4.31)$$

This can be translated into errors in $\hat{t}$ using the relationship (see Appendix B)

$$\sigma_{\hat{t}}^2 = \frac{\sigma_{\hat{W}}^2}{\left|dW/dt\right|^2} \quad (4.32)$$

This gives the predicted variance of $\hat{t}$ as

$$\sigma_{\hat{t}}^2 = \left(\frac{\alpha^2}{m}\left(\psi'(t^{-1}) + \frac{\pi^2}{6} - 1\right)\right.$$
$$+16\frac{(1-\alpha)^2}{m}\frac{\Gamma^4(t^{-1})}{\pi^3 t^3 \Gamma^6\left(\frac{1}{2}+t^{-1}\right)}\left(16\Gamma^2(t^{-1}) - \pi t^2\Gamma^2\left(\tfrac{1}{2}+t^{-1}\right) - 5\pi t\Gamma^2\left(\tfrac{1}{2}+t^{-1}\right)\right)$$
$$\left.+4\frac{\alpha(1-\alpha)}{m}\frac{\Gamma^2(t^{-1})}{\pi t\Gamma^2\left(\frac{1}{2}+t^{-1}\right)}\left(-6+8\ln 2 + t + 4\psi(t^{-1}) - 4\psi\left(\tfrac{1}{2}+t^{-1}\right)\right)\right) \qquad (4.33)$$
$$*\left[\frac{4(1-\alpha)\Gamma^2(t^{-1})}{t^3\pi\Gamma^2\left(\frac{1}{2}+t^{-1}\right)}\left(-t + 2\psi\left(\tfrac{1}{2}+t^{-1}\right) - 2\psi(t^{-1})\right) - \frac{\alpha}{t^2}\left(\psi'(t^{-1}) - t\right)\right]^{-2}$$

The weight variable α controls the error obtained with this parameter estimator. For $\alpha = 1$, (4.30) simplifies to a normalised log estimator and the predicted error is then same as (4.24) given in Section 4.3.2. For $\alpha = 0$, (4.30) reduces to a contrast of amplitude estimator. For other values of α the estimator has a varying fraction of the two moments of the texture image combining together to obtain $\hat{t}$ values. The weight α can be optimised to give the minimum variance on the estimates of t by solving the partial derivative

$$\frac{\partial \sigma_{\hat{t}}^2}{\partial \alpha} = 0 \qquad (4.34)$$

for α. Figure 4.8 shows the plot of optimum α over a range of t values. For large t the value of α which would minimise the error approaches a constant value. However, for t =0.5, the optimum α becomes extremely large and for values of t <0.5, α does not have a stable value which minimises $\sigma_{\hat{t}}^2$. Thus the value of α which gives the optimum estimate depends on the value of t, which is the quantity that is being estimated in the first place.

The simplest form for the hybrid normalised log estimator would be to use a fixed value for α. The texture measure is then calculated using a linear combination of the two moments which yields the t parameter estimate from (4.30). Figure 4.9 shows the resultant standard deviation for $\hat{t}$ when several different values of α have been used to obtain the texture measure. The plot shows that different values

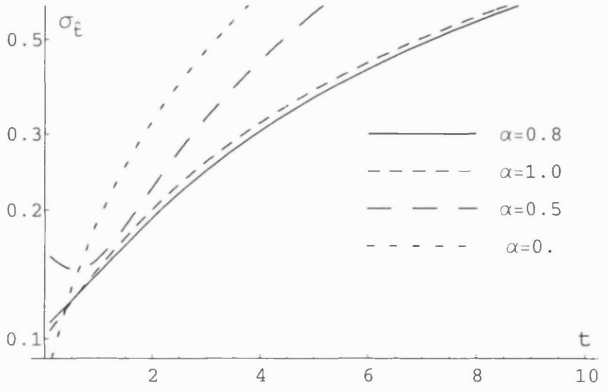


Figure 4.9: Standard deviation of $\hat{t}$ using hybrid normalised log estimator for m =256 samples and fixed values of α.

of α give better estimates in different regions of the t parameter space. For α =0.8, the error is the lowest for all t >0.5. This is significant, since over this range of t values, the error is lower than the normalised log estimator (α =1 case). The lowest error over the range t <0.5 is obtained when α =0, for which the estimator is equal to the amplitude contrast texture measure. It emerges that α =0.8 is the best compromise value which would give the lowest error over the widest range of t values.

Using the fixed value of α =0.8 to write the texture measure of the hybrid estimator, (4.30) takes the form

$$W \equiv 0.8\left(\langle \ln I \rangle - \ln\langle I \rangle\right) + 0.2\left(\frac{\langle I \rangle}{\langle \sqrt{I} \rangle^2} - 1\right) \tag{4.35}$$

Figure 4.10 compares the theoretical error obtained with this estimator with those of the normalised log estimator. The new estimator shows a clear improvement on the standard deviation in $\hat{t}$ for values of t >0.5. Simulation results obtained by estimating the t parameter from uncorrelated K-distribution noise samples showed

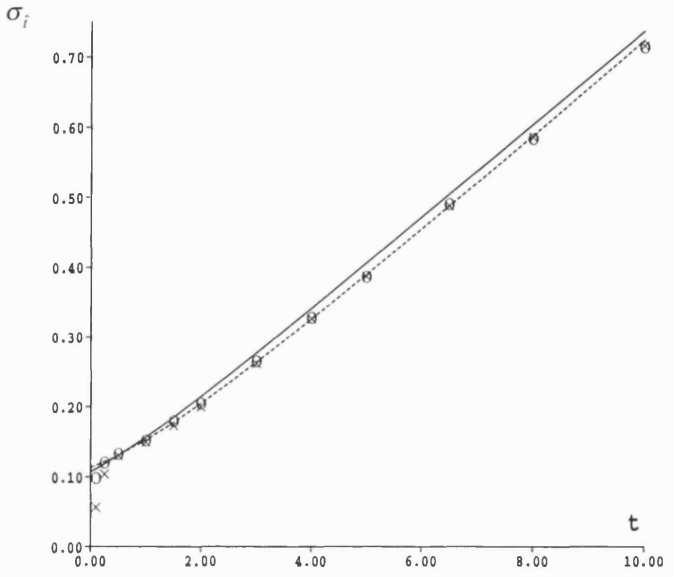


Figure 4.10: Comparison of standard deviation of $\hat{t}$ obtained with parameter estimators over m =256 samples; (—) normalised log prediction; (- - -) hybrid normalised log prediction, (0) normalised log simulation, 10^4 trials; (×) ML simulation, 10^4 trials

good agreement between theory and simulation. This modified estimator when compared with simulation results obtained from numerical ML solution shows that the errors are indistinguishable over the range t >0.5. Thus using a fixed weighted combination of the normalised log and the contrast of amplitude moment and performing a single texture measure inversion, an estimator for t is obtained which is near optimal over a wide range of values of the t parameter. This estimator has been able to achieve this improvement at the cost of just doubling the computational load. For values of t close to the origin, the estimator is however sub-optimal. This is so because the values of α which would minimise $\sigma_{\hat{t}}$ in this region are fairly unstable and a constant α is insufficient to give the best parameter estimate.

The neural nets were able to combine the two moments to give an improved estimate over a wide range of values of the t parameter (see Figure 4.7). The nets, therefore, must be adjusting the value of α in some form of adaptive manner to minimise the error in the parameter estimate. With the analytical estimator an optimal value for α has to be used in order to ensure minimum error on the estimate for t. Therefore, the texture measure estimator requires an optimum estimate for both t and α. Figure 4.11 shows the plot of the predicted errors in the estimate for t when the hybrid normalised log estimator uses optimum α values. The optimality of this estimator is assessed by comparing the errors with the ML solution. Although, a closed form does not exist for the ML solution, Blacknell [11] derived an analytical expression for the standard deviation of the order parameter estimates when using large sample sizes. Figure 4.11, therefore, also shows the optimal error on t estimates as given by the ML solution. The performance of the hybrid normalised log estimator is so close to the ML result that the two graphs are barely separable. It is only for $t > 8$ that there is a noticeable difference between the performance of the two estimators. From this comparison it can be seen that the hybrid normalised log estimator produces near optimal estimates for the t parameter.

A practical implementation of the hybrid normalised log estimator would require a two dimensional search in both t and α. A simple iterative scheme is used to implement this two dimensional search. The estimation process begins with an estimate of t with $\alpha = 1$. This estimated value of t is used to obtain the value of optimum α from the graph shown in Figure 4.8. The texture measure value is recalculated with this new α and a more accurate estimate of t is obtained. This process is repeated until t value stabilises to within a certain accuracy. The error results for the hybrid normalised log estimator implemented using this iterative scheme on simulated data is shown in Figure 4.11 as cross marks. The simulations match the predicted performance over all values of t except for those close to the origin. This is because for t values near the origin the α values are unstable and it is not always easy to find the optimal value of α. In such cases the value of α is forced to be equal to one and the estimator then simplifies to the normalised log

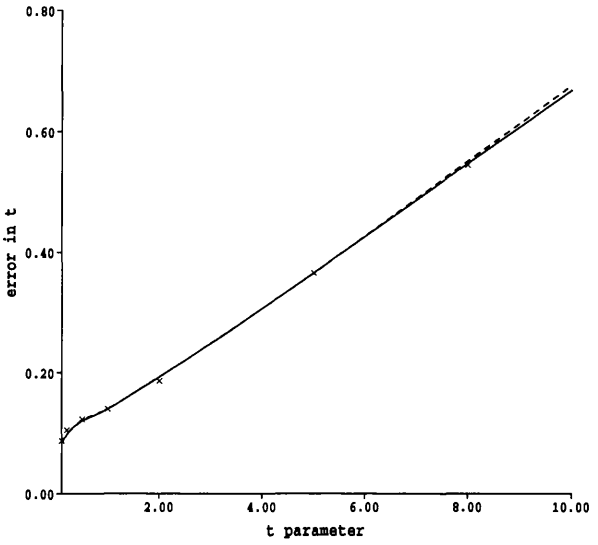


Figure 4.11: Comparison of the hybrid normalised log error in $\hat{t}$ with the ML performance for m =256 samples; ML predictions (___); hybrid normalised log predictions using optimum α (_ _ _); hybrid normalised log simulations, 10^4 trials (×).

texture measure estimator. The closeness of the match of the simulated errors to the predicted performance gives us confidence in the iteration procedure used for implementing the hybrid normalised log estimator. Thus it is shown that an adaptive weighting scheme can be utilised to combine the two texture measures to achieve improved parameter estimates.

The errors in $\hat{t}$ values are very close to the ML performance and for all practical purposes may be considered the same. However, it has not been shown that a combination of the two statistics can equal the ML solution. The advantage of the hybrid normalised log estimator over the ML estimator is that it gives a close approximation to the ML solution using a computationally manageable simple form, albeit over a restricted range of order parameter values. This restriction in the range of parameter values for which the estimator gives the best error performance is not a disadvantage as such since it is conceivable to have estimators which are tuned to different regions of the t parameter space and have a mechanism for se-

lectively choosing between them such that the resultant errors are a minimum over the whole of the range of values of interest.

4.4 Summary

In this chapter:

- Methods for estimating the two parameters that describe the K-distribution have been discussed.
- The ML estimates of the mean intensity and order parameter of K-distributed intensity were derived but were found to be too complex to be used in practice.
- It was shown that the average intensity, which is an optimum estimator for negative exponential mean, is also an unbiased estimator for K-distributed mean intensity. The variance of the sub-optimal mean estimator, however, increases with increased data spikiness.
- Two existing estimators for the order parameters of the K-distribution, the intensity contrast and the normalised log estimator, were both shown to be sup-optimal.
- It was shown that the analysis of the second parameter of the K-distribution is simplified by considering the reciprocal order parameter $t = 1/\nu$ which has a more stable inversion from the texture measures.
- A novel neural net based method was used to estimate the t parameter. It was shown that the neural net performance is strongly influenced by the type of pre-processing performed on the input data. The best performance for the estimate of t was obtained with a net trained on the normalised log and amplitude contrast moment of K-distributed intensity data.
- A new hybrid estimator, which uses adaptive weighting to combine the normalised log and the contrast of amplitude texture measure, was used to estimate t parameter values. Using theoretical predictions and simulation results it was shown that the new estimator achieves almost optimum performance.

Chapter 5

5. Weak scattering models

The statistical models studied in Chapter 3 assume a uniformly distributed phase for the detected field. The resultant phase is the sum of the phase contribution from individual scatterers within a resolution cell. The phase return from a scatterer is the sum of the phase of the backscattering coefficient and the two-way path from the scatterer to the sensor. Therefore, the orientation and the position of the individual scatterer determine its phase and hence the phase distribution of the detected field. When the cell dimensions are several wavelengths in size then the phase overlaps many times as the co-ordinate values of individual scatterers span from the near to far range and from left to right azimuth of a range cell. In such strong scattering regime the phase of the detected field will be uniformly distributed and the detected field can be adequately represented by the statistical models described in Chapter 3. It was shown that when the effective number of scatterers fluctuates according to a negative binomial distribution the resultant field statistics are K-distributed. However, when the resolution cell is only a few wavelengths in depth the scatterers will not have a uniformly distributed phase. Also if there is a bright target immersed in a resolution cell otherwise containing scatterers with uniformly distributed phase the net effect will again be a non-uniformly distributed phase for the detected field. In the random walk representation of the scattering process the non-uniform phase manifests as a bias in the random walk. This type of scattering is termed 'weak scattering'.

When the model deviates from strong scattering the K-distribution no longer accurately describes the single point statistics of the observed data. Blake *et al.* [12]

showed that for very high resolution X-band SAR data, where the resolution cell size is of the order of the wavelength, the single point statistics deviate from a K-distribution. Griffiths *et al.* [33] showed a similar non K-statistics behaviour for high resolution sonar reverberation from sea beds. Blake *et al.* [12] found that a mixture of K-distributions provides a reasonable fit to the high resolution X-band data. However, such empirical fits lack any physical justification. It is therefore necessary to be able to consider models that have some physical basis for representing the weak scattering process.

Baraket [4] showed that by modelling the weak scattering as a directional biased random walk for which the phase distribution is nonuniform, the generalised K-distribution is obtained. Jakeman and Tough [40] demonstrated that a compound representation of the generalised K-distribution is a Rice distribution with both the mean-square noise component and the coherent amplitude varying according to a gamma distribution in a correlated manner. Although this is an attractive new process, the random walk from which it derives does not correctly model the effects generated by weak scattering media. Another model for weak scattering is the homodyned-K distribution process, which is modelled as a coherent sum of a constant vector with a K-distributed vector. Jakeman [38] proposed this model for the statistics of laser light scattered by a localised turbulent thermal plume. The compound model for the homodyned-K distribution is a Rice distribution with a mean-square noise component varying according to a gamma distribution but with a constant-amplitude coherent component. The homodyned-K model is supported by some experimental data. For some situations it has a plausible phenomenological view for example the case of a single bright scatterer among a large number of smaller scatterers. The homodyned-K distribution is, therefore, attracting increasing interest for modelling of weak scattering.

The introduction of new models that are suitable for weak scattering then necessitates the task of devising optimal parameter estimates. There have been very few published results on parameter estimation of the homodyned-K and the generalised-K distribution. The distributions have an additional offset parameter which is a measure of the degree of bias in the random walk. The third parameter

further complicates the expression for the distribution and there are no closed form solutions for the ML parameter estimate. Therefore, sub-optimal estimators have to be considered.

In this chapter the amplitude statistics for the weak scattering model are discussed, some simple parameter estimators are proposed and their performance evaluated. Firstly, Section 5.1.1 describes the amplitude statistics for homodyned-K distribution. A large ν expansion of the homodyned-K distribution is used to derive estimators for the offset and order parameter of the distribution and an error analysis is carried out using the small perturbation method in Section 5.1.2. The amplitude statistics for the generalised-K distribution are described in Section 5.2.1 followed by a discussion of the estimators for its parameters in Section 5.2.2.

5.1 Homodyned-K distribution

In this section the first of the two models proposed for the weak scattering is discussed. Initially the amplitude statistics for the homodyned-K distribution are described, followed by a discussion of sub-optimal estimators for the parameters of the homodyned-K distribution.

5.1.1 Homodyned-K amplitude statistics

The basis for strong scattering is a uniformly distributed phase. However, in high resolution systems the path variations associated with the position of the effective scatterers are small and the phase overlaps with low probability. The resulting phase distribution is no longer uniformly distributed and becomes markedly concentrated around a deterministic value. One way of modelling this non-biased phase in the detected field is to represent the detected field as a coherent sum of a constant vector and a random vector for which the phase is uniformly distributed. The expression for the random walk of (3.9) is then modified to

$$E = \sum_{i=1}^{N} \left(A \exp(j\Phi) + A_0 \exp(j\Phi_0) \right) \tag{5.1}$$

where A_0 is the amplitude and Φ_0 the phase of the constant vector. The magnitude of A_0 determines the degree of bias in the random walk. When the number fluc-

tuation follows a negative binomial distribution, (5.1) is equivalent to the coherent sum of a constant vector with a K-distributed process. The statistic of the resultant field is a homodyned-K distribution, the amplitude pdf of which is given by

$$f_A(A) = \frac{2}{\Gamma(\nu)}\sqrt{\frac{Ab}{\pi A_0}}\sum_{k=0}^{\infty}\Bigg((0,k)(2AA_0)^{-k} b^{\left(\nu-k-\frac{1}{2}\right)/2} \\ (-1)^k \left|A_0 - A\right|^{\nu+k+\frac{1}{2}} K_{\nu+k-\frac{1}{2}}\left(2\sqrt{b}\left|A_0 - A\right|\right)\Bigg) \tag{5.2}$$

where $b = \nu/\mu$ and $(0,k) = \Gamma(1/2+k)/\left(k!\Gamma(1/2-k)\right)$ is Hankel's symbol. The expression for the pdf of the homodyned-K distribution is very complex involving a sum over infinite terms. Jakeman and Tough [40] showed that a closed form expression for the homodyned-K pdf is obtained for integral values of ν. For $\nu = 1$ the pdf takes the form

$$\begin{aligned} f_A(A) &= 4AbI_0\left(2A_0\sqrt{b}\right)K_0\left(2A\sqrt{b}\right);\ A > A_0 \\ &= 4AbI_0\left(2A\sqrt{b}\right)K_0\left(2A_0\sqrt{b}\right);\ A_0 > A \end{aligned} \tag{5.3}$$

The pdf form for values of ν greater than 1, where $\nu = q$ can be obtained by recognising that $(1+x)^{-q} = \left((q-1)!\right)^{-1}\frac{d}{dy}^{q-1}\left(y+x\right)^{-1}\Big|_{y=1}$. For $\nu \le 1$, Tough [85] showed that the pdf can also be written in closed form. For $A > A_0$ the pdf is given as[12]

$$f_A(A) = \frac{4b^{\nu}\Gamma(1-\nu)}{\Gamma(\nu)}A^{\nu}A_0^{\nu-1} \\ \sum_{k=0}^{\infty}(2k+1-\nu)\left(\frac{(1-\nu)_k}{k!}\right)^2 I_{2k+1-\nu}\left(2A_0\sqrt{b}\right)K_{2k+1-\nu}\left(2A\sqrt{b}\right) \tag{5.4}$$

When $A < A_0$ the arguments of the modified Bessel functions of the first and third kinds in this expression are interchanged. Thus it is possible to study the pdf

[12] $(1-\nu)_k$ is the Pochammer symbol $\Gamma(1-\nu+k)/\Gamma(1-\nu)$.

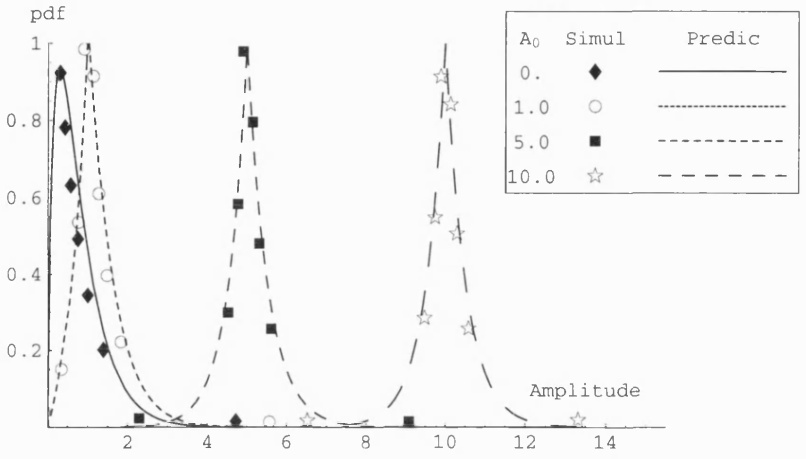


Figure 5.1: Comparison of the theoretical and simulated amplitude pdf of homodyned-K distribution for unity mean, $\nu = 1$ and $A_0 = \sqrt{\mu_0} = 0$, 1.0, 5.0, 10.0.

of the homodyned-K for processes which range from strongly non-Gaussian ($\nu < 1$) to effectively Gaussian ($\nu > 10$).

The homodyned-K distribution can also be represented in a compound form similar to that used for the K-distribution in Section 3.2.4. Thus the homodyned-K will have a speckle component arising from the coherent interference of many scatterers and a mean of the speckle, characterised by the underlying surface RCS, which is also be a random variable. The amplitude pdf of the homodyned-K distribution in the compound form is given as

$$f_A(A) = \int_0^\infty f_{A|\rho}(A|\rho) f_\rho(\rho) d\rho \tag{5.5}$$

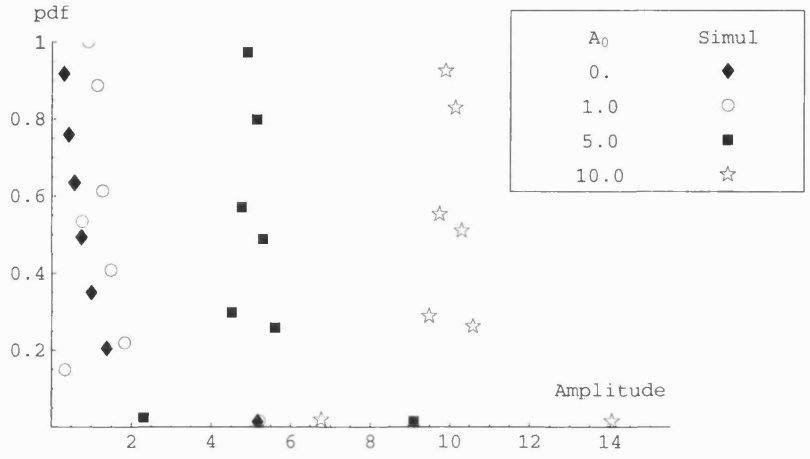
where f_ρ is the RCS component which gives the fluctuation in the mean-square noise level and $f_{A|\rho}$ is the speckle component. As the number fluctuation in the random walk of (5.1) is assumed to be a negative binomial process, the RCS component is modelled as a gamma distribution and is given by (3.47). For the homodyned-K distribution, since the phase is not uniformly distributed, the statistics of

the speckle deviate from a pure Rayleigh amplitude and are given by a Rice distribution as

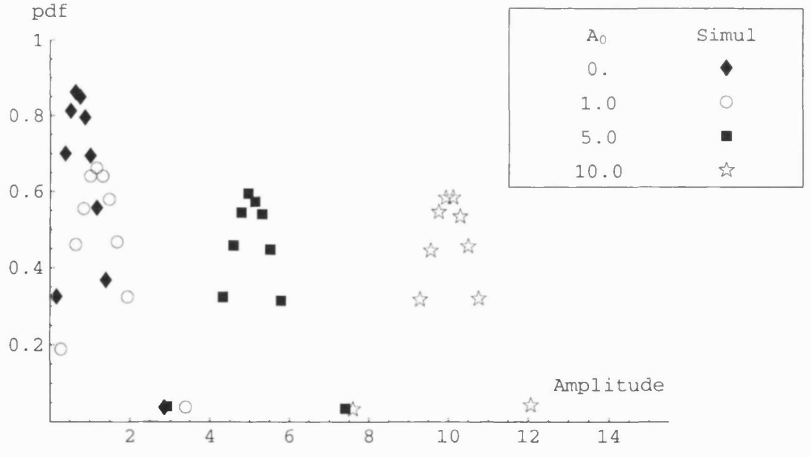
$$f_{A|\rho}(A|\rho) = \left(\frac{2A}{\rho}\right) I_0\left(\frac{2A\sqrt{\mu_0}}{\rho}\right) \exp\left(-\frac{(\mu_0 + A^2)}{\rho}\right) \tag{5.6}$$

where the intensity offset $\mu_0 = A_0^2$ and the mean-square noise component varies according to a gamma distribution.

The product model representation of the homodyned-K distribution also facilitates the generation of simulated data. Thus homodyned-K data of unity mean, given order parameter value and offset value can be obtained by generating uncorrelated gamma distributed noise of unity mean and defined order parameter and multiplying it by Rician generated noise of defined offset value and half mean-square noise value. The Rician data itself is obtained by generating zero mean half standard deviation Gaussian complex data and coherently adding the complex offset and taking the modulus of the resultant. Homodyned-K distribution data for a range of different parameter values were simulated and the histogram measured. Figure 5.1 shows the pdf obtained on simulated data for unity mean, $\nu = 1$ and a range of A_0 values. The corresponding theoretical pdf plots are also obtained, using (5.3), and are shown as line graphs. There is a good match between the theoretical pdf and the simulated histogram, which validates the simulation process. As the offset value increases the pdf plot is shifted up along the amplitude axis. Figure 5.2 shows the measured histogram for the homodyned-K distribution for values of ν equal to 0.5 and 10 respectively. Here again the pdf plots are shifted in amplitude as the offset value increases. The shape of the histogram becomes Gaussian in the limit of large order parameter values. The results of the theoretical pdfs are not shown in Figure 5.2 since the expressions were too computationally cumbersome to calculate. However, Figure 5.1 suggests that the results would have been a close match.



(a)



(b)

Figure 5.2: Amplitude pdf of the homodyned-K distribution for unity mean, $A_0 = \sqrt{\mu_0} = 0, 1.0, 5.0, 10.0$ and ν equal to (a) 0.5 and (b) 10.0.

So for it has been shown that the homodyned-K can be written in a compound form and the pdf has a closed form under certain conditions. The expression for the even order moments of the homodyned-K can be written as

$$\langle A^{2n}\rangle = \frac{n!\Gamma(1+n)}{\Gamma(\nu)}\left(\frac{\mu}{\nu}\right)^n \sum_{k=0}^{n} \frac{\Gamma(n-k+\nu)}{k!(n-k)!\Gamma(1+k)}\left(\frac{A_0\nu}{A}\right)^k \tag{5.7}$$

The moments of the homodyned-K distribution cannot be solved directly to give a closed form solution for the optimal parameter estimates. In the next section alternative sub-optimal estimators are considered which are based upon a large ν expansion of the homodyned-K distribution.

5.1.2 Parameter estimators for homodyned-K distribution based on a large ν approximation.

The maximum likelihood solutions for the parameters of this distribution do not have simple expressions in terms of combinations of distribution moments. In the absence of such straight forward moment based estimators an approximate solution for the parameter values is considered. The approach taken is to perform an expansion appropriate to large ν values. In the limit of large ν, the gamma distribution in the product model tends to a delta function rendering a Rice distribution. This is more apparent once the distribution is re-parameterised in terms of ν and the mean power $\langle A^2\rangle$ of the homodyned-K noise process. Thus

$$f_A(A) = \frac{2A\nu^\nu}{\langle A^2\rangle\Gamma(\nu)}\int_0^\infty \frac{1}{u} I_0\left(\frac{2A\sqrt{\mu_0}}{\langle A^2\rangle u}\right)\exp\left(-\frac{(\mu_0+A^2)}{\langle A^2\rangle u}\right)u^{\nu-1}\exp(-\nu u)du \tag{5.8}$$

The equation in this form allows us to show the effect of large ν while keeping the mean power level of the noise process at a constant value. Setting this equal to unity causes no loss in generality, thus the integral becomes

$$f_A(A) = \frac{2A\nu^\nu}{\Gamma(\nu)}\int_0^\infty \frac{1}{u} I_0\left(\frac{2A\sqrt{\mu_0}}{u}\right)\exp\left(-\frac{(\mu_0+A^2)}{u}\right)u^{\nu-1}\exp(-\nu u)du \tag{5.9}$$

Evaluation of this integral for large ν yields [85]

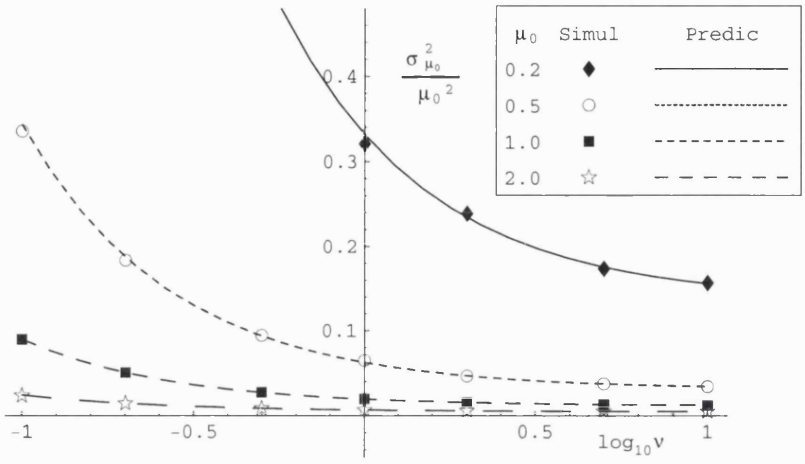


Figure 5.3: Comparison of the predicted and simulated relative variance of homodyned-K $\hat{\mu}_0$ parameter for m =256. Each simulated result obtained using 10^4 trials.

$$f_A(A) = \exp\left(-\left(\mu_0 + A^2\right)\right) \times \left\{2AI_0\left(2A\sqrt{\mu_0}\right) + \frac{1}{\nu}\left[A\left(2 + \left(\mu_0 + A^2\right)^2 + 4\left(\mu_0 A^2 - \left(\mu_0 + A^2\right)\right)\right)I_0\left(2A\sqrt{\mu_0}\right) + 2A^2\sqrt{\mu_0}\left(3 - 2\left(\mu_0 + A^2\right)\right)I_1\left(2A\sqrt{\mu_0}\right)\right]\right\} \tag{5.10}$$

From (5.10) the second and fourth order moments for the homodyned-K distribution are given as

$$\left\langle A^2 \right\rangle = 1 + \mu_0 \tag{5.11}$$

$$\left\langle A^4 \right\rangle = 2 + 4\mu_0 + \mu_0^2 + \frac{2}{\nu} \tag{5.12}$$

Rearranging these yields expressions for the estimate of the parameters of homodyned-K distribution. The viability of these estimators can then be demonstrated by performing an error analysis. Thus (5.11) gives an estimate for μ_0 as

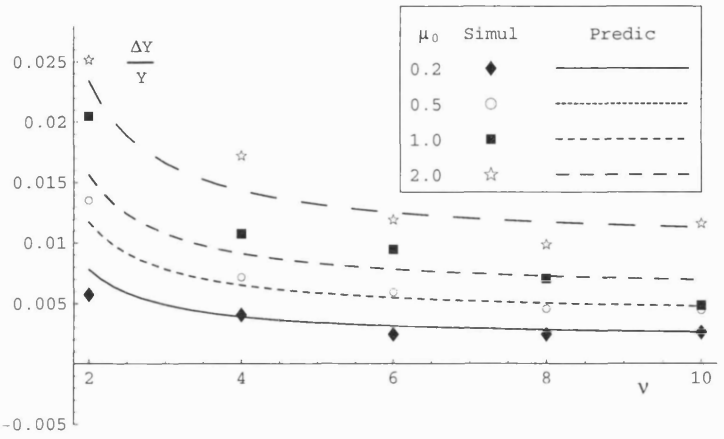
$$\hat{\mu}_0 = \hat{A^2} - 1 \tag{5.13}$$

The assessment of the error performance of this estimator for the offset value of the homodyned-K is carried out by performing a small perturbation analysis about the mean parameter value similar to that carried out in Section 4.2 and 4.3 for the K-distribution parameter estimators. (5.13) is an unbiased estimator for μ_0 and therefore its fractional bias is zero. Its relative variance is given as (see Appendix A.5)

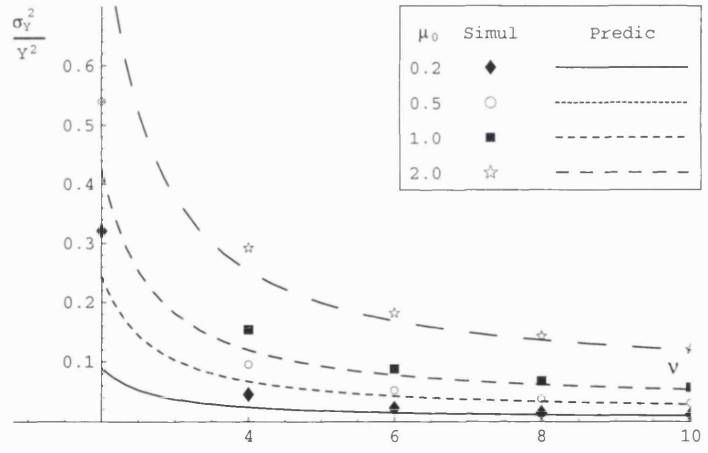
$$\frac{\sigma^2_{\hat{\mu}_0}}{\mu_0^2} = \frac{1}{m}\left(\frac{2+\nu+2\nu\mu_0}{\nu\mu_0^2}\right) \tag{5.14}$$

The relative variance for the offset parameter has a dependence on both μ_0 and ν values. Figure 5.3 shows the plot of the relative variance in the estimate of μ_0 as a function of ν for a number of μ_0 values. In order to keep the analysis consistent with the results shown in Chapter 4 all the plots for the weak scattering models are obtained for a sample size of m =256. The predicted error performance is also compared with simulations. The method adopted for producing the simulated results is similar to that used in the previous chapter. 256 samples of homodyned-K distributed data for unity mean, given μ_0 and ν values are generated and the μ_0 parameter estimated. The homodyned-K data is generated using the method outlined in Section 5.1.1. The variance in $\hat{\mu}_0$ is calculated using 10^5 trials. This is repeated for each set of parameter values for which the experimental error analysis is carried out. The results of the simulation are shown as marks of miscellaneous shapes in the same figure which contains the plot for the corresponding predicted error performance. Figure 5.3 shows that the error in $\hat{\mu}_0$ decreases with increasing ν and μ_0 and there is good agreement between simulation and prediction.

In order to obtain an expression for the estimate of the order parameter, the first step is to obtain an expression for a moment based texture measure which can be written in terms of the order parameter. This can be obtained by combining (5.11) and (5.12) and rearranging the equations such that μ_0 is eliminated. This leaves an expression for the texture measure which can be written as



(a)



(b)

Figure 5.4 Comparison of the predicted and simulated results for the (a) fractional bias and (b) relative variance, for the texture measure $\hat{Y}$ of the homodyned-K distribution for m =256. Each simulated result obtained using 10^5 trials.

$$Y \equiv \langle A^4 \rangle - \left(\langle A^2 \rangle + 1 \right)^2 = \frac{2}{\nu} - 2 \tag{5.15}$$

that is completely defined by the order parameter of the homodyned-K. This result is, however, obtained using simplified second and fourth moment expression

based upon a large ν expansion of the homodyned-K pdf. Inverting the estimate of the texture measure provides an estimate for the order parameter value. The fractional bias and the relative variance of the texture measure of (5.15) are given as (see Appendix A.6)

$$\frac{\Delta\hat{Y}}{Y} = \frac{2+\nu+2\nu\mu_0}{2m(\nu-1)} \tag{5.16}$$

$$\frac{\sigma^2_{\hat{Y}}}{Y^2} = \frac{-1+10\nu+\nu^2+28\nu\mu_0+4\nu^2\mu_0+3\nu\mu_0^2+2\nu^2\mu_0^2}{m(\nu-1)^2} \tag{5.17}$$

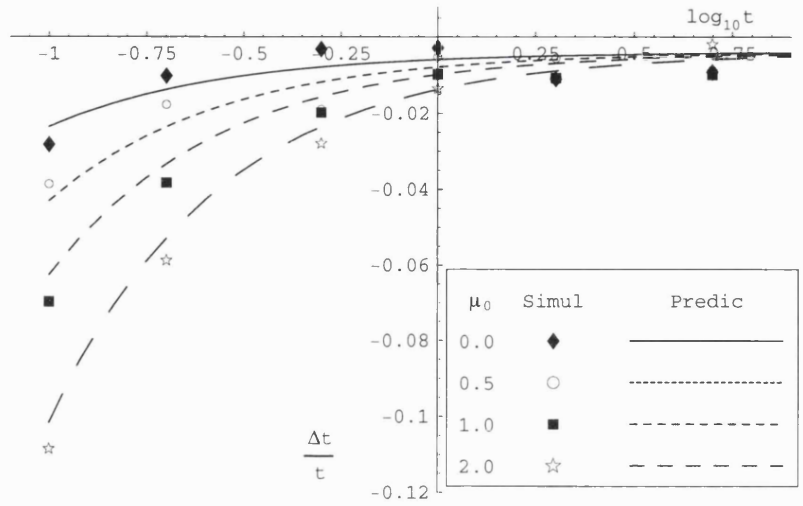
The fractional bias and the relative variance have a dependence on both the μ_0 and ν values. There is also a singularity at $\nu=1$. Figure 5.4a compares the predicted and simulated result for the fractional bias of the texture measure of the homodyned-K distribution whereas the corresponding results for relative variance of the texture measure are shown in Figure 5.4b. These results are again produced for a sample size of $m=256$ and using 10^5 trials for the simulations.

The results for the texture measure estimate show a good match between simulations and predictions for large ν. In the limit of large ν the error values in the texture estimates are reasonably low. However, as ν approaches 1 the errors get very large. The offset μ_0 scales the errors such that they increase with increasing μ_0 with the effect being more pronounced at smaller values of ν.

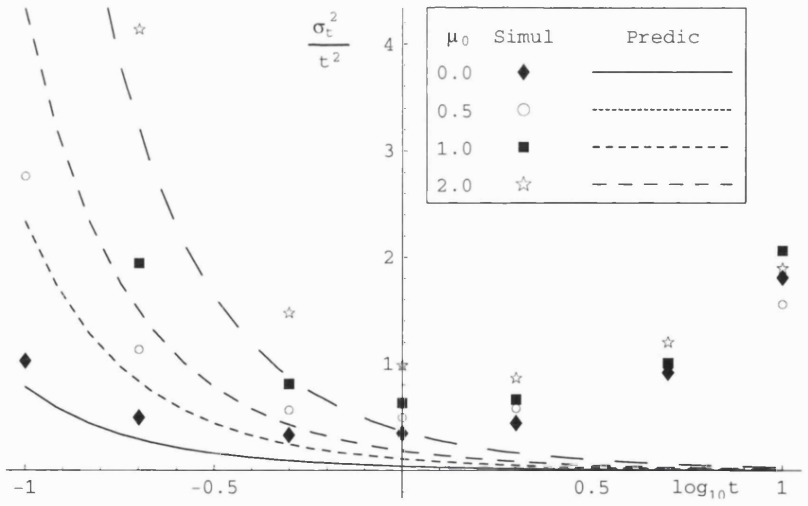
The required order parameter value of the homodyned-K distribution can be obtained by inverting the texture measure Y. However, it has been shown [50] that working in the space of $t \equiv 1/\nu$ results in better error performance for the parameter value. Thus using (5.15) an expression can be obtained for $\hat{t}$ in terms of the estimates for the texture measure V as

$$\hat{t} = \frac{1}{2}\hat{Y} + 1 \tag{5.18}$$

Using (5.16) and the relationship $\Delta t = \Delta Y/(dY/dt)$ the fractional bias for $\hat{t}$ of the homodyned-K distribution is obtained to first-order approximation in m as



(a)



(b)

Figure 5.5: Comparison of the predicted and simulated results for the (a) fractional bias and (b) relative variance, for $\hat{t}$ of the homodyned-K distribution for m =256. Each simulated result obtained using 10^5 trials.

$$\frac{\Delta\hat{t}}{t} = -\frac{1}{2m}\frac{\left(1+2t+2\mu_0\right)}{t} \tag{5.19}$$

The corresponding relative variance for $\hat{t}$ is obtained from (5.17) and using the relationship $\sigma_t^2 = \sigma_Y^2 / \left| dY/dt \right|^2$ as

$$\frac{\sigma_{\hat{t}}^2}{t^2} = \frac{1}{m} \frac{\left(1 + 4\mu_0 + 2\mu_0^2 + 10t + 28t\mu_0 + 3t\mu_0^2 - t^2\right)}{t^2} \tag{5.20}$$

The plots for the predicted and simulated fractional bias of $\hat{t}$ for the homodyned-K distribution are shown in Figure 5.5a. The fractional bias decreases as t increases. The fractional bias is not as small as the fractional bias of the texture measure Y from which the t parameter is derived. The effect of the μ_0 parameter is to increase the fractional bias of $\hat{t}$. The effect of this increase is more noticeable for smaller values of t.

Figure 5.5b shows the predicted and simulated results for the relative variance of $\hat{t}$ for the homodyned-K distribution. According to the theoretical predictions, the relative variance of $\hat{t}$ reduces as t increases. However, according to the simulated results, the errors are decreasing only for values of t <1. For t >1, the errors in the estimate for t are increasing instead of decreasing as predicted by (5.20). The deviation from prediction highlights the limitation of the small perturbation analysis when the large ν approximation is violated. Since these values correspond to spikier data, which is often the region of interest, the large errors limit the effectiveness of this estimator for obtaining the parameter values of the homodyned-K. In addition to Figure 5.5b where there is a very strong disagreement between theory and prediction for t >1, differences are also noticeable between the theoretical predictions and the simulated results for the error performance given in Figure 5.4a, Figure 5.4b and Figure 5.5a. Although the differences are small and the general trend for the simulation results broadly follows those of the theoretical predictions, nevertheless the differences reflect that the assumptions made for the small perturbation analysis are inadequate.

The estimators for the offset and the shape parameter for the homodyned-K distribution described in this section are obtained using an expansion of the pdf of the distribution in the limits of large ν. However, the observed errors are not as

predicted especially for small ν values. At small ν values the error and, therefore, the uncertainty in ν values is comparable to the actual value of the parameter itself, which makes it unacceptably high. Therefore, further development is required in order to obtain more reliable estimators for the parameters of the homodyned-K distribution.

5.2 Generalised-K distribution

The generalised-K distribution is the second of the two models being considered for the weak scattering process. First the general model and its amplitude statistics are discussed followed by an analysis of parameter estimators for the generalised-K distribution.

5.2.1 Generalised-K amplitude statistics

The homodyned-K distribution is based on a random walk model that adds a constant offset to an unbiased random walk to represent a weak scattering process. An alternative is to use a truly biased random walk to represent a weak scattering process. If the number fluctuation is again assumed to be negative binomial this gives rise to a generalisation of the K-distribution and the amplitude pdf is given by

$$f_A(A) = \frac{2d}{\Gamma(\nu)}\left(\frac{2A\nu}{\mu d}\right)^{\nu} I_0\left(\frac{2A\sqrt{\mu_0}}{\mu}\right) K_{\nu-1}(dA) \tag{5.21}$$

where

$$d^2 = \frac{4\nu}{\mu} + \frac{4\mu_0}{\mu^2} \tag{5.22}$$

As in the case of the homodyned-K distribution, the generalised-K has an additional third parameter, μ_0 which is the measure of the degree of bias in the random walk. In the limit $\mu_0 \to 0$ the unbiased K-distribution result is regained. As μ_0 increases this adds an offset to the pdf until in the limit of $\mu_0/\mu \to \infty$ the equation reduces to a gamma distribution. The ν parameter here again determines the de-

gree of spikiness of the data and now in the limit of $\nu \to \infty$ the distribution reduces to the Rice model

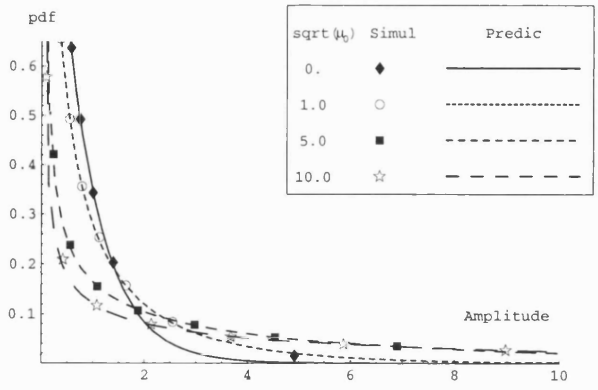
$$\lim_{\nu \to \infty} f_A(A) = \frac{2A}{\mu} I_0\left(\frac{2A\sqrt{\mu_0}}{\mu}\right) \exp\left(-\frac{\mu_0 + A^2}{\mu}\right) \tag{5.23}$$

The generalised-K distribution can also be represented in a compound form similar to that of (5.5). However, due to the difference in the way in which the bias in the random walk is modelled compared to the homodyned-K model, the speckle component amplitude statistics are given by a different Rician distribution which can be written as

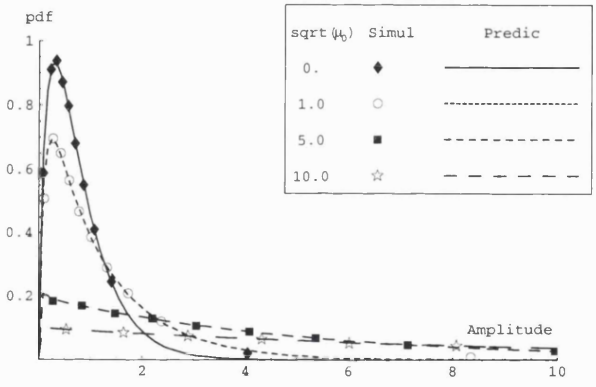
$$f_{A|\rho}(A|\rho) = \left(\frac{2A}{\rho}\right) I_0\left(\frac{2A\sqrt{\mu_0}}{\mu}\right) \exp\left(-\frac{2\nu\rho^2}{\mu^2} - \frac{A^2}{\rho}\right) \tag{5.24}$$

This Rician distribution has a mean-square equal to $\rho/2$ and a coherent amplitude given by $\rho\sqrt{\mu}_0/\mu$. The gamma distribution gives both the fluctuations in the mean-square noise level and the coherent amplitude in a correlated manner. The scattering process is thus a modulated Rice distribution.

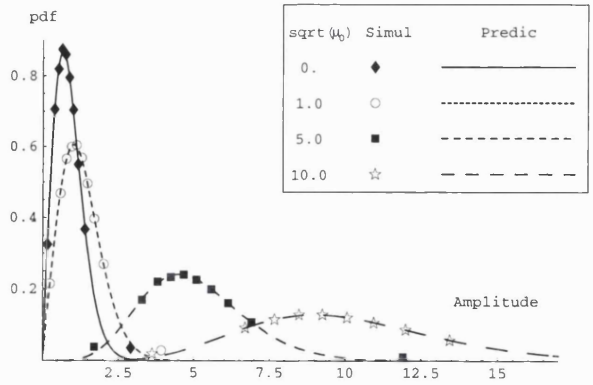
Simulated generalised-K distribution data can be generated using the product model representation. Zero mean, half standard deviation complex Gaussian data is generated and multiplied with gamma generated data of given mean and order parameter value. The complex offset value is also multiplied by a gamma generated data of given mean and order parameter value and normalised by the mean. The modulus of the coherent sum of the modulated complex offset and the modulated complex Gaussian then results in generalised-K data of the given mean, order parameter and offset value. In Figure 5.6 the theoretical and simulated pdfs for the generalised-K distribution over a range of parameter values are plotted. There is good agreement between the predicted pdf and the measured histogram which validates the simulation process. As the order parameter increases the generalised-K tends toward a Rician distribution. As the offset increases the peak of the plot shifts along the amplitude axis.



(a)



(b)



(c)

Figure 5.6: Comparison between the predicted and simulated amplitude pdf of the generalised-K distribution for unity mean, $\sqrt{\mu_0} = 0$, 1.0, 5.0, 10.0 and ν equals to (a) 0.5, (b) 1.0 and (c) 10.0.

The even order moments for this distribution can be expressed as

$$\langle A^{2n} \rangle = \left(\frac{\mu}{\nu} \right) \frac{n!\Gamma(n+\nu)}{\Gamma(\nu)} P_n^{(0,\nu-1)} \left(1 + \frac{4\mu_0}{\nu\mu} \right) \tag{5.25}$$

where

$$P_i^{(\beta,\epsilon)}(x) = 2^{-i} \sum_{k=0}^{i} \binom{i+\beta}{k} \binom{i+\epsilon}{i-k} (x-1)^{i-k} (x+1)^k \tag{5.26}$$

are the Jacobi polynomials. (5.26) does not have a closed form for the optimal parameter estimate for the generalised-K distribution. In the following section some approximate solutions for the parameter estimation of the generalised-K distribution are considered.

5.2.2 Parameter estimators for generalised-K distribution using even order moments

The first two even order moments of the generalised-K distribution can be written as

$$\langle A^2 \rangle = \mu + \mu_0 + \frac{\mu_0}{\nu} \tag{5.27}$$

and

$$\langle A^4 \rangle = \frac{(1+\nu)}{\nu^3} \left(2\mu^2\nu + 8\mu\nu\mu_0 + 4\mu\nu^2\mu_0 + 6\mu_0^2 + 5\nu\mu_0^2 + \nu^2\mu_0^2 \right) \tag{5.28}$$

The moments as they stand cannot be solved to extract the parameter values. There is not a large ν expansion which can be carried out for the generalised-K distribution as it was done for the homodyned-K model. However, for the extreme case of $\nu \to \infty$ the first two even-order moments reduce to the form

$$\langle A^2 \rangle = \mu + \mu_0 \tag{5.29}$$

and

$$\langle A^4 \rangle = 2\mu^2 + 4\mu\mu_0 + \mu_0^2 \tag{5.30}$$

Eliminating μ from (5.29) and (5.30) gives an expression for the estimate of μ_0 of the generalised-K distribution as

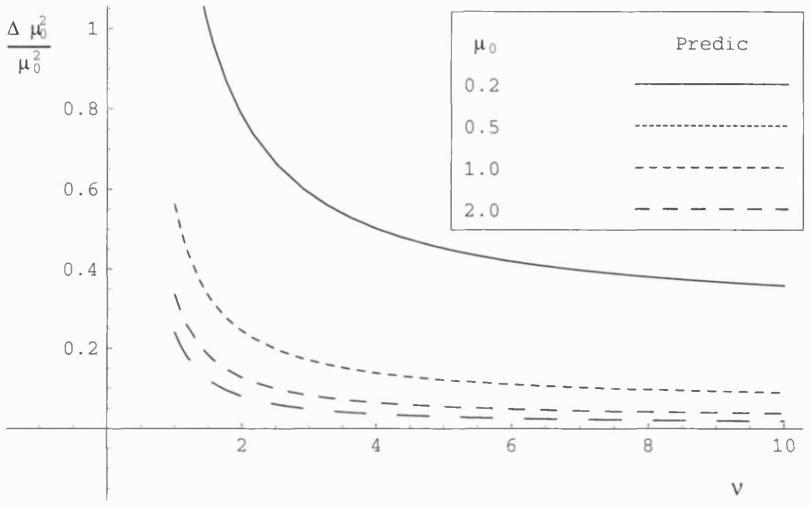
$$\hat{\mu}_0^2 = 2\hat{A}^{2^2} - \hat{A}^4 \tag{5.31}$$

Thus an estimate for the square of the offset parameter for the generalised-K is obtained in terms of the second and fourth order amplitude moments. An error analysis can be carried out on this estimator using the same small perturbation technique as carried out for the other estimators studied in this thesis. As before the small perturbation expansion is restricted to terms of order $O(1/m)$. The fractional bias and the relative variance are given as

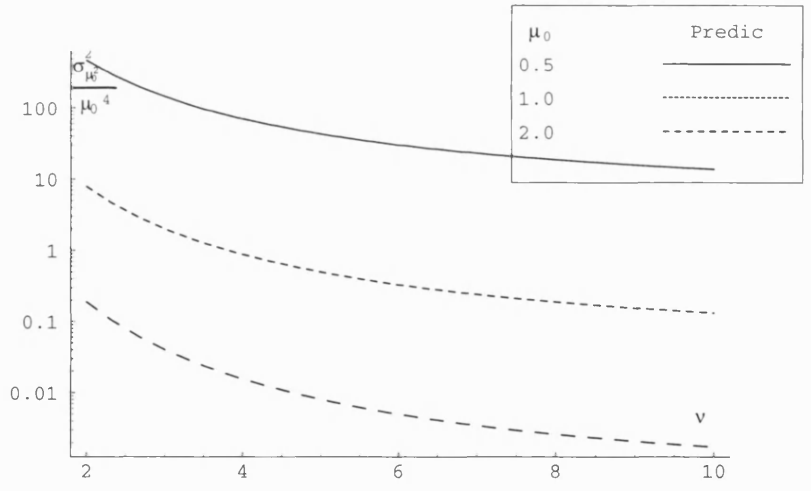
$$\frac{\Delta\hat{\mu}_0^2}{\mu_0^2} = \frac{2}{m}\left[\frac{6}{\nu^3} + \frac{1}{\nu^2}\left(10 + \frac{8}{\mu_0}\right) + \frac{1}{\nu}\left(3 + \frac{10}{\mu_0} + \frac{2}{\mu_0^2}\right) + \frac{1}{\mu_0^2} + \frac{2}{\mu_0}\right] \tag{5.32}$$

$$\frac{\sigma^2_{\hat{\mu}_0^2}}{\mu_0^4} = \frac{1}{m}\left[\begin{array}{l} \frac{1}{\nu^7}\frac{1260}{\mu_0^4} + \frac{1}{\nu^6}\left(\frac{2880}{\mu_0^5} + \frac{3018}{\mu_0^4}\right) + \frac{1}{\nu^5}\left(\frac{2160}{\mu_0^6} + \frac{6360}{\mu_0^5} + \frac{2498}{\mu_0^4}\right) \\ + \frac{1}{\nu^4}\left(\frac{576}{\mu_0^7} + \frac{4262}{\mu_0^6} + \frac{4656}{\mu_0^5} + \frac{780}{\mu_0^4}\right) + \frac{1}{\nu^3}\left(\frac{36}{\mu_0^8} + \frac{952}{\mu_0^7} + \frac{2617}{\mu_0^6} + \frac{1214}{\mu_0^5} + \frac{22}{\mu_0^4}\right) \\ + \frac{1}{\nu^2}\left(\frac{41}{\mu_0^8} + \frac{436}{\mu_0^7} + \frac{537}{\mu_0^6} + \frac{30}{\mu_0^5} - \frac{14}{\mu_0^4}\right) + \frac{1}{\nu}\left(\frac{10}{\mu_0^8} + \frac{64}{\mu_0^7} + \frac{23}{\mu_0^6} - \frac{6}{\mu_0^5} + \frac{4}{\mu_0^4}\right) \\ + \frac{1}{\mu_0^8} + \frac{4}{\mu_0^7} + \frac{1}{\mu_0^6} + \frac{2}{\mu_0^5} \end{array}\right] \tag{5.33}$$

Figure 5.7 shows the plot for the fractional bias and the relative variance for the estimate of the square of the offset parameter of the generalised-K distribution. From the plots it is clear that only in the limit of very large ν the error performance is anything like acceptable. For smaller ν values the error and, therefore, the uncertainty in the estimate of the offset is much larger than the value of the offset parameter rendering the estimate useless. The error increases both as ν and μ_0 decrease. Thus the estimator (5.31) is only useful over a very limited range of ν and μ_0. Since for real clutter data ν and μ_0 are likely to have a much wider range of values this estimator is only of academic interest with a very limited practical use.



(a)



(b)

Figure 5.7: Predicted error performance for the square of the offset parameter of the generalised-K distribution for a sample size of m =256 (a) fractional bias and (b) relative variance.

For the order parameter of the generalised-K distribution the moments can not be manipulated to arrive at a closed form for the estimate of the parameter. This poses a severe limitation in using the generalised-K for modelling real clutter data.

The investigation for the generalised-K distribution concludes that there are no simple moment based estimators for the generalised-K distribution which will give a good approximation to the ML solution of the parameter estimates. This is because the expressions for the moments of the generalised-K distribution are quite complex and there is no close form solution for the ML parameter estimate. A possible approach which could overcome this problem is to approximate the generalised-K distribution by a model which has a closed form for the ML solution of the parameter estimate. The estimates for the parameters of the generalised-K distribution would then by obtain by comparing the moments of the two distributions. The success of this approach would depend on how good the alternative model approximates the generalised-K distribution. One such distribution which has been proposed is the generalised gamma function [9]

$$f_y(y) = \frac{a_{g\Gamma}{}^{-\eta\zeta}\alpha}{\Gamma(\zeta)} y^{\eta\zeta-1} \exp\left[-\left(\frac{y}{a_{g\Gamma}}\right)^{\eta}\right] \tag{5.34}$$

which has three parameters $a_{g\Gamma}$, η and ζ. The ML solution for these three parameters of the generalised gamma function has a closed form. Using a method of moment comparison, the parameter values for the generalised-K distribution can then be deduced. An error analysis of the parameter estimates using this approximate model for the generalised-K distribution will show the usefulness of this approach. However, the parameter estimates obtained using this method would be sub-optimal. We are therefore still some distance away from obtaining efficient optimal estimators for the parameters of the generalised-K distribution.

5.3 Summary

In this chapter:

- It was shown that the condition for strong scattering, i.e. a uniformly distributed phase for the detected field, is violated when the size of the resolution cell is small compared to the wavelength or the reflecting target is very smooth.

- Two distributions were introduced which can model a weak scattering process; the homodyned-K distribution and the generalised-K distribution.

- It was shown that the random walk model for the homodyned-K distribution can be represented by the coherent sum of a constant vector and one which is a K-distribution process. The amplitude pdf of this distribution was shown to have a closed form under only certain ν parameter values. The compound representation, which is a Rice distribution with a mean-square noise component varying according to a gamma distribution and a constant-amplitude coherent component, was used to generate simulated data.

- Approximate moment based estimators were derived for the offset and the order parameter of the homodyned-K distribution using a large ν expansion of its pdf. The errors for the offset parameter estimate were shown to match the predicted performance over the whole range of order parameter values and over a wide range of offset values. The corresponding results for the t parameter were not so good, with the performance much worse for smaller values of the order parameter. It is concluded that further investigation is required to obtain good estimators for the parameters of the homodyned-K distribution.

- The model for the generalised-K distribution was described as a biased random walk such that the resultant phase becomes markedly concentrated around a deterministic value. Using the compound representation, which is a Rician distribution where both the mean-square noise component and the coherent amplitude vary according to a gamma distribution in a correlated manner, simulated data were generated and good match was obtained between theoretical pdfs and measured histograms.

- In the limit of $\nu \to \infty$ the first two even order moments of the generalised-K distribution were simplified to give an estimate for the offset parameter. However, its performance was shown to be very poor and further work is required in devising appropriate estimators for the parameters of the generalised-K distribution.

Chapter 6

6. Conclusion and future work

In this concluding chapter we shall examine to what extent the results presented in previous chapters answer the four main questions set out in the introduction regarding the clutter characterisation problem. We shall also highlight outstanding issues and make recommendations for future work.

6.1 Background theory

Before addressing the clutter characterisation problem the background theory needed to carry out the analysis of a coherent radar system was described in the first two chapters. In the introductory chapter a general description was given for a surveillance radar system. It was explained how such surveillance systems can be used either to detect ground moving targets or produce high resolution maps of the ground. For both of these applications it was shown that it is important to have a proper understanding of the clutter; in the former to get better clutter rejection and in the latter to attain optimum terrain classification.

The clutter statistics are measured on a pixel level. Each pixel is a single resolution cell corresponds to a specific footprint of the radar on the ground. The radar footprint has an important bearing on the clutter statistics since it determines the number of scatterers which are able to contribute toward the reflected energy observed in any given pixel. The size and shape of the radar footprint depend upon a number of system parameters which were described in Chapter 2. It was shown that for a SAR system the two spatial components that make up the two-dimensional radar image are range and azimuth. Range resolution was shown to be

inversely proportional to the bandwidth of the radar pulse. It was shown that higher range resolution can be obtained using a chirp pulse which gives a sinc function response to a point target. The azimuth resolution for a real beam system was shown to be inversely proportional to the aperture of the antenna expressed as a number of wavelengths. However, for a fully focused SAR it was shown that the azimuth resolution is equal to half the real aperture length and the system response to a point target is a sinc function. For non-imaging radar it was shown that the observed scattering is arranged into a range-Doppler map where Doppler data is obtained by a frequency transformation of the azimuth data. The Doppler resolution was shown to be equal to the reciprocal of the coherent integration time. It was also shown that the Doppler resolution determines the azimuth dimension of the radar footprint and is utilised in Doppler beam sharpening to obtain finer resolution clutter maps. Thus in MTI system where each pixel is a range-Doppler cell with a corresponding unique radar footprint, the clutter statistics are then a function of the range and Doppler resolution.

6.2 Forward problem

The first question in the introduction asked what statistical models can be used to describe the surface scattering. The most direct method for determining the characteristics of the observed scattering is by solving the forward problem. However, in Chapter 3 it was shown that there is still a considerable gap to bridge before realistic surfaces can be analysed using this approach. In forward problem, the approach taken is to define the surface height profile, specify the boundary conditions and solve the Maxwell equations in order to calculate the scattered field observed at a defined point in space. Considerable effort has been spent on devising approximate expressions for the scattering surface for which closed form analytical solutions are attainable. The Kirchhoff approximation was described which assumes a slowly varying height profile and is applicable for normally distributed rough surfaces. Real surfaces are, however, rarely normally distributed. For non-Gaussian rough surface the SPM was described as another suitable approach provided the surface roughness is small. The Maxwell equations can also be solved numerically using simulated surface height profiles. This enables the analysis of

non-ideal surfaces under realistic geometries, however, the computational load increases significantly with the complexity of the model, although that is less of a problem with ever increasing computer power.

The forward problem has been useful in building up an understanding of the relationship between physical parameters of the surface and the clutter characteristics. However, the techniques are still too underdeveloped to be used for devising optimum algorithms for target detection and clutter classification.

6.3 Strong scattering models

The latter part of Chapter 3 dealt with the question of how single or multiparameter statistical distributions can be used to describe the fluctuations in the detected intensities. It also looked at what physical justification can be associated with the statistical models. Starting from the assumption that the detected field is given by the sum of contributions from a large number of discrete elementary scatterers, it was shown that for strong scattering the phase will be uniformly distributed and independent. At low resolution the number of scatterers would tend to infinity and it was shown that for strong scattering the resultant field can be modelled as a Gaussian process with the intensity fluctuation given by a negative exponential distribution. However, it was observed that as resolution increases the clutter in both MTI and SAR images exhibit non-Gaussian statistics. The log-normal and Weibull distribution were not considered to be suitable models for describing the coherent clutter statistics despite empirical evidence from a number of researches since they cannot be represented by a discrete scatterer model. The K-distribution, on the other hand, was shown to be the result of a negative binomial fluctuation of the number of scatterers. This is quite a plausible model since as the size of the resolution cell decreases the number of scatterers will be finite and, given the complex nature of the clutter, the scatterer number will undoubtedly fluctuate. The K-distribution can also be expressed as product of two statistical processes; a negative exponential intensity fluctuation arising from the coherent interference of the scatterers and an intensity mean which fluctuates according to a gamma process associated with the underlying surface RCS fluctuation. It was shown that the or-

der parameter of the distribution can be related to the clutter spikiness. Chi-squared goodness-of-fit test and ML classification on MTI and SAR data gave empirical evidence to support the K-distribution as a reasonable descriptor for medium resolution clutter.

6.4 K-distribution parameter estimators

Having reached the conclusion that the K-distribution is a reasonable model for the coherent land clutter statistics the next task was to address the problem of optimal parameter estimation. This was investigated in Chapter 4. It was shown that the ML solution does not have a closed form and a numerical search over the mean intensity and order parameter of the K-distribution is computationally very inefficient. An ML estimate for the mean intensity is given by the average intensity provide the clutter is homogeneous and the image surface cross-section is constant which is the case for negative exponentially distributed intensity. For K-distributed intensity, where the surface cross-section fluctuates according to a gamma distribution, the average intensity can still be taken as a reasonable approximation for the mean intensity. A similar simple form for the optimum estimate for the order parameter is less forth coming. An example of a simple estimator for the order parameter of the K-distribution is the intensity contrast texture measure but it was shown to have quite high error values especially for small sample sizes. The normalised log texture measure was shown to be more reliable estimator. It was shown that better agreement is obtained between predicted errors and simulated results if the analysis is carried out in the t parameter space where $t \equiv 1/\nu$. However, overall the normalised log estimator was shown to be still sub-optimal compared to the ML estimator.

The texture measure estimators use the pdf of the statistical model to specify the relationship between the estimator and the model parameter. Neural net can obtain a similar connectivity between texture measure and model parameters through adaptive weighting. Thus a technique based on neural nets, in their capacity as functional estimators, was used to devise a t parameter estimator. A large number of experiments were carried out using a whole range of different neural

nets and the standard deviation of the t parameter estimates was measured. A number of lessons were learnt regarding the optimum arrangement of the neural net to carry out the task of the t parameter estimates. A few of the main point are summarised as follows.

1. A simple MLP with one hidden layer is sufficient to learn the problem. The weights connecting the input layer to the hidden layer nodes have to be non-linear to enable the neural net to have a non-linear mapping capability. On the other hand, for the neural net to give a real value output over a continuous range instead of the more common binary output, linear weights have to be used for connecting the hidden layer to the output node.

2. The output layer has just one node since only a single parameter estimate is required. The input layer will have number of nodes equal to the sample size or the number of moments used to calculate the estimate. The neural nets were found to be able to learn the parameter estimation problem with just 16 nodes in the hidden layer. A net with a larger hidden layer would require fewer epochs to reach the solution although each epoch would take longer since there are more weights to update. However, the net size cannot be increased indefinitely since it would then learn the training data exactly without any generalisation.

3. An estimator for a continuous function can be formed by training the neural net on images which contain just a few selected examples of the parameter values over the range of interest. For the t parameter it was found that the sampling space can be gradually increased as the t parameter of the simulated images increased. Thus for example, the difference in the t parameter values of the two images containing the data for the smallest parameter was 0.4 compared to the difference between the parameter values of the two images containing the highest values which were 2.0. The training process for the neural net utilised a smaller independent tuning data set to determine the peak on the convergence curve and halt the training proc-

ess. The testing of the neural net estimator showed that it is capable of interpolating between values of the parameter not included in the original training data set.

4. The convergence of the neural net is very sensitive to the dynamic range of the input data. The net failed to converge when trained directly on the intensity data. Using log intensity data also did not produce any fruitful results with the neural net. However, it was shown that if the neural nets are trained on texture measures of K-distribution data, they are capable of estimating the t parameter on unseen data when it is presented to the neural net with identical pre-processing as the training data.

The neural net research showed that good estimators for the t parameter of the K-distribution can be obtained provided the nets are trained on texture measures of the imaged data. The net found an optimal way of combining the normalised log intensity moment and the amplitude contrast of the data to give a very useful estimator for the parameter. This estimator out-performed the simple normalised log texture measure based estimator for a wide range of values. The net highlighted the importance of using combinations of moments to achieve more accurate texture measures.

The improvement in the t parameter estimate with the neural net estimator led us to propose a new hybrid estimator which combines the normalised log and the contrast of amplitude to provide a more accurate texture measure. The theoretical errors for this estimator were derived and it was demonstrated that its performance depends on the weight according to which the two moments are combined. It was shown that there is a constant value of this weight for which the hybrid normalised log estimator gives better performance over existing texture measure estimators over a wide range of t values which are of interest. However, for best performance an iterative scheme is used which selects the optimum α for any given value of t. It was shown that using a simple iterative scheme the hybrid normalised log estimator is able to obtain errors that are almost indistinguishable from the ML performance. The advantage of this moment based estimator is that it is computationally

more efficient than the numerical ML solution which makes it suitable for real time evaluation.

Neural nets have been able to obtain the parameter estimates without explicitly being told the ML criteria. This demonstrates the potential of neural nets to provide accurate parameter estimates in situations where it is not possible to write down the exact ML criteria. Thus for example in estimating the parameter values for correlated data where the exact pdf of the statistical model is not known or for estimating the line width in correlated data where it not clear what the quantity ought to be which is to be maximised, neural nets have the potential to be an effective estimator.

6.5 Weak scattering models

In Chapter 5 the clutter analysis is extended to higher resolution systems. The phase of the detected field can no longer be assumed to be uniformly distributed. This can be modelled by a weak scattering process for which the four clutter modelling questions posed in the introduction are answered in Chapter 5. It was shown that the partially developed speckle can be modelled as either a biased random walk which gives rise to a generalised-K distribution for the amplitude statistics when the number of steps fluctuates according to a negative binomial distribution. Alternatively, the weak scattering process can be modelled as a coherent sum of a constant vector with a K-distributed process which gives rise to a homodyned-K distribution. Both of these distributions can be expressed in a compound form similar to that of the K-distribution. It was shown that for the homodyned-K distribution it is given by a Rician distribution with a mean-square noise component varying according to a gamma distribution and a constant offset. The compound form for the generalised-K is also a gamma modulated Rice distribution but the gamma distribution gives both the fluctuations in the mean-square noise level and the coherent amplitude in a correlated manner. It was shown that the compound model can be used to generate simulated data for the weak scattering distributions. The simulation process was verified by getting a good match between the measured histograms and the predicted pdfs. However, the pdf comparison for the ho-

modyned-K was restricted to ν values equal to 1 due to a lack of a closed form for the theoretical pdf, even though histogram plots of the homodyned-K distribution for other ν values were shown.

The task of the optimal parameter estimation is complicated due to the introduction of a third offset parameter for the weak scattering models. For homodyned-K distribution a large ν expansion of the pdf is used to derive simplified expressions for the second and fourth moments. Using these expressions, an estimator is proposed for the homodyned-K offset parameter. An expression for the predicted error in the estimate of the offset parameter is derived using the small perturbation analysis. A good agreement is obtained between theoretical predictions and simulated results. An estimator is also proposed for the t parameter of the homodyned-K for which the error predictions are also derived. However, the shape parameter errors get unacceptably large as the assumption of large ν breaks down. For the generalised-K distribution an estimator for μ_0 in the limit of $\nu \rightarrow \infty$ is proposed but the errors render the estimator ineffective. Further investigations are required to obtain good estimates for the homodyned-K and the generalised-K distribution. Furthermore, the ML error performance for the parameters of the weak scattering models needs to be established either through numerical simulations or by deriving the Cramer- Rao bound to get a proper assessment of the sub-optimal nature of the proposed estimators for the weak scattering models.

6.6 Future Work

In Chapter 4 it was shown that near optimal estimators for the parameters of a non-Gaussian distribution can be obtained using a method of moment combination even though there is no closed form ML solution available. Although this showed only modest gains compared to previously known best texture measure estimator for the K-distribution, the main strength of the technique is that it is applicable to any theoretical distribution. Some initial results have been quoted in Chapter 5 for the estimators for the weak scattering models using this moments combination technique, however, further development has to be carried out before reliable estimators are obtained. It may require an investigation using neural nets to

determine the best combination of moments for the weak scattering parameters as it is reported in Chapter 4 for the K-distribution parameter. However, the training data set would need to contain a representative samples of both t and μ_0 parameters which could force it to be quite large. Furthermore, the assessment of the estimator performance would be greatly facilitated by a knowledge of the limit of the ML performance. This could either be obtained by performing a numerical ML or deriving the Cramer-Rao bound for the ML solution. Only once reliable estimators for the weak scattering models are developed can these distributions become practical for real radar systems. Further empirical evidence should also be gauged for the suitability of the weak scattering models for radar surface scattering and sonar seabed scattering.

In this thesis all parameter estimation has been carried out assuming independent data samples. However, in real radar data, clutter appears as correlated [62]. Using sub-sampling by an appropriate amount, independent samples can be obtained from correlated data, however, at the expense of loosing information. It is therefore, highly desirable to develop estimators for the parameters of correlated textures. This development has been hindered through the lack of suitable pdf models for correlated non-Gaussian textures. The neural net may prove very effective in developing suitable estimators for correlated non-Gaussian textures since they do not require to be given an exact ML criterion. They have the potential of inferring the relevent single and multi-point statistics from the data using non-linear weights. However, a closed form analytical expression for the correlated non-Gaussian pdf for the K-distribution or the weak scattering models will enable to carry out a small perturbation analysis on the parameter estimators which has been shown in this thesis to be a very useful analysis tool.

The main motivation for developing optimal estimators for the clutter distributions, as stated in the introduction, is to improve target detection and clutter classification which uses the clutter models. Historically such algorithms have been developed assuming Gaussian statistics even though considerable evidence has been generated to show that surface clutter exhibits non-Gaussian statistics. Where non-Gaussian clutter models have been used the full potential performance has

not been utilised due to sub-optimal estimators. The performance of such algorithms should be re-evaluated using the near optimal estimator proposed in this thesis for the K-distribution. This would then enable a comparison of the improvement gained in false alarm rate and clutter classification with the increased complexity of the algorithm. The same should be repeated for the weak scattering distributions once suitable optimal estimators have been developed. At higher resolution, land objects appear as targets extended over many pixels in a radar image. Therefore, for optimum target detection a multi-pixel detector has to be developed. This would enable to take proper account of the target correlations which can then be distinguished from any clutter correlations. However, this requires proper model for the higher order statistics of clutter and targets (since a multi pixel target can be envisaged to contain texture information) and optimum parameter estimators for these. This developed will hopefully lead to improved target detection performance in high resolution surveillance systems.

Appendix A

A. Evaluation of errors of estimated quantities using small perturbation analysis

The bias and variance of a function of estimated parameters can be approximated by expanding the function around the expected values of the estimates [64]. In Chapter 4 and 5 there were a number of quantities for which the predicted error performance was analysed using the small perturbation method. In the following sections the derivations for the theoretical expressions of the error performance for these quantities are shown.

A.1 Mean of variable x

The quantity and its estimate are defined by

$$\mu_x \equiv \langle x \rangle \tag{A.1}$$

$$\hat{\mu}_x \equiv \hat{x} = m^{-1} \sum_{i=1}^{m} x_i \tag{A.2}$$

We now expand the expression about its mean value. Thus (A.2) can be written as

$$\hat{\mu}_x \equiv \langle x \rangle (1 + \xi) \tag{A.3}$$

where $\xi = \hat{x}/\langle x \rangle - 1$. The expectation value of the function and the square of the function are then given as

$$\langle \hat{\mu}_x \rangle = \langle x \rangle + \langle x \rangle \langle \xi \rangle \tag{A.4}$$

$$\langle (\hat{\mu}_x)^2 \rangle = \langle x \rangle^2 + 2\langle x \rangle^2 \langle \xi \rangle + \langle x \rangle^2 \langle \xi^2 \rangle \tag{A.5}$$

The bias and the variance of the estimated quantity are defined respectively, as

$$\Delta\mu_x \equiv \langle \hat{\mu}_x \rangle - \mu_x = \langle x \rangle \langle \xi \rangle \tag{A.6}$$

$$\begin{aligned} \sigma^2_{\hat{\mu}_x} &\equiv \langle (\hat{\mu}_x)^2 \rangle - \langle \hat{\mu}_x \rangle^2 \\ &= \langle x \rangle^2 \langle \xi^2 \rangle - \langle x \rangle^2 \langle \xi \rangle^2 \end{aligned} \tag{A.7}$$

The expectation value for $\langle \xi \rangle$ is given as

$$\langle \xi \rangle = \left\langle \left(\frac{\hat{x}}{\langle x \rangle} - 1 \right) \right\rangle = 0 \tag{A.8}$$

and for $\langle \xi^2 \rangle$ it is given as

$$\langle \xi^2 \rangle = \left\langle \left(\frac{\hat{x}}{\langle x \rangle} - 1 \right)^2 \right\rangle = \frac{1}{m} \left(\frac{\langle x^2 \rangle}{\langle x \rangle^2} - 1 \right) \tag{A.9}$$

Substituting for $\langle \xi \rangle$ in (A.6) gives the bias as

$$\Delta\mu_x = 0 \tag{A.10}$$

which shows that the mean of a variable is an unbiased estimate. Substituting for $\langle \xi \rangle$ and $\langle \xi^2 \rangle$ in (A.7) gives the expression for the variance as

$$\sigma^2_{\hat{\mu}_x} = \frac{1}{m} \left(\langle x^2 \rangle - \langle x \rangle^2 \right) \tag{A.11}$$

Now if x is considered to be the intensity (i.e. $x = I$) then for an assumed negative exponential distribution the moments can be substituted from (3.30) in (A.11) that gives

$$\sigma^2_{\hat{\mu}_I} = \frac{1}{m}\mu_I^2 \tag{A.12}$$

which is the result (4.11) used in Chapter 4. The corresponding expression for the variance of the mean estimate of the intensity for an assumed K-distribution can be obtained using (3.49) to give

$$\sigma^2_{\hat{\mu}_I} = \frac{1}{m}\mu_I^2\left(1+\frac{2}{\nu}\right) \tag{A.13}$$

which is the result (4.12) used in Chapter 4.

A.2 Contrast V_x of the variable x

The measure and its estimate are defined by

$$V_x \equiv \frac{\langle x^2\rangle}{\langle x\rangle^2} - 1 \tag{A.14}$$

$$\hat{V}_x \equiv m^{-1}\sum_{i=1}^{m} x_i^2 \Big/ \left(m^{-1}\sum_{i=1}^{m} x_i\right)^2 - 1 \tag{A.15}$$

The denominator of the estimate is expanded about its mean value and only terms up to first order in $1/m$ are maintained. Using the same procedure as in Section A.1 it can be shown that the bias and variance are given as

$$\begin{aligned}\Delta V_x &\equiv \langle \hat{V}_x\rangle - V_x \\ &= -\frac{1}{m}\left[2X_3 + X_2 - 3X_2^2\right]\end{aligned} \tag{A.16}$$

$$\begin{aligned}\sigma^2_{\hat{V}_x} &\equiv \left\langle \left(\hat{V}_x\right)^2\right\rangle - \langle \hat{V}_x\rangle^2 \\ &= \frac{1}{m}\left[X_4 - 4X_3X_2 - X_2^2 + X_2^3\right]\end{aligned} \tag{A.17}$$

where $X_i \equiv \langle x^i\rangle / \langle x\rangle^i$ are the **normalised moments** of x. Once again considering the intensity variable, then for an assumed K-distribution, the moments are obtained using (3.49) and the bias is given as

$$\Delta\hat{V}_I = -2\frac{1}{m}\left(1+\frac{1}{\nu}\right)\left(1+\frac{6}{\nu}\right) \tag{A.18}$$

which leads to the fractional bias result (4.16) used in Chapter 4. Similarly the variance is given as

$$\sigma^2_{\hat{V}_I} = 4\frac{1}{m}\left(1+\frac{1}{\nu}\right)\left(1+\frac{4}{\nu}\right)\left(1+\frac{5}{\nu}\right) \tag{A.19}$$

which is the result (4.17). The first order approximation for the variance of ν from the measure V_x is obtained using the relationship (see Appendix B)

$$\sigma^2_{\hat{\nu}} = \frac{\sigma^2_{\hat{V}_I}}{\left|dV_I / d\nu\right|^2} \tag{A.20}$$

Therefore, for K-distributed intensity the predicted variance in the order parameter obtained using the intensity contrast estimator is gives as

$$\sigma^2_{\hat{\nu}} = \frac{\nu^4}{m}\left(1+\frac{1}{\nu}\right)\left(1+\frac{4}{\nu}\right)\left(1+\frac{5}{\nu}\right) \tag{A.21}$$

which is the result (4.19).

A.3 Normalised log U_x of the variable x

The measure and it estimate are defined by

$$U_x \equiv \langle \ln x \rangle - \ln\langle x \rangle \tag{A.22}$$

$$\hat{U}_x \equiv \frac{1}{m}\sum_{i=1}^{m} \ln x_i - \ln\left(\frac{1}{m}\sum_{i=1}^{m} x_i\right) \tag{A.23}$$

Again expanding about the mean value and retaining terms up to the first order in $1/m$ gives

$$\begin{aligned} \Delta U_x &\equiv \left\langle \hat{U}_x \right\rangle - U_x \\ &= \frac{1}{2m}\left[X_2 - 1\right] \end{aligned} \tag{A.24}$$

$$\begin{aligned}\sigma^2_{\hat{U}_x} &\equiv \left\langle \left(\hat{U}_x\right)^2 \right\rangle - \left\langle \hat{U}_x \right\rangle^2 \\ &= \frac{1}{m}\left[L_2 - L_1^2 + 2L_1 - 2X_1 L + X_2 - 1\right]\end{aligned} \tag{A.25}$$

for the bias and variance, respectively, where $X_i L \equiv \langle x^i \ln x \rangle / \langle x \rangle^i$ and $L_i \equiv \langle \ln^i x \rangle$. Using (3.49) for the intensity moments of the K-distribution gives the results quoted in Section 4.3.2. Using the relationship (B.6) the variance in ν or t can be obtained from (A.25) which gives the result (4.23) and (4.24) respectively.

A.4 Hybrid normalised log W_x of the variable x

The measure and its estimate are defined by

$$W_x \equiv \alpha\left(\langle \ln x \rangle - \ln\langle x \rangle\right) + (1-\alpha)\left(\frac{\langle x \rangle}{\langle \sqrt{x} \rangle^2} - 1\right) \tag{A.26}$$

$$\hat{W}_x \equiv \alpha\left(\frac{1}{m}\sum_{i=1}^{m} \ln x_i - \ln\left(\frac{1}{m}\sum_{i=1}^{m} x_i\right)\right) + (1-\alpha)\left(\frac{\frac{1}{m}\sum_{i=1}^{m} x_i}{\left(\frac{1}{m}\sum_{i=1}^{m} \sqrt{x_i}\right)^2} - 1\right) \tag{A.27}$$

Again expanding about the mean value and retaining terms up to the first order in $1/m$ gives the variance as

$$\begin{aligned}\sigma^2_{\hat{W}_x} &\equiv \left\langle \left(\hat{W}_x\right)^2 \right\rangle - \left\langle \hat{W}_x \right\rangle^2 \\ &= \frac{\alpha^2}{m}\left[L_2 - L_1^2 + 2L_1 - 2X_1 L + X_2 - 1\right] \\ &+ \frac{(1-\alpha)^2}{m}\left[X_{4\sqrt{\ }} - 4X_{3\sqrt{\ }}X_{2\sqrt{\ }} - \left(X_{2\sqrt{\ }}\right)^2 + 4\left(X_{2\sqrt{\ }}\right)^2\right] \\ &+ \frac{2\alpha(1-\alpha)}{m}\left[X_{2\sqrt{\ }}L - 2X_{2\sqrt{\ }}X_{\sqrt{\ }}L + X_{2\sqrt{\ }}L + 2X_{3\sqrt{\ }} - X_{2\sqrt{\ }} - \frac{X_{4\sqrt{\ }}}{X_{2\sqrt{\ }}}\right]\end{aligned} \tag{A.28}$$

where $X_{i\sqrt{\ }} \equiv \left\langle \sqrt{x^i} \right\rangle \Big/ \left\langle \sqrt{x} \right\rangle^i$ and $X_{i\sqrt{\ }}L \equiv \left\langle \sqrt{x^i} \ln x \right\rangle \Big/ \left\langle \sqrt{x} \right\rangle^i$. Again considering x to be the intensity and assuming a K-distribution, (3.49) can be used to expand the moments in terms of the parameter values that leads to the expression (4.31). Using (B.6) to convert (A.28) into errors in t leads to the result (4.33) used in Chapter 4.

A.5 Offset second moment μ_0 of the variable x

The measure and its estimate are defined by

$$\mu_0 = \left\langle x^2 \right\rangle - 1 \tag{A.29}$$

$$\hat{\mu}_0 = \frac{1}{m}\sum_{i=1}^{m} x_i^2 - 1 \tag{A.30}$$

Again expanding about the mean value and retaining terms up to the first order in $1/m$ gives the variance as

$$\begin{aligned} \sigma_{\hat{\mu}_0}^2 &= \left\langle \left(\hat{\mu}_0\right)^2 \right\rangle - \left\langle \hat{\mu}_0 \right\rangle^2 \\ &= \frac{1}{m}\left(\left\langle x^4 \right\rangle - \left\langle x^2 \right\rangle^2 \right) \end{aligned} \tag{A.31}$$

Assuming x to be homodyned-K distributed amplitude then the expressions for the second and fourth moment can be substituted from (5.11) and (5.12) respectively giving the result (5.14) for the relative variance in the offset parameter for the homodyned-K distribution.

A.6 Modified fourth central moment Y_x of the variable x

The measure and its estimate are defined by

$$Y_x \equiv \left\langle x^4 \right\rangle - \left(\left\langle x^2 \right\rangle + 1 \right)^2 \tag{A.32}$$

$$\hat{Y}_x \equiv \frac{1}{m}\sum_{i=1}^{m} x_i^4 - \left(\frac{1}{m}\sum_{i=1}^{m} x_i^2 + 1 \right)^2 \tag{A.33}$$

Again expanding about the mean value and retaining terms up to the first order in $1/m$ gives

$$\begin{aligned}\Delta Y_x &\equiv \langle \hat{Y}_x \rangle - Y_x \\ &= -\frac{1}{m}\left[\langle x^4 \rangle - \langle x^2 \rangle^2\right]\end{aligned} \tag{A.34}$$

$$\begin{aligned}\sigma_{Y_x}^2 &\equiv \left\langle \left(\hat{Y}_x\right)^2 \right\rangle - \langle \hat{Y}_x \rangle^2 \\ &= \frac{1}{m}\begin{pmatrix} \langle x^8 \rangle - 4\langle x^6 \rangle\langle x^2 \rangle - 5\langle x^4 \rangle^2 + 13\langle x^4 \rangle\langle x^2 \rangle^2 - 4\langle x^6 \rangle \\ +16\langle x^4 \rangle\langle x^2 \rangle - 16\langle x^2 \rangle^3 + 6\langle x^4 \rangle - 6\langle x^2 \rangle^2 - 8\langle x^2 \rangle^4 \end{pmatrix}\end{aligned} \tag{A.35}$$

for the bias and variance respectively. Again considering x to be the amplitude and assuming a homodyned-K distribution, the expression for the moments of the distribution can be substituted in (A.34) and (A.35) to arrive at the relationship (5.16) and (5.17) respectively. Using (B.6) to convert (A.34) and (A.35) into errors in t lead to the results (5.19) and (5.20) used in Chapter 5.

Appendix B

B. Calculating the variance of a parameter φ from a measure Q

Consider a statistic $\hat{Q}$ which is related to a parameter φ through a function g. The expectation value for $\hat{Q}$ is

$$Q = g(\varphi) \tag{B.1}$$

An estimate for φ, namely $\hat{\varphi}$, can be defined such that the measured statistic is related to the parameter estimate as

$$\hat{Q} = g(\hat{\varphi}) \tag{B.2}$$

In the asymptotic limit the measured statistic will take its expectation value and in this limit the inverse function of the statistic will be the true value of the parameter. The parameter estimate is thus asymptotically unbiased.

Using a Taylor series expansion

$$g(\hat{\varphi}) - g(\langle\hat{\varphi}\rangle) = (\hat{\varphi} - \langle\hat{\varphi}\rangle)\frac{\partial g}{\partial \hat{\varphi}}\Big|_{\langle\hat{\varphi}\rangle} + \frac{(\hat{\varphi} - \langle\hat{\varphi}\rangle)^2}{2}\frac{\partial^2 g}{\partial \hat{\varphi}^2}\Big|_{\langle\hat{\varphi}\rangle} + \cdots \tag{B.3}$$

In the asymptotic limit, $\langle\hat{\varphi}\rangle \to \varphi$ and substituting for the statistic $\hat{Q}$ in (B.3) gives

$$\hat{Q} - Q = (\hat{\varphi} - \langle\hat{\varphi}\rangle)\frac{\partial g}{\partial \hat{\varphi}}\Big|_{\langle\hat{\varphi}\rangle} + \frac{(\hat{\varphi} - \langle\hat{\varphi}\rangle)^2}{2}\frac{\partial^2 g}{\partial \hat{\varphi}^2}\Big|_{\langle\hat{\varphi}\rangle} + \cdots \tag{B.4}$$

Taking the square of both sides and then taking the expectation values gives

$$\left\langle \left(\hat{Q}-Q\right)^2 \right\rangle = \int \left(\hat{\varphi}-\langle\hat{\varphi}\rangle\right)^2 \left(\frac{\partial g}{\partial \hat{\varphi}}\Big|_{\hat{\varphi}} \right)^2 f_{\hat{\varphi}}(\hat{\varphi}) d\hat{\varphi} + \int \left(\hat{\varphi}-\langle\hat{\varphi}\rangle\right)^3 \frac{\partial g}{\partial \hat{\varphi}}\Big|_{\hat{\varphi}} \frac{\partial^2 g}{\partial \hat{\varphi}^2}\Big|_{\hat{\varphi}} f_{\hat{\varphi}}(\hat{\varphi}) d\hat{\varphi} + \cdots \tag{B.5}$$

Assuming that $\left\langle \left(\hat{\varphi}-\langle\hat{\varphi}\rangle\right)^j \right\rangle << \sigma_{\hat{\varphi}}^j$ where $j \le 3$, then the variance in $\hat{Q}$ is related to the variances in the estimated quantity $\hat{\varphi}$ as

$$\sigma_{\hat{Q}}^2 = \sigma_{\hat{\varphi}}^2 \left(\frac{\partial g}{\partial \hat{\varphi}}\Big|_{\hat{\varphi}} \right)^2 \tag{B.6}$$

Simulations were used to compare the third order central moments of $\hat{t}$ (and $\hat{v}$) values, obtained from the $\hat{V}$, $\hat{U}$ and $\hat{W}$ texture measures, to the variance in the estimated values and they were found to be an order of magnitude smaller. It will be assumed that the higher order central moments are also negligibly small. In this case the pdf of $\hat{t}$ (and $\hat{v}$) is very concentrated about its mean and the first order approximation used in (B.6) then holds true for the texture measures V, U and W.

Appendix C

C. Chi-squared significance test

This test is used to measure the goodness-of-fit. For a given set of random samples, we can test the hypothesis that the underlying distribution belongs to a specific family of distributions.

The hypothesis is tested by first binning the samples into M bins. The chi-squared statistic (χ^2) is then found by comparing the actual histogram entries with the predicted mean entry values

$$\chi^2 \equiv \sum_{i=1}^{M} \frac{\left(m_i - m\, f_i\right)^2}{m\, f_i} \tag{C.1}$$

where m_i is the number of entries in the histogram bin i, f_i is the pdf value for the data of bin i and m is the total number of data points in the sample. A probability is then assigned to this recorded difference given by

$$f_{\chi^2}(\chi^2) = \frac{2^{-\left(\frac{M-1}{2}\right)}}{\Gamma\left(\frac{M-1}{2}\right)} (\chi^2)^{\left(\frac{M-3}{2}\right)} \exp\left(-\chi^2/2\right) \tag{C.2}$$

A useful way of expressing the outcome of the chi-squared test is to observe the failure at the 5% level for the upper tail probabilities. The upper tail probability $f(\chi^2 > t)$ is given as

$$f(\chi^2 > t) = \int_t^{\infty} f_{\chi^2}(\chi^2) d\chi^2 \tag{C.3}$$

The chi-squared goodness-of-fit test was carried out on all three of the real radar images discussed in Chapter 3 (see Figure 3.6 and 3.7). For the MTI image, blocks of 91 range gates in each of the 64 frequency bins were selected for the chi-squared test. The data had been over-sampled by 2 and this was corrected by sub-sampling by two which gave 45 independent samples to be used for the goodness-of-fit test. A histogram was constructed from this data using five bins and setting the bin spacing to get equal probability of occupancy. This would give on average 9 samples per bin which were considered to be sufficient to measure the chi-squared statistics. The bin spacing for a negative exponential distribution with mean μ_I can be set analytically using the expression

$$x_{U_i} = -\mu_I \ln\left(1 - \frac{i}{M}\right) \tag{C.4}$$

for the upper bound of the ith bin. The appropriate bin spacing required for an equal probability of occupancy for a K-distribution with mean μ_I and order parameter ν has to be determined using an iterative process.

Furthermore, the chi-squared statistic nominally has $M-1$ degrees of freedom. However, since the parameters of the distribution under test have to be estimated using local statistics this effectively reduces the degrees of freedom for the chi-squared test. Using simulated data and measuring the failure rate at the 5% level on the upper tail probabilities it was found that the effective degree of freedom for the negative exponential distribution reduces to 3.0 and the K-distribution to 2.3. These were the values used for the goodness-of-fit results presented in Chapter 3.

The same procedure was carried out to perform the chi-squared test on the SAR images, the only difference being that a stepping window of 9×5 was used to select the independent samples. Thus for an original image of 512×512, a goodness-of-fit image of 102×56 was obtained.

Appendix D

D. Maximum likelihood classification

The ML classifier is used to select a model from a set of candidate distributions that best describes the data. For m independent samples $\{x_1, x_2, \ldots, x_m\}$, the total probability of occurrence P_T is given by

$$P_T(\{x_i\}|d_k)d\mathbf{x} = \prod_{i=1}^{m} f_x(x_i|d_k)dx \qquad \text{(D.1)}$$

where $f_x(x_i|d_k)$ is the pdf for the given distribution d_k. The ML classifier selects the distribution that yields the maximum value for $P_T(\{x_i\}|d_k)$. The ML classification is performed over small areas of the radar data to ensure that homogeneous regions are selected for this test. Where model parameters are not known *a priori* they are estimated from the data.

In Chapter 3 this test was used to select between negative exponential distribution and the K-distribution for the three examples of real images considered. The size of the stepping window used to select independent data samples for the ML classification test were kept identical to those used for the corresponding chi-squared test (see Appendix C). Thus for the MTI image (see Figure 3.6) the size of the stepping window used was 91 range cells by 1 frequency bin and the data was sub-sampled by 2. For the SAR images (see Figure 3.7) the corresponding size of the stepping window used was 5×9.

Appendix E

E. Relationship between the input and output distributions of basic binary operators

Given two independent positive random variables x_1 and x_2 with pdfs $f_{x_1}(x)$ and $f_{x_2}(x)$ respectively, the pdfs of their sum, difference, absolute difference, product, ratio and normalised ration are:

1. sum $y = x_1 + x_2$

$$f_y(y) = \int_0^y f_{x_1}(y-x) f_{x_2}(x) dx \tag{E.1}$$

which by the convolution theorem equals

$$= \frac{1}{2\pi} \int_0^\infty \Phi_{x_1}(\omega) \Phi_{x_2}(\omega) e^{-j\omega x} d\omega \tag{E.2}$$

where $\Phi_{x_1}(\omega)$ and $\Phi_{x_2}(\omega)$ are the characteristic functions of x_1 and x_2.

2. difference $y = x_1 - x_2$

$$f_y(y) = \int_{\max(-y,0)}^y f_{x_1}(y+x) f_{x_2}(x) dx \tag{E.3}$$

3. absolute difference $y = |x_1 - x_2|$

$$f_y(y) = \int_0^y \left[f_{x_1}(y+x) f_{x_2}(x) + f_{x_1}(x) f_{x_2}(y+x) \right] dx \tag{E.4}$$

4. product $y = x_1 x_2$

$$f_y(y) = \int_0^y f_{x_1}\left(\frac{y}{x}\right) f_{x_2}(x) \frac{1}{x} dx \tag{E.5}$$

5. ratio $y = \frac{x_1}{x_2}$

$$f_y(y) = \int_0^y f_{x_1}(yx) f_{x_2}(x) x dx \tag{E.6}$$

6. normalised ratio $y = \min\left(\frac{x_1}{x_2}, \frac{x_2}{x_1}\right)$

$$f_y(y) = \int_0^y \left[f_{x_1}(yx) f_{x_2}(x) + f_{x_1}(x) f_{x_2}(yx) \right] x dx \tag{E.7}$$

Appendix F

F. Glossary

This glossary briefly defines various mathematical functions and relationship referred to in the main body of the thesis. For further details on topics listed see; [1] for mathematical functions, [14] for relationships involving the Fourier transform, and [21, 63] for the statistics of random variables and processes. For ease of reference topics listed in the glossary are printed in **bold** in the main body of the thesis wherever the reader may require further explication of the topic (e.g., where it is first introduced), and when they are mentioned in the glossary under topic headings other than their own.

Autocorrelation and Autocovariance

The autocorrelation function (ACF) $R(\tau_1, \tau_2)$ of a random process $x(\tau)$ at τ_1 and τ_2 is defined by

$$R(\tau_1, \tau_2) = \langle x(\tau_1) x(\tau_2) \rangle \tag{F.1}$$

and the autocovariance $C(\tau_1, \tau_2)$ of a random process $x(\tau)$ at τ_1 and τ_2 is defined by

$$C(\tau_1, \tau_2) = \langle x(\tau_1) x(\tau_2) \rangle - \langle x(\tau_1) \rangle \langle x(\tau_2) \rangle \tag{F.2}$$

The normalised autocorrelation function $r(\tau_1, \tau_2)$ at τ_1 and τ_2 is derived by dividing the autocorrelation function at τ_1 and τ_2 by the expected values of $x(\tau)$ at τ_1 and τ_2

$$r(\tau_1,\tau_2)=\frac{R(\tau_1,\tau_2)}{\langle x(\tau_1)\rangle\langle x(\tau_2)\rangle}=\frac{\langle x(\tau_1)x(\tau_2)\rangle}{\langle x(\tau_1)\rangle\langle x(\tau_2)\rangle} \tag{F.3}$$

The normalised autocovariance function $c(\tau_1,\tau_2)$ is defined likewise

$$c(\tau_1,\tau_2)=\frac{C(\tau_1,\tau_2)}{\langle x(\tau_1)\rangle\langle x(\tau_2)\rangle}=\frac{\langle x(\tau_1)x(\tau_2)\rangle}{\langle x(\tau_1)\rangle\langle x(\tau_2)\rangle}-1 \tag{F.4}$$

The correlation coefficient is given by dividing the autocovariance by the variance.

Where $\tau_1=\tau_2=\tau_0$ the above functions generate the following first order statistics of $x(\tau)$ at τ_0. The autocorrelation function gives its second **moment**

$$R(\tau_0,\tau_0)=\langle x(\tau_0)^2\rangle \tag{F.5}$$

the autocovariance gives its **variance**

$$C(\tau_0,\tau_0)=\langle x(\tau_0)^2\rangle-\langle x(\tau_0)\rangle^2=\sigma^2_{x(\tau_0)} \tag{F.6}$$

the normalised autocorrelation function gives its second **normalised moment**

$$r(\tau_0,\tau_0)=\frac{\langle x(\tau_0)^2\rangle}{\langle x(\tau_0)\rangle^2}=x(\tau_0)^{(2)} \tag{F.7}$$

the normalised autocovariance function gives its **coefficient of variation**

$$c(\tau_0,\tau_0)=\frac{\langle x(\tau_0)^2\rangle-\langle x(\tau_0)\rangle^2}{\langle x(\tau_0)\rangle^2}=v^2_{x(\tau_0)} \tag{F.8}$$

When the random process $x(\tau)$ is **stationary** its autocorrelation function only depends on $\tau=\tau_1-\tau_2$, i.e.

$$R(\tau)=R(\tau_1,\tau_2)=R(\tau_3,\tau_4) \tag{F.9}$$

where $\tau = \tau_1 - \tau_2 = \tau_3 - \tau_4$.

Bessel functions

Bessel functions are solutions to the differential equations [1]

$$z^2 \frac{d^2 w}{dz^2} + z \frac{dw}{dz} + \left(a_B z^2 - v^2\right) w = 0 \tag{F.10}$$

where $a_B = \pm 1$. When $a_B = 1$ solutions include the Bessel function of the first kind $J_v(z)$, and when $a_B = -1$ solutions include the modified Bessel function of the second kind $K_v(z)$. Here we list the main properties of Bessel functions used in preceding chapters.

When $v = 0$ the Bessel function of the first kind has an integral representation [1]-(9.1.21)

$$J_0(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{jz\cos\theta} d\theta \tag{F.11}$$

and a series expansion [1]-(9.1.10)

$$J_0(z) = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{4^k (k!)^2} \tag{F.12}$$

The modified Bessel function of the second kind has an integral representation [1]-(9.6.24)

$$K_v(z) = \int_0^{\infty} e^{-z\cosh y} \cosh(vy) dy \tag{F.13}$$

where $|\angle z| < \pi/2$. The first derivative of $K_v(z)$ with respect to z is given by [1]-(9.6.26)

$$K'_{v-1}(z) = -\left(\frac{K_v(z) + K_{v-2}(z)}{2} \right) \tag{F.14}$$

Central limit theorem

The central limit theorem states that under certain conditions the sum $y = x_1 + \cdots x_m$ of m **independent** continuous random variables tends to being

Gaussian distributed with mean $\mu_y = \mu_1 + \cdots \mu_m$ and variance $\sigma_y^2 = \sigma_1^2 + \cdots \sigma_m^2$ as m increases.

Sufficient conditions for the central limit theorem to apply are [63]:

1. The variance of the sum must tend to infinity as the number of variables tends to infinity.
2. For some number $n > 2$ the nth moment of all the variables must be finite.

Characteristic function

The characteristic function $\Phi_x(\omega)$ of a random variable x is given by the **Fourier transform** of its **probability density function** $f_x(x)$

$$\Phi_x(\omega) = \int_{-\infty}^{\infty} f_x(x) e^{j\omega x} dx = \left\langle e^{j\omega x} \right\rangle \tag{F.15}$$

and the inverse Fourier transform of the characteristic function gives the pdf

$$f_x(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_x(\omega) e^{-j\omega x} d\omega \tag{F.16}$$

Substituting for $e^{j\omega x}$ by its series expansion given by [1]-(4.2.1)

$$e^{j\omega x} = \sum_{k=0}^{\infty} \frac{(j\omega x)^k}{k!} \tag{F.17}$$

the characteristic function of x is related to the **moments** by

$$\Phi_x(\omega) = \sum_{k=0}^{\infty} \frac{(j\omega)^k}{k!} \left\langle x^k \right\rangle \tag{F.18}$$

Thus the nth moment of x can be generated from the characteristic function by taking the nth derivative of its characteristic function

$$\frac{d^n \Phi_x(\omega)}{d\omega^n} = \sum_{k=0}^{\infty} \frac{j^k \omega^k}{k!} \left\langle x^{k+n} \right\rangle \tag{F.19}$$

and setting $\omega = 0$. The nth moment is then given by

$$\langle x^n \rangle = j^{-n} \frac{d^n \Phi_x(0)}{d\omega^n} \tag{F.20}$$

Circular symmetry

A bivariate function is termed to be circularly symmetric if its value only depends on the radial distance from the origin [63], i.e., the function $f(x, y)$ is circularly symmetric if

$$f(x_1, y_1) = f(x_2, y_2) \tag{F.21}$$

whenever

$$x_1^2 + y_1^2 = x_2^2 + y_2^2 \tag{F.22}$$

A complex random variable is said to be circularly symmetric if its **probability density function** is circularly symmetric; this will be the case if and only if the phase and amplitude are **independent**. The phase will then be uniformly distributed over the interval $(-\pi, \pi)$. The real and imaginary components of a circularly symmetric random variable are uncorrelated but they can only be independent if and only if they are **Gaussian**, i.e., if the amplitude is Rayleigh distributed.

Coefficient of variation

The coefficient of variation of a random variable x with **mean** μ_x and **variance** σ_x^2, is a measure of the width of its distribution relative to its mean value. It is given by

$$v_x^2 = \frac{\sigma_x^2}{\mu_x^2} \tag{F.23}$$

The square root of the coefficient of variation, i.e., standard deviation/mean, is normally used to measure contrast in an image degraded by multiplicative noise.

Convolution theorem

Given two functions $f_1(x)$ and $f_2(x)$ with **Fourier transforms** $F_1(\omega)$

and $F_2(\omega)$ respectively, the convolution theorem [63] states that their convolution $f(x) = f_1(x) * f_2(x)$ defined by

$$f(x) = \int_{-\infty}^{\infty} f_1(y) f_2(x - y) dy \tag{F.24}$$

has a Fourier transform $F(\omega)$ equal to the product of the Fourier transforms of the two functions

$$F(\omega) = F_1(\omega) F_2(\omega) \tag{F.25}$$

It follows from the convolution theorem that the characteristic function of the sum of two independent random variables will be given by the product of their **characteristic functions**. This is because the **probability density function** of the sum of two independent random variables is given by the convolution of their pdfs (E.1), and because the characteristic function of a random variable is just the Fourier transform of its pdf.

This result can be extended to the sum y of m independent random variables x_i, with the characteristic function of their sum $\Phi_y(\omega)$ being given by the product of the characteristic functions $\Phi_{x_i}(\omega)$ of the variables

$$\Phi_y(\omega) = \left\langle e^{j\omega y} \right\rangle = \left\langle e^{j\omega(x_1 + \cdots + x_m)} \right\rangle = \left\langle e^{j\omega x_1} \right\rangle \cdots \left\langle e^{j\omega x_m} \right\rangle = \prod_{i=1}^{m} \Phi_{x_i}(\omega) \tag{F.26}$$

Cumulative distribution function

The probability that a random variable x is less than y is given by its cumulative distribution function (cdf)

$$prob(x < y) = p_x(y) \tag{F.27}$$

The cdf of x is derived from its **probability density function** by integrating the latter from $-\infty$ to y

$$p_x(y) = \int_{-\infty}^{y} f_x(x) dx \tag{F.28}$$

Digamma function

The Digamma function [1] is defined

$$\psi(z) = \frac{d}{dz}\ln\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)} \tag{F.29}$$

where $\Gamma(z)$ is the **gamma function**. For positive integer values it is given by

$$\psi(k) = -\gamma_E + \sum_{i=1}^{k-1}\frac{1}{i} \tag{F.30}$$

when $k \geq 2$, and

$$\psi(1) = -\gamma_E \tag{F.31}$$

where $\gamma_E = 0.57722\ldots$ is **Euler's constant**. As $k \to \infty$

$$\psi(k) \to \ln k \tag{F.32}$$

The first derivative of Digamma function $\psi'(z)$ is known as the **Trigamma function**. For positive integer values it is given by

$$\psi'(k) = \frac{\pi^2}{6} - \sum_{l=1}^{k-1}\frac{1}{l^2} \tag{F.33}$$

when $k \geq 2$, and

$$\psi'(1) = \frac{\pi^2}{6} \tag{F.34}$$

As $k \to \infty$

$$\psi'(k) \to 0 \tag{F.35}$$

Ensemble average

The ensemble average (or expected value) of a function $u(x)$ of a continuous random variable x, is indicated by the operator $\langle\cdot\rangle$, and is defined

$$\langle u(x)\rangle = \int_{-\infty}^{\infty} u(x) f_x(x) dx \tag{F.36}$$

where $f_x(x)$ is the **probability density function** of x.

In the case of a discrete random variable k the ensemble average

$$\langle u(k) \rangle = \sum_{k=-\infty}^{\infty} u(k) f_k \tag{F.37}$$

where f_k is the **probability density function** of k.

Error function

The error function $erf(z)$ is the integral of the **Gaussian distribution** given by

$$erf(z) = \frac{2}{\pi} \int_0^z \exp(-y^2) dy \tag{F.38}$$

Estimation

Given a set of random variables $x_1, \ldots, x_m$ described by some unknown parameter φ, an estimate $\hat{\varphi}$ of φ based on the observations $x_1, \ldots, x_m$ is said to be **unbiased** if

$$\langle \hat{\varphi} \rangle = \varphi \tag{F.39}$$

and the estimate is said to be **consistent** if

$$\hat{\varphi} \rightarrow \varphi \tag{F.40}$$

as $m \rightarrow \infty$.

Fourier transform

The version of the Fourier transform $F(\omega)$ of a function $f(x)$ used in this thesis is defined

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-j\omega x} dx \tag{F.41}$$

and the inverse Fourier transform is defined

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega x} d\omega \tag{F.42}$$

Gamma function

The gamma function $\Gamma(z)$ is defined

$$\Gamma(z) = \int_0^{\infty} e^{-x} x^{z-1} dx \tag{F.43}$$

and has the property

$$\Gamma(z+1) = z\Gamma(z) \tag{F.44}$$

For positive integer values k

$$\Gamma(k+1) = k! = k \cdot (k-1) \ldots 3 \cdot 2 \cdot 1 \tag{F.45}$$

and for $z = \frac{1}{2}$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{F.46}$$

Gaussian distribution

A Gaussian (or normal) distributed random variable x with mean μ_x and **variance** σ_x^2 has **probability density function**

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{\frac{-(x-\mu_x)^2}{2\sigma_x^2}} \tag{F.47}$$

and central moments (i.e., with mean normalised to zero)

$$\left\langle (x-\mu_x)^n \right\rangle = \begin{cases} \Gamma\left(\frac{n+1}{2}\right) \Big/ \Gamma\left(\frac{1}{2}\right) & n \text{ even} \\ 0 & n \text{ odd} \end{cases} \tag{F.48}$$

Independence and Correlation

Two random variables x and y are statistically independent if and only if their **joint** pdf $f_{x,y}(x,y)$ is equal to the product of their **marginal** pdfs $f_x(x)$ and

$f_y(y)$

$$f_{x,y}(x,y) = f_x(x) f_y(y) \tag{F.49}$$

this is equivalent to the condition that for all **moments**

$$\langle x^{n_1} y^{n_2} \rangle = \langle x^{n_1} \rangle \langle y^{n_2} \rangle \tag{F.50}$$

When the above condition holds for the first moment (mean values)

$$\langle xy \rangle = \langle x \rangle \langle y \rangle \tag{F.51}$$

the random variables are said to be statistically uncorrelated. Thus independent random variables will also be uncorrelated; the reverse is only necessarily true if both random variables have **Gaussian distributions**.

Jacobian

Given two sets of m random variables $x = x_1, \ldots, x_m$ and $y = y_1, \ldots, y_m$ related by the transform $x_i = g_i(y)$, their **joint** pdfs $f_x(x)$ and $f_y(y)$ respectively are related by

$$f_y(y) = f_x(g_1(y), \ldots, g_m(y)) |J| \tag{F.52}$$

where J, the Jacobian of the transform $g(y)$ is given by

$$J = \frac{\partial x}{\partial y} = \begin{vmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_m}{\partial y_1} \\ \vdots & & \vdots \\ \frac{\partial g_1}{\partial y_m} & \cdots & \frac{\partial g_m}{\partial y_m} \end{vmatrix} \tag{F.53}$$

Joint and marginal distributions

The joint probability that $x_1 < y_1, \ldots, x_m < y_m$ is given by the joint **cumulative distribution function** $F_x(y)$, where $x = x_1, \ldots, x_m$ and $y = y_1, \ldots, y_m$. The joint **probability density function** of x is defined

$$f_x(y) = \frac{\partial^m F_x(y)}{\partial y_1 \ldots \partial y_m} \tag{F.54}$$

Thus

$$F_x(y) = \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_m} f_x(x) dx_1 \ldots dx_m \tag{F.55}$$

The marginal pdf of x_i is derived from the joint pdf of x by integrating it over $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m$

$$f_{x_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_x(x) dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_m \tag{F.56}$$

Mean

The mean value μ_x (or first **moment**) of a random variable x gives its expected value and is defined by its **ensemble average**

$$\mu_x = \langle x \rangle \tag{F.57}$$

Moments

The nth moment of a random variable x is defined

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n f_x(x) dx \tag{F.58}$$

where $f_x(x)$ is the **probability density function** of x. The moments of a random variable can also be generated from its **characteristic function**.

Normalised moments

The nth normalised moment of x is given by dividing its nth **moment** by its **mean** value raised to the power n

$$x^{(n)} = \frac{\langle x^n \rangle}{\langle x \rangle^n} \tag{F.59}$$

Probability density function

The probability density function (pdf) of x is given by differentiating its **cumulative density function**

$$f_x(x) = \frac{dF_x(x)}{dx} \tag{F.60}$$

Stationarity

A random variable $x(\tau)$ is considered to be strict-sense stationary if its statistical properties are invariant to a change of origin. The process is considered to be wide-sense stationary if the **mean** is constant and its **autocorrelation function** only depends on $\tau_1 - \tau_2$.

Variance and standard deviation

The variance σ_x^2 of a random variable x gives a measure of the width of its distribution and is defined

$$\sigma_x^2 = \left\langle \left(x - \mu_x\right)^2 \right\rangle = \left\langle x^2 \right\rangle - \left\langle x \right\rangle^2 \tag{F.61}$$

where $\mu_x = \langle x \rangle$ is the **mean** value of x. σ_x, the square-root of the variance, gives the standard deviation of x.

Appendix G

G. List of Symbols

α	weight ratio for combining the two moments in the hybrid normalised log texture measure
α_k	kth parameter of a distribution
β	parameter of the Jacobi polynomial
χ^2	chi-squared statistic
$\delta\Phi$	2-way phase difference
δR	2-way slant range path difference
Δh	height difference between two points on a surface S
$\Delta\mu_x$	bias in mean of x
$\Delta\mu_I$	bias in mean of intensity
$\Delta\theta$	angular width of a subaperture
ΔU, ΔU_x	bias of texture measure U_x
ΔV, ΔV_x	bias of texture measure V_x
ΔY, ΔY_x	bias of texture measure Y_x
ε	parameter of the Jacobi polynomial
Φ	phase of the detected field

$\Phi(\mathbf{r})$	phase of a wave at point $\mathbf{r}$
$\Phi(x,y)$	phase of a wave at point x and y
ϕ_0	depression angle
Φ_0	phase of constant vector
ϕ_k , $\phi_k(\mathbf{r})$	phase of the kth scatterer
γ	function of the components of $\mathbf{u}$
$\Gamma(z)$	gamma function with argument z
γ_E	Euler's constant
η	parameter of the generalised gamma distribution
φ	parameter of the statistic Q
ϑ	local incident angle
λ	wavelength
μ_0	intensity offset
μ_I	mean of intensity
μ_N	mean of number of steps in a random walk
μ_ρ	mean of RCS
μ_x	mean of random variable x
ν	order parameter
π	3.14159.....
θ	phase
θ_0	antenna azimuth angle
θ_i	incident angle of a plane wave to a surface

θ_i'	mirror incident angle of a plane wave to a surface
θ_k	azimuth angle between the start of the kth subaperture and the platform velocity vector
θ_m	azimuth angle between the middle of the subaperture and the platform velocity vector
θ_r	reflective angle of a plane wave from a surface
θ_{r_1}	reflective angle of a plane wave from a point x_1 on a surface
θ_{r2}	reflective angle of a plane wave from a point x_2 on a surface
ρ	surface radar cross-section
σ_E	standard deviation of detected field E
σ_h	standard deviation of surface height
σ_I	standard deviation of intensity
$\sigma_{\ln I}$	standard deviation of log intensity
σ_{μ_0}	standard deviation of the mean of intensity offset
σ_{μ_1}	standard deviation of the mean of intensity
σ_{μ_x}	standard deviation of the mean of x
σ_N	standard deviation of number of steps in a random walk
σ_v	standard deviation of v
σ_t	standard deviation of t
σ_U , σ_{U_x}	standard deviation of U
σ_V , σ_{V_x}	standard deviation of V
σ_W , σ_{W_x}	standard deviation of W

σ_x	standard deviation of x
σ_Y, σ_{Y_x}	standard deviation of Y
τ	time
τ_{int}	coherent integration time
τ_p	width of unmodulated pulse
τ_p'	width of modulated pulse
ω	dummy variable for frequency domain
ξ	error in the estimate of a quantity as a fraction of the expectation value
ψ	beamwidth
$\psi(z)$	digamma function with argument z
$\psi'(z)$	trigamma function with argument z
ψ_a	azimuth beamwidth
ψ_e	elevation beamwidth
ψ_S	surface term given by $\dfrac{\exp\left(j\dfrac{2\pi}{\lambda}r\right)}{r}$
ζ	parameter of the generalised gamma distribution
a	step length
A	amplitude
$A(\mathbf{r})$	amplitude of a wave at $\mathbf{r}$
$A(x,y)$	amplitude of a wave at point x and y

A_0	amplitude of constant vector
a_B	Bessel differential equation constant
$a_{g\Gamma}$	parameter of the generalised gamma distribution
a_k, $a_k(\mathbf{r})$	amplitude of the kth scatterer
b	the ratio ν/μ
B	bandwidth
$\mathcal{B}$	birth rate
b_W	scalar parameter of the Weibull distribution
c	speed of electromagnetic propagation through atmosphere, 3×10^8 ms^{-1}
$C(\tau)$	autocovariance function of a random process $x(\tau)$ at time τ
$c(\tau)$	normalised autocovariance function of a random process $x(\tau)$ at time τ
$C(u)$, $C(\underline{\mathbf{u}})$	characteristic function
c_W	skewness of the Weibull distribution
d	the expression $\sqrt{\frac{4\nu}{\mu}+\frac{4\mu_0}{\mu^2}}$
D	antenna length
$\mathcal{D}$	death rate
d_a	azimuth resolution
D_a	antenna azimuth length
D_e	antenna height
d_{gr}	ground range resolution

d_k	pdf of the kth distribution
d_{sr}	slant range resolution
E, $E(\mathbf{r})$	detected field
E_0	detected field at broadside
E_i	incident field
$E_{\mathcal{I}}$, $E_{\mathcal{I}}(\mathbf{r})$	imaginary component of detected field
E_i'	mirror incident field
E_P	scattered field observed at a point P
E_r	total scattered field
$E_{\mathcal{R}}$, $E_{\mathcal{R}}(\mathbf{r})$	real component of detected field
E_r^k	scattered field contribution from the h^k th term
E_S	field on surface S
$f(\tau)$	frequency at time τ
$f(x)$	function of the quantity x
$f(x,y,z)$	function of the quantities x, y and z
$F(x_a)$	detected field at azimuth lag x_a
$F(\omega)$	Fourier transform of a function $f(x)$
f_0	carrier frequency
f_A	amplitude pdf
f_c	clutter Doppler shift
f_{χ^2}	chi-squared probability
f_d	Doppler resolution

f_{d_k}	Doppler velocity of the kth subaperture
f_E	joint pdf of the detected field
f_i	predicted pdf corresponding to the ith histogram bin
f_I	intensity pdf
F_I	intensity cdf
f_N	pdf of the number of steps in a random walk
f_t	target Doppler frequency
f_x	pdf of random variable x
F_x	cdf of random variable x
$g(x)$	function of the variable x
h	surface height
$h(\tau)$	output of point target at time τ
$h(x,y)$	surface height at point x and y on surface S
h_1	surface height at point x_1
h_2	surface height at point x_2
$h_a(\tau)$	azimuth response of a point target at time τ
h_{ac}	platform altitude
$h_r(\tau)$	range response of a point target at time τ
i	loop counter, ith element of a list
I	intensity
I_i	intensity of the ith data sample
I_m	median value of intensity
j	loop counter, jth element in a list

j	$\sqrt{-1}$
J	Jacobian transform
$\mathcal{J}$	immigration rate
$J_o(\cdot)$	zeroth order Bessel function
k	loop counter, kth element of a list
$K_\nu(\cdot)$	modified Bessel function of the second kind of order ν
l	number of parameters in a distribution
L	likelihood function of a distribution with l parameters
L_i	log moments $\langle \ln^i x \rangle$
l_S	average slope of the surface S
L_S	synthetic aperture length
m	number of independent random samples
M	number of histogram bins
m_i	number of data entries in the ith histogram bin
n	number of moments
N	number of scatterers or steps in the random walk
N_p	number of pulses coherently integrated
n_S	normal to a surface S
P	observation point for the scattered field from a surface S
$p(\tau)$	transmitted pulse at time τ
P_{ac}	platform position
PRF	Pulse Repetition Frequency
P_T	total probability of occurrence of m independent data samples

Q	a statistic related to the parameter φ through the function g
r	distance between a point O on a surface S to an observation point P
$\mathbf{r}$	vector distance of an observation point from a point on a scattering surface
R	slant range
$r(\tau)$	normalised autocorrelation function of a random process $x(\tau)$ at time τ
$R(\tau)$	autocorrelation function of a random process $x(\tau)$ at time τ
R_0	slant range at broadside
R_{coeff}	reflectivity coefficient
R_g	ground range
r_τ	time resolution
s	dummy variable for integration
S	rough surface
S_{sr}	maximum slant range
t	reciprocal of ν
u	scalar variable
$\mathbf{u}$	vector variable
U, U_x	normalised log texture measure
V, V_x	intensity contrast texture measure
v_{ac}	platform velocity
v_I	coefficient of variation of intensity

v_t	target radial velocity
w	dummy variable
W, W_x	hybrid normalised log texture measure
$W_{in_j,k}$	neural net weight connecting the jth node of the input layer with the kth node of the hidden layer
$W_{k,out}$	neural net weight connecting the kth node of the hidden layer with the output layer node
x	random variable, Cartesian coordinate
X	target position on ground
$x^{(n)}$	normalised moments
x_a	azimuth lag
x_i	the ith data sample from m independent random variables
X_i	normalised moments $\langle x^i \rangle / \langle x \rangle^i$
X_iL	log normalised moments $\langle x^i \ln x \rangle / \langle x \rangle^i$
x_{in}	value of the input node of a neural net
x_{in_j}	value of the jth node in the input layer of a neural net
$X_{i\sqrt{\ }}$	normalised square root moments $\langle \sqrt{x^i} \rangle / \langle \sqrt{x} \rangle^i$
$X_{i\sqrt{\ }}L$	log normalised square root moments $\langle \sqrt{x^i} \ln x \rangle / \langle \sqrt{x} \rangle^i$
x_k	value of the kth node in the hidden layer of a neural net
x_{out}	value of the out node of a neural net
x_r	range lag
x_{U_i}	bin spacing for the upper bound of the ith histogram bin

y	random variable, Cartesian coordinate
Y, Y_x	texture measure for homodyned-k distribution
z	argument of a function, random variable, Cartesian coordinate

Bibliography

[1] M. Abramowitz and I. Stegun, '*Handbook of Mathematical Functions*', Dover, New York, 1965.

[2] E Bahar, 'Full-wave solutions for the scattered radiation fields from rough surfaces with arbitrary slope and frequency', *IEEE Trans. Antennas and Prop.*, Vol. 28, pp11-21, 1980.

[3] C. J. Baker, 'Coherent properties of K-distributed sea clutter', *In Proc. 16th Europ Microwave Conf.*, pages 311-316, (Dublin Ireland), 1986.

[4] R. Baraket, 'Weak-scatter generalisation of the K distribution in scattering experiments', *Waves in Random Media*, 401-409, 1986.

[5] F. G Bass and I. M. Fuks, '*Wave Scattering from Statistically Rough Surfaces*', Oxford Pergamon, 1979.

[6] P. Beckmann, 'A new approach to the problem of reflection from rough surfaces', *Acta Techn.*, Vol. 2, pp311-355, 1957.

[7] P. Beckmann and A. Spizzichino, '*The Scattering of Electromagnetic Waves from Rough Surfaces*', Pergamon Press, Oxford, 1963.

[8] P. Beckmann, 'Scattering by non-Gaussian surfaces', *IEEE Trans. Antennas and Prop.*, Vol. 21, pp169-175, 1973.

[9] M. Di Bisceglie, 'Statistical characterisation of high resolution sidescan sonar returns', *University College London internal report*, June 1997.

[10] D. Blacknell and R. G. White, 'Optimum classification of non-Gaussian processes using neural networks', *IEE Proc. Radar Sonar Navig.*, Vol. 141, pp56-66, 1994.

[11] D. Blacknell, 'Comparison of parameter estimation for K-distribution', *IEE Proc. Radar Sonar Navig.*, Vol. 141, pp45-52, 1994.

[12] A. R. Blake, D. Blacknell and C. J. Oliver, 'SAR clutter analysis and its resolution dependence', *In Proc. IEE Radar 97*, pages 124-128, (Edinburgh, UK), 1997.

[13] R. Boothe, '*The Weibull distribution applied to the ground clutter backscatter coefficients: Automatic detection and radar data processing*', Ed. D. C. Schleher. Artech House, London, 1980.

[14] R. N. Bracewell, '*The Fourier Transform and its Applications*', McGraw-Hill, New York, 1986.

[15] L. M. Brekhovskikh, 'The diffraction of waves by a rough surface' (in Russian), *Zh. Eksp. Teor. Fiz.*, Vol. 23, pp275-289, 1952.

[16] B. Brisco, F. T. Ulaby, F. Kouyate and T. H. L. Williams, 'Textural information in SAR images', *IEEE Trans. Geosci. Remote Sens.*, Vol. 24, pp235-245, 1986.

[17] W. G. Carrara, R. S. Goodman and R. M. Majewski, '*Spotlight Synthetic Aperture Radar: Signal Processing Algorithms*', Artech House, Norwood, MA, 1995.

[18] M. F. Chen, S. C. Wu and A. K. Fung, 'Regions of validity of common scattering models applied to non-Gaussian rough surfaces', *J. Wave-Material Interaction*, Vol. 2, pp9-27, 1987.

[19] H. T. Chuah and H. S. Tan, 'A radar backscatter model for forest stands', *Waves in Random Media*, Vol. 2, pp7-28, 1992.

[20] D. R. Cox and D V. Hinkley, '*Theoretical Statistics*', Chapman and Hall, 2nd edn., London, pp. 304, 1974.

[21] W. B. Davenport and W. L. Root, '*An Introduction to the Theory of Random Signals and Noise*', IEEE Press, New York, 1987.

[22] C. Deng and S. Haykins, 'Classification of radar clutter using neural networks', *IEEE Trans. Neural Netw.*, Vol. 2, pp589-600, 1993.

[23] L. N. Deryugin, 'Equations for the coefficients of reflection of waves from a periodically rough surface' (in Russian), *Dokl. Akad. Nauk SSSR*, Vol. 87, pp913-916, 1952.

[24] J. L. Ekstrom, 'The detection of steady targets in Weibull clutter', *In Proc. IEE Conf. Radar-Present and Future*, pages 221-226, (London, UK), 1973.

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

[26] A. Farina, A. Russo and F. A. Studer, 'Coherent radar detection in log-normal clutter', *IEE Proc. F*, Vol. 133, pp39-54, 1986.

[27] F. A. Fay, J. Clarke and R. S. Peters, 'Weibull distribution applied to sea clutter', *In Proc. IEE Radar 77*, pages 101-104, (London, UK), 1977.

[28] A. K. Fung and F. T. Ulaby, 'A scatter model for leafy vegetation', *IEEE Trans. GeoSci. Electron.*, Vol. 16, pp281-285, 1978.

[29] A. K. Fung and H. J. Eom, 'Note on the Kirchhoff rough surface solution in backscattering', *Radio Sci.*, Vol. 16, pp299-302, 1981.

[30] A. K. Fung and M. F. Chen, 'Numerical simulation of scattering from simple and composite random surfaces', *J. Opt. Soc. Am.*, Vol. 2, pp2274-2284, 1985.

[31] G. B. Goldstein, 'False-Alarm regulation in log-normal and Weibull clutter', *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 9, pp84-92, 1973.

[32] J. W. Goodman, 'Statistical properties of laser speckle patterns', In J. C. Dainty, editor, *Laser speckle and related phenomena*, pages 9-75. Springer, New York, 1975.

[33] H. D. Griffiths, J. Dunlop and R. Voles, 'Texture analysis of sidescan sonar imagery using statistical scattering models', *In Proc. SACLANTCEN Conf.*, pages 187-194, (Lerici, Italy), 1997.

[34] A Horne, 'Final report on SAR target modelling', *DRA report No. DRA/LS2/TR95001/1.0*, 1995.

[35] M. Jahangir, 'A study of coherent radar clutter for MTI radar systems', *DRA report No. DRA/CSS2/TR94039/2.0*, 1995.

[36] M. Jahangir, D. Blacknell and R. G. White. 'Accurate approximation to the optimum parameter estimate for K-distributed clutter', *IEE Proc. Radar Sonar Navig.*, Vol. 143, pp383-390, 1996.

[37] E. Jakeman and P. N. Pusey, 'A model for non-Rayleigh sea echo', *IEEE Trans. Antenna and Prop.*, Vol. AP-24, pp804-818, 1976.

[38] E. Jakeman, 'On the statistics of K-distributed noise', *J. Phys. A : Math. Gen.*, Vol. 13, pp31-48, 1980.

[39] E. Jakeman, 'Statistics of integrated gamma-lorentzian intensity fluctuations', *Optica Acta*, Vol. 27, pp735-741, 1980.

[40] E. Jakeman and R. J. A. Tough, 'Generalized K distribution: a statistical model for weak scattering', *J. Optical Soc. America A.*, Vol. 4, pp1764-1772, 1987.

[41] J. K. Jao, 'Amplitude distribution of composite terrain radar clutter and the K-distribution', *IEEE Trans. Antennas and Prop.*, Vol. AP-32, pp1049-62, 1984.

[42] I. R. Joughin, D. B. Percival and D. P. Winerbrenner, 'Maximum likelihood estimation of K distribution parameters for SAR data', *IEEE Trans. Geosci. Remote Sens.*, Vol. 31, pp989-999, 1992.

[43] I. Jouny, F. D. Garber and S. C. Ahalt, 'Classification of radar targets using synthetic neural networks', *IEEE Trans. Aerosp. Electon. Syst.*, Vol. 29, pp336-343, 1993.

[44] M. A. Karam, A. K. Fung and Y. M. M. Antar, 'Electromagnetic wave scattering from some vegetation samples', *IEEE Trans. GeoSci. Electron.*, Vol. 26, pp799-808, 1988.

[45] S. Kingsley and S. Quegan, '*Understanding Radar Systems*', McGraw-Hill, New York, 1992.

[46] D. E. Kreiten and G. G. Hogan, 'Statistical analysis of Ka-band sea clutter', *In Proc. OCEANS-91*, (Honolulu, Hawaii), 1991.

[47] E. O. LaCasce and P. Tamarkin, 'Underwater sound reflection from a corrugated surface', *J. Appl. Phys.*, Vol. 27, pp138-148, 1956.

[48] R. R. Lentz, 'A numerical study of electromagnetic scattering from ocean-like surfaces', *Radio Sci.*, Vol. 9, pp1139-1146, 1974.

[49] R. P. Lippmann, 'An introduction to computing with neural nets', *IEEE Int. Conf. Acoustics, Speech Signal Process*, pages 4-22, 1987.

[50] P. Lombardo and C. J. Oliver, 'Estimation of texture parameters in K-distributed clutter', *IEE Proc. F Radar Sonar Navig.*, Vol. 141, pp196-204, 1994.

[51] P. Lombardo, C. J. Oliver and R. J. A. Tough, 'Effect of noise on order parameter estimation for K-distributed clutter', *IEE Proc. Radar Sonar Navig.*, Vol. 142, pp33-40, 1995.

[52] M. Longo, E. Conte and M. Lops, 'Modelling and simulation of non-Rayleigh radar clutter', *IEE Proc.-F*, Vol. 138, pp121-130, 1991.

[53] R. J. Miller, 'A coherent model of radar weather clutter', *In Proc. IEE Coll. Radar clutter and Multipath Prop.*, pages 3/1-4, (London UK), 1989.

[54] H. Ogawa, M. Sekine, T. Musha, M. Aota, M. Ohi and H. Fukushi, 'Weibull distributed radar clutter reflected from sea ice', *Trans. IEICE Japan*, Vol. 70, pp116-120, 1987.

[55] J. A. Ogilvy, '*Theory of Wave Scattering from Random Rough Surfaces*', IOP Publ. Ltd, Bristol, 1991.

[56] C. J. Oliver, 'A model of non-Rayleigh scattering statistics', *Optica Acta*, Vol. 31, pp701-722, 1984.

[57] C. J. Oliver, 'The interpretation and simulation of clutter textures in coherent images', *Inv. Problems*, Vol. 2, pp481-518, 1986.

[58] C. J. Oliver, 'Synthetic aperture radar imaging', *J. Phy. D. Applied Phy.,* Vol. 22, pp871-890, 1989.

[59] C. J. Oliver and R. G. White, 'Radar clutter classification based on noise models and neural networks', *In Proc. IEEE Int. Radar Conf.*, pages 329-334, (Arlington VA, US), 1990.

[60] C. J. Oliver, 'Information from SAR images', *J. Phy. D. Applied Phy.,* Vol. 24, pp1493-1514, 1991.

[61] C. J. Oliver, 'Optimum Texture estimators for SAR clutter',. *J. Phys. D: Appl. Phys.*, Vol. 26, pp1824-1835, 1993.

[62] C. J. Oliver and S. Quegan, '*Understanding Synthetic Aperture Radar Images*', Artech House, Boston, 1998.

[63] A. Papoulis, '*Signal Analysis*', McGraw-Hill, New York, 1977.

[64] A. Papoulis, '*Probability, Random Variables and Stochastic Process*', McGraw-Hill, New York, 1991.

[65] D. Prichard, '*The Radar War: Germany's Pioneering Achievement 1904-45*', Patrick Stephens Ltd, Wellingborough, 1989.

[66] R. K. Raney, 'Considerations for SAR image quantification unique to orbital systems', *IEEE Trans. Geosci. Remote Sens.*, Vol. 29, pp754-760, 1991.

[67] Lord Rayleigh, '*The Theory of Sound*', Dover New York, 1945.

[68] S. O. Rice, 'Reflection of electromagnetic waves by slightly rough surfaces', *Commun. Pure Appl. Math.*, Vol. 4, pp351-378, 1951.

[69] A. W. Rihaczek, '*Principles of High-Resolution Radar*', McGraw-Hill. New York, 1969.

[70] A. W. Rihaczek and S. J. Hershkowitz, '*Radar Resolution and Complex-Image Analysis*', Artech House, Boston, 1996.

[71] D. E. Rumelhart and J. L. McCleland, '*Parallel distributed processing: Exploration in the microstructure of cognition, vol 1: Foundations*', MIT Press Massachusetts, 1986.

[72] G. C. Sarno, 'A model of coherent radar land backscatter', *In Proc. AGARD Target and clutter scattering.* pages 3/1-11, (Ottawo, Canada), 1991.

[73] D. C. Schleher, 'Radar detection in Log-Normal clutter', *In Proc. IEEE Int. Radar Conf. Radar* '75, pages 262-267, (Arlington VA, US), 1975.

[74] D. C. Schleher, 'Radar detection in Weibull clutter', *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AES-12, pp736-43, 1976.

[75] D. C. Schleher, '*MTI and Pulse Doppler Radar*', Artech House, MA, 1991.

[76] J. P. Schouten and A. T. De Hoop, 'Sur la reflexion d'une onde électromagnétique plane par une surface rugueuse parfaitement conductrice', *Ann. Télécomm.*, Vol. 12, pp211-214, 1957.

[77] M. Sekine, T. Musha, Y. Tomita, T. Hagisawa, T. Irabu and E. Kiuchi, 'Weibull distributed weather clutter', *IEEE Trans. Aerosp. Electon. Syst.*, Vol. 15, pp824-830, 1979.

[78] M. Sekine, T. Musha, Y. Tomita, T. Hagisawa, T. Irabu and E. Kiuchi, 'Log-Weibull distributed sea clutter', IEE Proc. F, Vol. 127, pp225-228, 1980.

[79] M. Sekine, S. Ohtani, T. Musha, T. Irabu, E. Kiuchi, T. Hagisawa and Y. Tomita, 'Weibull distributed ground clutter', *IEEE Trans. Aerosp. Electon. Syst.*, Vol. 17, pp596-598, 1981.

[80] W. T. Shaw and A. J. Dougan, 'Half-space Green's function and application to scattering of electromagnetic fields from ocean-like surfaces', *Waves from Random Media* Vol. 5, pp341-359, 1995.

[81] M. I. Skolnik, '*Introduction to Radar Systems*', McGraw-Hill, New York, 1962.

[82] M. I. Skolnik, '*Radar Handbook*', McGraw-Hill, 2nd edition, 1990.

[83] G. W. Stimson, '*Introduction to Airborne radar*', Second Ed., Scitech Publ., New Jersey, 1998.

[84] W. J. Szajnowski, 'Estimators of log-normal distribution parameters', *IEEE Trans. Aerosp. Electon. Syst.*, Vol. 13, pp533-536, 1977.

[85] R. J. A. Tough, 'New techniques for periscope detection', *DRA technical report No.* TWR/R005/01, 1996.

[86] G. V. Trunk and S. F. George, 'Detection of targets in non-Gaussian sea clutter', *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 6, pp620-628, 1970.

[87] V. Twersky, 'On the scattering and reflection of electromagnetic waves by rough surfaces', *Trans. I.R.E.*, Vol. 5, pp81-90, 1957.

[88] F. T. Ulaby and T. F. Bush, 'Fading characteristics of panchromatic radar backscatter from selected agriculture targets', *IEEE Trans. Geosci. Electron.*, Vol. 13, pp149-157, 1975.

[89] F. T. Ulaby, R. K. Moore and A. K. Fung, '*Microwave Remote Sensing*, volume 1', Addison Wesley, Reading, MA, 1981.

[90] F. T. Ulaby, R. K. Moore and A. K. Fung, '*Microwave Remote Sensing*, volume 2', Addison Wesley, Reading, MA, 1982.

[91] K. D. Ward, 'Compound representation of high resolution sea clutter', *Electron. Lett.*, Vol. 17, pp561-563, 1981.

[92] K. D. Ward, 'A radar sea clutter model and its application to performance assessment', *In Proc. IEE Int. Conf. Radar '82*, pages 203-207, (London UK), 1982.

[93] K. D. Ward, 'Application of the K-distribution to radar clutter', *In Proc. Int. Symp. Noise and clutter rejection in radars and imaging sensors.* pages 15-20, (Kyoto Japan), 1989.

[94] K. D. Ward, C. J. Baker and S. Watts, 'Maritime surveillance radar. Part 1: Radar scattering from the sea ocean surface', *IEE Proc.-F*, Vol. 137, pp51-62, 1990.

[95] M. P. Warden, 'An experimental study of some clutter characteristics', *AGARD Conf. Proc. On Advanced Radar Syst.*, No. 66, 1970.

[96] J. G. Watson and J. B. Keller, 'Rough surface scattering via the smoothing method', *J. Ac. Soc. Am.,* Vol. 75, pp1705-1708, 1984.

[97] J. W. Wright, 'A new model for sea clutter', *IEEE Trans. Antennas and Prop.*, Vol. 16, pp217-223, 1968.

[98] S. T. Wu and A. K. Fung, 'A noncoherent model for microwave emissions and backscattering from the sea surface', *J. Geophys. Res.*, Vol. 77, pp5917-5929, 1972.