

DISSERTATION SUBMITTED TO THE UNIVERSITY OF CAMBRIDGE
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Bayesian State-Space Modelling of Spatio-Temporal Non-Gaussian Radar Returns

Jacek Leszek NOGA
Corpus Christi College

December 1998



SIGNAL PROCESSING LABORATORY
Department of Engineering
University of Cambridge

To my parents.

Bayesian State-Space Modelling of Spatio-Temporal Non-Gaussian Radar Returns

Jacek Leszek NOGA
Corpus Christi College

December 1998

Radar backscatter from an ocean surface is commonly referred to as sea clutter. Any radar backscatter not due to the scattering from an ocean surface constitutes a potential target. This thesis is concerned with the study of target detection techniques in the presence of high resolution sea clutter.

In this dissertation, the high resolution sea clutter is treated as a compound process, where a fast oscillating speckle component is modulated in power by a slowly varying modulating component. While the short term temporal correlations of the clutter are associated with the speckle, the spatial correlations are largely associated with the modulating component. Due to the disparate statistical and correlation properties of the two components, a piecemeal approach is adopted throughout this thesis, whereby the spatial and the temporal correlations of high resolution sea clutter are treated independently.

As an extension of the previously reported work on target detection in coherent radar systems, a complex autoregressive process is proposed as the basis for characterisation of high resolution sea clutter spectra in incoherent radar systems. As no phase information is available in incoherent radar returns, the Gibbs sampler is used to facilitate sampling from the autoregressive process parameter posterior distribution, conditional on the observed amplitudes. To this end, the Hybrid Monte Carlo algorithm is employed to conditionally sample for the missing phases.

Based on birth-death migration arguments for the evolution of a population of scattering centres on an ocean surface, a conditional heteroscedastic (*i.e.* non-constant prediction error variance) model is proposed for the modulating component of high resolution sea clutter in the logarithm domain. However, based on the results obtained for a large database of sea clutter range profiles, it is shown that there appears to be no strong evidence for heteroscedasticity. Instead, contrary to the widely held beliefs, rather than being Gamma distributed, it is demonstrated that the modulating component of sea clutter is better modelled as being log-Normal distributed, and hence Gaussian distributed in the logarithm domain.

The findings presented in this dissertation are culminated in the context of Constant False Alarm Rate (CFAR) detection. In particular, a linearised state space model for the compound high resolution sea clutter in the logarithm domain is proposed. The state space model is used to obtain a *Maximum A Posteriori* (MAP) estimate of the underlying modulating component in the cell under test, based on which the CFAR detection threshold is set. The CFAR detection threshold is thus obtained as a simple weighted average of radar returns (in the logarithm domain) contained in the reference window.

Finally, based on a large database of high resolution sea clutter range profiles, it is demonstrated that the proposed state space CFAR detector, in conjunction with pulse integration techniques, achieves a near ideal CFAR detection performance, particularly in spiky and spatially correlated clutter environment.

Declaration

The research described in this dissertation was carried out by the author between October 1994 and September 1998. Except as indicated in the text, the contents are entirely original and are not the result of work done in collaboration. No part of this dissertation has been submitted to any other University. The dissertation contains no more than 65,000 words.

Jacek L. Noga
December 1998

Acknowledgements

Firstly, I would like to thank my supervisor, Dr WJ Fitzgerald, for all the help, advice and encouragement that he has provided me with over the last four years. I am also grateful to Dr S Luttrell for providing me with the sea clutter data that proved to be so vital for this research.

Special thanks go to Dr MD Wu, and Messrs J Vermaak, JCP Khoo and S Barker for their constant friendship and support throughout this long ordeal. This thesis could not be completed if it were not for the help of all the other members of the Signal Processing Laboratory, especially to Paul Walmsley, Tim Clapp and others responsible for the maintenance of the highly temperamental computer network.

Thanks to Jaco Vermaak, Mark Coates and Ben Bradshaw for proof reading parts of this thesis. An eternal debt of gratitude goes to Anna Piekarska for her self sacrifice in painstaking and thorough proof-reading of the final drafts. Unfortunately, any remaining mistakes are purely my own.

Finally, looking back and reflecting upon the four bygone years, I fondly remember my time spent in Cambridge. It was not just the place, but the people around me, who have made the place feel like home. I cherish the rare times spent with Ania, hectically traveling around the Continent; the words of wisdom received from my parents. Those times have now ended. It is now time to move on; to put the Cambridge experience behind me. It is time to write yet another chapter... though this time not for the thesis. I dedicate this new chapter to Ania and my dear parents.

This work was funded by the Commonwealth Scholarship Commission, and supported in part by DRA Agreement Number CSM3202905. My final year has largely been funded thanks to the generosity of the Department of Engineering, Corpus Christi College, Booz-Allen & Hamilton and, of course, my magnanimous parents.

Keywords

Sea Clutter; CFAR; Target Detection; State Space; Hybrid Monte Carlo.

Contents

1	Introduction	1
1.1	The Maritime Surveillance Radar	1
1.1.1	Practical Limitations of a Surveillance Radar System	3
1.1.2	The Range Profiles	5
1.2	Outline of the Dissertation	7
2	Review of Sea Clutter Modelling	9
2.1	Properties of Sea Clutter	9
2.1.1	High Resolution Sea Clutter	9
2.1.2	The Modulating Component	12
2.2	Deterministic Models for Sea Clutter	13
2.2.1	Dynamical System Reconstruction	14
2.2.2	Attractor Dimension and Lyapunov Characteristic Exponent	16
2.2.3	Non-linear Evolution Function	19
2.3	Non-Gaussian Models for Sea Clutter	21
2.3.1	Single Point Statistics Modelling	22
2.3.2	Single Point Distribution Parameter Estimation	24
2.3.3	Temporal Correlations	26
2.3.4	Spatial Correlations	30
2.4	Summary	34
3	Target Detection in the Presence of Sea Clutter	37
3.1	Introduction	37
3.2	Target Models	38
3.3	Single Pulse Fixed Threshold Detection	40
3.3.1	Non-fading Targets	41
3.3.2	Fluctuating Targets	41
3.4	Fixed Threshold Detection with Integration	42
3.4.1	Incoherent Integration	44
3.4.2	Binary Integration	46
3.5	Adaptive Threshold (CFAR) Detection	47
3.5.1	Range Profile Data	48

3.5.2	Effects of $p(\sigma)$ on Probability of False Alarm Prediction	50
3.5.3	Target Detection in the Presence of Correlated Clutter	52
3.5.4	Adaptive Threshold Estimation	56
3.5.5	Detection Performance Analysis of CA CFAR Detectors	58
3.5.6	Detection Performance of CFAR Detectors	60
3.6	Neyman-Pearson Detection	62
3.7	Target Detection in Coherent Radar Systems	64
3.7.1	Pre-detection Processing Techniques	65
3.7.2	Spectrum-based Target Detection Techniques	66
3.7.3	Extensions to Incoherent Radar Systems	66
3.8	Summary	67
4	Correlation Structure of Incoherent Speckle	71
4.1	Introduction	71
4.2	Autoregressive Modelling of Sea Clutter	72
4.2.1	Coherent Radar System	73
4.2.2	Incoherent Radar System	75
4.2.3	Sampling from $p(\mathbf{a} \mathbf{r})$	77
4.3	Hybrid Monte Carlo Algorithm	79
4.3.1	The Gibbs Distribution and Molecular Dynamics	82
4.3.2	Hybrid Monte Carlo Algorithm	83
4.3.3	Remarks about the HMC Algorithm	86
4.3.4	Symmetrical Momentum Tempering	88
4.4	Characterisation of Incoherent Clutter	90
4.4.1	Sampling Procedure	90
4.4.2	Synthetic $AR(1)$ Process	93
4.4.3	Synthetic $AR(2)$ Process	96
4.4.4	Synthetic $AR(3)$ Process	97
4.4.5	Incoherent Sea Clutter	100
4.4.6	Synthetic $AR(1)$ target in the Presence of Incoherent Sea Clutter	103
4.5	Discussion, Conclusion and Suggestions for Further Research	108
5	Modulating Component Modelling using Conditional Heteroscedastic Models	111
5.1	Introduction	111
5.2	Conditional Heteroscedastic Models	112
5.2.1	Discrete Time Conditional Heteroscedastic Models	114
5.3	Heteroscedasticity in Sea Clutter	120
5.3.1	The Form of Regression	121
5.3.2	Presence of Heteroscedasticity	124
5.4	Extensions to Range Profiles	126
5.4.1	Non-linearity in Range Profiles	127

5.4.2	Heteroscedasticity in Range Profiles	127
5.4.3	Symmetric Models	131
5.4.4	Implications for CFAR Detection	134
5.5	Discussion, Conclusion and Suggestions for Future Research	136
6	CFAR Target Detection in Spatially Correlated Sea Clutter	139
6.1	Introduction	139
6.2	Compound Sea Clutter Model	140
6.2.1	The Speckle Component	141
6.2.2	The Modulating Component	143
6.3	State Space Estimation of Local Clutter Power	144
6.3.1	LMAP Estimator for Local Clutter Power	145
6.3.2	Predictive LMAP Estimator for Local Clutter Power	146
6.3.3	Extensions to non-Homogeneous Clutter Environment	148
6.4	CHLMAP CFAR Detection Performance Analysis	149
6.4.1	Distribution of the Threshold	150
6.4.2	Target Detection Performance Analysis	151
6.5	Detection Performance in Spatially Correlated Clutter	154
6.5.1	Fixed Threshold and Ideal CFAR Detection Performance Limits	154
6.5.2	CA CFAR Detector Performance	157
6.5.3	(CH)LMAP CFAR Detector Performance	160
6.6	Discussion, Conclusion and Suggestions for Further Research	166
7	Summary, Conclusions and Suggestions for Further Research	169
A	Correlations in Rayleigh Clutter	173
B	Review of Markov Chain Theory	177
C	ML Parameter Estimation for AR-CH Models	181
C.1	AR-CH Models	181
C.1.1	Estimation of α	183
C.1.2	Estimation of β	184
C.1.3	Parameter Estimation Algorithm	185
C.2	Symmetric AR-CH Models	186
C.2.1	The Jacobian	187
C.2.2	Estimation of α	187
C.2.3	Estimation of β	188
D	Detection Threshold Setting	189
D.1	Introduction	189
D.2	Approximate Relationship of T and P_{fa}	189

D.2.1	The Inverse Gamma(α, β) Distribution	190
D.2.2	The Gamma(α, β) Distribution	190
D.2.3	The log-Normal(μ, σ) Distribution	191
E	(Inverse) Gamma-log Normal Distribution Matching using Moments	195
E.1	Introduction	195
E.2	Gamma-log Normal Distribution Matching	197
E.3	Inverse Gamma-log Normal Distribution Matching	198
	Bibliography	199

"It's not the cover that makes a thesis..."

"... it's the pretty looking pictures inside."

AEP
December 1998

1.1 THE MARITIME SURVEILLANCE RADAR

The concept of radar dates back to 1886, when Hertz discovered that metallic and dielectric objects reflect radio waves. Though substantial progress in radar technology has been made since, the most rapid development of radar systems has occurred during the Second World War, which is also the time when the term Radar (short for *Radio Detection and Ranging*) was coined by the US Navy. Although originally only meant for target detection and early warning, today radar is being used in a countless number of applications, such as acquisition, detection, height finding, homing, mapping, navigation, ranging, remote sensing, reconnaissance, search, speed measuring, surveillance, terrain avoidance, terrain following and tracking, to name but a few.

Many of the different radar applications have very disparate aims and design goals. A full treatment of the many applications and the types of radar systems in any detail would take many tomes. For reasons of space and brevity, only the most pertinent points relating to the maritime surveillance radar are discussed herein. The interested reader is referred to the excellent text by [Skolnik, 1980] for a more complete discussion of radar. Other noteworthy texts include [Cole, 1992] for an introductory treatment on radar, [Kingsley and Quegan, 1992] for a more comprehensive overview of radar and its applications and [Eaves and Reedy, 1987] for a comprehensive and practical discussion of radar systems. [Levanon, 1988] provides a lucid, mathematical overview of principles of radar signal processing and detection.

The aim of a surveillance radar is to detect targets by illuminating an area of interest with electromagnetic radiation. In a marine environment, the term *target* usually refers to any radar backscatter, which itself is not due to the scattering from an ocean surface (which is referred to as *sea clutter*). Any man-made (*e.g.* a ship, a periscope, an airplane, *etc.*) or natural (*e.g.* an iceberg) object may constitute a target.

Most surveillance radars operate in searchlight mode, whereby a radial beam of energy is transmitted and received by a rotating antenna, much the same as the beam of beacon light on a lighthouse. The backscattered energy is received and processed by the radar to detect any potential targets before being displayed on the operator's visual display unit (see figure 1.1 for a

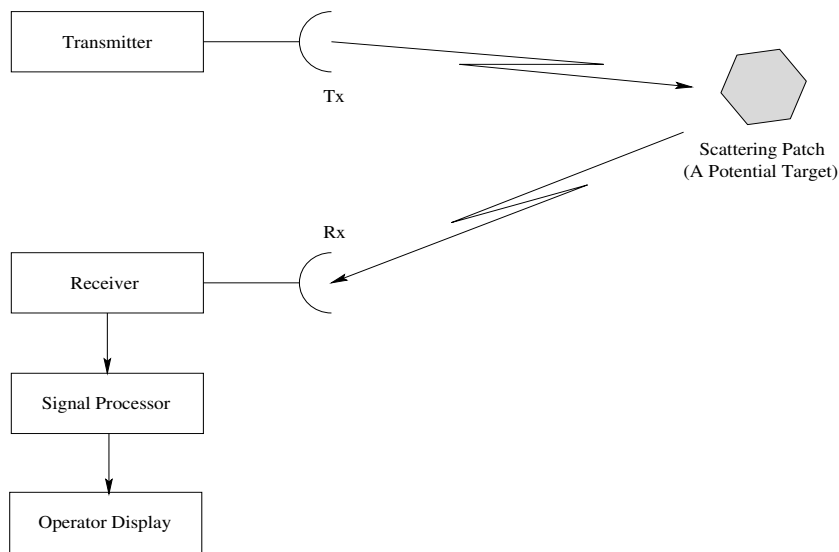


Figure 1.1: *Diagram of a Surveillance Radar System.*

schematic diagram of a surveillance radar system).

It has been found in practice that the operators are very sensitive to the number of alarms that they deal with and that these need to be kept within acceptable limits. It is therefore paramount that the radar's signal processing unit suppresses the number of *false alarms* to within an acceptable level, while keeping the number of *real target detections* to the maximum. Furthermore, in order to avoid target masking effects due to false alarm bunching, the average probability of false alarm should be kept as spatially uniform as possible.

In order to design 'optimal' detectors for a particular radar system operating in a marine environment, the properties of the backscattered clutter need to be understood. The properties of the backscattered radiation, as received by the radar, depend on a number of factors. These include the type of the transmitter and the receiver, the mode of operation (coherent, polarimetric, frequency agile, *etc.*), environmental conditions (sea state, viewing angle, extraneous sources of interference, presence of targets, *etc.*), range, resolution, sampling frequency, multi-path effects, *etc.*

The properties of sea clutter are discussed at length in chapter 2. The purpose of the following discussion is to give a flavour of typical constraints and trade-offs that have direct ramifications on the design of the signal processing unit of a maritime surveillance radar system.

1.1.1 Practical Limitations of a Surveillance Radar System

The most fundamental limit on the maximum range of a radar system is given by the *Radar Equation*. This equation, of the form

$$R_{max} = \left[\frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 k T B F_n L (SNR_o)_{min}} \right]^{1/4} \quad (1.1)$$

can be derived using some simple geometrical scattering arguments. The radar equation relates the maximum range of the radar, R_{max} , to the minimum detectable signal to noise ratio, $(SNR_o)_{min}$, at the output of the radar receiver. This signal to noise ratio corresponds to the signal reflected from an illuminated patch (*i.e.* including the sea clutter contribution, which for the purpose of the derivation of this equation is considered a part of the target returns), while the noise refers to the thermal noise, kT , associated with the antenna and the receiver. The effective bandwidth of the receiver is B , while the receiver noise factor is F_n . The radar equation is derived for a radar system operating at wavelength λ , transmitting a pulse of power P_t , with the antenna gain G , and system losses represented by L . The electromagnetic energy is considered to be backscattered from an illuminated patch with the effective radar scattering cross-section (RCS), σ .

The above limit is derived based on the minimum signal to noise ratio imposed by the transmission system, which is the SNR corresponding to thermal noise. From a signal processing (*i.e.* target detection) point of view, the situation is further exacerbated by the presence of sea clutter, which at high resolutions can have target like behaviour. Such issues are dealt with in the following chapter. For the time being, however, it is worth keeping in mind the distinction between thermal noise and sea clutter returns. While the former plays a prominent role at the transmitter/detector design stage, the latter becomes the foremost concern in the design and evaluation of optimal target detection schemes.

Referring to the radar equation (equation 1.1), it can be seen that in order to maximise the operating range of the radar, the designer can increase the transmitted pulse power, P_t , increase the radar wavelength, λ , increase the antenna gain, G , decrease the antenna and the receiver temperature, T , decrease the receiver noise factor, F_n , or decrease the receiver bandwidth, B . While T , F_n , G and λ are largely determined by the available hardware and the system specifications, the transmitted pulse power and the receiver bandwidth can be used to determine the radar's operating mode and resolution.

In order to keep the radar system cost low, the peak transmitted power should be kept as low as possible. This can be achieved with high duty cycle radars, such as the continuous wave (CW) radar. Unfortunately, although the maximum range of a CW radar can be large, such radar is inherently incapable of resolving targets in range. Another problem associated with high

duty cycle radars sharing a receiver and a transmitter antenna, is the masking of the returned energy by leakage from the transmitter. This results in blind regions, where the radar is unable to detect any potential targets.

The in-range resolution

The in-range resolution, δ_R , of a surveillance radar is determined by the transmitted pulse width, and is given by

$$\delta_R = \frac{c}{2B}$$

where $B = 1/T_P$ is the bandwidth associated with a pulse of width T_P , and c is the speed of light. Therefore, high resolution can in principle be achieved by employing narrow width pulses in a pulsed radar. Unfortunately, narrow pulses are associated with lower duty cycle, and hence higher peak to average power ratio is required for the transmitter. Furthermore, narrow pulse width requires larger bandwidth at the receiver, resulting in increased thermal noise, and therefore reduced effective range of the radar.

In order to overcome the difficulties associated with short duration pulses whilst maintaining the high range resolution of the radar, pulse compression techniques can be employed. By employing pulse compression, higher duty cycles and hence lower peak power pulses need to be transmitted, thus effectively increasing the useful range and decreasing the overall system cost of the radar. Pulse compression is achieved by frequency or phase modulating the transmitted signal to achieve a larger effective bandwidth, B , from which a pulse with effective duration $\tau = 1/B$ can be resynthesized using a matched filter. For some radar systems, the in-range resolution using pulse compression can be increased to several centimeters.

Pulsed radar systems (including pulse compression radars) can suffer from range ambiguities. Range ambiguity can arise as a result of too high pulse repetition frequency (PRF). In particular, if the time between pulses is shorter than the return time from the maximum range, then as is shown in figure 1.2, ambiguities in range arise. Furthermore, the matched filter relies on frequency (or phase) modulation of the transmitted waveform in order to determine the range. Unfortunately, the backscattered signal may be Doppler shifted due to the target or background clutter motion. Ambiguities in range and Doppler can thus arise at the matched filter stage. Such ambiguities can be optimised by a judicious choice of PRF and the coding scheme used for pulse compression.

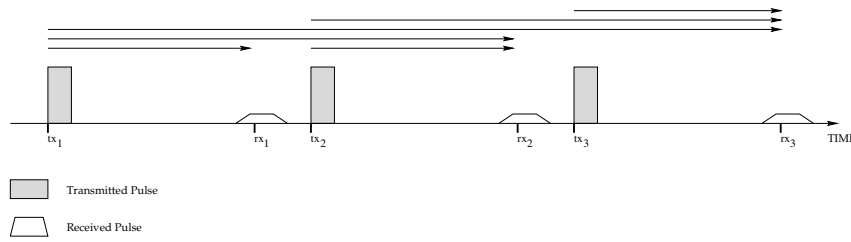


Figure 1.2: Range ambiguity in pulsed radar. The received pulse at time rx_2 could either be due to the pulse transmitted at time tx_1 or tx_2 . Likewise the received pulse at time rx_3 could either be due to the pulse transmitted at tx_1 , tx_2 or tx_3 .

The azimuth resolution

The width of the transmitted beam determines the azimuth resolution of the radar. In particular, for a radar beamwidth ϕ , the resolution at range R is given by

$$\delta_\phi = \phi R$$

Clearly, the azimuth resolution is range dependent. The beamwidth is determined by the radiation pattern of the antenna, and typically can range anywhere from 1° to 10° , which at 10km range translates into the across range resolution of 175m and 1.75km, respectively.

The effective radar aperture can be synthetically increased in the case of airborne radar by Doppler processing the echoes obtained by an aircraft sweeping the radar beam across an illuminated patch, while flying orthogonal to the line of sight of the target. Synthetic Aperture Radar (SAR) [Brown, 1967] [Brown and Porcello, 1969] can increase the across-range resolution to the same order of magnitude as the resolution achievable in-range using pulse compression techniques. However, the signal processing techniques required for target detection for SAR are different to those required for a surveillance radar operating in the searchlight mode. Therefore, SAR will not be considered further in this thesis.

1.1.2 The Range Profiles

As was discussed above, the azimuth resolution of a surveillance radar operating in searchlight mode is limited by the beamwidth. The beam is typically rotated at a fraction of a revolution per second, while the pulse repetition frequency is of the order of a few kHz (limited by the maximum required unambiguous range). The downconverted waveform at the output of the receiver is then sampled at a frequency which is limited by the bandwidth of the receiver (corresponding to the effective pulse width).

The sampled data at the output of the receiver is then passed into a matched pulse decompression filter in order to obtain the returns at any particular range. The resulting snapshot in time of the returns from the different ranges is referred to as a *range profile*. The returns at different ranges within each profile are referred to as *range cells*, and these can contain returns from any potential targets, background clutter and/or noise. The surveillance radars considered herein are considered to be operating in open ocean conditions. Land clutter and littoral effects [Branson, 1998] add an additional level of complexity and are not dealt with in this thesis.

The radar backscatter at any particular range can undergo Doppler shifts, f_D , due to a moving target or the background clutter. Such Doppler shift is manifested in the change of frequency between the transmitted, f_T and the received waveform, f_R , and is given by

$$f_D = f_R - f_T = -\frac{2\dot{R}}{\lambda}$$

where \dot{R} is the radial velocity component of the scattering centres within the illuminated patch, while the wavelength, λ , for an I-band radar (9.5-10GHz) is approximately 3cm. As will be discussed in chapter 3, such Doppler shifts can form the basis for target detection in the presence of high resolution sea clutter in coherent radar systems.

In a surveillance radar operating in searchlight mode, the range profiles are formed on the pulse by pulse basis. As the radar beam moves relatively slowly in azimuth, the successive range profiles can be considered to correspond to the same illuminated area of sea over a number of pulses. This fact can be used as the basis for analogue or binary integration, or for extracting Doppler spectra of the returns obtained from individual range cells at different ranges.

It is important to note that the statistics of the clutter returns can be heterogenous in range within a single range profile (*i.e.* correlation structure and spikiness of the returns may change within a single profile). This is particularly pronounced in airborne surveillance radar, where due to the height of the radar platform, the range of the radar and hence the number of range cells is large (as large as a few tens of thousands of range cells per range profile). The range cells within each profile correspond to a wide span of grazing angles and azimuth resolutions, which themselves are range dependent.

The large number of range cells within a single profile, and a large number of profiles per antenna scan, result in severe limitations on the complexity of the algorithms that can successfully be implemented for target detection in real time. For reasons of system cost and complexity, the memory available for storage of the data is very limited, hence only one profile at a time is normally processed. This can either be the raw single pulse profile, or a profile in some way averaged over a number of pulses. These issues will be addressed in more detail in the following chapters.

It is worth mentioning that a number of maritime surveillance radars are also capable of operating in polarimetric mode. In the polarimetric mode, the transmitted beam can either be horizontally (H) or vertically (V) polarised, while the receiver antenna can be tuned to receive the horizontally or vertically polarised backscatter. The full polarimetric scattering matrix, consisting of HH, HV, VH and VV components can thus be obtained and used for enhanced target detection in the presence of sea clutter.

Finally, coherent radar systems require a very stable transmitter and phase reference for the receiver, thus increasing the system cost. For this reason, a number of maritime surveillance radar systems are incoherent, in that no phase information is available at the output of the receiver. The range profiles obtained for such radars consist only of amplitude returns from different ranges, and hence no Doppler based processing is possible.

1.2 OUTLINE OF THE DISSERTATION

As was discussed in the previous section, the ultimate aim of the signal processing unit of a maritime surveillance radar system is to process the received and sampled data in such a way, so as to minimise the number of spurious and false alarms to within some acceptable limit, while keeping the number of true target detections to the maximum. To this end, the research undertaken as part of this thesis is concerned with the study and implementation of target detection schemes in the presence of sea clutter. The remainder of the dissertation is structured thus:

Chapter 2 presents a review of sea clutter modelling. The scope of the review encompasses a discussion on the properties of sea clutter under a range of operating and viewing conditions, as well as statistical and deterministic modelling of sea clutter.

Chapter 3 presents a review of target detection techniques in the presence of sea clutter. The scope of the review encompasses a discussion on fixed and adaptive threshold detection, optimal Neyman-Pearson detection, as well as Doppler based detection techniques.

Chapter 4 deals with parametric characterisation of high resolution sea clutter spectra in incoherent radar systems. In particular, a complex autoregressive process is proposed as the basis for modelling of the correlation structure of the incoherent clutter. Any Doppler information contained in the incoherent radar returns is expected to be manifested through the position of the poles of the underlying complex autoregressive process, of which only the amplitudes are observed. As no phase information is available in incoherent radar returns, the Gibbs sampler is used to facilitate sampling from the autoregressive process parameter posterior distribution,

conditional on the observed amplitudes. To this end, the Hybrid Monte Carlo algorithm is employed to conditionally sample for the missing phases [Noga and Fitzgerald, 1997].

Chapter 5 deals with a novel class of heteroscedastic (*i.e.* non-constant prediction error variance) models for the modulating component of sea clutter [Noga and Fitzgerald, 1998a]. Such models are motivated by a discrete generalisation of the birth-death migration process for the evolution of a population of scattering centres on the sea surface. Due to the non-constant, state dependent variance of the prediction errors, the conditional heteroscedastic models can potentially provide the basis for CFAR detectors with more spatially uniform probability of false alarm. However, as is shown in this chapter, the evidence for heteroscedasticity in sea clutter is not very compelling. Instead, contrary to the widely held belief, rather than being Gamma distributed, the modulating component of high resolution sea clutter appears to be better modelled as being log-Normal distributed.

Chapter 6 deals with constant false alarm rate (CFAR) target detection in the presence of high resolution sea clutter. To this end, a linearised state space model for sea clutter in the logarithm domain is proposed [Noga and Fitzgerald, 1998b]. The CFAR detection threshold is then set, based on the MAP estimate of the modulating component in the cell under test, conditional on the radar returns in the reference window. It is shown that, in conjunction with pulse integration techniques, the performance of the proposed CFAR detection scheme approaches that of the ideal CFAR detector, particularly for spiky and highly correlated clutter.

Chapter 7 presents a summary and some conclusions of the research contained in this dissertation. The conclusions are further supplemented by suggestions for potentially fruitful and interesting avenues of future research.

2.1 PROPERTIES OF SEA CLUTTER

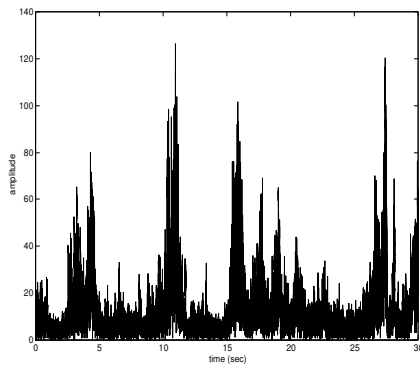
Radar backscatter from an ocean surface is commonly referred to as the *sea clutter*. The backscatter is the result of scattering of the electromagnetic energy from the capillary waves on the surface of the sea, which are in turn modulated by the swell. The first component, that due to the capillary waves, is referred to as the *speckle* and has experimentally been shown to decorrelate on the scale of a few milliseconds. The second component, that due to the swell, decorrelates on a much longer scale.

The sea clutter data used in this thesis has been obtained using an I-band (9.5-10GHz) instrumentation radar [Hirst and Baker, 1987], capable of simultaneous coherent and cross polarisation measurements. The system employs pulsed waveforms with 50MHz FM modulation within the pulse, which gives a range resolution of approximately 4m [Ward et al., 1990]. The pulse repetition frequency used is 1kHz, and pulse-to-pulse frequency agility is often used, where the transmitted frequency is stepped by the pulse bandwidth from pulse to pulse over the 500MHz radar bandwidth. The system is also capable of measuring the full coherent scattering matrix at a 500 Hz repetition frequency. An antenna giving an 1.2° beamwidth is used and, generally, measurements are made 'search-lighting' an area of sea. Range profiles are recorded, thereby allowing both the temporal and spatial correlation properties of the backscatter to be determined simultaneously.

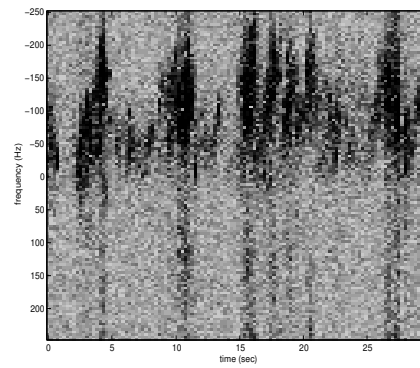
In the coherent mode, the radar operates coherently from pulse to pulse, and both the in-phase and quadrature components are recorded. In the polarimetric mode, pulse-to-pulse switching between orthogonal polarisation on the transmitter and a two-channel receiver is used to enable the recording of the full coherent polarisation scattering matrix. The effective pulse repetition frequency for each polarisation channel is 500Hz.

2.1.1 High Resolution Sea Clutter

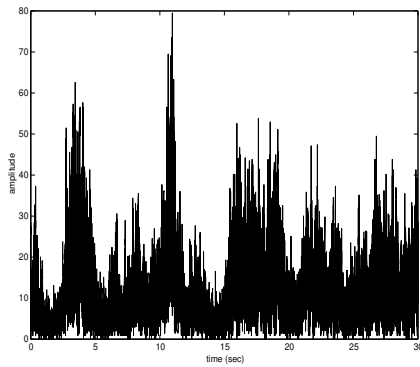
Thirty seconds of pulse by pulse time history of horizontally and vertically polarised amplitude returns obtained using a coherent radar operating in polarimetric mode are shown in figure



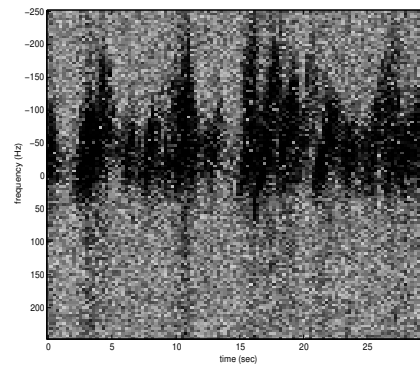
(a) Pulse by pulse amplitude (HH)



(b) Spectrogram (HH)



(c) Pulse by pulse amplitude (VV)



(d) Spectrogram (VV)

Figure 2.1: Pulse by pulse time history of the amplitude of (a) horizontally and (c) vertically polarised, coherent sea clutter from a single range cell, along with the corresponding spectrograms (see (b) and (d)) obtained using 128 point FFTs.

2.1, along with the corresponding spectrograms. The effective sampling frequency for each polarisation is 500Hz, and the spectrograms are obtained by performing a 128-point FFT on sequential windows of the coherent clutter. The pulse by pulse data was obtained in open ocean conditions at a look angle of 215.1° into the wind at 7.5km range. The resolution of each range cell is effectively 4m in range and 160m in azimuth.

The time history of amplitude of the coherent clutter returns from an individual range cell (figure 2.1(a) and 2.1(c)) clearly shows the two principal components of the high resolution sea clutter; namely the fast oscillating speckle component, modulated by the underlying component due to swell. The overall behaviour of the returns in the two polarisations is similar, although the detailed structure is different, indicating that different scatterers are contributing to the scattering process. Furthermore, characteristic spikiness of the horizontally polarised clutter compared with vertically polarised clutter is evident in figure 2.1.

By comparing fixed frequency and frequency agile records, [Ward et al., 1990] found that the fast fluctuation decorrelates from pulse to pulse with frequency agility and is correlated for between 5 and 10ms with fixed frequency. The underlying, modulating structure is unaffected by frequency agility and generally fluctuates on the time scale of the order of seconds. As [Ward et al., 1990] point out, the temporal and spatial correlation of the modulating component is very dependent on the radar parameters, the viewing angles and the environmental conditions.

Spectrograms of the coherent clutter returns from an individual range cell (figure 2.1(b) and 2.1(d)) show the evolution of the Doppler spectrum of the coherent clutter, for both the horizontally and vertically polarised returns. It is evident that the changes in the Doppler spectra for both polarisations are very similar, and that they correspond to the underlying modulating process due to the swell. This is indeed in agreement with expectations, in that the instantaneous velocity of the wave front changes as the swell rises and falls periodically, thus resulting in changes in the Doppler spectra, as well as in the intensity of the returns.

Although the Doppler information in coherent radars can be used to design powerful detectors [Nohara and Haykin, 1993], a large proportion of radars use only the envelope of the received signal in their processing [Ward et al., 1990]. As the phase of the received signal is not utilised, the corresponding radar systems can be incoherent from pulse to pulse, and can therefore use simple transmitters, such as pulsed magnetrons. However, particularly at high resolutions and at low grazing angles, the amplitude of the clutter returns has frequently been found to be 'spiky'. That is, the probability of the signal crossing a threshold is much higher than would be expected for a simple Rayleigh distributed clutter for the given clutter power. This is illustrated in figure 2.2, where the histograms of the amplitudes of a hundred thousand samples of clutter returns from an individual range cell are compared to a Rayleigh distribution for which the parameters were estimated by the method of Maximum Likelihood (ML). It can be seen from the figure that the high resolution sea clutter data has much heavier tails than the Rayleigh distribution. This is particularly pronounced in the case of horizontally polarised clutter.

By plotting cumulative pulse by pulse distributions of the amplitude returns from individual range cells taken over 240ms on Weibull paper, [Ward et al., 1990] show that the amplitude of sea clutter is *locally* Rayleigh distributed. This is an important observation, which suggests that the non-Rayleigh nature of the amplitude distribution of high resolution clutter is associated with bunching of scatterers by the sea wave structure, rather than being due to a small number of effective scatterers. This observation forms the basis for compound models of sea clutter.

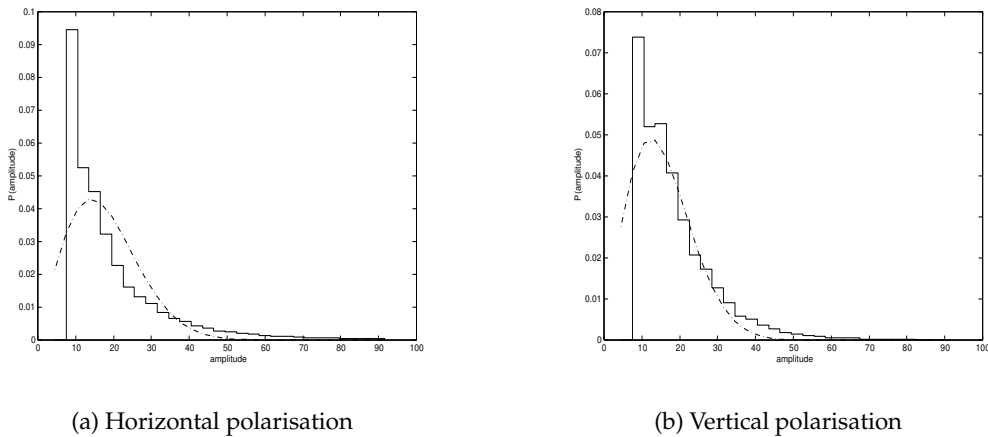


Figure 2.2: Histograms of (a) horizontally and (b) vertically polarised clutter amplitude returns from an individual range cell, along with the corresponding ML Rayleigh distribution (— · —).

2.1.2 The Modulating Component

The modulating component associated with the swell can be obtained by averaging the pulse by pulse amplitude data over relatively short intervals (typically, a few hundred milliseconds). This ensures that the speckle component, which has the decorrelation time of the order of 10ms, is removed. The range-time amplitude profiles for horizontally and vertically polarised clutter are shown in figure 2.3. As can be seen from the figure, the in-range resolution of the range cells for a high resolution radar is smaller than the wavelength of the underlying swell.

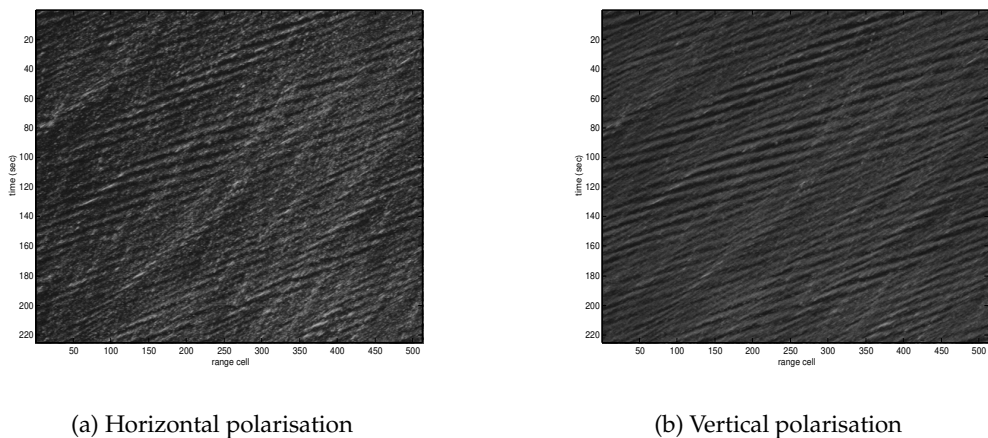


Figure 2.3: 220 seconds of range-time amplitude profiles for (a) horizontally and (b) vertically polarised coherent clutter, integrated over 250 pulses.

The swell structure is clearly visible in both the horizontal and vertical polarisation. However, as [Ward et al., 1990] point out, the actual observed swell structure in the range time profiles depends on radar polarisation, sea state and the grazing angle. The observed swell structure

also depends on the viewing aspect. In particular, looking down the swell, a slow and well defined wave pattern is observed. On the other hand, looking across the swell direction breaks up the wave pattern. This can easily be explained in terms of the radar resolution, which for the high resolution radar considered here is 4m in range, and 160m in azimuth at 7.5km. Looking into the swell resolves structure greater than 4m. On the other hand, looking across the swell only resolves structure greater than 160m, and that is not in range, but rather in time as the swell moves through the beam. Hence the wave-like pattern of the into-the-swell view is not observed.

The range-time profiles in figure 2.3 illustrate that the modulating component in any one range cell is not independent from its neighbors. The spatial correlation properties of the modulating component turn out to be highly resolution dependent and need to be taken into account in the context of target detection in the presence of sea clutter [Watts and Ward, 1987]. [Ward et al., 1990] synthesize lower resolution returns from high resolution radar clutter, with the effect of the short wavelength components getting averaged out, leaving progressively longer wavelengths. In the limit of low resolution radar, the modulating component becomes constant, and the radar returns become Rayleigh distributed.

By matching the first four moments of the modulating component obtained over many airborne sea clutter measurements, [Ward et al., 1990] demonstrate that the Gamma distribution provides a superior fit to the sea clutter data over the log-Normal distribution, for both horizontally and vertically polarised clutter. These results provide an experimental basis for the compound K-distribution (*i.e.* Rayleigh distributed speckle modulated in power by a Gamma distribution) as a statistical model of sea clutter. The significance of this model is that the non-Rayleigh amplitude distribution of sea clutter may be separated into two components which, because of their spatial and temporal correlation properties, need to be introduced at different stages in the evaluation of signal processing performance.

2.2 DETERMINISTIC MODELS FOR SEA CLUTTER

For a number of decades, sea clutter has been modelled as a stochastic process arising from random scattering from the sea surface. It is widely accepted that at low resolutions the amplitude of the backscatter is Rayleigh distributed, while at higher resolutions the clutter becomes increasingly spiky. The Gaussian scattering assumption has been found to be no longer adequate to characterise the clutter at high resolutions, and other distributions, such as the K-distribution have since been proposed.

Although the K-distribution has semi-empirical origin, based on a random walk modulated by a birth-death migration process [Jakeman and Tough, 1988], some authors have argued that sea clutter may actually be deterministic in origin [Leung and Haykin, 1990]. The evidence for this is compelling, and some serious consideration needs to be given to the possibility and utility of modelling sea clutter as a deterministic process.

In theory, given the angle of incidence of the radar wave and the geometry of the ocean surface, it should be possible to predict the exact trajectory of the backscattered wave. But for a fixed surface geometry, a slight deviation in the angle of incidence may produce a large change in the backscattering process. This very sensitivity of the scattering process to the initial conditions is suggestive of the possibility of sea clutter being chaotic scattering phenomenon [Leung and Lo, 1993].

According to the theory of non-linear dynamics, a very low number of degrees of freedom is sufficient to create highly irregular 'dynamic chaos' [Leung and Haykin, 1990]. The resulting process may appear to the naked eye as being random in origin. In fact, random process theory is an empirical technique for coping with inadequate information about the physical sources responsible for generation of the processes. Therefore, if the sea clutter truly is deterministic in origin, then stochastic modelling unnecessarily assumes too many degrees of freedom and cannot fully extract the information contained in the experimental data.

It is not the purpose of this section to fully investigate the utility of using non-linear dynamics for the purpose of sea clutter modelling, but rather to provide a critique of the work presented on the subject in the literature. Most notably, based on a number of chaotic system invariants, [Haykin and Li, 1995] advocates the use of non-linear models for sea clutter modelling. However, as [Davies, 1998] points out, such invariants cannot be used to diagnose chaos by themselves, and their use makes sense only in the context of the data being deterministic to start with. This point will be illustrated in the present section by demonstrating that the dynamical system invariants as used by [Haykin and Li, 1995] can falsely diagnose a purely stochastic and linear process as being chaotic. Also, [Haykin, 1996] claim that the detection performance of an incoherent radar can approach that of a coherent radar by using non-linear dynamics. Such claims rely on the pulse by pulse returns being chaotic. Some plausible reasons for the potential improvement in the detection performance are given in section 2.2.3, while more in-depth discussion on this subject is presented in chapter 4.

2.2.1 Dynamical System Reconstruction

The dynamical concepts introduced in this section are very basic and are kept simple for the purpose of this exposition. For a more complete and in-depth review of non-linear dynamics and chaos theory, the interested reader is referred to the papers by [Parker and Chua, 1987] and [Haykin and Li, 1995].

Let $x_0(t)$ represent the time series associated with backscatter returns from a single range cell. In the case of incoherent radar $x_0(t)$ are the amplitudes of the returns, while for coherent radar $x_0(t)$ are the complex (*i.e.* in-phase and quadrature) samples of the backscatter. The dynamics of the electromagnetic wave scattering are then described by the set of N variables $\mathbf{x}(t) = \{x_0(t), x_1(t), \dots, x_{N-1}(t)\}$ evolving in an N dimensional space according to the system of first order differential equations, written in terms of the state vector $\mathbf{x}(t)$ as

$$\frac{d}{dt}\mathbf{x}(t) = \phi(\mathbf{x}(t))$$

where the vector function ϕ is, in general, non-linear. Note that only a low dimensional subspace of the N dimensional space is observed in the form of the clutter $x_0(t)$.

An important issue associated with the study of non-linear dynamical systems is that of *reconstruction of dynamics*, *i.e.* the ability to reconstruct the underlying dynamics of a non-linear dynamical system, that is known to be chaotic, from a time series made up of physical observations. A theorem independently proposed by Takens and Packard, known as the *Takens Embedding Theorem* provides a mathematical basis for the dynamic reconstruction problem. It effectively states that model reconstruction of a non-linear dynamical system from observations of *just a single* dimensional subspace of the system should succeed to a certain extent, and the reconstruction is independent of which signal component is used. Because of unavoidable imprecision in building the model and noise contained in the observed data, such model reconstruction cannot be expected to produce a model exactly the same as the original system. However, the reconstruction will have the same behaviour as the original system up to a diffeomorphism (*i.e.* to within a smooth curvilinear change of coordinates) [Haykin and Li, 1995].

The potential implications of the Takens Embedding Theorem for radar are profound. What the theorem implies is that it does not make a difference whether only the amplitudes, or the complex, coherent data is used for the reconstruction of the underlying system dynamics. In other words, if sea clutter is indeed chaotic, the performance of the incoherent radar should approach that of coherent radar, since the dynamics of the underlying system should in theory be reconstructed equally well from both. In fact, results obtained for complex data in [Haykin and Li, 1995] have been found to give results similar to those obtained from the amplitudes alone.

Without undue mathematical rigor, the idea behind applying the Takens Embedding Theorem to dynamical system reconstruction from a single dimensional time series can be explained as follows. Consider an N -dimensional chaotic dynamical process described by the difference equation

$$\mathbf{x}(k+1) = \phi(\mathbf{x}(k))$$

where ϕ is a non-linear function. Given an initial state $\mathbf{x}(0)$, the system trajectory is a sequence of points $\{\mathbf{x}(k) = \phi^k(\mathbf{x}(0)), k = 1, 2, \dots\}$. Suppose now that a sequence of measurements $\{y(t) = f(\mathbf{x})(t), t = 0, 1, \dots\}$ is made, where f is a smooth observation map of the dynamical system. It can then be shown (see [Haykin and Li, 1995]) from the Takens Embedding Theorem that with an *embedding dimension* $D_E \geq 2N + 1$, there exists a function F such that

$$y(t + D_E\tau) = F(y(t), y(t + \tau), \dots, y(t + (D_E - 1)\tau)) \quad (2.1)$$

where F is a non-linear function that predicts the next entry in the observed time series and τ represents the delay used for regression. Therefore, in principle, the reconstructed phase space can be approximated by using a time series observed from the original system. Unfortunately, Takens' Embedding Theorem merely asserts the existence of such a function, but does not specify its form.

Therefore, not only can the attractor of the original N -dimensional dynamical system be embedded in a D_E -dimensional space reconstructed from delayed observations,

$$\Phi(t) = \{y(t), y(t + \tau), \dots, y(t + (D_E - 1)\tau)\}$$

but also the observed time series $y(t)$ is a deterministic process evolving according to equation 2.1. However, although based on Takens' Embedding Theorem, the framework for the dynamical system reconstruction has been firmly established, the choice of the embedding dimension D_E , delay τ and the form of the non-linear function F has not yet been addressed. These issues are dealt with in the following sections.

2.2.2 Attractor Dimension and Lyapunov Characteristic Exponent

There are a number of invariants used in studying chaotic systems. One such invariant is the attractor dimension. There are also many kinds of attractor dimension, but these are all defined using a somewhat similar concept. The basic idea of an attractor dimension can be explained by considering a point in the embedding phase space, \mathbf{x} , and a d -dimensional sphere of radius r centered at \mathbf{x} , $B_r(\mathbf{x})$. The attractor dimension measures the rate at which the number of points in $B_r(\mathbf{x})$ decreases as r is reduced [Leung, 1995].

Another dynamical system invariant that is widely used in chaos theory, is the Lyapunov exponent [Leung and Lo, 1993]. A Lyapunov exponent qualitatively shows the sensitive dependence on initial conditions, and also gives a quantitative measure of the average rate of separation or attraction of nearby trajectories on the attractor.

Both of the above mentioned invariants have been used in the literature to demonstrate that

sea clutter can indeed be considered to be chaotic [Haykin and Li, 1995]. However, as argued by [Davies, 1998], such invariants ought not be used when testing for chaos, but rather ought to be used solely to extract information about the chaotic system. It is the purpose of the present section to illustrate that application of the above mentioned dynamical system invariants for the purpose of the detection of chaos can lead to misleading conclusions.

In order to illustrate this claim, ten thousand samples (*i.e.* 20 seconds) of horizontally polarised sea clutter data shown in figure 2.1(a) are used. In addition, ten thousand samples of synthetic clutter data are generated, such that the spectrum of the synthetic data matched the spectrum of the first 128 samples of the real clutter data (only the spectrum of the first 128 samples was used for reasons of clutter stationarity). The amplitude of the synthetically generated data, being a complex Gaussian process, is Rayleigh distributed, with the correlation structure matching that of the original sea clutter. It is of interest to compare the attractor dimension and the Lyapunov exponent of the real clutter and the synthetically generated data. As will soon be demonstrated, large similarities exist between the two sets of results, casting serious doubt about the suitability of using the dynamical system invariants for the purpose of diagnosing chaos.

Only the time series of amplitudes for both the real clutter and the complex Gaussian synthetic data are used. In order to reconstruct the D_E dimensional phase portrait as discussed above, the delay τ first needs to be determined. The choice of τ needs due care. If τ is too small, then $y(t)$ and $y(t + \tau)$ are too close, and hence not uncorrelated enough to serve as independent coordinates. If, on the other hand, τ is too large, chaos makes $y(t)$ and $y(t + (D_E - 1)\tau)$ disconnected. There are a number of different criteria for choosing τ presented in the literature. [Leung, 1992] propose to choose τ corresponding to the first zero crossing of the autocorrelation function of $y(t)$. Alternatively, τ can be chosen on the basis of some mutual information measure [Haykin and Li, 1995]. Based on the autocorrelation function of the data, τ was chosen to be approximately equal to 5 for the real clutter data, while it was only 2 for the synthetically generated data. The difference stems from the fact that, due to stationarity concerns, the synthetic data was generated to match the autocorrelation structure of only the first 128 samples of the real data.

Having selected τ , the embedding phase space consists of the collection of D_E dimensional vectors $\Phi(t) = \{y(t), y(t + \tau), \dots, y(t + (D_E - 1)\tau)\}$. The embedding dimension D_E can in principle be estimated from the *correlation dimension*, which is one of a number of different measures that can be used to estimate the attractor dimension [Haykin and Li, 1995].

In order to evaluate the correlation dimension, the cumulative correlation

$$C(r) = \frac{1}{T^2} \sum_{k,j=1; k \neq j}^T H(r - |\Phi(k) - \Phi(j)|)$$

needs to be evaluated over a range of values of r , for all pairs of T samples of the data. The Heaviside function H is defined as

$$H(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{otherwise} \end{cases}$$

and the norm $|\Phi(k) - \Phi(j)|$ is arbitrary, though the standard Euclidean norm was used in the following. The correlation dimension is then defined by

$$D_C = \lim_{r \rightarrow 0} \frac{\log(C(r))}{\log r}$$

The logarithm of the cumulative correlation as a function of the logarithm of r for $T = 10^4$ samples of real sea clutter data is shown in figure 2.4(a), while the results for the synthetic data are shown in figure 2.5(a). The results were obtained for different values of the embedding dimension D_E , ranging from 5 to 50. The region in which the power law is obeyed by $\log(C(r))$ vs $\log r$ appears as a straight line, and the slope, which is used as an estimate of the correlation dimension, is found by fitting a least squares line to this part of the graph. The attractor dimension is estimated from the resulting correlation dimension D_C , provided the sequence of estimates converges for increasing values of the embedding dimension D_E .

The estimates of the correlation dimension obtained for a range of values of embedding dimension D_E are shown in figures 2.4(b) and 2.5(b) for the sea clutter data and the synthetic data, respectively. It can be seen that for real clutter the correlation dimension saturates to give an estimate of the attractor dimension $D_C \approx 10$, corresponding to the embedding dimension $D_E \approx 40$. Note, however, that the correlation dimension also saturates for the synthetically generated *stochastic* data, to give an estimate of the attractor dimension of $D_C \approx 9$, corresponding to the embedding dimension of $D_E \approx 25$.

Clearly, no conclusion can be drawn about sea clutter being chaotic from the finite correlation dimension alone. Indeed, [Leung and Lo, 1993] acknowledge that the statistical colored random process may exhibit saturation in the above analysis, although white Gaussian noise is expected not to produce such effects.

Another test used to establish whether sea clutter is chaotic is the Lyapunov exponent, which measures the sensitivity of the system to the initial conditions. In order to estimate the largest Lyapunov exponent from the time series data, an algorithm presented in [Haykin and Li, 1995] was used. The algorithm amounts to calculating the log ratio of distances between two closest phase space points one sample period apart. If the distance between the points increases, the points diverge, indicating a positive Lyapunov exponent. The formula used for the estimation

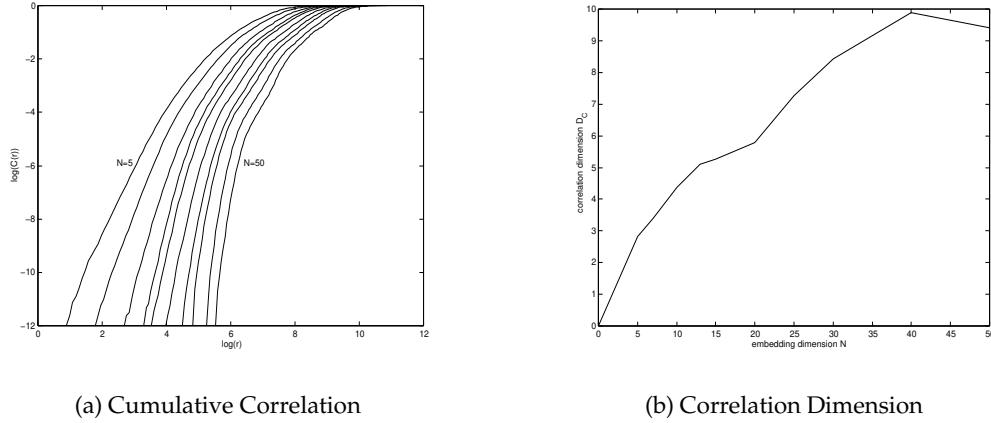


Figure 2.4: (a) Cumulative Correlation as a function of the logarithm of r and (b) Correlation Dimension (D_C) as a function of the embedding dimension ($N \equiv D_E$) for 10^4 samples of horizontally polarised amplitude data.

of the largest Lyapunov exponent is

$$\lambda = \frac{1}{K} \sum_{k=1}^K \log \frac{|\Phi(k+1) - \Phi(k'+1)|}{|\Phi(k) - \Phi(k')|}$$

where k' is the index of the point nearest $\Phi(k)$ in the phase space, and where the averaging is done for K consecutive phase space points. It was found that for both the real sea clutter data and the synthetic data, the largest Lyapunov exponent is positive. Yet again, no certain conclusion can be reached as to whether the sea clutter is chaotic or not, based on the two invariants above.

2.2.3 Non-linear Evolution Function

In the previous section some doubts were cast as to the validity of the claim of sea clutter being chaotic. Yet the results presented by a number of authors ([Haykin, 1996], [Haykin and Li, 1995], [Leung and Lo, 1993]) seem to support the notion that by using dynamical system models, detection performance of incoherent radar clutter approaches that of coherent radars. The purpose of this section is to briefly suggest some reasons for this apparent paradox.

The results presented in [Haykin, 1996], [Haykin and Li, 1995] and [Leung and Lo, 1993] have one thing in common. They all use neural networks to model the non-linear evolution function F . The embedding dimension D_E is determined from the correlation dimension estimate, while the delay τ is trivially determined either from the autocorrelation function, or some other mutual information measure. The main rationale for using neural networks to model F is sim-

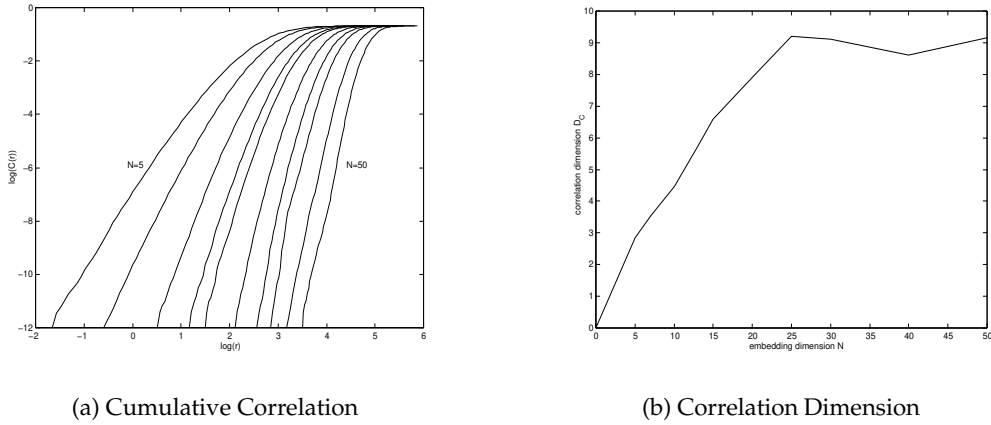


Figure 2.5: (a) Cumulative Correlation as a function of the logarithm of r and (b) Correlation Dimension (D_C) as a function of the embedding dimension ($N \equiv D_E$) for 10^4 samples of synthetically generated amplitudes of complex data.

ple. Since the form of the non-linearity F is unknown, a flexible class of non-linear functions capable of capturing wide ranging set of dynamics is required, and it is widely accepted that neural networks are ideally suited to such situations. The potential difficulty with neural networks in practice is that they tend to be rather slow to train, and may not be particularly useful for on-line applications. However, it is not the purpose of this thesis to discuss the merits of neural networks or otherwise.

Consider the clutter to be a complex Gaussian process, *i.e.* resulting from random scattering off a rough surface, where the correlations are associated with effects such as Doppler shifts due to wave motion. Making a simplifying assumption that the scattering process is stationary (*i.e.* Doppler shifts are constant) and ignoring the modulating process due to the swell, the complex clutter $x(t) = [x_I(t), x_Q(t)]$ can be considered to be well modelled by a complex autoregressive process given by

$$x_I(t) = \sum_i^N [a_I(i)x_I(t-i) - a_Q(i)x_Q(t-i)] + \sigma e_I(t)$$

$$x_Q(t) = \sum_i^N [a_Q(i)x_I(t-i) + a_I(i)x_Q(t-i)] + \sigma e_Q(t)$$

where $[e_I(t), e_Q(t)]$ is a complex i.i.d. Gaussian innovation process. The additional complications associated with the swell and non-stationarity of the scattering process will only add to the complexity of the model, and for the purpose of exposition here, need not be considered.

As can be seen from the equations above, a simple linear model is optimal for the case of coherent clutter, where both the in-phase and quadrature components are observed. The amplitudes of the coherent process evolve according to

$$\begin{aligned}
 y(t) &= \sqrt{x_I(t)^2 + x_Q(t)^2} & (2.2) \\
 &= f(x_I(t-1), x_Q(t-1), \dots, x_I(t-N), x_Q(t-N), e_I(t), e_Q(t)) \\
 &\approx F(y(t-1), \dots, y(t-D_E), \nu(t))
 \end{aligned}$$

where f is well defined. For incoherent radar the in-phase and quadrature components are not directly observed, and the form of regression on previous observations is therefore not readily obvious. The observed amplitude at time t , $y(t)$, is clearly not independent of the previous observations $\{y(t-1), \dots, y(t-D_E)\}$, and is expected to be related to them through some *non-linear* function F , where the noise process $\nu(t)$ is related to $\{[e_I(t-k), e_Q(t-k)]; k = 1, \dots, D_E\}$ in some non-trivial way.

The results presented in the previous section indicate that the time series obtained from equation 2.2 has a finite embedding dimension, and a positive largest Lyapunov exponent. This can be very misleading, as it suggests that the time series may in fact be chaotic. However, although clearly not deterministic, from the discussion above it is apparent that non-linear regression is expected to perform substantially better than any linear regression model for $y(t)$. Exactly how much of the system dynamics can be captured using non-linear regression is not clear, but as results using neural networks in [Haykin and Li, 1995] indicate, the detection performance based on the amplitudes alone approaches that of coherent radar.

2.3 NON-GAUSSIAN MODELS FOR SEA CLUTTER

Whether or not sea clutter is chaotic still remains wide open to debate. As was shown above using a number of dynamical system invariants, correlated Gaussian speckle can misleadingly be diagnosed as being chaotic. At another level, there is always the possibility that it is the underlying modulating component of sea clutter that evolves according to some high dimensional non-linear system dynamics. The point remains that to date no *simple* deterministic model for sea clutter has been proposed.

In the light of the uncertainty about the actual mechanism responsible for generating sea clutter, the potential utility associated with the modelling of sea clutter as a non-linear dynamical process is greatly diminished by the apparent model complexity required. It is for this very pragmatic reason that a statistical framework for sea clutter modelling is adopted throughout this thesis and will be now discussed.

2.3.1 Single Point Statistics Modelling

The single point statistics of low resolution sea clutter amplitudes, r , have experimentally been found to be Rayleigh distributed, *i.e.*

$$p(r|\sigma) = \frac{2r}{\sigma} \exp(-r^2/\sigma) \quad (2.3)$$

while the intensity of the received clutter, $I = r^2$, is exponentially distributed, *i.e.*

$$p(I|\sigma) = \frac{1}{\sigma} \exp(-I/\sigma)$$

This observation can be readily justified by the Central Limit Theorem applied to Gaussian scattering from a large population of scattering centres within a patch illuminated by a low resolution surveillance radar. In fact, only Rayleigh distributed amplitudes are compatible with the Gaussian scattering hypothesis [Azzarelli, 1995]. The average power of the clutter returns, σ , is related to the radar cross section (RCS) of the scattering patch.

Similarly, assuming uncorrelated returns for an L -look radar (*i.e.* where the returns are integrated over L consecutive pulses), the distribution for the intensity of the integrated clutter, I_L , can be obtained by convolving the distributions of the individual returns [Levanon, 1988], and can be shown to be of the form [Oliver, 1993]

$$p(I_L|\sigma) = \frac{1}{I_L} \left(\frac{LI_L}{\sigma} \right)^L \frac{1}{\Gamma(L)} \exp\left(-\frac{LI_L}{\sigma}\right) \quad (2.4)$$

where $\Gamma(\cdot)$ is the Gamma function. The independent samples from pulse to pulse for an L -look radar can be obtained using techniques such as frequency agility.

At high resolutions and low grazing angles, the clutter has been found to be more 'spiky', in that the amplitude distribution exhibits a higher probability of the signal crossing the threshold than would be expected for a simple Rayleigh distributed clutter. To model such behaviour, heavier tailed two-parameter distributions, such as the log-Normal distribution

$$p(r|a, b) = \frac{1}{2r\sqrt{\pi \ln b}} \exp\left(-\frac{1}{4 \ln b} \ln^2(r/a)\right) \quad (2.5)$$

and the Weibull distribution

$$p(r|a, b) = abr^{b-1} \exp(-ar^b) \quad (2.6)$$

were introduced. Although still used by some for detection performance prediction in the presence of sea clutter [Rifkin, 1994], such models are clearly not satisfactory, in that they do not provide any insight into the actual mechanism responsible for generating the radar backscatter

in a marine environment.

The non-Rayleigh nature of the clutter amplitude distribution could potentially suggest that the Central Limit Theorem may not be applicable for the returns from a single range cell of a high resolution radar. However, as was discussed in section 2.1, the high resolution clutter is locally Rayleigh distributed. It is only the average power of the returns that changes with range and time. This suggests that the non-Rayleigh nature of the amplitude distribution of high resolution clutter is associated with bunching of scatterers by the sea wave structure, rather than being due to a small number of effective scatterers. This observation is applicable to the current high resolution radars (range cell size of a few meters in range, and a few hundred meters in azimuth). The number of scattering centres may no longer be sufficiently large at higher resolutions (of the order of a few centimeters in the radar systems of the future) to warrant the use of the Central Limit Theorem, and the compound Rayleigh models discussed below may no longer be appropriate [Ward et al., 1990].

In high resolution radars, provided that the individual resolution cell size is large enough for the Central Limit Theorem to hold (*i.e.* the effective number of scattering centres is large enough), the amplitude returns have been found to be locally Rayleigh distributed (see equation 2.3), with the local average power of the returns denoted as σ . It has also been found that the local power variations are associated with the swell, and are well modelled by a Gamma distribution

$$p(\sigma|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \sigma^{\alpha-1} \exp(-\beta\sigma) \quad (2.7)$$

As will be argued in chapters 3 and 5, the Gamma distribution is not necessarily the best distribution for modelling the statistics of the modulating component. To a large extent, the Gamma distribution stems from the legacy of the search for 'reasonable' distributions that have some properties (such as scaling) of the empirically observed clutter [Jakeman and Tough, 1988]. Some credibility to the Gamma distribution for the modulating component has been given by the birth-death migration process model for the evolution of a population of scatterers on the ocean surface.

The single point statistics of the high resolution clutter are obtained by compounding the locally distributed Rayleigh speckle with the Gamma distributed modulating component due to swell, giving rise to the celebrated K-distribution [Jakeman and Tough, 1988]

$$\begin{aligned} p(r|\alpha, \beta) &= \frac{4\beta r}{\Gamma(\alpha)} (\beta r^2)^{(\alpha-1)/2} K_{\alpha-1} \left(\sqrt{4\beta r} \right) \\ p(I|\alpha, \beta) &= \frac{2\beta}{\Gamma(\alpha)} (\beta I)^{(\alpha-1)/2} K_{\alpha-1} \left(\sqrt{4\beta I} \right) \end{aligned} \quad (2.8)$$

where $K_\alpha(\cdot)$ is the Modified Bessel function [Abramowitz and Stegun, 1970].

The K-distribution is a member of a family of compound distributions of the form

$$p(r|\alpha, \beta) = \int_0^\infty p(r|\sigma)p(\sigma|\alpha, \beta)d\sigma \quad (2.9)$$

The compound form of the distribution for the clutter returns has a number of distinct advantages over the empirical log-Normal and Weibull distributions. Firstly, there are good physical grounds for treating the speckle and the modulating components separately. While the former is associated with the random scattering from a large population of scatterers, the latter is associated with the fluctuations in the radar cross section due to the swell, which is resolved by a high resolution radar. As was discussed in section 2.1, the two components have very different correlation properties, which also allows the temporal and the spatial correlations in sea clutter to be treated separately. This becomes particularly useful in target detector design and analysis (see chapter 3).

As a final note on the single point statistical modelling of sea clutter, it is interesting to note that other extensions to the compound models are possible, whereby the scatterer population is divided into groups of scatterers with similar properties (*e.g.* same correlation structure), and where the number of elements in each group is a random variable [Azzarelli, 1995]. Also, other effects such as the thermal noise [Farina and Lombardo, 1994], forward scattering [Ward et al., 1990], as well as littoral effects [Branson, 1998], influence the form of the clutter amplitude distribution. However, the treatment of such effects is beyond the scope of this thesis, and therefore these are not dealt with any further.

2.3.2 Single Point Distribution Parameter Estimation

Estimation of the clutter distribution parameters is of great importance for a variety of radar detection and estimation problems. In particular, the knowledge of the parameters becomes crucial in threshold estimation for target detection (including CFAR and coherent target detection) in the presence of sea clutter.

The Maximum Likelihood estimates, which are also efficient, are easily obtained for delta correlated Rayleigh, log-Normal, as well as Weibull distributed clutter. The correlations in the clutter, which are dealt with in the ensuing sections, are presently ignored for the purpose of estimating the single point distribution parameters.

For Rayleigh distributed clutter (equation 2.3), the parameter corresponding to the radar cross section, σ , is trivially obtained by Maximum Likelihood from either the clutter amplitude, r , or the clutter intensity, I , as

$$\sigma = \frac{4}{\pi} \langle r^2 \rangle = \langle I \rangle$$

while the Weibull distribution (equation 2.6) Maximum Likelihood parameter estimates can be obtained by solving

$$-\frac{1}{b} + \frac{\langle r^b \ln r \rangle}{\langle r^b \rangle} = \langle \ln r \rangle$$

$$a = \frac{1}{\langle r^b \rangle}$$

for a and b . The parameters of the Gamma distribution (equation 2.7) are also readily obtained by Maximum Likelihood, by solving

$$\psi^{(0)}(\alpha) - \ln \alpha = \langle \ln \sigma \rangle - \ln \langle \sigma \rangle$$

$$\beta = \frac{\alpha}{\langle \sigma \rangle}$$

for α and β , where $\psi^{(0)}(\cdot)$ is the digamma function.

The Maximum Likelihood equations for the K-distribution parameters (equation 2.8) cannot be written in a simple form. The parameters of the K-distribution can either be obtained by a computationally expensive, two dimensional optimization procedure, or alternatively, an approximate scheme for parameter estimation can be employed. Any such approximate scheme is required to be as unbiased, and have as small a variance of the parameter estimates, as possible for a small clutter sample size.

Using the fact that the underlying modulating process, σ , remains approximately constant over a number of looks, [Oliver, 1993] demonstrates that in the limit of large L for an L -look radar, the distribution of the intensity can be approximated by

$$p_I(I) \approx p_\sigma(I) \left[1 + \frac{\alpha}{2L} \left(1 + \alpha - \frac{2I\beta}{\alpha}(1 + \alpha) + \frac{I^2\beta^2}{\alpha} \right) \right]$$

where $p_\sigma(\cdot)$ is given in equation 2.7. In other words, in the limit of a large number of looks, the distribution of the integrated intensity coincides with that of the Gamma distributed underlying modulating component σ . This observation has led [Oliver, 1993] to recommend the use of the mean of the logarithm, and the mean of the intensity to estimate the K-distribution parameters, α and β , respectively.

[Raghavan, 1991a] proposes to estimate the Maximum Likelihood (ML) parameters for a Gamma distribution, from which the K-distribution parameters are obtained by matching the arithmetic and the geometric mean (*i.e.* the mean in the intensity and in the logarithm domain, respectively) of the distributions. [Raghavan, 1991a] has found that this method provides lower variance in the estimates than by matching the two moments of the K-distribution and the data directly, and that the estimates thus obtained are approximately equal to the numerically evaluated ML estimates.

[Lombardo and Oliver, 1994] further investigate a number of parameter estimation schemes based on texture (*i.e.* clutter RCS) measures, including matching the intensity contrast measure, normalised log measure and the variance of log measure. On comparing the predicted bias errors of the estimators, the normalised log estimator provided a near-optimum measure on which to base the estimation of α .

[Blacknell, 1994a] performs a comparison of the different parameter estimation techniques for K-distributed clutter and finds that *the best overall estimation performance is provided by an estimation scheme that uses the mean of the data and mean of the log of the data as the two moments when applied to single-look data*. Based on these results, the K-distribution parameters can be obtained from the two above mentioned intensity moments by solving

$$\begin{aligned}\psi^{(0)}(\alpha) - \ln(\alpha) &= \langle \ln I \rangle - \ln \langle I \rangle - \psi^{(0)}(1) \\ \beta &= \frac{\alpha}{\langle I \rangle}\end{aligned}$$

for the K-distribution parameters α and β . The parameter α is often referred to as the *order parameter*. The order parameter controls the spikiness of the distribution, while the mean of the distribution is determined by the ratio α/β . Typically encountered values of α in sea clutter range from about 0.1 for very spiky clutter to 10 for near Rayleigh distributed clutter.

Finally, [Lombardo et al., 1995] study the effects of thermal noise on the order parameter estimation. They have found that the log texture measure was the most susceptible to bias, especially for small values of α (*i.e.* spiky clutter), while the contrast measure suffers from being less sensitive to variations in the order parameter, which renders it comparatively useless for order parameter estimation. They recommend the use of the amplitude contrast measure for order parameter estimation, particularly in an intermediate region of the PDF. They find that this compromise measure yields estimation errors not very different from that of normalised log and has an acceptable robustness to noise.

2.3.3 Temporal Correlations

So far, only the single point statistics of sea clutter has been considered. In particular, log-Normal, Weibull and compound K are the distributions of choice for high-resolution sea clutter. However, the knowledge of the single point statistics alone is insufficient, as the clutter correlations can adversely affect the detection performance prediction in the presence of sea clutter.

[Szajnowski, 1976] presents an algorithm for generation of correlated, incoherent (*i.e.* amplitude only) log-Normal clutter samples, while [Conte and Longo, 1987b] extend the argument to a coherent model for log-Normal clutter. Similarly, [Szajnowski, 1977] presents an algorithm for

generation of correlated incoherent Weibull clutter samples, while that algorithm is extended by [Li and Yu, 1989] to simulation of correlated, coherent Weibull clutter. Unfortunately, merely extending the single point log-Normal and Weibull distributions to include the correlations does not fully take into account the correlation properties of sea clutter.

As was discussed in section 2.1, sea clutter is a compound process, where the two components (the speckle and the modulating component) arise due to different physical processes and exhibit different correlation structures. [Marier, 1995] presents a method for generation of correlated, incoherent K-distributed clutter samples, based on the compound form of the K-distribution (equation 2.9). The correlations of the two components are modelled independently, and the effects of the correlations in the two components on the correlation structure of the overall clutter are studied for different values of the order parameter (*i.e.* spikiness) of the K-distribution.

In his analysis, [Marier, 1995] does not differentiate between the spatial and the temporal correlations associated with the two components of the compound K-distribution model. However, the approach of treating the two clutter components independently to model the overall correlations of the clutter is very attractive and also reflects the observed properties of the clutter, which were discussed in section 2.1. In fact, the piecemeal approach to modelling the spatial and temporal correlations in sea clutter is adopted in this thesis. The present section deals with the temporal correlations largely associated with the speckle component, while section 2.3.4 deals with the spatial correlations associated with the modulating component.

Coherent Clutter

The pulse to pulse clutter samples obtained using a coherent search radar from a single range cell, denoted by $\mathbf{x} = [x_{I1} + jx_{Q1}, \dots, x_{IT} + jx_{QT}]'$, are widely agreed to be locally Gaussian distributed, *i.e.*

$$p(\mathbf{x}|\sigma) = \frac{1}{(\pi\sigma)^T |A|^{1/2}} \exp\left(-\frac{1}{\sigma} \mathbf{x}' \mathbf{A}^{-1} \mathbf{x}\right) \quad (2.10)$$

where the matrix \mathbf{A} is the *normalised* covariance matrix of the complex clutter (*i.e.* the diagonal entries of the covariance matrix are set to unity). In the case of sea clutter, the temporal correlations are largely due to the speckle (with the modulating component, σ , being effectively constant over the decorrelation time of the speckle), while the speckle is mostly decorrelated in range.

Based on the above observation, [Conte and Longo, 1987a] have proposed to model the coherent clutter as a Spherically Invariant Random Process (SIRP). In particular, a vector $\mathbf{x} = [x_1, \dots, x_n]'$ with mean μ and covariance matrix \mathbf{M} is called a Spherically Invariant Random

Vector (SIRV) if its distribution takes on the form

$$p_n(\mathbf{x}) = |\mathbf{M}|^{-1/2} (2\pi)^{-n/2} h_n(q)$$

where $q = (\mathbf{x} - \boldsymbol{\mu})' \mathbf{M}^{-1} (\mathbf{x} - \boldsymbol{\mu})$. The distribution of a SIRV $p_n(\mathbf{x})$ is a generalisation of an n^{th} order Gaussian distribution. The class of admissible functions is defined as

$$h_n(q) = \int_0^\infty \sigma^{-n/2} \exp\left(-\frac{q}{\sigma}\right) f(\sigma) d\sigma$$

where $f(\sigma)$ is the characteristic distribution function (CDF). A SIRP $x(t)$ is defined so that every sample taken from it is a SIRV with one and the same CDF.

Both the Weibull (for $b \leq 2$) and the compound K distributions belong to the family of SIRPs. However, the log Normal distribution does not. Furthermore, the SIRPs have a number of attractive properties. In particular, a SIRV \mathbf{x} is completely characterised by the knowledge of its mean $\boldsymbol{\mu}$, its covariance matrix \mathbf{M} and its first-order distribution $p_1(\mathbf{x})$ (or, alternatively, the characteristic distribution function $f(\sigma)$). Also, SIRVs are closed with respect to linear transformations. The latter point is rather attractive, in that for the purposes of detection, a SIRV can, in principle, be pre-whitened by a linear transformation, with the resulting vector still being a SIRV. The properties of SIRPs are further discussed in papers by [Conte and Longo, 1987a], while [Rangaswamy et al., 1993] discuss the properties of a number of members of the SIRP family. The issues associated with generating SIRVs are discussed in [Rangaswamy et al., 1995] and [Conte et al., 1991a].

SIRPs form a flexible family of distributions capable of modelling the temporal correlation structure of the coherent sea clutter. One shortcoming of such models is that the correlation properties of the modulating component σ are not taken into account. This does not matter for the purposes of modelling the temporal correlations, as σ is effectively constant for any SIRV \mathbf{x} , but the spatial correlations of σ become important in the context of CFAR detection. The issues associated with modelling the spatial correlation structure of sea clutter are treated in more detail in the section 2.3.4.

Incoherent Clutter

As was discussed in the previous section, the temporal correlations in coherent sea clutter are readily modelled using SIRPs, which are simply an extension of a multivariate Gaussian distribution. However, in the case of incoherent clutter, the situation becomes a little more interesting. In particular, the autocorrelation function of incoherent clutter contains no phase information and, furthermore, the Rayleigh distribution being non-Gaussian is not necessarily defined uniquely by the univariate marginal distribution and the second order statistics.

No simple analytical form exists for a multivariate (of order greater than two) correlated Rayleigh distribution corresponding to the amplitudes of a complex Gaussian process. This does not mean that multivariate Rayleigh distributions do not exist (see [Marshall and Olkin, 1967] and [Singpurwalla and Youngren, 1993]), but rather that such distributions are not necessarily compatible with the complex Gaussian process hypothesis.

Due to the difficulty of modelling the joint distribution of the correlated Rayleigh clutter, the pulse to pulse samples from a single range cell are usually assumed to be uncorrelated. This is not an unreasonable assumption, as frequency agility can be employed in incoherent radars to remove the temporal correlations associated with the speckle, while leaving the underlying modulating component, σ , unchanged.

However, although it proves to be difficult to model the joint statistics of correlated Rayleigh clutter directly, it is relatively easy to synthetically generate correlated Rayleigh distributed samples [Szajnowski, 1977]. As is discussed in appendix A, the correlated Rayleigh samples can be obtained from a spherically symmetric, zero-mean complex Gaussian process. Denote by $\rho_r(k)$ the *normalised* autocovariance function of the Rayleigh amplitudes, and by $\rho_{x_I x_I}(k)$ the *normalised* autocorrelation of the real component (and imaginary component, $\rho_{x_Q x_Q}(k) = \rho_{x_I x_I}(k)$, by symmetry) of the complex Gaussian process, and by $\rho_{x_I x_Q}(k)$ the *normalised* cross-correlation of the in-phase and quadrature components of the complex Gaussian process. It can be shown that the amplitudes of the complex Gaussian process are Rayleigh distributed, and the correlations of the amplitudes are related to the correlations of the complex Gaussian process through the following identity

$$\rho_r(k) = \frac{\pi}{4 - \pi} \left({}_2F_1[-1/2, -1/2, 1, \rho_{x_I x_I}(k)^2 + \rho_{x_I x_Q}(k)^2] - 1 \right)$$

where ${}_2F_1[\cdot]$ is a Gauss hypergeometric function [Abramowitz and Stegun, 1970].

As expected, the phase relationship in the autocorrelation function of a complex Gaussian process (which contains the Doppler information) is lost in the autocovariance of the Rayleigh distributed amplitudes in an incoherent radar. However, as will be discussed in chapter 3, the correlation structure of the amplitudes may still be useful for target detection in the Moving Target Indicator (MTI) framework.

Furthermore, as was already stated, the Rayleigh distribution is non-Gaussian, and hence is not completely defined by its marginal and second order statistics. However, in the case of sea clutter, the Rayleigh distributed amplitudes are the result of observing the envelope of a complex Gaussian speckle component. It is this observation that provides the motivation for the work undertaken in chapter 4, where a full probabilistic model for the correlated clutter is employed in conjunction with Markov Chain Monte Carlo techniques in order to investigate the underlying complex Gaussian model parameter structure from the observed Rayleigh distributed amplitudes.

2.3.4 Spatial Correlations

The previous section has dealt with the short term temporal correlation structure in sea clutter, which is largely associated with the speckle component. The short term statistics of coherent clutter returns are readily modelled using Spherically Invariant Random Processes, while the joint statistics of the correlated amplitude returns in incoherent radar do not have a simple analytical form. Furthermore, the short term temporal correlations due to the speckle can be removed by the use of frequency agility.

The spatial correlations in sea clutter arise mostly from the underlying modulating component (associated with the swell). As the speckle is largely decorrelated in range, the spatial correlation structure of sea clutter is largely independent of whether the clutter is coherent or not, or whether frequency agility is employed. Furthermore, the modulating component is also correlated in time, but on a scale much longer than the speckle component. The modulating component can therefore be considered constant on a pulse to pulse basis, but not so on a scan to scan basis.

As was mentioned in section 2.1, the modulating component of high resolution sea clutter is widely agreed to be Gamma distributed (this point is further discussed in chapters 3 and 5). [Watts and Ward, 1987] provide an excellent discussion on the dependence of the order parameter of spatially correlated clutter distribution, on the radar range resolution. In particular, they observe that the reflectivity (*i.e.* local power) of the returns at lower range resolution can be expressed as the sum of a number of correlated Gamma distributed variates, which have also been found to be Gamma distributed, but with a different shape parameter. The implications of this observation are profound.

Given a high resolution range profile, where the underlying modulating component, σ_H , is Gamma distributed (as in equation 2.7) with order parameter α_H and with correlation coefficient of σ_H given by $\rho_\sigma(k)$ for samples spaced apart by k range sample intervals, a low resolution range profile can be simulated by averaging blocks of N adjacent samples of the high resolution range profile.

The clutter reflectivity, σ_L , of the resulting low resolution range profile is also Gamma distributed, with order parameter α_L . By using arguments based on the first two moments of the sum of the correlated high resolution variates, [Watts and Ward, 1987] show that the order parameter of the low resolution clutter is related to the order parameter of the high resolution clutter through

$$\frac{\alpha_H}{\alpha_L} = \frac{N + 2(N-1)\rho_\sigma(1) + 2(N-2)\rho_\sigma(2) + \dots + 2\rho_\sigma(N-1)}{N^2}$$

which reduces to $\alpha_L = N\alpha_H$ for spatially uncorrelated clutter. As expected, by reducing

the resolution (*i.e.* increasing N), the order parameter of the resulting lower resolution clutter increases and in the limit the modulating component becomes constant. Furthermore, as [Watts and Ward, 1987] point out, the distribution of the clutter speckle component remains unchanged with radar resolution. Therefore, the distribution of the overall clutter envelope will still be K-distributed and in the limit of large N will tend to the Rayleigh distribution (as expected for low resolution clutter).

[Watts and Ward, 1987] found that the above result for the change in order parameter is in close agreement with the observed data. This finding has profound implications on the target detection strategy. In particular, there is a trade off in choosing between higher resolution, which is associated with a decrease in the order parameter (*i.e.* increased clutter spikiness), and simple binary integration for the purpose of target detection.

Spatial correlations in the modulating component of sea clutter also influence the performance of adaptive target detectors (such as the CFAR detector), when the detection threshold for the range cell under test is adapted according to the clutter samples in the neighbouring range cells. The threshold itself is therefore very sensitive to the spatial correlations in the modulating component. These issues will be discussed in more detail in chapter 3.

There is great utility in obtaining a multivariate distribution to capture the joint statistics of the underlying modulating component of sea clutter. Unfortunately, a Gamma distributed process is not uniquely defined by its single point marginal and second order statistics. Also, to the best of the author's knowledge, despite the efforts of many researchers in this field, no simple analytical expression has been found for a multivariate Gamma distribution beyond the second order (*i.e.* bivariate case).

[Armstrong and Griffiths, 1991b] derive an expression for a bivariate Gamma distribution with half-integer order parameter, based on the sum of squares of correlated Gaussian variates. This approach amounts to passing the correlated Gaussian variates through a square law non-linearity, and then summing the resulting variates to obtain samples of a correlated Gamma process. Using Laguerre polynomials, [Raghavan, 1991b] derives a similar expression for a bivariate Gamma distribution with arbitrary order parameter, but which reduces to the distribution presented by [Armstrong and Griffiths, 1991b] for half-integer values of the order parameter. The form of the bivariate Gamma distribution for σ_l and σ_m spaced k samples apart, with the single point marginal given by equation 2.7 (*i.e.* $\sigma \sim \Gamma(\alpha, \beta)$) and with the normalised covariance $\rho_\sigma(k)$, is given by

$$\begin{aligned}
 p(\sigma_l, \sigma_m | \alpha, \beta, \rho_\sigma(k)) &= \Gamma(\sigma_l | \alpha, \beta) \Gamma(\sigma_m | \alpha, \beta) \left[\frac{\Gamma(\alpha)}{1 - \rho_\sigma(k)} (\sigma_l \sigma_m \rho_\sigma(k) \beta^2)^{-(\alpha-1)/2} \right] \\
 &\times \exp \left[-\frac{\beta \rho_\sigma(k)}{1 - \rho_\sigma(k)} (\sigma_l + \sigma_m) \right] I_{\alpha-1} \left[\frac{2\beta \sqrt{\sigma_l \sigma_m \rho_\sigma(k)}}{1 - \rho_\sigma(k)} \right] \quad (2.11)
 \end{aligned}$$

where $m = l + k$, $0 \leq \rho_\sigma(k) \leq 1$, and where the standard notation for the Gamma distribution has been adopted, *i.e.*

$$\Gamma(x|\alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta x)$$

Based on the bivariate form of the distribution for σ_l and σ_m , the conditional distribution $p(\sigma_l|\sigma_m)$ is easily obtained from the knowledge of the marginal $p(\sigma_m)$. [Bucciarelli et al., 1996] propose a multivariate, exponentially correlated, Gamma distribution based on the first order Markov decomposition, *i.e.*

$$p(\sigma_1, \dots, \sigma_N|\alpha, \beta, \rho_\sigma) = p(\sigma_1|\alpha, \beta) \prod_{n=1}^{N-1} p(\sigma_{n+1}|\sigma_n, \alpha, \beta, \rho_\sigma)$$

Clearly, due to the constraint on the resulting correlation structure, such a simple model may not be very adequate for most practical cases of interest.

Although often desirable, the multivariate distribution is not the only way of describing the statistics of the data. In the present case, where the spatial correlations of the modulating component of sea clutter are of interest, the modulating component can be considered to be the result of a stochastic process. [Jakeman and Tough, 1988] have suggested using a Stochastic Differential Equation (SDE) based model, which is based on semi-phenomenological arguments for the evolution of a population of scattering centres on the ocean surface. However, this is not the only possible model for a correlated process with single point Gamma statistics. [Shephard, 1994], [Singpurwalla and Youngren, 1993] and [Lewis et al., 1989] present some other first order Markov Gamma process models. Unfortunately, being first order Markov, such models suffer from the shortcoming in that they are only capable of modelling a very limited range of correlation functions.

However, as will be argued in chapters 3 and 5, the modulating component of sea clutter need not be Gamma distributed. In fact, the Gamma distribution assumption may result in overly optimistic detection performance prediction for target detection in the presence of sea clutter. For this reason, a family of non-constant variance (*i.e.* heteroscedastic) models is proposed in chapter 5 for modelling the statistics of the modulating component of sea clutter. As will be argued in chapter 6, it is the conditionally Gaussian form of the predictive distribution, and the ability to model a wide range of correlations, that makes such models potentially very attractive from a CFAR detection point of view.

Generation of Correlated Gamma Variates

For the purpose of target detection performance prediction in the presence of sea clutter using Monte Carlo (*i.e.* simulation) methods, it may be necessary to generate spatially correlated samples with the desired single point statistics (usually assumed to be Gamma distributed for the modulating component). This issue has been dealt with at length by a number of authors.

[Marier, 1995] employs the scheme proposed by [Armstrong and Griffiths, 1991b], to generate correlated Gamma samples with half-integer order parameter, by summing correlated Gaussian variates, which are processed by a square law zero memory non-linearity. The resulting samples are Gamma distributed, with the joint two-point statistics given by equation 2.11. The disadvantage of this scheme is that the samples can be generated only for a Gamma distribution with half-integer order parameter, and with a non-negative autocorrelation function.

[Watts, 1996] proposes a scheme to generate Gamma distributed samples of any order parameter, but with an exponentially decaying autocorrelation function, using a method based on transforming correlated Gaussian variables using a zero memory non-linearity.

[Blacknell, 1994b] proposes a method based on the properties of the characteristic functions of correlated Gamma distributions. With this method, correlated Gamma samples are obtained from uncorrelated Gamma variates by using a moving average filter. Although samples with any non-negative autocorrelation structure and order parameter can be obtained in this way, the method suffers from the shortcoming in that it becomes increasingly difficult to estimate the coefficients of the moving average filter as the correlation length increases.

[Oliver and Tough, 1986] discuss two techniques for obtaining correlated, Gamma distributed samples. The first method is based on Euler integration, and simulation of a Stochastic Differential Equation with Gamma single point statistics. The usefulness of this approach is severely limited by the restrictions on the correlation structure of the resulting samples.

An alternative technique proposed by [Oliver and Tough, 1986] is similar to that proposed by [Blacknell, 1994b], in that the correlated samples are obtained as a linear, weighted average of Gamma distributed samples. The weighting coefficients are evaluated such that the resulting samples have the required first and second order statistics. However, unlike the technique proposed by [Blacknell, 1994b], the samples obtained by simple averaging of the uncorrelated Gamma variates need not themselves be Gamma distributed. Therefore, the technique proposed by [Oliver and Tough, 1986] is useful only insofar as the samples with roughly the correct single-point statistics, but with the correct second order statistics, are generated.

As a final note, [Lombardo and Oliver, 1995] present a discussion on the issues associated with estimating the correlation properties of K-distributed clutter. In particular, they find that the correlation properties estimated in the logarithm or the amplitude domain are significantly better than those obtained in the intensity domain.

2.4 SUMMARY

The main purpose of this chapter was to present a review of the field of sea clutter modelling. To this end, the properties of high resolution sea clutter were discussed in section 2.1. From empirical observations, the compound form of the clutter is apparent not only in the single point statistics, but also in the correlation structure.

Sea clutter is effectively a compound process, where the fast oscillating speckle is modulated in power by an underlying process associated with the sea swell. While the speckle decorrelates on the scale of a few milliseconds, it is effectively uncorrelated spatially. Furthermore, the speckle can be decorrelated by using frequency agility. The speckle has been observed to be locally Rayleigh distributed, and the overall clutter distribution is well modelled as being K-distributed.

The modulating component, which is associated with the swell structure of the sea, is usually spatially correlated. The modulating component decorrelates on the scale of a few seconds, and has been found to be unaffected by frequency agility. It is the very disparate correlation properties of the two sea clutter components, which allow the clutter to be treated in a piece-meal fashion, whereby the spatial and the temporal correlations of high resolution sea clutter are treated independently for the purpose of target detection analysis in the presence of high resolution sea clutter.

In section 2.2, the issues associated with non-linear modelling of sea clutter were discussed. Although no conclusions were drawn as to whether the sea clutter is deterministic or not, it was noted that the tests used by some authors to diagnose non-linearities and chaos in the sea clutter data can lead to misleading conclusions. However, some authors have published very favorable detection performance results for an incoherent radar, approaching the performance achievable using a fully coherent radar. A plausible reason for this was given, based on an autoregressive model for the speckle component. The issue of target detection in correlated incoherent clutter is dealt with further in chapter 4.

Probabilistic models for the statistics of high resolution sea clutter were discussed in section 2.3. In particular, while modelling of the single point statistics of high resolution sea clutter is

well understood, due to the non-Gaussian nature of high resolution sea clutter, the modelling of the temporal and the spatial correlations proves to be more of a challenge. In particular, as was discussed in section 2.3.3, while it is relatively simple to model the temporal correlations of the coherent speckle (conditional on the modulating component), the temporal correlations of the incoherent speckle are not readily modelled. In fact, for incoherent radars, the speckle is often assumed to be uncorrelated from pulse to pulse, the condition which can be readily achieved by the use of frequency agility. The issues associated with the modelling of the temporal correlations of incoherent speckle are dealt with further in chapter 4.

Finally, in section 2.3.4, the issues associated with the modelling of the spatial statistics of sea clutter were discussed. The spatial correlations in sea clutter are mainly associated with the modulating component, which is widely agreed to be Gamma distributed. The multivariate Gamma distribution is not uniquely defined by the single point marginal and the second order statistics. The issues associated with the modelling of the spatial correlations of the modulating component are dealt with further in chapter 5.

Target Detection in the Presence of Sea Clutter

3

3.1 INTRODUCTION

The aim of a surveillance radar is to detect targets by illuminating an area of interest with electromagnetic radiation. In a marine environment, the term *target* usually refers to any radar backscatter that is not due to the scattering from an ocean surface. This may include backscatter from some man-made object (such as a ship, a periscope or an airplane), or some other natural object (such as an iceberg).

Target detection can be looked upon as a decision-theoretic problem [Bernardo and Smith, 1994]. Within this framework, given a number of backscattered samples (which could include targets in addition to the background clutter), there are two possible actions: declare the presence of a target, or alternatively, declare that no target is present in the received signal. Associated with each of the actions, are two consequences:

- if a target is declared, but none is present, the consequence is referred to as a *false alarm*
- if a target is declared, and the target is present, the consequence is a successful target detection
- if no target presence is declared, and a target is present in the received clutter, the consequence is a missed target detection
- if no target presence is declared, and no target is present in the received clutter, the consequence is a successful non-detection

One of the design objectives of a surveillance radar system is to suppress the number of false alarms to within an acceptable level, while maximising the number of real target detections. This indeed is a very risk averse design criterion, where the expected utility is a function of the probability of detection, subject to a maximum acceptable level of the probability of false alarm.

The action of declaring a detection can be based on any number of different criteria. For example, if the clutter is known to be deterministic, detection can be declared when the prediction error exceeds some pre-defined threshold [Haykin, 1996]. Alternatively, of more interest for the present discussion, the action of declaring a detection within a statistical framework is based on

some signal statistic (*e.g.* amplitude, spectrum, likelihood) exceeding some pre-defined threshold. The action of setting the threshold constitutes the action of declaring a detection, and hence the threshold has to be set such that the overall expected utility is maximised.

Statistical target detection is possible only if both the background clutter and any potential targets can be characterised and differentiated by their statistical properties. Thus far in this dissertation, only the statistical properties of the background sea clutter have been discussed. No mention as yet has been made of the target statistics. This issue is dealt with briefly in section 3.2. Fixed threshold detection based on the single point amplitude statistics of the radar returns is dealt with in section 3.3, while in section 3.4 the discussion is extended to fixed threshold target detection following pulse integration. Adaptive threshold detection techniques are reviewed in section 3.5, a brief review of Neyman-Pearson detection techniques is presented in section 3.6, while in section 3.7 techniques for target detection in coherent radar systems are briefly discussed. A summary of the main points raised in this chapter is presented in section 3.8.

3.2 TARGET MODELS

Complex bodies, such as ships and aircraft, have been mapped to yield their radar cross sections in various planes as a function of aspect angle. Because of their motion relative to the radar, the aspect angle changes, and it may therefore be more practical to describe them in terms of the probability distribution function of their cross section, $p(A)$. Furthermore, there may be a myriad of potentially different targets (*i.e.* signals other than those due to thermal noise or the backscatter from an ocean surface), each with different statistical properties. Therefore, unless a particular target is being sought, the utility of using a specific target model may not be very great.

Swerling observed that targets may be grouped into a number of rather general classes, based on their apparent scattering characteristics. Broadly speaking, targets have been observed to be either fast or slow fluctuating. In the former case, the target returns are observed to vary on a pulse to pulse basis, while in the latter case, the targets remain essentially fixed from pulse to pulse, while fluctuating on the scan to scan basis. Alternatively, the targets can be non-fluctuating (*i.e.* constant amplitude returns).

With regard to the amplitude statistics of the target returns, the backscatter from a target can be considered to arise due to scattering from many independent scattering centres of approximately equal strength. Alternatively, it can arise from scattering from a dominant scatterer, along with a number of other smaller scattering centres. Where there are many independent

scatterers of equal strength, the intensity (*i.e.* power) of the returns is exponentially distributed, *i.e.*

$$p(I|\sigma_T) = \frac{1}{\sigma_T} \exp(-I/\sigma_T)$$

while the corresponding amplitude returns are Rayleigh distributed

$$p(A|\bar{A}) = \frac{A}{\bar{A}^2} \exp\left(\frac{-A^2}{2\bar{A}^2}\right) \quad (3.1)$$

with the average power $\sigma_T = 2\bar{A}^2$. For the case of one dominating, and many independent scattering centres, the intensity is chi-square distributed with $k=2$, *i.e.*

$$p(I|\sigma_T) = \frac{4I}{\sigma_T^2} \exp(-2I/\sigma_T) \quad (3.2)$$

Based on the above properties, Swerling proposed four widely adopted statistical target models, classified according to the rate of fluctuations and the single point statistics:

1. **Swerling I** refers to a slowly fluctuating target with Rayleigh distributed amplitude statistics (equation 3.1).
2. **Swerling II** refers to a fast fluctuating target with Rayleigh distributed amplitude statistics (equation 3.1).
3. **Swerling III** refers to a slowly fluctuating target with chi-square ($k=2$) distributed intensity statistics (equation 3.2).
4. **Swerling IV** refers to a fast fluctuating target with chi-square ($k=2$) distributed intensity statistics (equation 3.2).

A non-fluctuating (*i.e.* constant amplitude) target is often referred to as **Swerling 0**.

It is interesting to note that the Swerling models are not the only possible target models (*e.g.* see [Shnidman, 1995]). It is the simplicity of the Swerling models that most often makes them the models of choice for target detection performance analysis. For optimal detection of a particular target (*e.g.* a helicopter hovering above the sea surface), more detailed statistical models, specific to the particular target, may need to be used.

3.3 SINGLE PULSE FIXED THRESHOLD DETECTION

Target backscatter is always accompanied by noise and other extraneous clutter. It is the goal of a decision circuit to detect the presence of any potential targets, while suppressing the number of false detections to within some acceptable level. Denote by r the amplitude of the radar return. The amplitude of the return can either be due to the noise (including clutter) alone, or due to the noise plus a target. It is the difference between the amplitude statistics of the two cases that forms the basis for statistical detection of targets in the presence of noise and clutter.

The amplitude statistics of sea clutter were discussed in section 2.3. In particular, it was shown that a target and thermal noise free clutter amplitude is locally Rayleigh distributed, *i.e.*

$$p(r|\sigma) = \frac{2r}{\sigma} \exp(-r^2/\sigma)$$

with the underlying modulating component, σ , being Gamma distributed.

In addition to the background clutter, the amplitude statistics of the radar backscatter is also affected by the presence of thermal noise. However, in a marine environment, the backscatter from an ocean surface usually dominates any thermal noise contributions. For reasons of space and brevity, the thermal noise effects are not dealt with any further in the present treatment. Instead, the interested reader is referred to the excellent discussion on the subject in [Watts, 1987].

The combined amplitude statistics of sea clutter and a target with amplitude A is Rice distributed [Levanon, 1988], *i.e.* the combined radar return, r , is distributed as

$$p(r|\sigma, A) = \frac{2r}{\sigma} \exp\left(-\frac{r^2 + A^2}{\sigma}\right) I_0\left[\frac{2rA}{\sigma}\right] \quad (3.3)$$

where $I_0[\cdot]$ is the modified Bessel function of order zero. The target amplitude itself may fluctuate on different time scales, according to some distribution $p(A|\bar{A})$ (*e.g.* Swerling cases). Similarly, the sea clutter cross section, σ , though effectively constant over short time intervals, fluctuates according to some distribution $p(\sigma|\alpha, \beta)$. Therefore, the unconditional clutter plus target amplitude distribution is given by

$$p(r|\bar{A}, \alpha, \beta) = \int_{\sigma} \int_A p(r|\sigma, A) p(A|\bar{A}) p(\sigma|\alpha, \beta) dA d\sigma$$

where the target mean power $\sigma_T = 2\bar{A}^2$, and where α and β are the clutter RCS (*i.e.* modulating component) distribution parameters.

The probability of false alarm is given by the probability of a target-free amplitude return

(i.e. $A = 0$) exceeding some predefined threshold, V_T , i.e.

$$P_{fa}(V_T|A = 0) = p(V_T \leq r|A = 0) = \int_{V_T}^{\infty} p(r|A = 0)dr$$

Similarly, the probability of detection is given by the probability of a radar return, which includes a target of amplitude A , exceeding some predefined threshold, V_T , i.e.

$$\begin{aligned} P_d(V_T|\bar{A}) &= \int_A^{\infty} P(V_T \leq r|A)p(A|\bar{A})dA \\ &= \int_{V_T}^{\infty} p(r|\bar{A})dr \end{aligned}$$

The above analysis is only appropriate for the case of a single pulse detection. The issues associated with evaluation of the probability of false alarm and detection following pulse integration are dealt with by [Watts et al., 1990], and are discussed further in section 3.4.

3.3.1 Non-fading Targets

In the case of a non-fading target (i.e. Swerling 0) with constant amplitude, A , the single pulse probability of false alarm, P_{fa} , and the probability of detection, P_d , for threshold level V_T can be obtained by integrating equation 3.3 to give

$$\begin{aligned} P_{fa}(V_T|\sigma) &= \exp\left(-\frac{V_T^2}{\sigma}\right) \\ P_d(V_T|\sigma, A) &\approx \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{V_T}{\sqrt{\sigma}} - (SCR)^{1/2}\right) \right] \end{aligned} \quad (3.4)$$

where the approximation for the probability of detection is valid for the signal to clutter ratio $SCR = A^2/\sigma \gg 1$. The above probabilities are conditional on the underlying modulating process power, σ . To obtain the *average* values for the probabilities of detection and of false alarm, $P_{fa}(V_T|\sigma)$ and $P_d(V_T|\sigma, A)$ should be averaged over all values of σ . However, that proves to be no trivial task, and following [Watts et al., 1990], simulations based on real averaged range profiles are used throughout this thesis to estimate the *average* probabilities for real clutter data.

3.3.2 Fluctuating Targets

The amplitude of the target returns for a Swerling 0 target is constant. For Swerling I and Swerling II targets, the amplitude of the target returns fluctuates, and is Rayleigh distributed. To obtain the overall distribution of the returns due to a fluctuating target plus clutter, the target plus clutter distribution in equation 3.3 has to be averaged over all the possible values of

A , resulting in the overall radar return amplitude being distributed as

$$\begin{aligned} p(r|\sigma, \bar{A}) &= \int p(r|\sigma, A)p(A|\bar{A})dA \\ &= \frac{2r}{\sigma + 2\bar{A}^2} \exp\left(\frac{-r^2}{\sigma + 2\bar{A}^2}\right) \end{aligned} \quad (3.5)$$

Therefore, the probability of false alarm and the probability of detection, conditional on the clutter power, σ , are given by

$$\begin{aligned} P_{fa}(V_T|\sigma) &= \exp\frac{-V_T^2}{\sigma} \\ P_d(V_T|\sigma, \bar{A}) &= \exp\frac{-V_T^2}{\sigma + 2\bar{A}^2} \end{aligned} \quad (3.6)$$

To obtain the *average* values for the probabilities of detection and of false alarm, $P_{fa}(V_T|\sigma)$ and $P_d(V_T|\sigma, \bar{A})$ should be averaged over all values of σ . Closed form expressions for the average probability of false alarm can be derived when σ is Gamma or Inverse Gamma distributed (see appendix D). However, if σ is log-Normal distributed, an approximate scheme relating the threshold to the probability of false alarm can be obtained. Polynomial approximations relating the detection threshold V_T to the *average* probability of false alarm when σ is Gamma, Inverse Gamma or log-Normal distributed are presented in appendix D.

3.4 FIXED THRESHOLD DETECTION WITH INTEGRATION

The previous section has dealt with target detection based on a single pulse return. Clearly, such detection may not be very satisfactory, as the single point distributions for the clutter and the target may heavily overlap. Some means of separating the two distribution may be required for improved detection performance.

One standard technique to separate the target and the clutter distributions is to process a number of consecutive radar returns before declaring a detection. High detection probability can then be obtained even when the single pulse signal to clutter ratio is near or below unity. The process by which the pulses are combined is known as integration.

There are a number of ways of performing the integration, depending on the stage of radar processing at which the pulses are integrated. There are principally three such stages:

1. *Coherent integration* before envelope detection, where the pulse to pulse returns are added and averaged coherently.
2. *Incoherent integration* after envelope detection, where the pulse to pulse amplitude re-

turns are averaged incoherently.

3. *Binary integration* after the single pulse detection circuit, where the detection is based on the binary output of a single pulse detection circuit.

The merits of the different schemes are discussed in [Eaves and Reedy, 1987], and presently only a brief exposition on the estimation of the probabilities of detection and false alarm is given for the incoherent and the binary integration schemes. The discussion herein is limited to the Swerling I and II target models, with the interested reader referred to the excellent paper by [Shnidman, 1995] for a more comprehensive review of the subject.

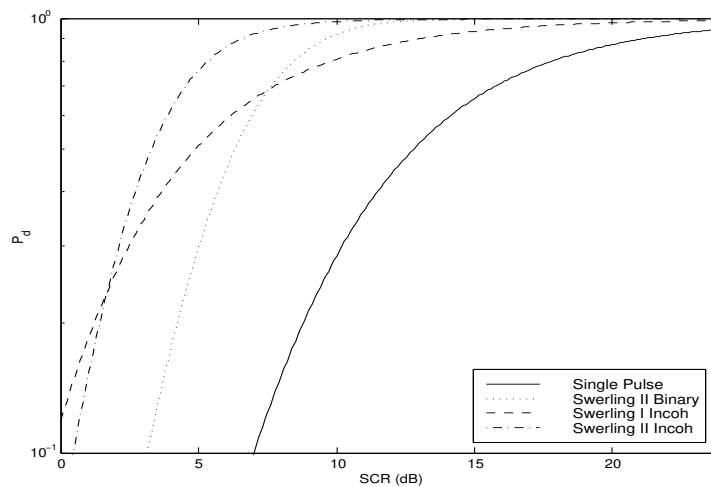


Figure 3.1: Probability of detection as a function of Signal to Clutter Ratio (SCR) in dB. Results shown for single pulse detection, as well as detection employing 10 pulse incoherent and binary integration. The results are shown for Swerling I and II target models in uncorrelated Rayleigh clutter, with $P_{fa} = 10^{-6}$.

Figure 3.1 compares the target detection performance of detectors employing 10 pulse incoherent and binary integration, to target detection performance of single pulse detector. The results in the figure are shown for Swerling I and Swerling II targets in uncorrelated Rayleigh distributed clutter, with $P_{fa} = 10^{-6}$. The clutter is assumed to be independent from pulse to pulse. Swerling I target remains constant, while Swerling II target fluctuates on pulse to pulse basis. Binary integration is only shown for the fast fluctuating Swerling II target.

One typically adopted measure of detection performance improvement is based on the change in the signal to clutter ratio (SCR) required in order to achieve a particular probability of detection (typically $P_d = 0.5$ or 0.9) for the specified probability of false alarm [Eaves and Reedy, 1987]. Based on this measure, it can be seen from figure 3.1 that integrating 10 pulses provides 6dB improvement for binary integration for Swerling II target, 8dB improvement for incoherent integration for Swerling I target, and 10dB improvement for incoherent detection for Swerling II target over single pulse detection with $P_{fa} = 10^{-6}$ and $P_d = 0.5$.

The two Swerling cases considered in figure 3.1 represent the two extreme cases of fully correlated and uncorrelated target models. The interested reader is referred to the paper by [Kanter, 1986] for the detection probability analysis of partially correlated targets in uncorrelated Rayleigh distributed clutter, using incoherent integration for different numbers of integrated pulses.

It is interesting to note that regardless of whether the target is correlated or not, pulse integration, be it incoherent or binary, provides an improvement over single pulse detection in terms of the SCR ratio required to achieve the desired probability of detection for the given probability of false alarm. The incoherent and the binary integration schemes are now briefly discussed.

3.4.1 Incoherent Integration

Coherent integration, where complete phase information from pulse to pulse is available, provides the largest increase in signal to clutter ratio for targets in uncorrelated clutter. In fact, for coherent integration of M pulses, an increase in signal to clutter ratio by up to a factor of M can be expected.

In addition to integration, other potentially powerful strategies are possible for coherent target detection in the presence of background clutter (*e.g.* by making the full use of the Doppler information). These techniques are briefly discussed in section 3.7. The present section is concerned with the effect of integration after the envelope detection stage of a radar system, *i.e.* when only the amplitude, and no phase information is available.

[Watts et al., 1990] present an in-depth discussion on the detection performance analysis of target detectors employing incoherent integration. In particular, [Watts et al., 1990] show that the analysis depends not only on the target and clutter marginal statistics, but on their respective fluctuating characteristics as well. [Watts, 1987] shows that the analysis is further complicated by the presence of thermal noise in the clutter returns.

The present discussion is limited to Swerling 0, I and II targets in completely uncorrelated clutter (which can be obtained through the use of frequency agility). For the analysis of partially correlated targets, the reader is referred to the paper by [Kanter, 1986], while [Shnidman, 1995] provides a more in-depth analysis of the detection performance including incoherent integration for a wide variety of clutter and target models.

Consider M consecutive radar return amplitudes from a single range cell, denoted by $\{r_k : k = 1 \dots M\}$, from which an integrated return, denoted as $z = \sum_k r_k^2$, is obtained. Each of the returns r_k can consist of clutter as well as targets, and each return is individually distributed

as in equation 3.3 if the target is slow fluctuating, and as in equation 3.5 if the target is fast fluctuating.

The modulating component of sea clutter, σ , can be considered to be effectively constant over the integration period, while the speckle is uncorrelated from pulse to pulse. The distribution of the sum of M independent variates can be obtained as an M -fold convolution of the individual return distributions, and hence the distribution of the incoherently integrated, target free speckle can be shown to be of the form [Levanon, 1988]

$$p(z|\sigma, A = 0) = \frac{1}{z} \left(\frac{z}{\sigma}\right)^M \frac{1}{\Gamma(M)} \exp(-z/\sigma)$$

The probability of false alarm for a detector employing M pulse incoherent integration with the detection threshold for z set to V_T^2 , can thus be shown to be given by

$$\begin{aligned} P_{fa}(V_T|\sigma) &= p(V_T^2 \leq z|\sigma, A = 0) \\ &= \frac{\Gamma(M, V_T^2/\sigma)}{\Gamma(M)} = 1 - \frac{1}{\Gamma(M)} \gamma(M, V_T^2/\sigma) \end{aligned} \quad (3.7)$$

where $\gamma(\cdot, \cdot)$ is the incomplete Gamma function. In order to obtain the *average* probability of false alarm, $P_{fa}(V_T|\sigma)$ should be averaged over all possible values of σ , but again, this proves to be a rather daunting task analytically.

The probability of detection in uncorrelated clutter depends on the rate of fluctuation of the target. In particular, for a *non-fluctuating* target (Swerling 0), the target amplitude is constant over the integration period, while the clutter contributions are independent from pulse to pulse. The distribution of the integrated return, z , can therefore be obtained as an M -fold convolution of the the individual return distributions as given by equation 3.3 for each of the individual returns, r_k . An approximate expression for the probability of detection for $M > 100$ and $R_P < 0.6$ can then be shown to be given by [Levanon, 1988]

$$P_d(V_T|\sigma, \bar{A}) \approx \phi \left[\phi^{-1}[P_{fa}(V_T|\sigma)] - \sqrt{M} R_P / 2 \right]$$

where $R_P = 2A^2/\sigma$, and $\phi(x) = (2\pi)^{-1/2} \int_x^\infty \exp -\frac{1}{2}t^2 dt$.

For a *slow fluctuating* target (Swerling I), the target amplitude A is effectively constant over the integration period, and hence the distribution for the integrated return z can be obtained as an M -fold convolution of the individual return distributions as given by equation 3.3 for each of the individual returns r_k . The result is the distribution $p(z|A, \sigma)$, which then needs to be averaged over the slow fluctuating target distribution $p(A|\bar{A})$, from which the average

probability of detection can be approximately expressed for $M \geq 10$ as

$$P_d(V_T|\sigma, \bar{A}) \approx \exp \left[\frac{2\phi^{-1}[P_{fa}(V_T|\sigma)]}{\sqrt{M}\bar{R}_P} \right]$$

where $\bar{R}_P = 4\bar{A}^2/\sigma$.

For a *fast fluctuating* target (Swerling II), both the clutter and the target are uncorrelated from pulse to pulse, and hence the distribution for the integrated return is obtained as an M -fold convolution of the individual return distributions as given by equation 3.5. This results in the distribution $p(z|\bar{A}, \sigma)$, from which the probability of detection can be approximately expressed for $M \geq 10$ as

$$P_d(V_T|\sigma, A) \approx \phi \left[\frac{2\phi^{-1}[P_{fa}(V_T|\sigma)] - \sqrt{M}\bar{R}_P}{\bar{R}_P + 2} \right]$$

Clearly, yet again, none of the above expressions for the probability of detection is easily averaged analytically over all the values of the modulating component, σ , in order to obtain the unconditional average probability of detection in compound clutter environment.

3.4.2 Binary Integration

Although not as powerful as incoherent integration, binary integration following a single pulse detection circuit is amenable to simple analysis, particularly for constant and fast fluctuating targets. Binary integration has also been shown to be very suitable for practical implementation of pulse to pulse integration when large radar pulse bandwidths are involved [Watts, 1985].

A binary integration detector uses the binary output (*i.e.* detection or non-detection) of a single pulse detection circuit. Detection is declared for M integrated pulses if the number of individual detections within an M pulse block exceeds some pre-set limit (which in the case of thermal noise is shown to be optimal for 6 individual detections within a block of 10 pulses [Watts, 1985]).

For the case of constant, or fast fluctuating targets (*e.g.* Swerling 0 and II, respectively) in the presence of uncorrelated clutter, it can be shown that for non-overlapping blocks of M pulses with detection threshold set at K detections out of M samples, the probability of false alarm (detection) for the binary integrator is related to the single pulse probability of false alarm (detection) through

$$P_n(V_T|M) = \sum_{k=K}^M \binom{M}{k} [P_n(V_T|M=1)]^k [1 - P_n(V_T|M=1)]^{M-k}$$

where $P_n(V_T|M)$ refers to either the probability of detection $n \equiv d$, or the probability of false alarm $n \equiv fa$, for an M pulse binary detector. Similarly, $P_n(V_T|M = 1)$ corresponds to the probability of false alarm or detection for a single pulse detection circuit.

Finally, it is interesting to note that, unlike the case of the constant and fast fluctuating targets, the target detection performance analysis is somewhat complicated for slow fluctuating targets (e.g. Swerling I). This is because in the case of the slow fluctuating targets it is necessary to average the integrated probability of detection over all the possible target amplitudes A .

3.5 ADAPTIVE THRESHOLD (CFAR) DETECTION

The target detection schemes discussed so far were either based on a single radar return, or on a return integrated over a number of consecutive pulses. The detection threshold was fixed across a range profile at a level ensuring that the correct *average* probability of false alarm is achieved. The speckle component was assumed to be independent on pulse to pulse basis (which can be achieved by use of frequency agility), while the underlying modulating component was assumed to be effectively constant over the integration period. Both the speckle and the modulating component were effectively treated as being uncorrelated spatially.

The fixed threshold detection is indeed optimal if both components of sea clutter are spatially uncorrelated and homogeneous. However, although the speckle component tends to be spatially uncorrelated, this is seldom true of the modulating component of high resolution sea clutter, which itself is associated with the swell. The consequences of using a fixed threshold detector in a compound clutter environment, when the clutter modulating component is spatially correlated, are twofold. Firstly, spatial variations in false alarm rate arise due to the spatial variations in the modulating component, and secondly, a detection loss is incurred as a result of not taking the spatial correlations of the modulating component into account when setting the detection threshold.

The inherent limitations on the performance of the adaptive threshold techniques for target detection in compound clutter environment are determined by the statistics of the modulating component. These limitations are now illustrated for a database of averaged range profiles extracted from a high resolution sea clutter profile database [Hirst and Baker, 1987].

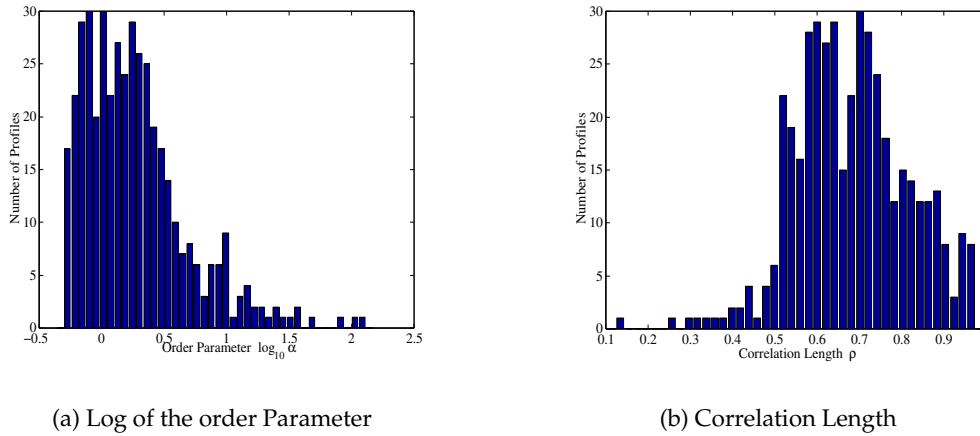


Figure 3.2: Histogram of (a) the logarithm of the Gamma distribution order parameter, $\log_{10}\alpha$, and (b) the correlation length, ρ , for the database of 440 range profiles.

3.5.1 Range Profile Data

The averaged range profiles used in this dissertation were gathered using a high resolution radar, operating at 1kHz pulse repetition frequency. The profiles were then obtained using 200 sample block averaging of the pulse by pulse data in order to remove the speckle component. Four hundred and forty adjacent range profile pairs are used throughout this thesis. These were randomly selected from an extensive database of averaged sea clutter range profiles [Hirst and Baker, 1987]. Each pair of profiles can be used as a test/training set for the purpose of parameter estimation and target detection performance assessment in real clutter conditions.

The collection of the selected profiles is expected to be representative of a wide range of environmental and viewing conditions. To illustrate the statistical diversity of the profiles, histograms of the number of profiles with particular features in common are shown in figures 3.2(a) and 3.2(b). In particular, figure 3.2(a) shows the histogram of the modulating component order parameter, α , while figure 3.2(b) shows the histogram of the correlation length, ρ (*i.e.* the rate of decay of the autocorrelation function), for the database of averaged range profiles. Whilst the order parameter is indicative of the spikiness of the distribution of the profile, the correlation length is indicative of the amount of correlation present within a profile ($\rho = 1$ corresponds to fully correlated data). It can be seen from the figure that most of the profiles are partially correlated, with ρ in the range of about 0.5 and 0.9. Likewise, most of the profiles have the order parameter in the range of 0.5 for spiky clutter, and 10 for nearly Rayleigh distributed clutter.

Figure 3.3 illustrates the number of profiles as a function of log-likelihood deviation from log-Normal and Gamma distributions. In particular, for each of the profiles, Maximum Likelihood parameter estimates were obtained for Gamma, Log-Normal and Inverse Gamma distribu-

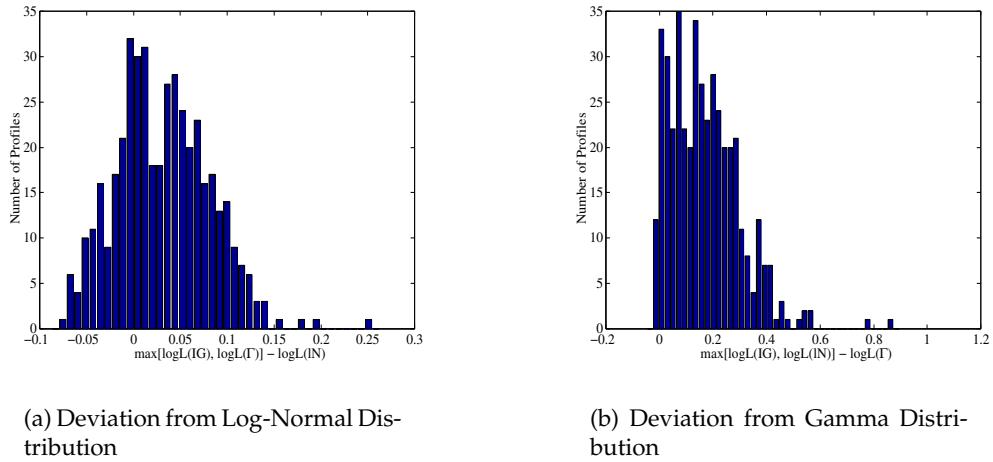


Figure 3.3: Histogram of the log-likelihood deviation from (a) log-Normal (lN) and (b) Gamma (Γ) distributions, for the database of 440 range profiles.

tions, and the corresponding likelihoods evaluated. The likelihoods give an indication of the probability with which the data was produced by the particular distribution, with the parameters at their ML values.

Figure 3.3(a) illustrates the deviation of the data (in the log-likelihood sense) from the log-Normal distribution, which is measured by how much the likelihood corresponding to either the Inverse Gamma or the Gamma distribution exceeds the log-Normal distribution likelihood for the range profile. Similarly, figure 3.3(b) illustrates the deviation of the data (in the log-likelihood sense) from the Gamma distribution, which is measured by how much the likelihood corresponding to either the Inverse Gamma or the log-Normal distribution exceeds the Gamma distribution likelihood of the data.

Based on these measures, the results presented in figure 3.3 show that while approximately half of the range profile data appears to be log-Normal distributed, hardly any profiles are well modelled as being Gamma distributed. However, the likelihood is not a particularly good measure for comparing models. It merely serves as an illustrative measure here, used to quantify the deviation of the profiles from log-Normality. No conclusions ought to be drawn on the basis of the likelihood alone. More pertinent measures, from the target detection point of view, are employed in the following section in order to further illustrate that the modulating component of sea clutter may, in fact, be better modelled as being log-Normal distributed, and not Gamma distributed as is reported in the literature [Ward et al., 1990].

3.5.2 Effects of $p(\sigma)$ on Probability of False Alarm Prediction

[Armstrong and Griffiths, 1991a] investigate the effects of the compound form of the K-distributed clutter on adaptive threshold detection. While the K-distributed model assumes the modulating component to be Gamma distributed, to the best of the author's knowledge, no study has been undertaken to analyse the effect of the deviation of the modulating component statistics from the Gamma distribution assumption in compound clutter environment.

As the results in the previous section indicate, based on the log-likelihood deviation from the log-Normal and the Gamma distributions for a large number of range profiles, the modulating component of sea clutter *may* better be modelled as being log-Normal distributed. The argument is now strengthened by estimating the probability of false alarm for a range profile, based on a number of different distributions for the statistics of the modulating component.

A horizontally polarised averaged range profile (see section 2.1) was used to obtain the Maximum Likelihood parameter estimates (using 10^5 averaged clutter data samples) for the Gamma, the Inverse Gamma and the Log-Normal distribution. The log-likelihood for the three distributions is given in the legend of figure 3.4. Based on the likelihood, the Gamma distribution is the least likely, while the Inverse Gamma distribution is marginally more likely than the log-Normal distribution to correspond to the data.

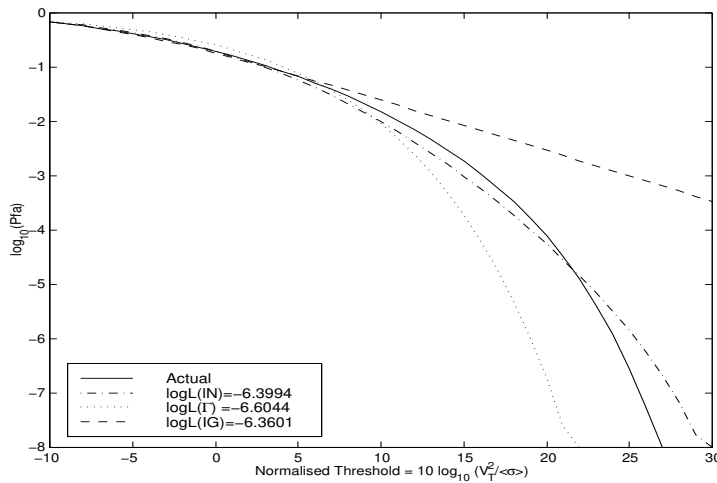


Figure 3.4: Effect of the form of $p(\sigma)$ on the probability of false alarm prediction for horizontally polarised sea clutter. Prediction results for Log-Normal (IN), Gamma (Γ) and Inverse Gamma (IG) distributions with ML parameter estimates are compared to the actual false alarm performance for the real data.

From a target detection perspective, it is the accuracy of the estimate of the probability of false alarm that provides a more pertinent measure of the deviation of the statistics of the clutter data from an assumed distribution. In particular, for a range profile $\{\sigma_n; n = 1, \dots, N\}$, using

fixed threshold detection with the threshold set to V_T , the *average* probability of false alarm is given by

$$P_{fa} = \frac{1}{N} \sum_{n=1}^N \exp(-V_T^2/\sigma_n)$$

For the three distributions used to model the statistics of the modulating component (*i.e.* Gamma, Inverse Gamma and log-Normal), the estimates of the probability of false alarm are shown in figure 3.4. These estimates were obtained using the ML estimates for the distribution parameters, and are compared to the actual probability of false alarm, as obtained from the knowledge of the actual values of the modulating component, $\{\sigma_n\}$.

The results presented in figure 3.4 illustrate that the Gamma distribution severely underestimates the probability of false alarm for a given threshold. This is particularly pronounced at low values of P_{fa} . For example, a threshold setting based on the Gamma distribution model for the modulating component predicts that $P_{fa} = 10^{-6}$, when the actual probability of false alarm is larger by a couple of orders of magnitude, *i.e.* $P_{fa} = 10^{-4}$. Therefore, modelling the modulating component of high resolution sea clutter as being Gamma distributed can result in an overly optimistic estimate of the probability of false alarm. Figure 3.4 also illustrates that when the modulating component is assumed to be Inverse Gamma distributed, the resulting estimate of the probability of false alarm tends to be rather conservative, while the estimate based on the log-Normal distribution assumption provides the best fit to the actual probability of false alarm over a wide range of thresholds.

The above analysis was extended to the database of 440 averaged range profiles. For each of the profiles in the database, a detection threshold was estimated using the range profile data, such that the actual average probability of false alarm for each profile, $P_{fa} = 10^{-6}$. For each of the profiles, based on the threshold setting above, the probability of false alarm was estimated assuming the modulating component to be Gamma, Inverse Gamma and Log-Normal distributed. The distribution parameters were estimated from the range profile data by Maximum Likelihood.

The probability of false alarm estimates obtained for each of the three distributions are shown in Figure 3.5. Figure 3.5(a) shows that the Gamma distribution assumption for the modulating component consistently results in the actual probability of false alarm being *underestimated*. Similarly, figure 3.5(c) shows that the Inverse Gamma distribution assumption for the modulating component consistently results in the actual probability of false alarm being *overestimated*. Finally, figure 3.5(b) shows that the log Normal distribution assumption for the modulating component of sea clutter consistently results in the most accurate estimate for the probability of false alarm.

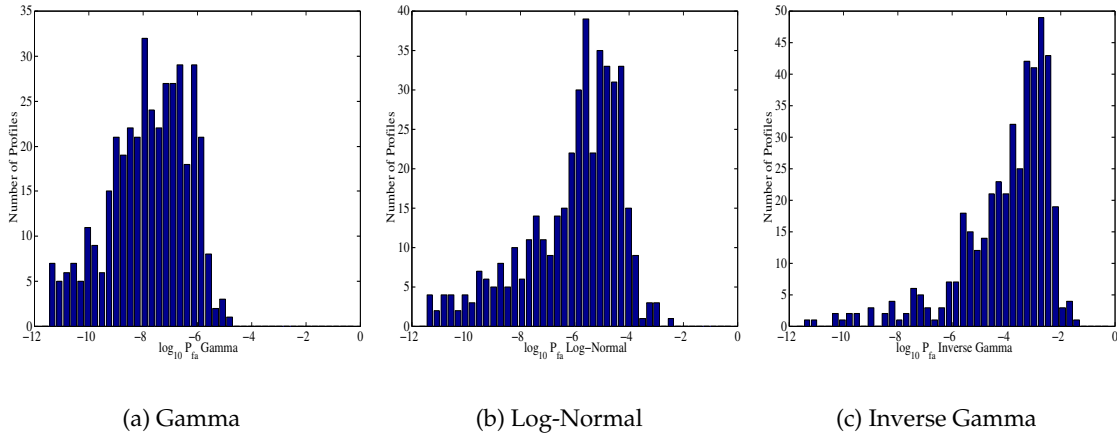


Figure 3.5: Histogram of the probability of false alarm estimates based on (a) Gamma, (b) log-Normal and (c) Inverse Gamma distributions for a threshold set so that the actual $P_{fa} = 10^{-6}$, for the database of 440 range profiles.

The above observations appear to contradict the findings in [Ward et al., 1990], where the authors have found the Gamma distribution to be the best model for the data. While the analysis presented by [Ward et al., 1990] is based on the first four normalised moments of the modulating component, σ , the present analysis is based on a more relevant moment, of the form $E[\exp(-V_T^2/\sigma)]$, which is of more practical interest in threshold setting and detection performance prediction. It is therefore very compelling to conclude that models other than the Gamma distribution may need to be sought for the modulating component of sea clutter. Furthermore, it may be the case that no single distribution captures the clutter statistics well, in which case a more flexible family of models would be desirable.

3.5.3 Target Detection in the Presence of Correlated Clutter

The previous section has dealt with the effects of the marginal statistics of the modulating component of sea clutter on false alarm rate prediction in compound clutter environment. The present section deals with the potential benefits of incorporating the knowledge of the spatial correlation structure of the modulating component, on target detection performance in compound clutter environment.

Figure 3.6 shows the spatial variations in the probability of false alarm and the probability of detection for a number of different detection schemes (to be subsequently discussed), for the *spatially correlated* range profile shown in figure 3.6(a). In each case, the average probability of false alarm across the profile is 10^{-4} and the average probability of detection is 0.5. Following [Armstrong and Griffiths, 1991a], the analysis presented herein is limited to single ‘pulse’ detection of Rayleigh fluctuating (*i.e.* Swerling I and II) targets, where the term ‘pulse’ refers here

to a single detection opportunity.

Figures 3.6(c) and 3.6(d) show the variations of local probability of false alarm and the probability of detection across the range profile for a fixed threshold detector (*i.e.* the threshold is constant across the range profile). Although the required probability of false alarm has been achieved across the range profile on *average*, *locally* the probability of false alarm can substantially exceed the specified level, sometimes by a few orders of magnitude. The high false alarm rates in some of the range cells result in an undue decrease in the false alarm rates in the remaining range cells of the profile, such that the overall desired average probability of false alarm is attained.

The effect of a fixed threshold in correlated clutter is that, while in some regions the effective threshold is locally set too low for the given clutter power (resulting in locally high P_{fa}), in other regions the threshold is locally set too high, resulting in an overall increase in the average signal to clutter ratio (SCR) required to achieve the desired average probability of detection for a given probability of false alarm.

It is interesting to note that the local probability of false alarm for an integrated clutter return in the n^{th} range cell, with the threshold set to $V_T(n)$ and the local clutter power level σ_n , is given by equation 3.7, *i.e.*

$$P_{fa}(V_T(n)|\sigma_n) = \frac{\Gamma(M, V_T(n)^2/\sigma_n)}{\Gamma(M)}$$

Clearly, to ensure spatially constant probability of false alarm, the threshold for the n^{th} range cell should be set to be proportional to the local clutter power level, *i.e.*

$$V_T(n) = G\sqrt{\sigma_n}$$

where G is adjusted for the required *average* probability of false alarm across the range profile. A detector which exactly adapts the threshold to the local clutter power level achieves a spatially constant rate of false alarm and is referred to as an *ideal constant false alarm rate* (ideal CFAR) detector [Watts, 1985].

The ideal CFAR detector assumes the exact knowledge of the local clutter power, which seldom is available in real radar systems. In fact, the fixed threshold detector and the ideal CFAR detector represent the two extremes of detection performance achievable in compound clutter environment. Both the fixed threshold and the ideal CFAR detectors are optimal (and identical) in spatially uncorrelated Rayleigh distributed clutter, while the ideal CFAR represents the best achievable performance in spatially correlated clutter [Watts, 1996].

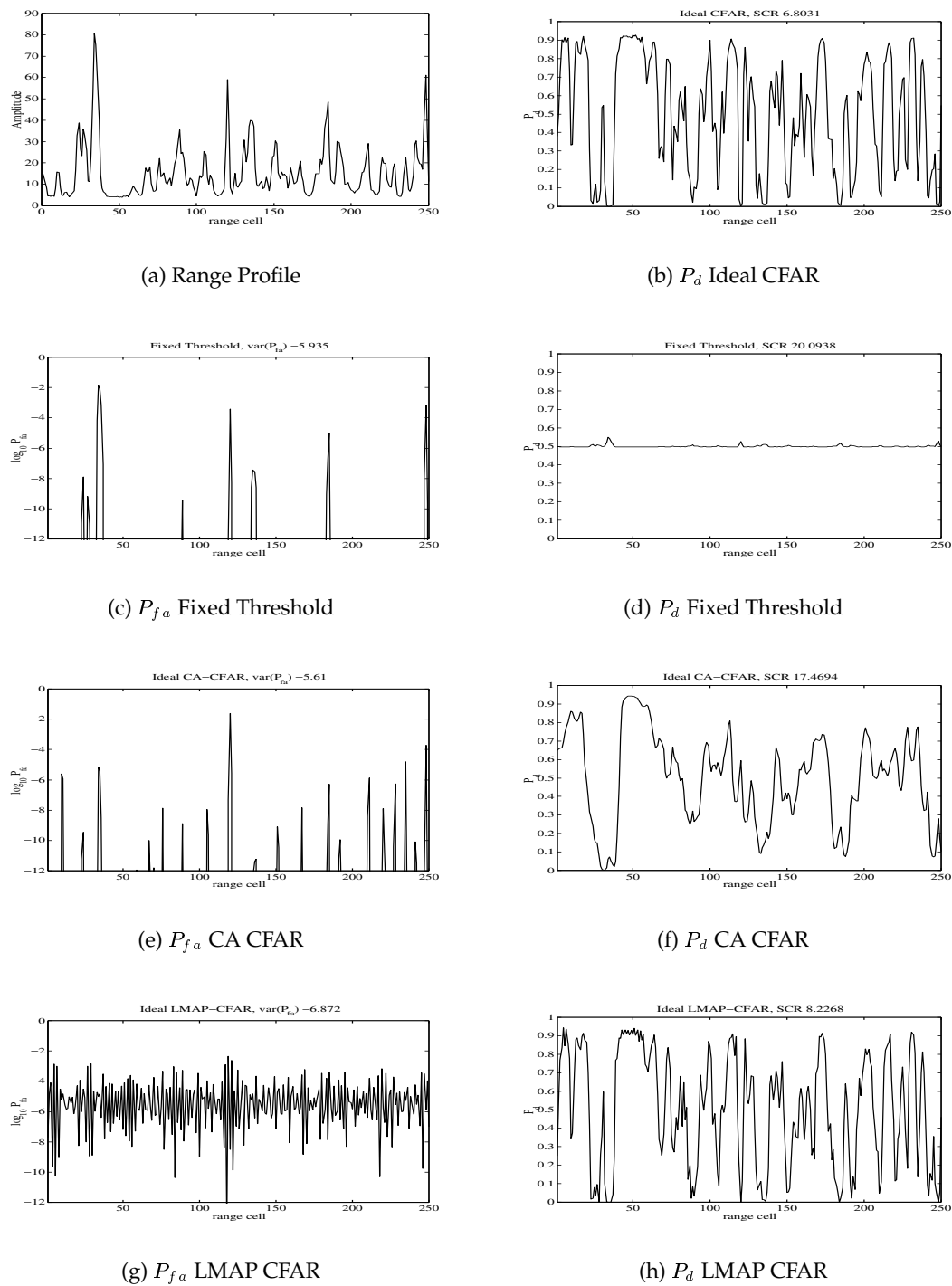


Figure 3.6: Spatial variations in probability of false alarm and detection across a range profile with partially correlated modulating component (order parameter $\alpha = 0.7$ and correlation length $\rho = 0.59$). The results are shown for a number of different detection schemes with the average $P_{fa} = 10^{-4}$ and the average $P_d = 0.5$, assuming single pulse detection and Swerling I and II targets.

Figure 3.6(d) and 3.6(b) show the spatial variation of the probability of detection for the range profile shown in figure 3.6(a), for the fixed threshold and the ideal CFAR detectors, respectively. The superior detection performance of the ideal CFAR over the fixed threshold detector is demonstrated by noting that the SCR required to achieve $P_d = 0.5$ with $P_{fa} = 10^{-4}$ is 7dB for the ideal CFAR and 20dB for the fixed threshold detector.

The large discrepancy in the detection loss (*i.e.* the increase in the SCR required to achieve a given average probability of detection) is further illustrated in figure 3.7(a). The SCR required to achieve an average probability of detection of 0.5 with an average probability of false alarm set to 10^{-4} is shown as a function of the compound distribution order parameter, α . The figure illustrates the SCR required for both the fixed threshold and the ideal CFAR detection schemes. The estimates have been obtained numerically, based on both the Gamma and the log-Normal distribution models for the statistics of the modulating component of sea clutter.

Figure 3.7(a) shows that the average SCR required to achieve the given probability of false alarm for a fixed threshold detector, relative to an ideal CFAR detector increases for compound clutter with decreasing order parameter (*i.e.* spiky clutter). For example, for the Gamma distributed modulating component with order parameter $\alpha = 10^{-0.5}$ the potential gain of 14dB in the required SCR can be achieved by using ideal CFAR instead of the fixed threshold detector. The potential gain is 24dB when the modulating component is assumed to be log-Normal distributed.

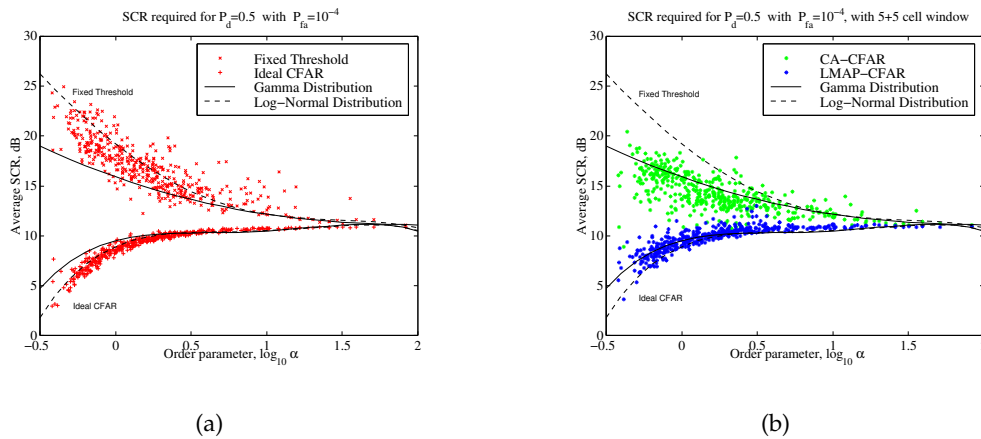


Figure 3.7: SCR required for $P_{fa} = 10^{-4}$ and $P_d = 0.5$ assuming single pulse detection and Swerling I and II targets. (a) Results obtained for 440 range profile database using a Fixed Threshold and an Ideal CFAR detectors are compared with the predictions based on the Gamma and the Log-Normal distribution assumption for the modulating component. (b) Results obtained for 440 range profile database using a 5+5 range cell reference window CA CFAR and a LMAP CFAR, when the reference window contains speckle free modulating component samples.

The previously mentioned results in figure 3.7(a) represent the theoretical performance achievable by the ideal CFAR and the fixed threshold detectors, under the assumption of the modulating component being Gamma and log-Normal distributed. Figure 3.7(a) also shows the *actual* SCR required to achieve $P_d = 0.5$ with $P_{fa} = 10^{-4}$ for a database of 440 averaged range profiles, for both the fixed threshold and the ideal CFAR detection schemes. The results presented in figure 3.7(a) illustrate that the SCR required to achieve the given detection performance for the range profiles in the database corresponds more closely to the numerical estimates obtained for the compound clutter with log-Normal distributed modulating component, rather than to the numerical estimates obtained for the compound clutter with Gamma distributed modulating component.

This observation further supports the findings of section 3.5.2, where the widely accepted assumption of the Gamma distributed modulating component was found to lead to overly optimistic estimates of the probability of false alarm in compound clutter environment.

3.5.4 Adaptive Threshold Estimation

In the previous section it was shown that potentially considerable improvements in detectability can be obtained in spatially correlated clutter by locally adjusting the threshold according to the local clutter power level. The maximum improvement is limited by the case of ideal CFAR detection, whereby the threshold is set based on the exact knowledge of the local clutter power.

In practice, the exact level of the local clutter power, σ_n , is not known and has to be inferred from the radar returns instead. A consequence of inferring the local clutter power from the radar returns (which can consist of targets, as well as the speckle modulated in power by the modulating component, σ) is that the threshold itself becomes a random variable. Therefore, in addition to the local fluctuations in the clutter power, the fluctuations in the threshold estimate need to be taken into account when evaluating the detection performance of an adaptive threshold detector. Any correlations between the threshold estimate and the radar return in the range cell being thresholded need to be taken into account in the analysis.

In practice, an estimate of the local clutter power for the range cell under test (denoted as range cell n) is obtained from the radar returns in a *reference window*, consisting of L range cells on either side of the range cell under test. In order to ensure statistical independence of the threshold estimate and the radar returns in the range cell under test, a small number, l , of *guard cells* may be inserted between the reference window and a range cell under test.

More concisely, denote a range profile of N (possibly pre-integrated) radar return amplitude samples by $\mathbf{r} = [r_1, \dots, r_N]'$. The range profile may consist of targets, as well as the background clutter, for which the corresponding local power level across the profile is given by

$\sigma = [\sigma_1, \dots, \sigma_N]^t$. Consider the radar return in the n^{th} range cell, r_n . Denote by \mathbf{r}_{-n} the $2(L-l)$ samples in the *reference window* (i.e. a range window immediately surrounding the n^{th} range cell under test, but *excluding* the cell under test), where

$$\mathbf{r}_{-n} = [r_{n-L}, \dots, r_{n-l-1}, r_{n+l+1}, \dots, r_{n+L}]^t$$

with l denoting the number of guard cells excluded from the reference window, in order to ensure the independence of the threshold from the sample under test.

Figure 3.8 illustrates a schematic diagram of a constant false alarm rate (CFAR) detector, where the threshold for the n^{th} range cell, is based on the local clutter power estimate, $\hat{\sigma}_n$, derived from the radar returns in the reference window. There are a number of issues that need to be considered when choosing the form of the local clutter power estimator.

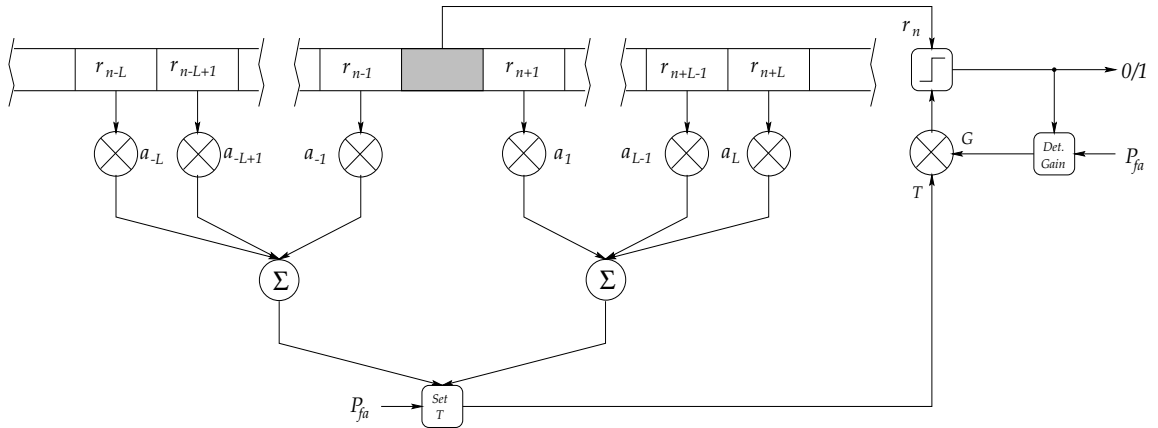


Figure 3.8: Schematic diagram of a CFAR detector. Detection threshold $V_T(n) = G \times T$ (where $T \equiv \sqrt{\hat{\sigma}_n}$) is estimated from the radar returns in the reference window. No guard cells are shown explicitly.

In particular, the local clutter power estimate, $\hat{\sigma}_n$, is obtained from the radar returns in the reference window, \mathbf{r}_{-n} . Such returns may not be target free, or the background clutter may not be homogeneous. For example, a number of interfering or spatially extensive targets may be present within the reference window, thus contaminating the background clutter observations, resulting in a potentially biased estimate of the background clutter power. Also, effects such as clutter edges (i.e. change of clutter reflectivity within the reference window) will similarly affect the estimates of the background clutter power.

If the surrounding environment is at least locally homogeneous (i.e. if the samples in the reference window are i.i.d. exponential variates), then an *efficient* estimator of the noise power is the sample mean. The corresponding CFAR detector, usually referred to as the cell-averaging or CA CFAR, is optimum in the sense of minimizing the detectability loss given homogeneity.

Unfortunately, CA CFAR does not prove to be particularly robust against interfering targets and clutter edges. While the presence of interferers inside the reference window generally leads to an overestimate of the actual noise power, the clutter edge may result in either over or underestimation, depending on whether the cell under test is in clear or in clutter. Either effect is undesirable, in that overestimation results in masking of legitimate targets, while underestimation results in false alarm rate inflation.

[Gandhi and Kassam, 1988] review a number of schemes that deal with some of the previously mentioned effects, including GO (greatest of), SO (smallest of), TM (trimmed mean) and OS (ordered statistics) CFAR. GO and SO detectors estimate the local power for the range cell under test by symmetrically partitioning the reference window about the test range cell. The local power estimate is then obtained as the greater and the smaller of the two estimates obtained from the two individual windows, respectively. These detectors are designed to deal with the problem of clutter edges, although they do not fare well in a multiple target scenario.

The trimmed mean detector obtains the local power estimate from the censored mean of the samples in the neighbouring range cells. The ordered statistics detector obtains the local power estimate by ordering the neighbouring samples, with the k^{th} smallest sample forming the estimate. Such estimators prove particularly robust in situations with multiple target interference. However, the ordered statistics estimators tend to ignore any potential correlation structure present in the clutter. [Lops and Willett, 1994] propose to use L1-filters to weigh the samples in the reference window not only by the rank, but also by the relative proximity of the sample to the cell under test. Other variants of the ordered statistics scheme also exist [He, 1994], but will not be reviewed here any further.

3.5.5 Detection Performance Analysis of CA CFAR Detectors

In the present section, a brief exposition on detection performance analysis of a CA CFAR detector is given. The discussion is limited to target-free, homogeneous background clutter (*i.e.* no clutter edges). The detection performance analysis is extended to other adaptive threshold detectors in chapter 6.

Based on the radar return samples in the reference window, r_{-n} , the CA CFAR detection threshold is obtained as a linear average of the amplitude returns within the reference window, *i.e.*

$$V_T(n) = G \sum_{k=-L, k \neq [-l, \dots, l]}^L a_k r_{n+k} \quad (3.8)$$

where G is the threshold gain, adjusted to ensure the correct *average* probability of false alarm. For a CA CFAR detector, all of the weights for the samples within the reference window are

equal (*i.e.* $a_k = 1/2(L - l)$), although sample weighting is possible, for potentially improved detection performance [Bucciarelli et al., 1996].

Since the radar returns in the reference window, \mathbf{r}_{-n} , are Rayleigh distributed random variables with the corresponding local power, σ_{-n} , the threshold derived from the samples in the reference window is itself a random variable with the distribution, conditional on σ_{-n} , denoted by $p(V_T(n)|\sigma_{-n}, \mathbf{a})$, where $\mathbf{a} = \{a_k\}$. Although the distribution $p(V_T(n)|\sigma_{-n}, \mathbf{a})$ is easily obtained in closed form for Rayleigh distributed clutter [Levanon, 1988], the form of this distribution is not easily obtained in the case of compound clutter, where the local power (*i.e.* the modulating component) is not spatially constant.

The *average* probability of false alarm and detection for range cell n can be obtained by averaging $P_{fa}(n|V_T(n), \sigma_n)$ and $P_d(n|V_T(n), \sigma_n, \bar{A})$ over the distribution of the threshold, *i.e.*

$$\begin{aligned}\hat{P}_{fa}(n|\sigma, \mathbf{a}) &= \int_{V_T(n)} P_{fa}(n|V_T(n), \sigma_n) p(V_T(n)|\sigma_{-n}, \mathbf{a}) dV_T(n) \\ \hat{P}_d(n|\sigma, \bar{A}, \mathbf{a}) &= \int_{V_T(n)} P_d(n|V_T(n), \sigma_n, \bar{A}) p(V_T(n)|\sigma_{-n}, \mathbf{a}) dV_T(n)\end{aligned}\quad (3.9)$$

where $P_{fa}(n|V_T(n), \sigma_n)$ and $P_d(n|V_T(n), \sigma_n, \bar{A})$ are the same as for the fixed threshold case, discussed in section 3.4. Both $\hat{P}_{fa}(n|\sigma, \mathbf{a})$ and $\hat{P}_d(n|\sigma, \bar{A}, \mathbf{a})$ need to be further averaged over all the possible σ , but, in practice, this can be achieved by studying the detection performance on averaged range profiles with known local clutter power [Watts et al., 1990]. Alternatively, Monte Carlo sampling techniques can be used to study the detection performance of the adaptive threshold detectors [Armstrong and Griffiths, 1991a].

To facilitate the analysis of CA CFAR target detection in spatially correlated, compound clutter environment, [Watts, 1996] proposes to approximate the distribution of the threshold by a Gaussian distribution. This can be justified by invoking the Central Limit Theorem, since the threshold is a weighted sum of random variates with finite variance. [Watts, 1996] argues that the following approximation

$$p(V_T(n)|\sigma_{-n}, \mathbf{a}) \approx N(\mu_V, \sigma_V)$$

works well, even for reference windows with as few as five range cells.

The mean and the variance of the Gaussian distribution can be obtained analytically for the case of Rayleigh distributed speckle, locally modulated in power by σ_{-n} . In particular, it can

be readily shown that

$$\begin{aligned}\mu_V &= G\sqrt{\pi/2} \sum_{k=-L, k \neq [-l, \dots, l]}^L a_k v_{n+k} \\ \sigma_V^2 &= G^2(2 - \pi/2) \sum_{k=-L, k \neq [-l, \dots, l]}^L a_k^2 v_{n+k}^2\end{aligned}$$

where $v_n = \sqrt{\sigma_n/2}$. Based on the Gaussian approximation to the distribution of the threshold, [Watts, 1996] derives an approximate expression for the probability of single pulse detection of fluctuating Swerling I and II targets in compound clutter environment as

$$\begin{aligned}P_d(n|\sigma, \bar{A}, \mathbf{a}) &= \frac{1}{2\sqrt{\alpha}} \exp\left[\frac{m_V^2}{2\sigma_V^2} \left(\frac{1}{\alpha} - 1\right)\right] \left(1 + \operatorname{erf}\left[\frac{m_V}{\sqrt{2\alpha\sigma_V^2}}\right]\right) \\ \alpha &= 1 + \frac{\sigma_V^2}{A^2 + v_n^2}\end{aligned}\quad (3.10)$$

3.5.6 Detection Performance of CFAR Detectors

The two main reasons for using an adaptive threshold for target detection were discussed in the previous section. Firstly, the threshold is adapted to the spatial variations in the local clutter power in order to obtain a more spatially constant probability of false alarm. Secondly, by adapting to the local clutter level, potential improvements in target detectability can be obtained in spatially correlated clutter. Maximum improvements are obtained using ideal CFAR detection, where the local clutter power is assumed to be known exactly.

As was discussed above, a number of schemes for local clutter power estimation have been proposed in the literature, for different clutter and target scenarios. In fully uncorrelated and spatially homogeneous clutter, all such schemes suffer from a *CFAR detection loss* compared to fixed threshold detection. The detection loss (*i.e.* increase in SCR required for a given level of P_d) arises from the finite size of the reference window, from which the clutter parameters, and hence the threshold setting is obtained. However, as is shown in figure 3.7(a), in spatially correlated clutter, potential detection gain, approaching that of the ideal CFAR detector, is possible, particularly in the case of spiky clutter.

[Armstrong and Griffiths, 1991a] undertake a study of the detection loss, relative to the ideal CFAR detection, of three different CFAR detectors. Their study encompasses the cell averaging (CA), greatest of (GO) and order statistics (OS) CFAR, under a range of correlated, compound clutter conditions. They find that in the uncorrelated clutter, all three detectors suffer a large

detection loss relative to ideal CFAR, particularly in the case of spiky clutter (this is indeed evident from figure 3.7(a), since the detectors effectively reduce to fixed threshold detection in the case of uncorrelated clutter). [Armstrong and Griffiths, 1991a] also find that the OS detector suffers notably higher loss than the CA and GO detectors, particularly for spiky clutter.

For completely correlated clutter, [Armstrong and Griffiths, 1991a] find that a significant detection gain can be obtained in spiky clutter, relative to detection in Rayleigh distributed clutter. This is particularly pronounced for moderate to low values of P_d . However, for moderately spiky clutter and most practical degrees of correlation, [Armstrong and Griffiths, 1991a] have found that the detectability loss relative to the ideal CFAR is close to that for completely uncorrelated clutter modulation. In short, the detectability loss suffered by the simple CFAR detectors studied by [Armstrong and Griffiths, 1991a] is close to the worst case loss, that is, the loss under conditions of the clutter modulation process being decorrelated from one range cell to the next.

[Bucciarelli et al., 1996] study a number of different schemes, which attempt to take the spatial correlation structure of the clutter into account when estimating the local clutter power from the reference window. They have found that the estimator that linearly averages the radar returns in the logarithm domain, where the weights are based on the covariance matrix of the modulating component, outperforms the CA CFAR scheme. Furthermore, for high enough correlations, the detection curve of their scheme (referred to as LMAP CFAR), is always close to the ideal CFAR and higher than the fixed threshold, while that of the CA CFAR detection scheme does not have these properties.

Figure 3.7(a) illustrates the SCR required for the given probability of detection in the case of the fixed threshold and the ideal CFAR detection. While in the ideal CFAR case the exact knowledge of the local clutter power in the range cell under test is assumed, it is interesting to consider another 'ideal' case, whereby the exact knowledge of the clutter power in the reference window, but not in the range cell under test, is available. Basically, the reference window is considered to contain *speckle free clutter*, based on which, an estimate of the local clutter power in the range cell under test is obtained.

Figure 3.7(b) shows the results for the database of 440 averaged range profiles, obtained using an ideal ¹ CA CFAR and an ideal LMAP CFAR detectors, with 5+5 range cell reference windows. The figure shows that while for most of the range profiles in the database only a small gain in the required SCR is obtained for the ideal CA CFAR detector relative to the fixed threshold detection, the gain obtained for the ideal LMAP CFAR detector approaches the gain

¹'ideal' refers to the idealised case of speckle free samples from the reference window, from which the local clutter power is estimated. This is in contrast to the ideal CFAR, where the local clutter power is known exactly, not only in the reference window, but also in the range cell under test.

achieved by the ideal CFAR detector for most of the range profiles in the database.

What this rather simplistic analysis shows is that simple cell averaging, even when no noise is present, provides only a marginal improvement in detectability relative to the fixed threshold detection. This is simply because simple cell averaging does not constitute a particularly good estimator for the local clutter power variations, unlike the LMAP detector, which is 'tuned' to the clutter spatial correlation structure.

To further illustrate the point, figures 3.6(e) and 3.6(g) illustrate the spatial variations of the probability of false alarm for the ideal CA CFAR and the ideal LMAP CFAR detectors, respectively, with the average $P_{fa} = 10^{-4}$ for the range profile illustrated in figure 3.6(a). The figure shows that the CA CFAR detector does not improve substantially in terms of the spatial variations in P_{fa} relative to the fixed threshold detector (figure 3.6(c)). In fact the log variance of P_{fa} for the ideal CA CFAR is -5.6, which is larger than the corresponding log variance of P_{fa} of -5.9 obtained using a simple fixed threshold detector. The log variance of P_{fa} of -6.8 is obtained using the ideal LMAP CFAR detector. The smaller variance is clearly evident in the reduced spatial variations of P_{fa} in figure 3.6(g) for the ideal LMAP CFAR, compared to either the fixed threshold, or the ideal CA CFAR detectors.

The spatial variations in the probability of detection for the ideal CA CFAR and the ideal LMAP CFAR are shown in figures 3.6(f) and 3.6(h), respectively. The SCR required for $P_d = 0.5$ is 17dB for the CA CFAR and 8dB for the LMAP CFAR, compared to 7dB required for the ideal CFAR and 20dB required for the fixed threshold detector.

It is evident from figures 3.6 and 3.7 that the CA CFAR is only capable of providing a rather modest improvement in terms of the SCR required for the given probability of detection, relative to the fixed threshold detector. The LMAP CFAR detector, on the other hand, is potentially capable of improvements approaching those of ideal CFAR, and substantially larger than those obtained using CA CFAR for most range profiles considered. This point is discussed further in chapter 6.

3.6 NEYMAN-PEARSON DETECTION

Neyman-Pearson (NP) Detection Rule [Kazakos and Papantoni-Kazakos, 1990] is the statistically optimal decision rule for target detection subject to maximum false alarm rate constraints. NP detection requires the full knowledge of both the background clutter and the target plus clutter likelihoods.

Consider N temporal radar return samples from a single range cell, denoted as $\mathbf{z} = [z_1, \dots, z_N]$. These returns can either consist of the background clutter alone (hypothesis H_0), or alternatively, the returns can consist of the target returns contaminated by the background clutter (hypothesis H_1). In particular, the two hypothesis can be stated as

$$H_0 : \quad \mathbf{z} = \mathbf{c}$$

$$H_1 : \quad \mathbf{z} = \mathbf{c} + \mathbf{p}$$

with \mathbf{c} denoting N clutter samples, while \mathbf{p} denotes the target contributions contained in the radar returns. Denoting by $f_0(\mathbf{z})$ and $f_1(\mathbf{z})$ the PDFs of \mathbf{z} conditioned to the hypothesis H_0 and H_1 , the optimal Neyman-Pearson detection test is given in the form of a likelihood ratio test [Pentini et al., 1992]

$$\lambda(\mathbf{z}) = \frac{f_1(\mathbf{z})}{f_0(\mathbf{z})}$$

where a detection is declared (*i.e.* H_1 is true) when $\lambda(\mathbf{z})$ exceeds some predefined threshold T . Furthermore, defining a decision region D in \mathbf{z} -space as the set of points \mathbf{z} such that $\lambda(\mathbf{z}) > T$, *i.e.*

$$D = \{\mathbf{z} | \lambda(\mathbf{z}) > T\}$$

the probability of false alarm and probability of detection are, respectively, given by

$$P_{fa} = \int_D f_0(\mathbf{z}) dD$$

$$P_d = \int_D f_1(\mathbf{z}) dD$$

with $dD = dz_1 \dots dz_N$. The detection threshold T is adjusted such that the required average probability of false alarm is achieved.

The Neyman-Pearson detection strategy discussed above is optimal in the sense that for a fixed value of the probability of false alarm, the probability of detection is maximised. It is interesting to note that the single pulse detection based on the amplitude of the returns in compound clutter environment need not be optimal in the Neyman-Pearson sense. This is particularly evident if the distribution of the background clutter (or the joint distribution of the background clutter and any potential targets) happens to be multi-modal.

The difficulty in using the Neyman-Pearson detection scheme stems from the need for the distribution for the likelihood of the background clutter, as well as the joint likelihood of a target and the background clutter. As was pointed out in chapter 2 no simple expression exists for the joint distribution of incoherent clutter, while in section 3.4 it was argued that no simple ex-

pression exists for a fluctuating Swerling target in partially correlated compound clutter environment. For these very reasons, the suboptimal target detection schemes discussed in sections 3.3-3.5 are the most commonly employed target detection techniques in radar systems.

A number of authors have proposed alternative suboptimal detection schemes based on the optimal Neyman-Pearson detector for target detection in coherent radar systems. In particular, [Conte et al., 1991b] propose to use SIRP models (discussed in chapter 2) to model the joint statistics of the background clutter, while the target is assumed to be known (and deterministic) to within an unknown initial phase and possibly amplitude. However, due to the difficulties of marginalising the unknown phase and amplitude, [Conte et al., 1991b] propose to use a suboptimal Generalised Likelihood Ratio Test (GLRT), whereby point estimates maximising the likelihood are used for the initial phase and amplitude. The properties of the proposed detector are further discussed by [Conte et al., 1994]. Despite a number of attractive features (including CFAR detection), the main criticism of the GLRT scheme is that it assumes a prior knowledge of the (deterministic) target, which renders such a technique overly restrictive for most practical situations.

[Farina and Lombardo, 1994] and [Gini et al., 1995] propose to extend the GLRT scheme to target detection in the presence of a mixture of K-distributed clutter and Gaussian (*e.g.* thermal) noise. [Lombardo and Farina, 1996] propose similar extensions to coherent detection of targets in the presence of compound clutter with partially correlated texture. Unfortunately, due to the complexity of the proposed schemes, the discussion presented by the authors is solely theoretical and no detection performance analysis is presented. Such complex schemes are not pursued any further in this thesis. Instead, a discussion of alternative techniques for target detection in coherent radar systems is now briefly presented.

3.7 TARGET DETECTION IN COHERENT RADAR SYSTEMS

Single pulse fixed threshold target detection based on the amplitude statistics of radar returns was discussed in section 3.3. As was shown in section 3.4, a substantial improvement in target detectability can be obtained by using pulse integration techniques, while the discussion in section 3.5 has shown that further improvements can be obtained by accounting for the spatial correlation structure of sea clutter.

There are many situations in which the amplitude distributions of the background clutter and the desired target heavily overlap, or alternatively, if the target is small, target masking by the background clutter may occur (*e.g.* icebergs in arctic waters). In such cases, in order to achieve performance beyond that achievable with adaptive threshold detection, techniques

such as polarisation diversity may be employed. It may also be that the relative motion between the desired target and its background permits the exploitation of differential Doppler shifts to isolate the desired target responses from large dominating (in amplitude) background returns.

Pulse integration provides a very crude way of exploiting the temporal correlation information in the background clutter, as the Doppler content of the returns is not fully utilised. Neyman-Pearson detection techniques discussed in section 3.6 constitute an optimal detection strategy in correlated clutter. However, as was pointed out in section 3.6, Neyman-Pearson detection techniques tend to be computationally expensive and require the exact knowledge of target statistics. In order to enhance the target detection performance in coherent radar systems without the computational complexity of the optimal Neyman-Pearson detector, a number of different techniques are possible. Chiefly, such techniques fall into two categories:

- pre-detection processing of the radar returns, followed by amplitude statistics-based detection techniques discussed in sections 3.3-3.5.
- detection based solely on the temporal correlation structure of radar returns (*i.e.* detection utilising only the Doppler, and not the amplitude, information contained in the returns).

3.7.1 Pre-detection Processing Techniques

The basic idea behind the pre-processing of the radar returns prior to target detection is to filter out the signal that is due to the background clutter, so as to increase the effective signal-to-clutter ratio at the target detection stage of the radar system. The schemes that use filtering to remove the background clutter contributions, and hence to enhance the moving target returns, are known as Moving Target Indicators (MTIs). In particular, an MTI radar is defined as one that uses simple band reject filtering to reject the return from fixed (stationary) targets. Enhanced detection and display of moving target responses are thereby achieved [Eaves and Reedy, 1987].

The Doppler filter is structured to have a desirable passband such that a relatively narrow bandwidth clutter is rejected. Since the target velocity is unknown *a priori*, the filter has a broad passband. The outputs of the filter are envelope detected and supplied as inputs to the post-Doppler processing stages, such as the incoherent integration and the adaptive threshold detection techniques discussed in sections 3.3-3.5. It is interesting to note that once the background clutter contributions have mostly been removed through the use of Doppler filtering, simple incoherent pulse integration will suffice to further increase the signal-to-clutter ratio for enhanced detection performance.

Many techniques are available for realising the MTI filter. Two particularly advantageous implementations are the delay line canceler and the range gate and filter [Eaves and Reedy, 1987].

They differ mainly by their implementation, though the functions they perform are essentially identical. In particular, the delay line canceler is implemented at the raw Intermediate Frequency (IF) signal sampling stage, where the clutter contribution at the current sample is canceled by subtracting a linear combination of samples stored in a delay line, with sample spacing being an integral multiple of the inter pulse separation. In the case of range gate and filter, the IF signal is first sampled into individual range cells, and then the clutter in each range cell is filtered to remove the frequency components associated with the background clutter.

3.7.2 Spectrum-based Target Detection Techniques

There exist alternatives to MTI and pre-detection filtering techniques in coherent radar systems. In particular, a pulsed Doppler radar can be used to look for targets moving at particular speeds. This is achieved by using a bank of filters for each range gate, each of which is then thresholded in order to look for moving targets [Eaves and Reedy, 1987]. In fact, this is a form of coherent integration mentioned in section 3.4 [Bucciarelli et al., 1996], except that the integrated returns are weighted so as to achieve Doppler frequency selectivity.

[Nohara and Haykin, 1993] present another notable detection scheme utilising the Doppler information contained in coherent clutter returns. In particular, [Nohara and Haykin, 1993] propose to model the temporal evolution of the radar returns as a complex autoregressive process. They find that the amplitudes of the poles of the complex autoregressive process corresponding to the growlers (small icebergs) lie closer to the unit circle than the amplitudes of the process poles corresponding to the background clutter (*i.e.* the Doppler shifts due to sea clutter are not as well defined as in the case of the solid target). This technique was shown by [Nohara and Haykin, 1993] to outperform an incoherent pulse integrated fixed threshold detector for a number of sea trial datasets.

The parametric approach of [Nohara and Haykin, 1993] is a medium-dwell time detection technique, most useful in the cases when the targets are small relative to the large dominating (in amplitude) background returns. This is particularly true when the relative motion between the desired target and its background permits the exploitation of differential Doppler shifts to isolate the desired target from the background clutter.

3.7.3 Extensions to Incoherent Radar Systems

Although the radar returns in incoherent radar systems do not contain the phase information necessary for Doppler processing techniques discussed above, it is possible to adapt an incoherent pulsed radar to exploit the Doppler shift so as to provide detection of targets that are moving relative to their background [Eaves and Reedy, 1987]. A number of such techniques

will now be briefly reviewed.

Radars that are classically configured as incoherent systems are sometimes used to detect and display targets that are in motion relative to their background. For example, over several inter-pulse periods, phenomenon known as the 'butterflies' [Eaves and Reedy, 1987] appears on a video display as a result of the fluctuating amplitude of the sum of the returns from both the background and the target. This observation forms the basis for an incoherent MTI technique, also known as a clutter-referenced or externally coherent MTI system. Incoherent MTI radars have, for many applications, been demonstrated to extend adequately the background clutter constrained detection limit. A number of constraints on the detection performance of the incoherent MTI radar include:

- limitations due to ambiguous range responses, which can seriously degrade clutter rejection performance.
- limitations due to signal-to-clutter and clutter-to-noise ratios.
- limitations due to a radar system inter-pulse amplitude stability.

In short, greater processing losses, relative to coherent, exist in incoherent MTI radars.

Other examples of techniques used in incoherent radar systems include phase-sensitive and coherent-on-receive schemes, which attempt to reconstruct a phase reference from the received signal (for a detailed discussion of such schemes, the interested reader is referred to the excellent text by [Eaves and Reedy, 1987]). Alternatively, detection schemes based on non-linear modelling of the temporal dynamics of sea clutter have been proposed in the literature (see for example [Haykin, 1996]).

As was discussed in chapter 2, the apparent success of target detection schemes based on the non-linear models of temporal dynamics of sea clutter could largely be due to the non-linearities arising from observing only the amplitudes of a complex, linear stochastic process (*e.g.* an autoregressive process). Following this lead, the complex autoregressive model-based scheme for coherent radar systems (as discussed in the previous section) is extended to incoherent radar systems in the following chapter.

3.8 SUMMARY

The main purpose of this chapter was to present a review of target detection techniques in compound clutter environment. However, in addition to a review of statistical target detection techniques, an interesting result was presented in this chapter. In particular, based on

a database of averaged range profiles, it was shown that the modulating component of high resolution sea clutter, contrary to widely held beliefs, is better modelled as being log-Normal distributed, rather than being Gamma distributed.

A general class of statistical target models, known as the Swerling target models, was reviewed in section 3.2. Single pulse target detection in the presence of uncorrelated background clutter was reviewed in section 3.3 for the case of Swerling 0, I and II targets. As was pointed out in section 3.3, the single point amplitude statistics of the target and the background clutter may heavily overlap. This results in an undesirably large signal to clutter ratio required for the given probability of detection.

A number of pulse integration schemes were reviewed in section 3.4. In particular, it was shown that a substantial reduction in the SCR required for the given detection level can be obtained by employing incoherent integration prior to the detection circuit, or by employing binary integration after the single pulse detection circuit. It was also shown that when pulse integration is employed, the target detection performance is highly dependent on the fluctuation characteristics of the targets and the background clutter.

As was pointed out in section 3.5, a potentially large increase in detectability (as measured by the SCR required for the given probability of detection) can be obtained in compound clutter environment by adapting the detection threshold to the local clutter power level. This observation forms the basis for adaptive threshold detection, also known as the constant false alarm rate (CFAR) detection. The maximum attainable detectability is determined by the exact knowledge of the local clutter power level, and is referred to as ideal CFAR.

The extent of the improvement in detectability achievable by employing adaptive threshold detection schemes is also determined by the extent of spatial correlations and the weight of the tails of the distribution of the background clutter. In fact, it was shown in section 3.5 that the improvement in detectability achievable using ideal CFAR over fixed threshold detection is larger in the case of the modulating component being log-Normal distributed, than in the case of the modulating component being Gamma distributed. Furthermore, based on the results obtained for a large database of high resolution sea clutter range profiles, it was shown in section 3.5 that, contrary to the widely held belief, rather than being Gamma distributed, the modulating component of sea clutter is better modelled as being log-Normal distributed.

Another fundamental result, as borne out by the analysis in section 3.5, is that the target detection performance of a simple CA CFAR detector is inherently incapable of approaching the ideal CFAR detector target detection performance. Instead, simple linear weighting (as determined by the correlation structure of the modulating component in the logarithm domain) of radar returns in the reference window was shown to provide a substantial improvement over

the CA CFAR, approaching the ideal CFAR target detection performance. A further study of adaptive threshold detection techniques in compound clutter environment is undertaken in chapter 6.

Finally, in sections 3.6 and 3.7, a brief review of target detection techniques in coherent radar systems was presented. These broadly fall into three areas. Firstly, the radar returns can be pre-processed to reduce the effects of the clutter (*e.g.* by filtering), followed by target detection based on the envelope of the filtered returns. Secondly, an optimal Neyman-Pearson detector can be employed to detect targets in coherent clutter. Unfortunately, the usefulness of such a scheme is limited by the need for a specific target model. Lastly, spectral properties of the radar returns can be used for the purpose of target detection in coherent clutter. The spectrum-based target detection provides the motivation for the work undertaken in chapter 4.

Correlation Structure of Incoherent Speckle

4.1 INTRODUCTION

High resolution sea clutter is a compound process consisting of the speckle component, modulated in power by the underlying modulating component associated with the swell. The speckle is largely spatially uncorrelated and exhibits short term temporal correlations. The modulating component, although spatially correlated, can be considered to be effectively constant on the scale of a few hundred milliseconds. The correlation structure of either of the two components is largely independent of the other.

As a result of the disparate correlation structure of the two components, the speckle and the modulating component of high resolution sea clutter can be treated independently. To this end, the discussion in the present chapter is limited in scope to the modelling of medium dwell time (*i.e.* on the scale of less than one second) temporal correlations of high resolution sea clutter, which are largely associated with the speckle. The treatment of the spatial correlations, associated with the modulating component, is deferred to later chapters.

The motivation for modelling the correlation structure of clutter returns stems from the observation that sea clutter returns generally have larger spectral widths than the spectral widths of some target returns (*e.g.* growler returns). [Nohara and Haykin, 1993] propose to use a complex autoregressive process to characterise the Doppler spectra of sea clutter returns in coherent radar systems. In particular, it is the poles of a complex autoregressive process that form the basis for a design of a coherent target detector.

The research presented in this chapter is a direct extension of the parametric target detection approach reported by [Nohara and Haykin, 1993]. In particular, a complex autoregressive process is proposed as the basis for characterisation of high resolution sea clutter spectra in incoherent radar systems. Since no phase information is available, any Doppler information contained in the incoherent radar returns is expected to be manifested through the position of the poles of the underlying complex autoregressive process. It is the aim of the present chapter to investigate whether, as is the case in coherent radar systems, the poles corresponding to the underlying complex autoregressive process can potentially be used for characterisation of medium dwell time incoherent high resolution sea clutter.

The remainder of the present chapter is structured as follows. In section 4.2, a model based on the complex autoregressive process is proposed for high resolution *incoherent* sea clutter. Hybrid Monte Carlo algorithm is briefly reviewed in section 4.3, while in section 4.4, the Hybrid Monte Carlo algorithm is employed within a Gibbs sampling framework in order to explore the posterior distribution of the incoherent clutter model parameters. Finally, the summary of the results is presented in section 4.5, along with some conclusions and suggestions for future research.

4.2 AUTOREGRESSIVE MODELLING OF SEA CLUTTER

Figure 4.1 shows the pulse by pulse time history of the amplitudes of horizontally and vertically polarised coherent sea clutter from a single range cell, along with the corresponding spectrograms obtained using 128 point FFTs. In particular, figures 4.1(a) and 4.1(c) clearly illustrate the presence of the fast fluctuating speckle component, and the slower fluctuating modulating component. The spectrograms shown in figures 4.1(b) and 4.1(d) illustrate the evolution of the Doppler spectra of the background clutter, associated with the short term fluctuations of the speckle component. The Doppler spectra are usually very dispersed, as illustrated in figures 4.1(b) and 4.1(d). Furthermore, the Doppler spectra change over time in line with the wave motion, as determined by the swell (*i.e.* the modulating component).

As was explained in chapter 3, the Doppler spectra can be utilised for enhanced target detection in one of two ways:

- pre-detection processing of the radar returns, followed by amplitude statistics-based detection techniques discussed in sections 3.3-3.5.
- detection based solely on the temporal correlation structure of radar returns (*i.e.* detection utilising only the Doppler, and not the amplitude, information contained in the returns).

The pre-detection filtering effectively enhances the signal to clutter power ratio, thus potentially resulting in an improved target detection performance. Unfortunately, the pre-detection filtering may also remove some of the energy associated with the target returns. Target detection based on the amplitude statistics following the pre-detection filtering may thus become infeasible, particularly in the case of small targets buried in strong background clutter.

As an alternative, target detection can also be based purely on the temporal correlation structure of radar returns. It is the temporal correlation structure based approach that is adopted in this chapter. To this end, section 4.2.1 deals with the case of a coherent radar system, while the analysis is extended to an incoherent radar system in section 4.2.2

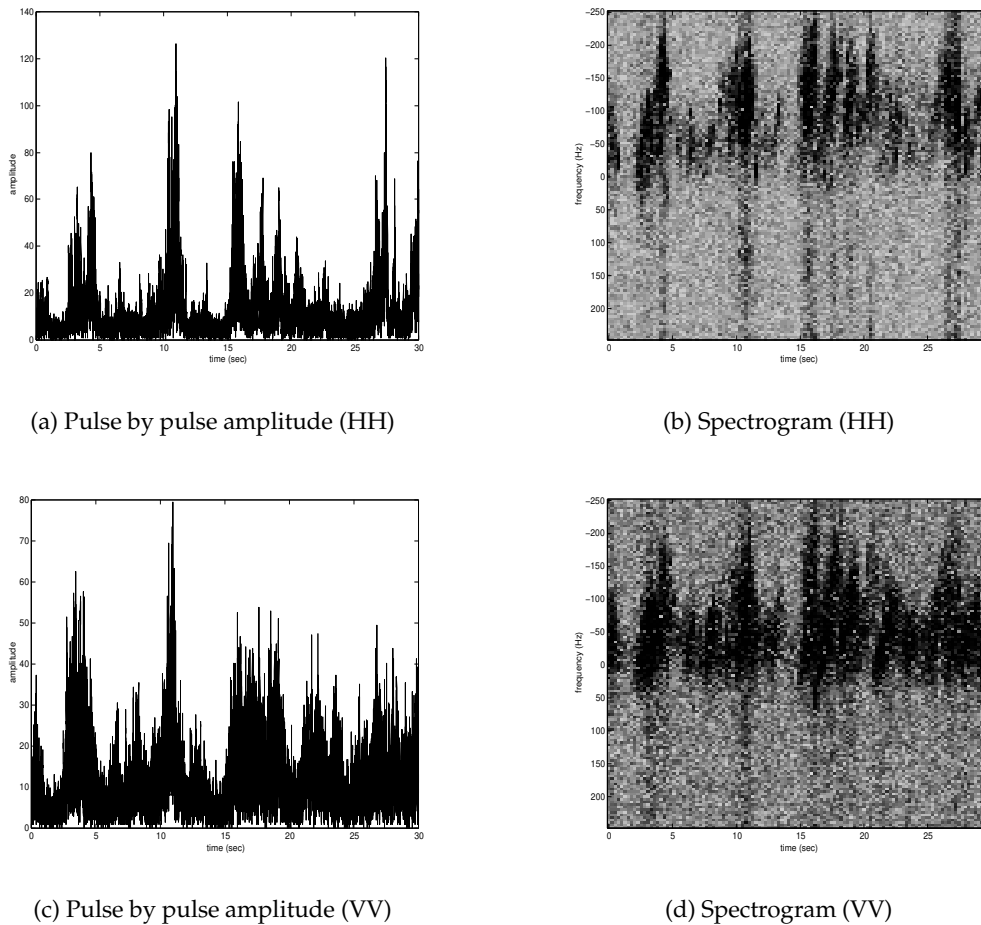


Figure 4.1: Pulse by pulse time history of the amplitude of (a) horizontally (HH) and (c) vertically (VV) polarised coherent sea clutter from a single range cell, along with the corresponding spectrograms (see (b) and (d)) obtained using 128 point FFTs.

4.2.1 Coherent Radar System

Both the phase and the amplitude information is contained in coherent radar returns. The Moving Target Indicator (MTI) scheme [Eaves and Reedy, 1987] utilises the phase information by linearly band-passing the raw IF radar signal prior to detection in order to reduce the spectral contributions from the background clutter. Instead of using linear filtering (as in the MTI scheme), [Haykin and Li, 1995] and [Haykin, 1996] advocate to use a non-linear predictor to better remove the background clutter contributions from the radar returns.

As an alternative to pre-detection filtering of the radar returns, [Nohara and Haykin, 1993] proposes to use a complex autoregressive process to characterise the medium dwell-time Doppler spectra of sea clutter returns with just a few parameters. These parameters can then be incorporated into the design of a coherent detector, as discussed at length in [Nohara and Haykin, 1993]. Such a detector attempts to take an advantage of the spectral width differences between the tar-

get and the clutter spectra. This technique has proven to be particularly useful for the detection of small growlers in the presence of sea clutter.

The present discourse is limited in scope to parametric modelling of the Doppler spectra of radar returns, while the issues associated with target detection based on the model parameters are left as an exercise for future research. To this end, denote a complex radar return sample at time n by $\mathbf{x}(n) = x_I(n) + x_Q(n)$ with the corresponding radar return amplitude $r(n) = \sqrt{x_I(n)^2 + x_Q(n)^2}$ and phase $\theta(n) = \tan x_Q(n)/x_I(n)$. The complex autoregressive process for the radar return $\mathbf{x}(n)$ can be written as

$$x_I(n) = \sum_i (a_{Ii}x_I(n-i) - a_{Qi}x_Q(n-i)) + \sigma e_I(n) \quad (4.1)$$

$$x_Q(n) = \sum_i (a_{Qi}x_I(n-i) + a_{Ii}x_Q(n-i)) + \sigma e_Q(n) \quad (4.2)$$

or, in more concise notation ¹

$$\mathbf{x}(n) = \sum_{i=1}^p \mathbf{a}_i \mathbf{x}(n-i) + \sigma \mathbf{e}(n) \quad (4.3)$$

where $\{\mathbf{a}_i = a_{Ii} + a_{Qi}; i = 1, \dots, p\}$ are the complex process parameters, σ^2 is the variance of the innovation process, \mathbf{e} , and $\mathbf{e}(n)$ is an i.i.d. zero mean, unity variance, spherically symmetric complex Gaussian process.

Based on the uniform priors for \mathbf{x} and \mathbf{a} and the uninformative Jeffrey's prior for σ^2 (i.e. $p(\sigma^2) \propto 1/\sigma^2$), the conditional densities corresponding to N complex process samples $\mathbf{x} = [\mathbf{x}(1), \dots, \mathbf{x}(N)]'$ and to the complex autoregressive process parameters $\mathbf{a} = \{\mathbf{a}_i\}$ and σ^2 , can be shown to be given by ²

$$p(\mathbf{x}|\mathbf{a}, \sigma^2) = N(\mathbf{0}, \Sigma_x) \quad (4.4)$$

$$p(\mathbf{a}|\mathbf{x}, \sigma^2) = N(\mu_a, \Sigma_a)$$

$$p(\sigma^2|\mathbf{x}, \mathbf{a}) = IG(\alpha_\sigma, \beta_\sigma) \quad (4.5)$$

$$\Sigma_x = \sigma^2(\mathbf{A}^H \mathbf{A})^{-1} \quad (4.5)$$

$$\mu_a = (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{x} \quad (4.6)$$

$$\Sigma_a = \sigma^2(\mathbf{X}^H \mathbf{X})^{-1} \quad (4.7)$$

$$\alpha_\sigma = N \quad (4.8)$$

$$\beta_\sigma = \frac{1}{2} \mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x} \quad (4.9)$$

¹In the vector notation adopted throughout this chapter \mathbf{x}^* denotes an element-wise complex conjugate of the vector \mathbf{x} . Similarly, \mathbf{x}^H denotes the Hermitian operator (i.e. a conjugate transpose operation), while \mathbf{x}' denotes the usual transpose operation on \mathbf{x} .

²This follows from $\sigma \mathbf{e} = \mathbf{x} - \mathbf{X} \mathbf{a} = \mathbf{A} \mathbf{x}$.

In the above equations \mathbf{A} is an $N \times N$ matrix of complex AR process parameters, while \mathbf{X} is an $N \times p$ matrix of the complex radar return samples³.

The *Maximum A Posteriori* (MAP) estimate of the complex autoregressive process parameters, \mathbf{a} , can be obtained from equation 4.4. Once the autoregressive parameters are calculated, the autoregressive spectrum estimate can be computed using the following equations

$$S(\omega) = \frac{\sigma_\nu^2}{\left|1 - \sum_{k=1}^p \mathbf{a}_k e^{-ik\omega}\right|^2} \quad (4.10)$$

$$\sigma_\nu^2 = \rho(0) - \sum_{k=1}^p \mathbf{a}_k^* \rho(k)$$

where $-\pi \leq \omega \leq \pi$, and where $\rho(k)$ corresponds to the autocorrelation of the process at k^{th} lag.

The autoregressive spectrum can similarly be obtained from the poles corresponding to the complex autoregressive process parameters, \mathbf{a} . In particular, the closer a pole is to the unit circle, the larger the peak of the spectrum at the frequency determined by the phase of the pole. [Nohara and Haykin, 1993] has found that the spectral width of the growler returns is usually much narrower than the spectral width of the background clutter. In fact, the spectral width seems to be the only characteristic of the Doppler spectrum that uniquely distinguishes between clutter and growler spectra. Since the spectral width is independent of the mean frequency and the growler-to-clutter ratio, it is an ideal statistic upon which to base a detector.

While a complex autoregressive process-based estimate of the Doppler spectrum is relatively simple to obtain for a coherent radar system, this is no longer the case when no phase information is available. The case of modelling of the incoherent clutter correlation structure is dealt with further in the following section.

4.2.2 Incoherent Radar System

It was argued in chapter 2 that the high resolution coherent radar clutter is locally complex Gaussian distributed. The principle difference between an incoherent and a coherent radar system lies in the availability of a stable phase reference. In an incoherent radar system, only the amplitudes, r , (and not the phases, θ) of the returns are observed. In the present chapter, the incoherent radar clutter is considered to correspond to the amplitudes of a complex Gaussian distributed process. Therefore, given the covariance matrix of the underlying complex Gaussian process, Σ_x , upon change of coordinates from quadrature to polar, the marginal

³For notational simplicity, the end effects are not included in the present chapter.

distribution of the amplitudes of the complex Gaussian process is given by

$$p(\mathbf{r}|\Sigma_x) = \int p(\mathbf{r}, \theta|\Sigma_x) d\theta$$

It is shown in appendix A that the single point statistics of \mathbf{r} are Rayleigh distributed, while the autocovariance of the amplitudes, \mathbf{r} , is related to the autocorrelation of the underlying complex Gaussian process, \mathbf{x} , through

$$\rho_r(k) \approx \rho_{II}^2(k) + \rho_{IQ}^2(k) \quad (4.11)$$

where $\rho_{II}(k)$ and $\rho_{IQ}(k)$ are the real and imaginary parts of the autocorrelation function of the complex Gaussian process, \mathbf{x} , at k^{th} lag.

The power spectrum is related to the autocorrelation function of the returns through a Fourier transform. As a consequence of the relationship in equation 4.11, the Doppler frequency information contained in the power spectrum of the underlying complex Gaussian process is largely lost when only the amplitudes of the returns are observed. Furthermore, the second order statistics are insufficient to fully characterise the (Rayleigh) distribution of the amplitudes corresponding to the complex Gaussian clutter. However, as will now be discussed, this does not imply that the incoherent returns contain no information about the spectrum of the underlying complex Gaussian process.

In particular, using equation 4.11 it is easy to show that although the *absolute* Doppler frequency information is lost, the *relative* frequency contributions contained in the coherent clutter Doppler spectrum are preserved in the spectrum of incoherent clutter. Furthermore, equation 4.11 can also be used to show that in the case of a single pole complex autoregressive process, the information about the magnitude of the complex process pole is preserved in the spectrum of the amplitudes. In fact, this is best illustrated by the worst case scenario (from the second order statistics point of view), *i.e.* when the power spectrum of the coherent clutter contains just a single peak. In this case no Doppler frequency information is contained in the power spectrum of the amplitudes of the clutter. However, one vital piece of information still remains in this extreme case, in that no peaks in the power spectrum of incoherent clutter indicate a process pole of unity magnitude, but of unknown phase.

In the remainder of this chapter, a complex autoregressive process-based characterisation of high resolution sea clutter spectra is adopted for incoherent radar systems. This results in a non-linear state space model for the observed locally Rayleigh distributed amplitudes, \mathbf{r} , as

given by

$$\begin{aligned}
 \text{Measurement Equation} \quad & r(n) = \sqrt{x_I(n)^2 + x_Q(n)^2} & (4.12) \\
 \text{State Transition Equation} \quad & \begin{cases} x_I(n) = \sum_i (a_{Ii}x_I(n-i) - a_{Qi}x_Q(n-i)) + \sigma e_I(n) \\ x_Q(n) = \sum_i (a_{Qi}x_I(n-i) + a_{Ii}x_Q(n-i)) + \sigma e_Q(n) \end{cases}
 \end{aligned}$$

It is the magnitude of the underlying complex autoregressive process poles that correspond to the spectral widths of the peaks in the Doppler spectrum of the incoherent returns. It was already mentioned that in the case of the first order autoregressive process the information about the magnitude of the process pole is preserved in the process amplitudes. However, based on the second order moment arguments, ambiguities in the innovation process variance are expected even for the first order autoregressive process, due to the unknown phase of the process poles. Unfortunately, due to the non-Gaussian nature of the incoherent clutter returns, moments higher than of the second order need to be considered when trying to establish process identifiability conditions for complex autoregressive processes of higher orders. This, however, proves to be no easy task.

Finally, it is worth noting that the state space model can be utilised for target detection in the presence of sea clutter in one of two ways:

- following [Haykin, 1996], detection can be based on the predictive distribution for the incoherent clutter, *i.e.* $p(r(n)|\mathbf{r}(n-1))$.
- following [Nohara and Haykin, 1993], detection can be based on the parameter structure of the underlying complex Gaussian process, *i.e.* $p(\mathbf{a}|\mathbf{r})$.

Unfortunately, in either case, due to the non-linear nature of the observation equation, the distributions of interest are not available in closed form. For this reason, a sampling based approach is adopted in this chapter. Furthermore, for practical reasons, the computational expense associated with sampling from the full predictive conditional distribution at each radar return sample (at 1kHz in real time) prohibits any further consideration of the predictive distribution based approach. It is the latter, medium dwell time block-based approach of [Nohara and Haykin, 1993] that remains the focus of the work undertaken in the remainder of this chapter. To this end, samples of the autoregressive process parameters, \mathbf{a} , are obtained from $p(\mathbf{a}|\mathbf{r})$ in order to characterise the high resolution sea clutter spectrum in incoherent radar system.

4.2.3 Sampling from $p(\mathbf{a}|\mathbf{r})$

The state space model in equation 4.12 has been developed in terms of the in-phase and quadrature components for the underlying complex Gaussian distributed process. However, in order

to better facilitate the analysis, it proves to be beneficial to transform the problem into polar components, *i.e.* in terms of the observed amplitudes, \mathbf{r} , and missing phases, Θ .

Under the polar parametrisation, the conditional structure of the model is given by

$$p(\Theta|\mathbf{r}, \mathbf{a}, \sigma^2) \propto p(\mathbf{r}, \Theta|\mathbf{a}, \sigma^2) \quad (4.13)$$

$$p(\mathbf{a}|\mathbf{r}, \Theta, \sigma^2) = N(\mu_a, \Sigma_a) \quad (4.14)$$

$$p(\sigma^2|\mathbf{r}, \Theta, \mathbf{a}) = IG(\alpha_\sigma, \beta_\sigma) \quad (4.15)$$

where uniform priors were assumed for the phases, Θ , and the parameters, \mathbf{a} , while an uninformative Jeffrey's prior was used for the process variance, σ^2 . The distribution parameters are defined as in equations 4.5-4.9, and the conditional distribution for Θ is in non-standard form, given by

$$p(\Theta|\mathbf{r}, \mathbf{a}, \sigma) \propto \frac{1}{(2\pi\sigma^2)^N \prod_n r_n} \exp\left(-\frac{1}{2\sigma^2} \sum_{n=1}^N (f_I(n) + f_Q(n))\right) \quad (4.16)$$

where

$$f_I(n) = \left(r_n \cos \theta_n - \sum_i (a_I(i)r_{n-i} \cos \theta_{n-i} - a_Q(i)r_{n-i} \sin \theta_{n-i}) \right)^2$$

$$f_Q(n) = \left(r_n \sin \theta_n - \sum_i (a_I(i)r_{n-i} \sin \theta_{n-i} + a_Q(i)r_{n-i} \cos \theta_{n-i}) \right)^2$$

Unfortunately, as was pointed out in section 4.2.2, no simple expression exists for the posterior distribution of the process parameters, $p(\mathbf{a}|\mathbf{r})$. Therefore, a sampling-based approach is adopted here in order to explore the parameter posterior. This is most easily achieved by noting that convergence to the joint ensures convergence to the marginal. To this end, samples from $p(\mathbf{a}|\mathbf{r})$ are obtained by sampling for $[\mathbf{a}, \sigma^2, \Theta]$ from the joint posterior distribution $p(\mathbf{a}, \sigma^2, \Theta|\mathbf{r})$. Based on the conditional distribution representation in equations 4.13-4.15, the Gibbs sampler [Geman and Geman, 1984] provides means for sampling from the joint posterior distribution.

The Gibbs sampler is one of the most widely used sampling algorithms (see [Wu, 1998] for an extensive review of sampling techniques). Its popularity stems from the relative simplicity of implementation. In particular, it is often simpler to sample from the conditional distributions, than it is to sample directly from the joint. In the present case, the Gibbs sampling algorithm for sampling from the joint posterior distribution of $[\mathbf{a}, \sigma^2, \Theta]$ can be stated as

-
1. At iteration t , given a sample $[\Theta^{(t-1)}, \mathbf{a}^{(t-1)}, (\sigma^2)^{(t-1)}]$ from the previous iteration, consecutive sample $[\Theta^{(t)}, \mathbf{a}^{(t)}, (\sigma^2)^{(t)}]$ is obtained as

$$\begin{aligned}\mathbf{a}^{(t)} &\sim p(\mathbf{a}|\mathbf{r}, \Theta^{(t-1)}, (\sigma^2)^{(t-1)}) \\ (\sigma^2)^{(t)} &\sim p(\sigma^2|\mathbf{r}, \Theta^{(t-1)}, \mathbf{a}^{(t)}) \\ \Theta^{(t)} &\sim p(\Theta|\mathbf{r}, \mathbf{a}^{(t)}, (\sigma^2)^{(t)})\end{aligned}$$

2. Increment t and go to step 1
-

Upon completion of the Gibbs sampling algorithm, the samples of \mathbf{a} from $p(\mathbf{a}, \sigma^2, \Theta|\mathbf{r})$ are retained as the samples from the marginal posterior distribution, $p(\mathbf{a}|\mathbf{r})$.

While the conditional distributions for \mathbf{a} and σ^2 are in standard form, the conditional distribution for Θ is not. Therefore, to circumvent the difficulty associated with the sampling from the conditional distribution of the missing phases, $p(\Theta|\mathbf{r}, \mathbf{a}, \sigma)$, the Hybrid Monte Carlo algorithm will now be discussed in section 4.3 and employed in section 4.4 for sampling from the joint posterior within the Gibbs sampling framework.

4.3 HYBRID MONTE CARLO ALGORITHM

Within the Gibbs sampling framework discussed in section 4.2.3, samples of the missing phases are required from the conditional distribution $p(\Theta|\mathbf{r}, \mathbf{a}, \sigma^2)$. Unfortunately, the conditional distribution for Θ (see equation 4.16) is in non-standard form and hence cannot be sampled from directly. Furthermore, the dimensionality of the missing phases vector, Θ , is the same as the dimensionality of the vector of observed amplitudes, \mathbf{r} . In the present chapter, this corresponds to 150 samples, though in real radar systems, blocks of as many as 500 samples may need to be considered.

Sampling from high dimensional distributions is a problem that is encountered in a wide range of applications, such as missing data and state space problems. Unfortunately, only in very few cases can the specific problem structure be exploited for efficient sampling from a high dimensional distribution (*e.g.* state sampling in a Gaussian state space model). Instead, the two most popular techniques for sampling from high dimensional distributions include the Gibbs sampler [Geman and Geman, 1984] and the Metropolis-Hastings algorithm [Metropolis et al., 1953] [Hastings, 1970].

Both the Gibbs sampler and the Metropolis-Hastings algorithm belong to a family of Markov Chain Monte Carlo (MCMC) sampling algorithms. The MCMC sampling algorithms utilise a Markov Chain with the desired invariant probability distribution in order to obtain the samples from the distribution of interest. A brief review of Markov Chain theory is presented in appendix B, while the interested reader is referred to an excellent exposition by [Wu, 1998] for a more extensive review of Markov Chain Monte Carlo methods.

For notational convenience, denote the distribution of interest as $p(\Psi)$, where Ψ is the high dimensional vector of interest (*e.g.* the missing phases, Θ). The Gibbs sampler relies on iteratively sampling the conditional distributions of the form

$$\Psi_n \sim p(\Psi_n | \Psi_{-n})$$

in order to obtain the samples from the joint distribution, $p(\Psi)$, where Ψ_n represent subvectors of the high dimensional vector Ψ . The Gibbs sampler relies on the fact that sampling from the lower dimensional distributions is easier than sampling from the original high dimensional distribution. Unfortunately, particularly when correlations between the different subvectors exist, the Gibbs sampling algorithm tends to be very inefficient [Roberts and Sahu, 1997].

As an alternative to Gibbs sampling, the Metropolis-Hastings algorithm provides the means for jointly sampling from the high dimensional distribution. To this end, the transition kernel for the Metropolis-Hastings algorithm is given by

$$T(\Psi' | \Psi) = q(\Psi' | \Psi) \alpha(\Psi' | \Psi) + \delta(\Psi' - \Psi) \left[1 - \int \alpha(\Psi' | \Psi) q(\Psi' | \Psi) d\Psi' \right]$$

where $q(\Psi' | \Psi)$ is the *proposal* density, and $\alpha(\Psi' | \Psi)$ is the *acceptance* probability, while $\delta(\cdot)$ is the indicator function.

The popularity of the Metropolis-Hastings sampling algorithm stems from its simplicity and flexibility. While the form of the proposal density determines the dynamics of the resulting Markov chain, the acceptance probability ensures the correct invariant distribution. In fact, the Metropolis-Hastings kernel satisfies detailed balance with respect to the required invariant distribution $p(\Psi)$, if and only if

$$p(\Psi) q(\Psi' | \Psi) \alpha(\Psi' | \Psi) = p(\Psi') q(\Psi | \Psi') \alpha(\Psi | \Psi')$$

Furthermore, an irreducible Metropolis kernel is Harris recurrent [Wu, 1998].

The most common form of the acceptance probability for the transition kernel, which satisfies the detailed balance condition, is given by

$$\alpha(\Psi'|\Psi) = \min \left[1, \frac{p(\Psi')q(\Psi|\Psi')}{p(\Psi)q(\Psi'|\Psi)} \right] \quad (4.17)$$

As [Wu, 1998] points out, although there exist several alternative forms of the acceptance probability function, the Hastings version is optimal for quite an extensive range of choices, chiefly because it rejects proposals less often than the others. Furthermore, with the acceptance probability given by equation 4.17, the actual acceptance rate and the Markov chain dynamics are largely determined by the choice of the proposal distribution $q(\Psi'|\Psi)$. There is a myriad of choices available for the proposal density, the most popular of which are discussed in [Gilks et al., 1996].

The general form of the Metropolis-Hastings sampling algorithm for obtaining samples from an invariant distribution $p(\Psi)$ can be summarised as follows:

-
1. At iteration t of the Metropolis-Hastings algorithm, given a sample $\Psi^{(t-1)}$ from the previous iteration, generate a candidate sample Ψ' from the proposal distribution $q(\Psi'|\Psi^{(t-1)})$, *i.e.*

$$\Psi' \sim q(\Psi'|\Psi^{(t-1)})$$

2. Accept the proposed candidate (*i.e.* $\Psi^{(t)} = \Psi'$) with probability

$$\alpha(\Psi'|\Psi^{(t-1)}) = \min \left[1, \frac{p(\Psi')q(\Psi^{(t-1)}|\Psi')}{p(\Psi^{(t-1)})q(\Psi'|\Psi^{(t-1)})} \right]$$

and retain the previous sample (*i.e.* $\Psi^{(t)} = \Psi^{(t-1)}$) otherwise

3. Increment t and go to step 1
-

In addition to the ratio of the invariant distribution $p(\Psi')/p(\Psi^{(t-1)})$, the acceptance step of the Metropolis-Hastings algorithm also requires the ratio of the proposal distributions, $q(\Psi^{(t-1)}|\Psi')/q(\Psi'|\Psi^{(t-1)})$. For this reason, the proposal distributions are generally local in nature, in that a candidate sample, Ψ' , is drawn from a distribution centered around the current sample, $\Psi^{(t-1)}$. Such a procedure is inherently very inefficient, in that the candidate samples are local and undirected, resulting in many, possibly better, potential candidate samples remaining undiscovered.

In order to improve the efficiency of the proposal distribution, [Duane et al., 1987] have proposed to utilise the gradient information contained in the invariant distribution. In particular, the authors have proposed to use a deterministic transition kernel based on a discretised

form of Molecular Dynamics. The resulting algorithm, which will be discussed in more detail in the following sections, is referred to as the Hybrid Monte Carlo algorithm. The algorithm has been popularised by [Neal, 1992], with further reviews presented by [Wu, 1998] and [Ó Ruanaidh and Fitzgerald, 1996].

4.3.1 The Gibbs Distribution and Molecular Dynamics

In classical mechanics, a dynamic system of N particles is characterised at any fixed time instant by a set of $2N$ coordinates, corresponding to the positions ($\Psi = [\psi_1, \dots, \psi_N]$) and the momenta ($\mathbf{p} = [p_1, \dots, p_N]$) of the particles. At any time instant, the positions and the momenta of the particles provide a complete description of a *microstate* of the system. Furthermore, the positions and the momenta can be represented by a point in a $2N$ dimensional space, spanned by a Cartesian reference frame of $2N$ mutually orthogonal axes. This space is referred to as the *phase space*.

Every possible microstate of the physical system, $[\Psi, \mathbf{p}]$, has an associated energy, $H(\Psi, \mathbf{p})$, referred to as the *Hamiltonian* of the system. For the system of N particles, in the absence of an external field, the system Hamiltonian can be written as a sum of two terms

$$\begin{aligned} H(\Psi, \mathbf{p}) &= \beta_E E(\Psi) + \beta_K K(\mathbf{p}) \\ K(\mathbf{p}) &= \frac{1}{2} \sum_i p_i^2 \end{aligned}$$

where $E(\Psi)$ corresponds to the potential energy, while $K(\mathbf{p})$ corresponds to the kinetic energy of the physical system. Furthermore, β_E is the inverse temperature of the system, while β_E/β_K is analogous to the mass of the particles in the physical system⁴.

At any given time, the microstate represents a single realisation of the physical system, where the microstate is distributed according to the *canonical* (or the *Gibbs*) distribution

$$p(\Psi, \mathbf{p}) = \frac{1}{Z(\Psi)Z(\mathbf{p})} \exp -H(\Psi, \mathbf{p})$$

where $Z(\Psi)$ and $Z(\mathbf{p})$ are the normalising constants, also known as the *Partition Functions*.

In classical mechanics, in absence of external fields, the Hamiltonian of the system is conserved, *i.e.*, $\frac{dH(\Psi, \mathbf{p})}{dt} = 0$. The gradient of the potential energy with respect to the particle positions, $\partial E(\Psi)/\partial \Psi_i$, defines a force that acts to change the system configuration via its effects on the

⁴The momenta temperature β_K is used here instead of the usual particle masses. The two parametrisations are identical, although the parametrisation adopted presently sheds some light on the intricacies of the Hybrid Monte Carlo algorithm.

system's momentum. In particular, the positions and the momenta of the particles evolve according to Hamilton's equations of motion (*i.e. Hamiltonian Dynamics*), as given by

$$\begin{aligned}\frac{\partial \Psi_i}{\partial t} &= \sqrt{\beta_K} p_i \\ \frac{\partial p_i}{\partial t} &= -\frac{1}{\sqrt{\beta_K}} \frac{\partial E(\Psi)}{\partial \Psi_i}\end{aligned}$$

where, for reasons subsequently explained, β_E is set to unity. Hamiltonian Dynamics has a number of very important properties. In addition to conserving the Hamiltonian of the system, Hamiltonian Dynamics is also time reversible. However, most importantly, Hamiltonian Dynamics *preserves phase space volume*. This result is known as the *Liouville's theorem*.

Molecular Dynamics method simulates the evolution of a dynamical system by discretising Hamilton's Equations. To this end, by far the most popular discretisation scheme is the *leapfrog scheme* given by

$$\begin{aligned}p_i(t + \frac{\epsilon}{2}) &= p_i(t) - \frac{\epsilon}{2\sqrt{\beta_K}} \frac{\partial E}{\partial \Psi_i}(t) \\ \Psi_i(t + \epsilon) &= \Psi_i(t) + \epsilon \sqrt{\beta_K} p_i(t + \frac{\epsilon}{2}) \\ p_i(t + \epsilon) &= p_i(t + \frac{\epsilon}{2}) - \frac{\epsilon}{2\sqrt{\beta_K}} \frac{\partial E}{\partial \Psi_i}(t + \epsilon)\end{aligned}$$

where ϵ is the discrete time step size. Indeed, as was the case with Hamiltonian Dynamics, the leapfrog scheme, denoted by

$$\{\Psi(t + \epsilon), \mathbf{p}(t + \epsilon)\} = L(\Psi(t), \mathbf{p}(t)|\epsilon, N_l)$$

is not only time reversible, but also preserves the state space volume. However, due to discretisation errors the Hamiltonian need no longer be conserved.

4.3.2 Hybrid Monte Carlo Algorithm

The Hybrid Monte Carlo (HMC) algorithm is a global sampling algorithm that incorporates the Molecular Dynamics features into the Metropolis-Hastings algorithm framework. To this end, the sampling is performed on an extended *phase space* $\{\Psi, \mathbf{p}\}$, where the original variables, Ψ , are augmented by fictitious *generalised momenta*, \mathbf{p} . The generalised momenta are conjugate to Ψ and of the same dimensionality as Ψ , with a microstate $[\Psi, \mathbf{p}]$ being distributed according to the Gibbs distribution, *i.e.*

$$p(\Psi, \mathbf{p}) = \frac{1}{Z(\Psi)Z(\mathbf{p})} \exp -H(\Psi, \mathbf{p}) \quad (4.18)$$

where the Hamiltonian for the augmented system is given by

$$H(\Psi, \mathbf{p}) = \beta_E E(\Psi) + \beta_K K(\mathbf{p})$$

with the Kinetic energy of the system defined as

$$K(\mathbf{p}) = \frac{1}{2} \sum_i p_i^2$$

It is interesting to note that, provided $\beta_E = 1$ and $E(\Psi) = -\ln p(\Psi)$, the marginal distribution for Ψ in equation 4.18 and the high dimensional distribution of interest, $p(\Psi)$, coincide. Also, the Gibbs distribution factorises into

$$\begin{aligned} p(\Psi, \mathbf{p}) &= p(\Psi)Q(\mathbf{p}) \\ Q(\mathbf{p}) &= \frac{1}{Z(\mathbf{p})} \exp -\beta_K \frac{1}{2} \sum_i p_i^2 \end{aligned}$$

Sampling from the marginal distribution for Ψ can be performed by sampling from the Gibbs distribution (see equation 4.18). This is possible since convergence to the joint, implies convergence to the marginal, distribution. The fictitious momenta simply facilitate the sampling from the coordinates, Ψ , via the means of Molecular Dynamics. In that respect, HMC algorithm belongs to the class of algorithms which make use of auxiliary variables to induce global moves.

The HMC transition kernel for Ψ is given by

$$\begin{aligned} T(\Psi^*|\Psi) &= \int d\mathbf{p}d\mathbf{p}^* Q(\mathbf{p}) \delta(L(\Psi, \mathbf{p}) \rightarrow (\Psi^*, \mathbf{p}^*)) A(\Psi^*|\Psi, \mathbf{p}, \mathbf{p}^*) \\ &+ \delta(\Psi^* - \Psi) \left[1 - \int d\mathbf{p}d\mathbf{p}^* d\Psi^* A(\Psi^*|\Psi, \mathbf{p}, \mathbf{p}^*) Q(\mathbf{p}) \delta(L(\Psi, \mathbf{p}) \rightarrow (\Psi^*, \mathbf{p}^*)) \right] \end{aligned} \quad (4.19)$$

where $L(\Psi, \mathbf{p})$ is a deterministic leapfrog function, and $\delta(\cdot)$ is the indicator function. In other words, sampling of Ψ from $p(\Psi)$ using the *Hybrid Monte Carlo* algorithm proceeds as follows:

-
1. Generate momenta $\mathbf{p}^t \sim Q(\mathbf{p}) = N(\mathbf{0}, 1/\beta_K \mathbf{I})$
 2. Starting with Ψ^t and freshly generated \mathbf{p}^t , iterate the following leapfrog equations N_l times to get new proposed state $(\Psi^*, \mathbf{p}^*) = L(\Psi^t, \mathbf{p}^t)$

$$\begin{aligned} p_i(t + \frac{\epsilon}{2}) &= p_i(t) - \frac{\epsilon}{2\sqrt{\beta_K}} \frac{\partial E}{\partial \Psi_i}(t) \\ \Psi_i(t + \epsilon) &= \Psi_i(t) + \epsilon \sqrt{\beta_K} p_i(t + \frac{\epsilon}{2}) \\ p_i(t + \epsilon) &= p_i(t + \frac{\epsilon}{2}) - \frac{\epsilon}{2\sqrt{\beta_K}} \frac{\partial E}{\partial \Psi_i}(t + \epsilon) \end{aligned}$$

3. Set $\Psi^{t+1} = \Psi^*$ with probability

$$A(\Psi^*|\Psi^t, \mathbf{p}^t, \mathbf{p}^*) = \min [1, \exp(-(H(\Psi^*, \mathbf{p}^*) - H(\Psi^t, \mathbf{p}^t)))]$$

and $\Psi^{t+1} = \Psi^t$ otherwise

4. Go to step 1 to refresh the momenta

The transition kernel (see equation 4.19) for the HMC algorithm shows the importance of sampling the generalised momenta (*i.e.* $\mathbf{p} \sim Q(\mathbf{p})$) at every iteration of the algorithm, in order to obtain the samples from the marginal distribution, $p(\Psi)$. It is the refreshing of the momenta that provides the element of randomness into an otherwise deterministic algorithm⁵. It is also interesting to note that the marginal distribution for the generalised momenta, $Q(\mathbf{p})$, need not be Gaussian. Other distributions, such as the Cauchy distribution can be used, provided the Hamiltonian and the acceptance probability are modified appropriately [Wu, 1998].

Once a fresh sample of the generalised momenta is obtained from $Q(\mathbf{p})$, a candidate sample for $[\Psi^*, \mathbf{p}^*]$ is obtained by iterating the Molecular Dynamics equations. To this end, the leapfrog scheme is used presently, though other discretisation schemes are also possible [Neal, 1992]. Finally, in order to ensure the correct invariant distribution and to compensate for any discretisation errors inherent in the Molecular Dynamics simulation, the candidate state is accepted with probability $A(\Psi^*|\Psi^t, \mathbf{p}^t, \mathbf{p}^*)$. By Liouville's Theorem, based purely on physical arguments, the phase space volume is preserved following Molecular Dynamics evolution, and hence the ratio of the proposal distributions is unity. The acceptance probability is therefore a function of the energy (*i.e.* probability) at the beginning and the end of the Molecular Dynamics trajectory.

Given the Liouville's Theorem and the time reversibility of the leapfrog scheme, it is relatively easy to show that the basic Hybrid Monte Carlo algorithm satisfies detailed balance, with the correct invariant distribution, $p(\Psi)$ [Wu, 1998]. The apparent non-linearities and the coordinate-momenta couplings present in the Molecular Dynamics are inconsequential, as a result of the Liouville's Theorem. However, for the Markov Chain induced by the HMC algorithm to be ergodic, it has to be irreducible and aperiodic, in addition to leaving the desired distribution invariant. Whether these conditions are satisfied is dependent on the nature of the potential energy function, $E(\Psi)$ (and hence $p(\Psi)$) and the specifics of the HMC scheme employed. This point will now be discussed in more detail in the following section.

⁵Without refreshing of the momenta, the Molecular Dynamics essentially evolves along an iso-energy surface.

4.3.3 Remarks about the HMC Algorithm

The usefulness of the HMC algorithm stems from the use of the gradient information of the invariant distribution. In particular, the gradient information acts as a force field guiding the particles in a fictitious physical system towards regions of low energy (*i.e.* high probability) through the evolution of the Molecular Dynamics equations. Unfortunately, Molecular Dynamics evolves the dynamic system on an iso-energy surface in the extended phase space.

As [Wu, 1998] shows, the HMC algorithm satisfies detailed balance. However, in order for the Markov Chain induced by the HMC to be irreducible, it has to be possible to eventually reach *any* other microstate in the extended phase space, from any microstate to which the invariant distribution assigns positive probability measure. To some extent, refreshing the momenta at every iteration of the HMC algorithm ensures that different iso-energy surfaces are explored, and hence that any point in the phase space can eventually be reached. However, as will now be shown, due to the nature of the basic HMC algorithm, the change in the Hamiltonian of the system can be very small from one iteration of the algorithm to the next. This may result in extremely slow exploration of the different iso-energy surfaces, as well as a possibility of the Markov Chain not being irreducible.

In the basic HMC algorithm discussed in the previous section, the Hamiltonian is of the form

$$H(\Psi, \mathbf{p}) = E(\Psi) + \beta_K K(\mathbf{p})$$

Furthermore, the Hamiltonian for the Molecular Dynamics (the *Guiding Hamiltonian*) and the Hamiltonian in the acceptance probability (the *Acceptance Hamiltonian*) are identical. In addition, the corresponding marginal distribution for the generalised momenta is a Gaussian distribution with zero mean, and variance $1/\beta_K$. Therefore, at each iteration of the HMC algorithm, the momenta are refreshed as $\mathbf{p} \sim N(\mathbf{0}, 1/\beta_K \mathbf{I})$. However, regardless of the value of β_K , the effects of β_K are simply nulled in the leapfrog scheme. In particular, upon re-normalisation of \mathbf{p} to unity (*i.e.* let $\tilde{\mathbf{p}} = \sqrt{\beta_K} \mathbf{p}$), the first two steps of the HMC algorithm become

-
1. Generate momenta $\tilde{\mathbf{p}}^t \sim Q(\tilde{\mathbf{p}}) = N(\mathbf{0}, \mathbf{I})$
 2. Starting with Ψ^t and freshly generated $\tilde{\mathbf{p}}^t$, iterate the following leapfrog equations N_l times to get new proposed state $(\Psi^*, \tilde{\mathbf{p}}^*) = L(\Psi^t, \tilde{\mathbf{p}}^t)$

$$\begin{aligned} \tilde{p}_i(t + \frac{\epsilon}{2}) &= \tilde{p}_i(t) - \frac{\epsilon}{2} \frac{\partial E}{\partial \Psi_i}(t) \\ \Psi_i(t + \epsilon) &= \Psi_i(t) + \epsilon \tilde{p}_i(t + \frac{\epsilon}{2}) \\ \tilde{p}_i(t + \epsilon) &= \tilde{p}_i(t + \frac{\epsilon}{2}) - \frac{\epsilon}{2} \frac{\partial E}{\partial \Psi_i}(t + \epsilon) \end{aligned}$$

Therefore, the only *free* parameters are the number of leapfrog iterations, N_l , and the discrete step size, ϵ . The consequences of this observation are profound. In particular, as long as the guiding and the acceptance Hamiltonians are identical, the effects of momenta in Molecular Dynamics are independent of the momenta temperature, β_K , and consequently of the particle masses. Therefore, whether $\frac{\partial E}{\partial \Psi_i}(t)$ is equal to 10^0 or 10^{10} , p_i will always be such that $p_i \sim N(0, 1)$.

Despite the contributions of the momenta being independent of the particle masses (and β_K), there is still plenty of scope for ‘tweaking’ the HMC Algorithm. Tricks involving random sampling of the step size, ϵ , and number of leapfrog iterations, N_l , are well known and tend to improve the mixing of the sampler [Neal, 1992]. Random sampling of these parameters also ensures that the Markov Chain induced by the HMC algorithm is aperiodic. However, more importantly, it is also known that the *guiding Hamiltonian*, H' , and the *acceptance Hamiltonian*, H , need not be the same [Duane et al., 1987]. In fact, it may sometimes prove fruitful to make the two Hamiltonians different in order to improve the performance of the sampler.

The acceptance Hamiltonian is largely fixed by the required invariant distribution for Ψ . However, the guiding Hamiltonian can be made different from the acceptance Hamiltonian for one of two reasons. Firstly, the guiding Hamiltonian can be constructed in such a way so as to compensate for any discretisation errors induced by Molecular Dynamics, thus potentially resulting in an increased acceptance rate. Secondly, the guiding Hamiltonian can be constructed in such a way so as to increase the ‘mobility’ of the candidate state, such that the phase space is better explored (possibly at the expense of lower acceptance rate).

Provided the two Hamiltonians are identical, *i.e.* $H' = H$, and no discretisation errors are present, the HMC algorithm results in 100% acceptance rate. However, the leapfrog scheme results in discretisation errors, which also limit the mobility of the Markov Chain. Furthermore, as was pointed out above, the relative contributions of the sampled momenta can be small compared to the potential field effects, possibly resulting in the Markov Chain not being irreducible.

One possible scheme for improved exploration of the phase space is to temper the momenta temperature, β_K . To this end, consider the acceptance Hamiltonian $H = E(\Psi) + K(\mathbf{p})$ (note that without loss of generality, the temperature β_K has been arbitrarily set to one), while the guiding Hamiltonian $H' = E(\Psi) + \beta_K K(\mathbf{p})$. In short, the two Hamiltonians differ in the temperature of the momenta (or equivalently, the inverse masses of the fictitious particles). Using the fact that Molecular Dynamics conserves the Hamiltonian, ie

$$E(\Psi) + \beta_K K(\mathbf{p}) = E(\Psi^*) + \beta_K K(\mathbf{p}^*)$$

and provided the leapfrog scheme introduces no discretisation errors, it can be shown that the

acceptance probability is given by

$$A(\Psi^*|\Psi, \mathbf{p}, \mathbf{p}^*) = \min [1, e^{-\Delta H}]$$

$$\Delta H = \left(1 - \frac{1}{\beta_K}\right)(E(\Psi^*) - E(\Psi))$$

Therefore, as β_K is increased (*i.e.*, the masses of the fictitious particles are decreased), the particles in the Molecular Dynamics evolution become more mobile, resulting in faster exploration of the phase space, at the expense of lower acceptance rate. Thus, a simple form of tempering is achieved by a mere change of the momenta temperature in the guiding Hamiltonian. Furthermore, by randomly selecting the temperature, β_K , in the guiding Hamiltonian, many different iso-energy surfaces are explored, thus ensuring the irreducibility of the Markov Chain.

4.3.4 Symmetrical Momentum Tempering

The above argument can now be extended further. The momenta temperature, β_K , can be tempered in a symmetric fashion with each leapfrog step. To this end, β_K is increased for the first half of the leapfrog steps, *i.e.* $\beta_K^0 < \beta_K^1 < \dots < \beta_K^{T/2}$, while β_K is decreased for the latter half of the leapfrog steps, *i.e.* $\beta_K^{T/2} > \dots > \beta_K^0$. Time reversibility is preserved at each iteration of the momentum-based symmetric tempering algorithm, as is the phase space volume. Therefore, provided the acceptance probability is modified to account for the intermediate steps [Neal, 1994], the detailed balance for the momentum-based tempering algorithm is preserved. Furthermore, based on the conservation of the Hamiltonian at every leapfrog iteration, the acceptance probability is approximately given by ⁶

$$A(\Psi^*|\Psi, \mathbf{p}, \mathbf{p}^*) = \min \left[1, \exp(-(\check{F} - \hat{F})) \right]$$

$$\hat{F} = \sum_{i=0}^{T/2-1} \left(\frac{1}{\beta_i} - \frac{1}{\beta_{i+1}} \right) E(\hat{\Psi}_i)$$

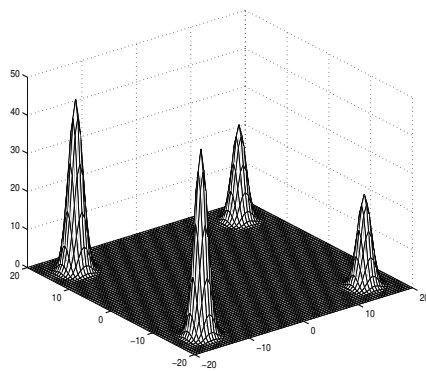
$$\check{F} = \sum_{i=0}^{T/2-1} \left(\frac{1}{\beta_i} - \frac{1}{\beta_{i+1}} \right) E(\check{\Psi}_i)$$

where the notation of [Neal, 1994] has been adopted. It is also interesting to note the analogy between the acceptance probability in the momentum-based tempering scheme, and the tempered transition scheme proposed by [Neal, 1994]. However, unlike the Tempered Transition algorithm, no special sampling procedure is required for sampling from the invariant distribution at higher temperatures. Instead, it is the tempering of the particle momenta in the evolution of the Molecular Dynamics that results in enhanced 'mobility' of the candidate states.

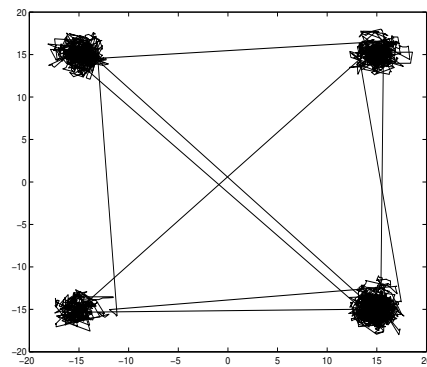
⁶Due to the discretisation errors in Molecular Dynamics, the Hamiltonian may not be exactly conserved. The expression for the acceptance probability is only an approximate guide to illustrate the effects of tempering on the acceptance rate.

In principle, changing of the momenta temperature in the guiding Hamiltonian allows the algorithm to better explore the phase space. In particular, by first increasing β_K (*i.e.* decreasing the fictional particle masses) better mobility is obtained, and then as β_K is decreased, the Molecular Dynamics converges to a local Energy minimum (*i.e.* probability maximum). Unfortunately, as is the case with the Tempered Transitions algorithm, the acceptance probability tends to be rather low. Therefore, in order to improve the mixing of the sampler, it is recommended that the momentum-based tempered run be performed every few iterations of the standard HMC algorithm.

Figure 4.2 illustrates the increased mobility of a HMC algorithm incorporating the momentum-based tempering. In particular, the HMC algorithm was employed to sample from a mixture of four Gaussians shown in figure 4.2(a). The standard HMC algorithm, once settled in one of the modes, would be very unlikely indeed to find another mode in this distribution. However, as figure 4.2(b) shows, all modes are discovered when tempering runs are included as part of the HMC algorithm. In particular, the results shown in figure 4.2(b) correspond to 10000 iterations of the HMC algorithm, with a tempering run performed every 10^{th} iteration. The step size, ϵ , number of leapfrog steps, N_l , and the momenta temperatures, β_K^0 and $\beta_K^{T/2}$, were chosen randomly from bounded log-uniform distributions.



(a) Multi-modal Distribution



(b) Samples from the Distribution

Figure 4.2: (a) Multi-modal distribution (a mixture of four Gaussians) and (b) 10000 samples from the distribution obtained HMC algorithm, with a momentum-based tempered run performed every 10^{th} iteration.

In short, the results shown in figure 4.2 illustrate that tempering of momenta helps the Hybrid Monte Carlo algorithm move around a multimodal distribution. However, unlike conventional tempering algorithms, tempering of momenta does not require the knowledge of the normalising constants at different temperatures.

4.4 CHARACTERISATION OF INCOHERENT CLUTTER

The sea clutter data studied in the present chapter was obtained in open ocean conditions at a look angle of 215.1° into the wind at 7.5km range at the effective sampling frequency of 500Hz for each polarisation channel. The resolution of each range cell is effectively 4m in range and 160m in azimuth. For the purposes of this chapter, both the real and synthetic sea clutter data is limited to 150 consecutive samples from a single range cell, which corresponds to 300ms of raw data. The reasons for this are twofold. Firstly, the statistics of sea clutter may no longer be stationary on time scales any longer than a few hundred milliseconds. Secondly, the algorithm considered in the present chapter becomes computationally prohibitive for data lengths much larger than 150 samples.

A non-linear state space model developed in section 4.2 is utilised presently for the purpose of characterisation of sea clutter using just a few parameters (*i.e.* the poles of the complex autoregressive state transition process). To this end, the Hybrid Monte Carlo algorithm (see section 4.3) is used within a Gibbs sampling framework in order to explore the posterior distribution of the complex autoregressive process parameters, when only the amplitudes of the process are observed. The sampling procedure specific to the model introduced in section 4.2.2 is discussed in section 4.4.1, while the sampling algorithm is applied to real and synthetic sea clutter data in sections 4.4.2-4.4.6.

4.4.1 Sampling Procedure

As discussed in section 4.2.3, convergence to the joint ensures convergence to the marginal distribution. In particular, the Gibbs sampler can be employed in order to obtain samples from the parameter posterior, $p(\mathbf{a}|\mathbf{r})$ (see section 4.2.2 for notation). The Gibbs sampler exploits the conditional structure of the state space model, as given by equations 4.13-4.15. To this end, the following algorithm has been adopted in the present chapter for exploring the distribution of the complex autoregressive process parameters, conditional on the observed amplitudes:

-
- Initialise the missing phases and the process parameters, *i.e.* $[\Theta^{(0)}, \mathbf{a}^{(0)}, (\sigma^2)^{(0)}]$.
 - Run the Gibbs sampler for N_{itns} iterations:
 1. At iteration t , given a sample $[\Theta^{(t-1)}, \mathbf{a}^{(t-1)}, (\sigma^2)^{(t-1)}]$ from the previous iteration, consecutive sample $[\Theta^{(t)}, \mathbf{a}^{(t)}, (\sigma^2)^{(t)}]$ is obtained as

$$\begin{aligned} \mathbf{a}^{(t)} &\sim p(\mathbf{a}|\mathbf{r}, \Theta^{(t-1)}, (\sigma^2)^{(t-1)}) \\ (\sigma^2)^{(t)} &\sim p(\sigma^2|\mathbf{r}, \Theta^{(t-1)}, \mathbf{a}^{(t)}) \\ \Theta^{(t)} &\sim p(\Theta|\mathbf{r}, \mathbf{a}^{(t)}, (\sigma^2)^{(t)}) \end{aligned}$$

2. Increment t and go to step 1

- Disregard the first $N_{burn-in}$ samples.
- Retain the remaining samples of the process parameters, \mathbf{a} and σ^2 .

The above algorithm was implemented in Matlab 5.x, and run on Pentium II 266+ based computers. With $N_{itns} = 10000$, each complete run of the algorithm took, on average, in excess of 15 hours to complete. Based on the convergence diagnostics criteria discussed below, it was found that $N_{burn-in} = 500$ iterations were usually sufficient for the sampler to converge to the invariant distribution. The details of initialisation, sampling from the conditional distributions and convergence diagnostics are discussed in more detail below.

Initialisation

Initial estimates of Θ , \mathbf{a} and σ^2 are required in the above algorithm. These can either be obtained randomly, or alternatively, potentially better estimates can be obtained by utilising the information contained in the observed amplitudes of the sea clutter data.

A very simple approach is followed in the present chapter, whereby the underlying complex Gaussian process is assumed to be real and equal to the observed time series of the process amplitudes. Therefore, the following procedure has been adopted for the purpose of automatic initialisation of the sampler:

-
- Set $[\mathbf{a}^{(0)}, (\sigma^2)^{(0)}]$ to the MAP estimates of an autoregressive process parameters corresponding to the time series given by $\{r(n) - \langle r(n) \rangle : n = [1, \dots, N]\}$.
 - Set $\Theta^{(0)}$ to randomly drawn variates from the uniform distribution over $[0, 2\pi]$. In order to get 'better' initial estimates, an initial run of the HMC algorithm is then performed, conditional on the initial $[\mathbf{a}^{(0)}, (\sigma^2)^{(0)}]$ estimates.
-

Sampling from $p(\mathbf{a}|\mathbf{r}, \Theta, \sigma^2)$ and $p(\sigma^2|\mathbf{r}, \Theta, \mathbf{a})$

From equations 4.14 and 4.15, the conditional distributions for \mathbf{a} and σ^2 are given by

$$\begin{aligned} p(\mathbf{a}|\mathbf{r}, \Theta, \sigma^2) &= N(\mu_a, \Sigma_a) \\ p(\sigma^2|\mathbf{r}, \Theta, \mathbf{a}) &= IG(\alpha_\sigma, \beta_\sigma) \end{aligned}$$

where the distribution parameters are defined as in equations 4.5-4.9.

Both of the conditional distributions are in standard form, and are therefore sampled from easily. An added complication in the case of the conditional distribution for \mathbf{a} stems from the stability condition required of the complex autoregressive process parameters. In particular, the process poles associated with the the process parameters, \mathbf{a} , need to lie within a unit circle, in order for the complex autoregressive process to be stable. This condition is achieved in the work presented in this chapter by employing a rejection sampling step in order to reject any samples, \mathbf{a} , that do not satisfy the stability criterion.

Sampling from $p(\Theta|\mathbf{r}, \mathbf{a}, \sigma^2)$

The conditional distribution for the missing phases, Θ , is given by equation 4.16. In more concise notation, the logarithm of the distribution in equation 4.16 can be written as

$$\begin{aligned} \ln p(\Theta|\mathbf{r}, \mathbf{a}, \sigma^2) &= \text{const} - \frac{1}{2\sigma^2} \mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x} \\ \mathbf{x} &= [r(1)e^{i\theta(1)}, \dots, r(N)e^{i\theta(N)}]' \end{aligned}$$

where \mathbf{A} is the matrix of complex parameters, \mathbf{a} (*c.f.* \mathbf{A} in equations 4.5-4.9).

Since the conditional distribution, $p(\Theta|\mathbf{r}, \mathbf{a}, \sigma^2)$, is non-standard form, the HMC algorithm (see section 4.3) is employed in this chapter to obtain samples of the missing phases, Θ . To this end, the derivative of the logarithm of the conditional distribution (as required by the HMC algorithm) can be shown to be given by

$$\frac{\partial \ln p(\Theta|\mathbf{r}, \mathbf{a}, \sigma^2)}{\partial \Theta} = \frac{1}{\sigma^2} \Im \{ \mathbf{x}' .* [\mathbf{A}^H \mathbf{A} \mathbf{x}] \}$$

where $\Im\{\cdot\}$ is the imaginary part, and $a.*b$ represents the vector elementwise product (*c.f.* Matlab vector elementwise multiplication function).

At every iteration of the Gibbs sampler, a sample of the missing phases, Θ , is obtained from $p(\Theta, \mathbf{a}, \sigma^2|\mathbf{r})$. To this end, a single iteration of the HMC algorithm is performed. However, in order to improve the mixing of the sampled phases, a tempering run is performed every 10th iteration of the Gibbs sampler. Therefore, unless stated otherwise, the following schedules have been adopted in this chapter for the HMC algorithm:

-
- During a *non-tempering* run set the HMC parameters to the following:
 - randomly set the step size, dt , such that $\log_{10} dt \sim \text{Uniform}[-3, 0]$
 - randomly set the number of leapfrog steps, N_l , such that $N_l \sim \text{Uniform}[50, 500]$
 - During a *tempering* run set the HMC parameters to the following:

- randomly set the step size, dt , such that $\log_{10} dt \sim \text{Uniform}[-3, 0]$
- randomly set the number of leapfrog steps, N_l , such that $N_l \sim \text{Uniform}[50, 500]$
- randomly set the maximum and the minimum momentum 'temperatures', β_{Kmax} and β_{Kmin} , such that

$$\begin{aligned}\log_{10} \beta_{Kmax} &\sim \text{Uniform}[0, 3] \\ \log_{10} \beta_{Kmin} &\sim \text{Uniform}[-1, 0]\end{aligned}$$

Convergence Diagnostics

The main objective of the present chapter is to explore the distribution of the underlying complex autoregressive process parameters, conditional on the observed amplitudes (*i.e.* with the phase information missing). To this end, since the complete knowledge of the *complex* time series is also available, a very simple convergence diagnostic is presently adopted.

In particular, a number of Markov Chains are run concurrently on the same data set. Half of the Markov Chains are initialised using the actual process phases and with process parameters estimated by Maximum Likelihood from the coherent clutter. The other half of the chains are initialised using an automated procedure outlined above. At each iteration of the Gibbs sampler, a polar plot of the process poles is displayed for each of the chains. Convergence is declared when the samples of the process poles for the randomly initialised sampler reach the same area of the complex plane, as that reached by the samples of the process poles obtained using the Markov Chains initialised from the MAP estimates.

The above procedure assumes that once two chains reach the same point, both chains have converged to the invariant distribution. The convergence, as diagnosed by the above criterion, was usually found to occur within 100-200 iterations of the Gibbs sampler. Unless stated otherwise, a 'burn-in' period of 500 iterations has been adopted for all the subsequent simulations presented in this chapter.

4.4.2 Synthetic AR(1) Process

For illustrative purposes, 150 samples of synthetic data were generated from a complex autoregressive process with a *single* complex pole (*i.e.* AR(1) process) at $0.8\angle 0.4\pi$ and the innovation process variance set to unity. The magnitude of the pole was chosen such that the spectral width of the data is relatively large, as is observed in high resolution sea clutter returns. The phase of the process pole was chosen rather arbitrarily to reflect the typical Doppler shifts observed in the spectrograms shown in figure 4.1.

An $AR(2)$ model is used to model the $AR(1)$ data in order to better facilitate the sampling of the complex autoregressive process parameters (by avoiding getting 'stuck' in local modes) and also to investigate the sensitivity of the sampling scheme to model misselection. For this reason, models of order greater than those used to generate the synthetic data are used in the present chapter. However, the issues associated with model selection will not be addressed in this dissertation and are left as an exercise for future research.

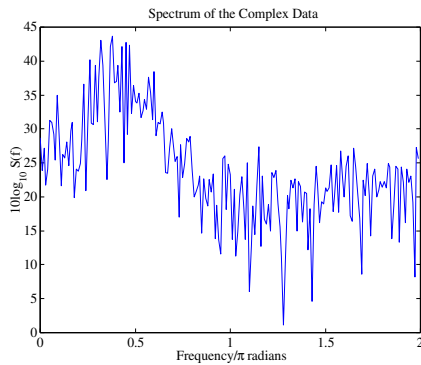
Figure 4.3(a) illustrates the presence of a spectral peak at a frequency of 0.4π (corresponding to the phase of the process pole) in the log of the power spectrum of the complex $AR(1)$ process. However, no such spectral peaks are present in the log of the power spectrum obtained from the *amplitudes* of the complex $AR(1)$ process, as illustrated by figure 4.3(b). This observation is very much in line with the discourse presented in chapter 3 and section 4.2.2. In particular, figures 4.3(a) and 4.3(b) show that the absolute frequency information (*e.g.* single sided spectral peaks) contained in the coherent (*i.e.* complex) signal is largely lost when only the amplitudes of the complex process are observed. In fact, it is only the frequencies corresponding to the sum and the difference of the frequency components contained in the coherent clutter spectrum that are retained in the spectrum of incoherent clutter.

Figure 4.3(c) shows the complex $AR(2)$ process poles estimated by Maximum Likelihood (ML) from the $AR(1)$ process data. The ML estimates of the $AR(2)$ model poles are also presented in table 4.1. As expected, the magnitude and phase of the dominant $AR(2)$ pole is close to the true value for the original $AR(1)$ process (*i.e.* $0.8\angle 0.4\pi$), while the magnitude of the smaller pole is an order of magnitude smaller than the dominant pole. As the complex process amplitudes do not contain the information about the *absolute* phase of the $AR(1)$ process pole, the smaller pole in the $AR(2)$ model effectively serves as a phase reference. To this end, table 4.1 also shows the *relative* sum and difference components of the phases of the two $AR(2)$ process poles.

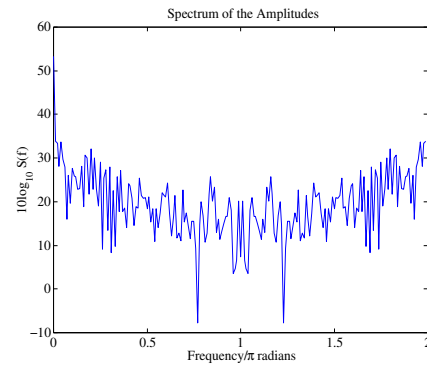
Process Poles		Relative Phases	
k	ML est. ($\rho_k \angle \phi_k$)	[i, j]	$ \phi_i \pm \phi_j $
1	$0.81 \angle 0.40\pi$	[1, 2]	0.33π 0.47π
2	$0.08 \angle 0.07\pi$		

Table 4.1: Complex $AR(2)$ process poles and their relative phases. Complex $AR(2)$ process parameters estimated by ML from 150 samples of a complex $AR(1)$ process with the true process pole at $0.8\angle 0.4\pi$.

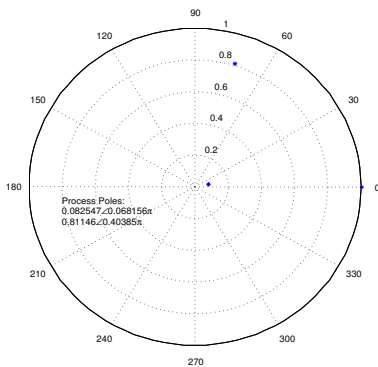
Figures 4.3(d)-4.3(f) show the results of 10000 iterations of a Gibbs sampler, with the first 500 samples discarded as 'burn-in'. As discussed in section 4.4.1, the Gibbs sampler was used to sample the complex $AR(2)$ model parameters, \mathbf{a} and σ^2 , conditional on the observed amplitudes of the synthetic complex $AR(1)$ process data. In particular, figure 4.3(d) shows a histogram of the relative phases of the sampled process poles. Clearly, the information about the



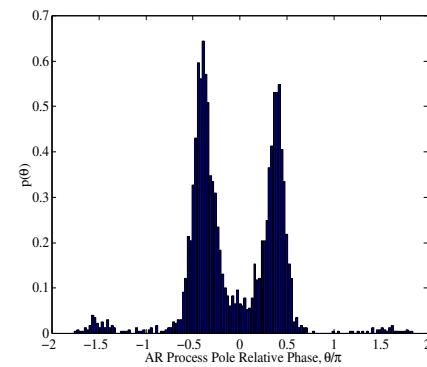
(a) Spectrum of Coherent Clutter



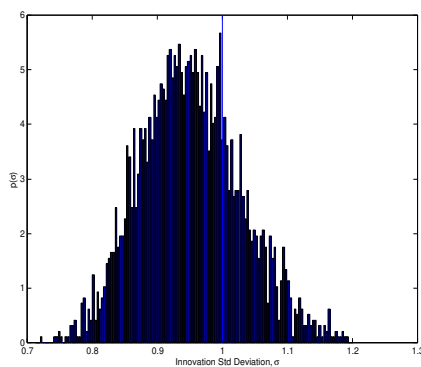
(b) Spectrum of Incoherent Clutter



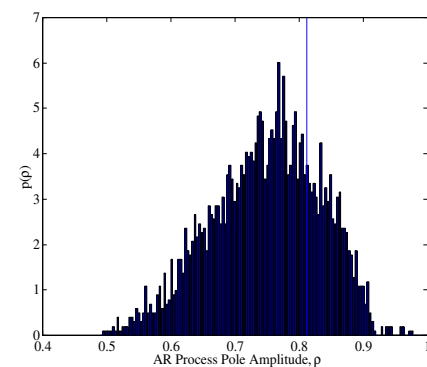
(c) Process Poles



(d) Relative Process Pole Phases



(e) Inn. Process Std. Dev.



(f) Process Pole Magnitude

Figure 4.3: (a) Coherent and (b) Incoherent spectrum of 150 samples of a synthetic complex AR(1) process, with the process pole at $0.8 \angle 0.4\pi$. (c) ML estimates of complex AR(2) process parameters obtained from the complex data, along with the samples of (d) relative phases, (e) innovation process standard deviation and (f) process pole magnitudes obtained using 10000 iterations of the Gibbs sampler with random initial conditions with 500 discarded burn-in samples. ML estimates indicated by vertical lines.

relative phases of the process poles is preserved and is in line with the relative phases expected from the ML estimates obtained from the coherent data (see table 4.1)⁷. Similarly, figure 4.3(e) shows that the samples of the innovation process standard deviation are also in line with the actual value of unity.

The histogram of the sampled magnitudes of the dominant pole of the $AR(2)$ model is shown in figure 4.3(f), along with the ML estimate obtained from the coherent data. The result shown in figure 4.3(f) seems to confirm the assertion made in section 4.2.2, in that although the absolute phase information about the complex process poles is lost when only the magnitudes are observed, the information about the magnitudes of the complex process poles appears to be preserved. It is now of interest to investigate whether the information about the magnitude of the process poles is also preserved for complex autoregressive models of order larger than unity, when only the amplitudes of the process are observed.

4.4.3 Synthetic $AR(2)$ Process

Following the analysis in the previous section, 150 samples of synthetic data were generated from a complex autoregressive process of order two, with the complex process poles at $0.8\angle 0.4\pi$ and $0.9\angle 0.6\pi$. The innovation process variance was once again set to unity.

Figure 4.4(a) illustrates the presence of a spectral peak at a frequency of 0.6π (and a smaller peak at 0.4π) in the log power spectrum of the complex $AR(2)$ process. Once again, no clear peak is visible in the corresponding incoherent log power spectrum, shown in figure 4.4(b). Figure 4.4(c) shows the complex $AR(3)$ process poles estimated by ML from the $AR(2)$ process data. The two dominant poles corresponding to the $AR(2)$ process are clearly visible, with a smaller pole also present at $0.03\angle 0.56\pi$. The ML estimates of the $AR(3)$ process parameters are also shown in table 4.2, along with the *relative* sum and difference components of the phases of the three $AR(3)$ process poles.

Process Poles		Relative Phases	
k	ML est. ($\rho_k\angle\phi_k$)	$[i, j]$	$ \phi_i \pm \phi_j $
1	$0.92\angle 0.61\pi$	[1, 2]	0.23 π 0.99 π
2	$0.79\angle 0.38\pi$	[1, 3]	0.05 π 0.83 π
3	$0.03\angle 0.56\pi$	[2, 3]	0.18 π 0.94 π

Table 4.2: Complex $AR(3)$ process poles and their relative phases. Complex $AR(3)$ process parameters estimated by ML from 150 samples of a complex $AR(2)$ process with the true process poles at $0.8\angle 0.4\pi$ and $0.9\angle 0.6\pi$.

⁷The sum and the difference phase contributions are limited to lie in the range of $[-\pi, \pi]$. Therefore, spectrum ‘folding’ occurs in the plots of the samples of the process pole phases shown in this chapter, since the relative phases have been plotted on $[-2\pi, 2\pi]$.

Figures 4.4(d)-4.4(g) show the results of 10000 iterations of a Gibbs sampler, with the first 500 samples discarded as 'burn-in'. As discussed in section 4.4.1, the Gibbs sampler was used to sample the complex $AR(3)$ model parameters, \mathbf{a} and σ^2 , conditional on the observed amplitudes of the synthetic complex $AR(2)$ process data. In particular, figure 4.4(d) shows a histogram of the relative phases of the sampled process poles. The information about the relative phases of the two *dominant* process poles appears to be preserved and to be in line with the relative phases expected from the ML estimates obtained from the coherent data (see table 4.2). Figure 4.4(e) shows that the samples of the innovation process standard deviation are centered around 1.8, and thus overestimate the true value of unity.

The histograms of the sampled magnitudes of two dominant poles of the $AR(3)$ model are shown in figures 4.4(f) and 4.4(g), along with the ML estimates obtained from the coherent data. Figures 4.4(f) and 4.4(g) illustrate that the samples obtained for the magnitudes of the two dominant poles of the $AR(3)$ process slightly underestimate the pole magnitudes obtained by ML from the complex data. However, the magnitudes of the two dominant poles still appear to be in correct proportions (0.7 and 0.85 for the sampled, *c.f.* 0.79 and 0.92 by ML), with the discrepancy potentially justified on the grounds of the overestimation of the innovation process variance. Indeed, in the light of the lack of the phase information, there may exist ambiguities between the relative contributions of the innovation process variance and the magnitudes of the process poles to the overall signal power. As pointed out in section 4.2.2, the issues associated with the identifiability of higher order autoregressive processes deserve further consideration.

4.4.4 Synthetic $AR(3)$ Process

Following the analysis in the previous sections, 150 samples of synthetic data were generated from a complex autoregressive process of order three, with the complex process poles at $0.7\angle 0.6\pi$, $0.8\angle 0.4\pi$ and $0.9\angle -0.2\pi$. The innovation process variance was once again set to unity.

Figure 4.5(a) illustrates the presence of spectral peaks at the frequencies of -0.2π and 0.4π , corresponding to the two dominant poles (and a smaller peak at 0.6π) in the log power spectrum of the complex $AR(3)$ process. Similarly, figure 4.5(b) illustrates peaks at 0.2π and 0.6π in the incoherent log power spectrum, corresponding to the sum and the difference of the peak frequencies present in the coherent power spectrum. Figure 4.5(c) shows the complex $AR(4)$ process poles estimated by ML from the $AR(3)$ process data. The three dominant poles corresponding to the $AR(3)$ process are clearly visible, with a smaller pole also present at $0.02\angle -0.15\pi$. The ML estimates of the $AR(4)$ process parameters are also shown in table 4.3, along with the *relative* sum and difference components of the phases of the three dominant $AR(4)$ process poles.

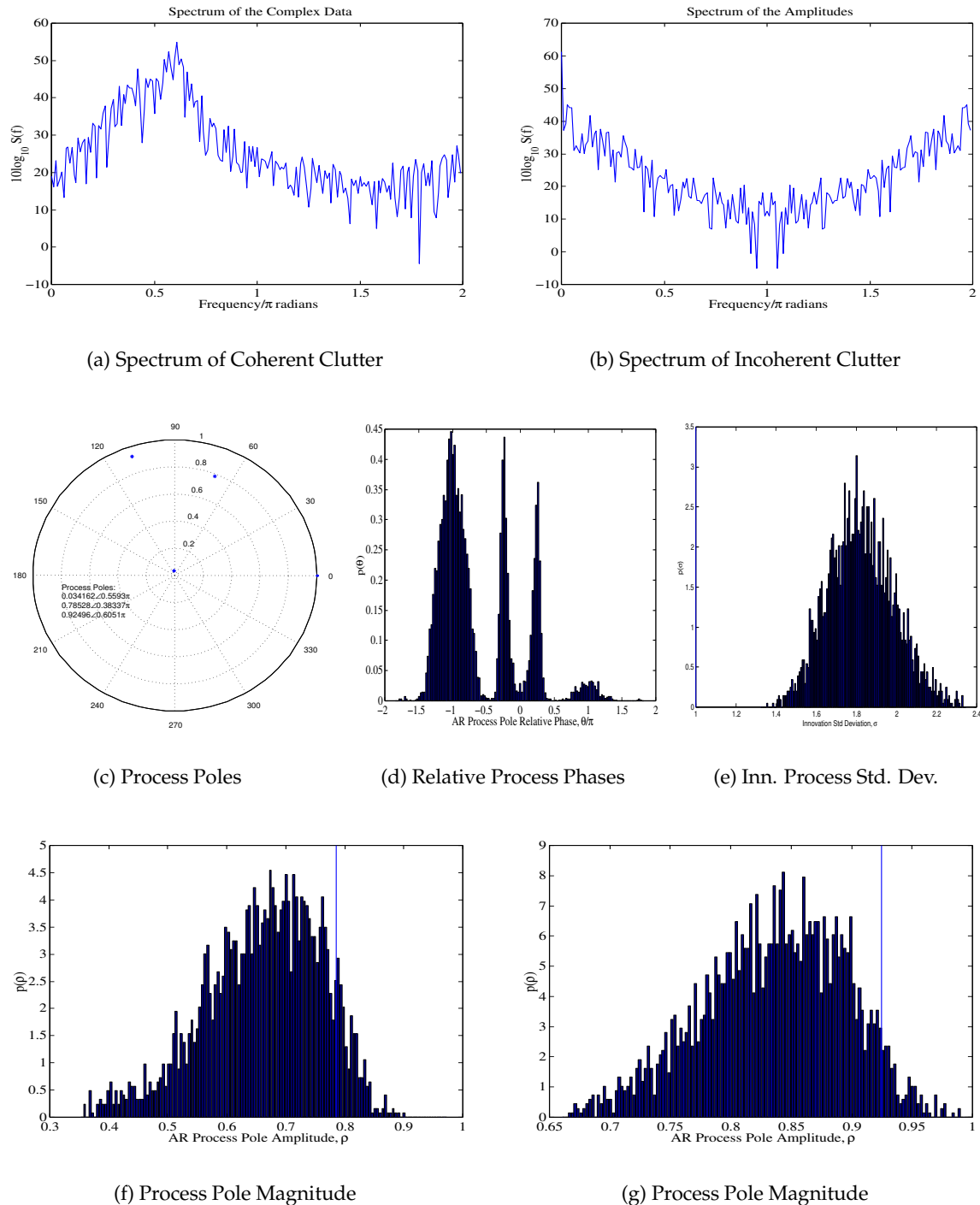


Figure 4.4: (a) Coherent and (b) Incoherent spectrum of 150 samples of a synthetic complex AR(2) process, with the process poles at $0.8 \angle 0.6\pi$ and $0.9 \angle 0.4\pi$. (c) ML estimates of complex AR(3) process parameters obtained from the complex data, along with the samples of (d) relative phases, (e) innovation process standard deviation and (f)-(g) process pole magnitudes obtained using 10000 iterations of the Gibbs sampler with random initial conditions with 500 discarded burn-in samples. ML estimates indicated by vertical lines.

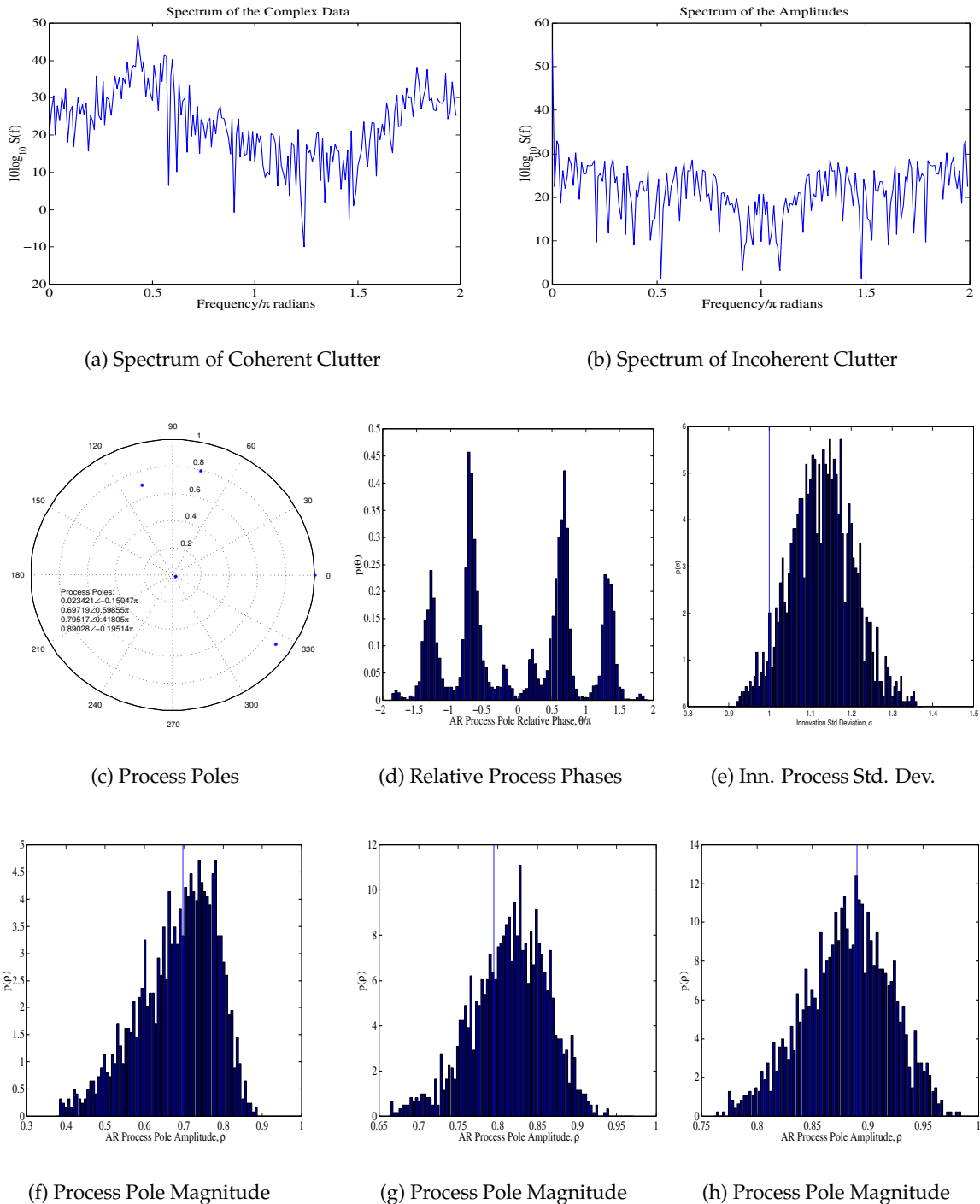


Figure 4.5: (a) Coherent and (b) Incoherent spectrum of 150 samples of a synthetic complex AR(3) process, with the process poles at $0.7 \angle 0.6\pi$, $0.8 \angle 0.4\pi$ and $0.9 \angle -0.2\pi$. (c) ML estimates of complex AR(4) process parameters obtained from the complex data, along with the samples of (d) relative phases, (e) innovation process standard deviation and (f)-(h) process pole magnitudes obtained using 10000 iterations of the Gibbs sampler with random initial conditions with 500 discarded burn-in samples. ML estimates indicated by vertical lines.

Process Poles		Relative Phases	
k	ML est. ($\rho_k \angle \phi_k$)	$[i, j]$	$ \phi_i \pm \phi_j $
1	$0.89 \angle -0.20\pi$	[1, 2]	0.22π 0.62π
2	$0.80 \angle 0.42\pi$	[1, 3]	0.4π 0.8π
3	$0.70 \angle 0.60\pi$	[2, 3]	0.18π 0.98π
4	$0.02 \angle -0.15\pi$		

Table 4.3: Complex $AR(4)$ process poles and their relative phases. Complex $AR(4)$ process parameters estimated by ML from 150 samples of a complex $AR(3)$ process with the true process poles at $0.7 \angle 0.6\pi$, $0.8 \angle 0.4\pi$ and $0.9 \angle -0.2\pi$.

Figures 4.5(d)-4.5(h) show the results of 10000 iterations of a Gibbs sampler, with the first 500 samples discarded as 'burn-in'. As discussed in section 4.4.1, the Gibbs sampler was used to sample the complex $AR(4)$ model parameters, \mathbf{a} and σ^2 , conditional on the observed amplitudes of the synthetic complex $AR(3)$ process data. In particular, figure 4.5(d) shows a histogram of the relative phases of the sampled process poles. The information about the relative phases of the two *dominant* process poles appears to be preserved and to be in line with the relative phases expected from the ML estimates obtained from the coherent data (see table 4.3). Figure 4.5(e) shows that the samples of the innovation process standard deviation are centered around 1.15, and thus slightly overestimate the true value of unity.

The histograms of the sampled magnitudes of three dominant poles of the $AR(4)$ model are shown in figures 4.5(f)-4.5(h), along with the ML estimates obtained from the coherent data. Figure 4.5(h) illustrates that the samples obtained for the magnitudes of the dominant pole of the $AR(4)$ model are centered around the ML estimate obtained from the coherent data. However, despite a slightly overestimated innovation process variance, figures 4.5(f) and 4.5(g) show that the samples obtained for the magnitudes of the second and the third largest poles are also slightly overestimated, compared to the pole magnitudes obtained by ML from the complex data.

4.4.5 Incoherent Sea Clutter

It was shown in sections 4.4.2-4.4.4 that the information about the magnitudes of the poles of a complex autoregressive process is largely preserved when only the amplitudes of the process are observed. Presently, the analysis is extended to 150 samples of vertically polarised high resolution sea clutter data. The ML parameter estimates of a complex $AR(5)$ model fitted to the coherent sea clutter data are shown in table 4.4, along with the *relative* sum and difference components of the phases of the three dominant $AR(5)$ process poles. The ML complex $AR(5)$ process pole estimates obtained from the coherent clutter data are also shown graphically in figure 4.6(c).

Process Poles		Relative Phases	
k	ML est. ($\rho_k \angle \phi_k$)	$[i, j]$	$ \phi_i \pm \phi_j $
1	$0.78 \angle -0.09\pi$	[1, 2]	0.21π 0.39π
2	$0.73 \angle -0.30\pi$	[1, 3]	0.85π 0.95π
3	$0.39 \angle -0.96\pi$	[2, 3]	0.66π 0.74π
4	$0.38 \angle 0.57\pi$		
5	$0.18 \angle -0.22\pi$		

Table 4.4: Complex $AR(5)$ process poles and their relative phases. Complex $AR(5)$ process parameters estimated by ML from 150 samples of vertically polarised high resolution sea clutter data.

Figure 4.6(a) illustrates the presence of a spectral peak at the frequencies around -0.3π (corresponding to one of the dominant poles) in the log power spectrum of the coherent clutter data. However, as illustrated by figure 4.6(b), no spectral peaks appear to be present in the log of the power spectrum obtained from the *amplitudes* of the coherent data. This observation further reinforces the observation that the second order statistics are insufficient to characterise the incoherent sea clutter data.

Figures 4.6(d)-4.6(h) show the results of 10000 iterations of a Gibbs sampler, with the first 500 samples discarded as 'burn-in'. As discussed in section 4.4.1, the Gibbs sampler was used to sample the complex $AR(5)$ model parameters, \mathbf{a} and σ^2 , conditional on the incoherent clutter data. In particular, figure 4.6(d) shows a histogram of the relative phases of the sampled process poles. The information about the relative phases of the two *dominant* process poles appears to be preserved and to be in line with the relative phases expected from the ML estimates obtained from the coherent data (see table 4.4). Figure 4.6(e) shows that the samples of the innovation process standard deviation are centered around 5.2, and thus underestimate the ML value of 6.3.

The histograms of the sampled magnitudes of three dominant poles of the $AR(5)$ model are shown in figures 4.6(f)-4.6(h), along with the ML estimates obtained from the coherent data. Figures 4.6(f)-4.6(h) illustrate that the samples obtained for the magnitudes of the three dominant poles of the $AR(5)$ model overestimate the estimates obtained by ML from the coherent clutter data. As was the case for the $AR(2)$ data in section 4.4.3, compared to the ML estimates, the samples of the process poles magnitudes are overestimated at the expense of the underestimated innovation process variance samples. In the present case, the pole magnitude-process variance ambiguity can further be exacerbated by the relatively small magnitude of the dominant poles (0.73 and 0.78) for the coherent sea clutter data.

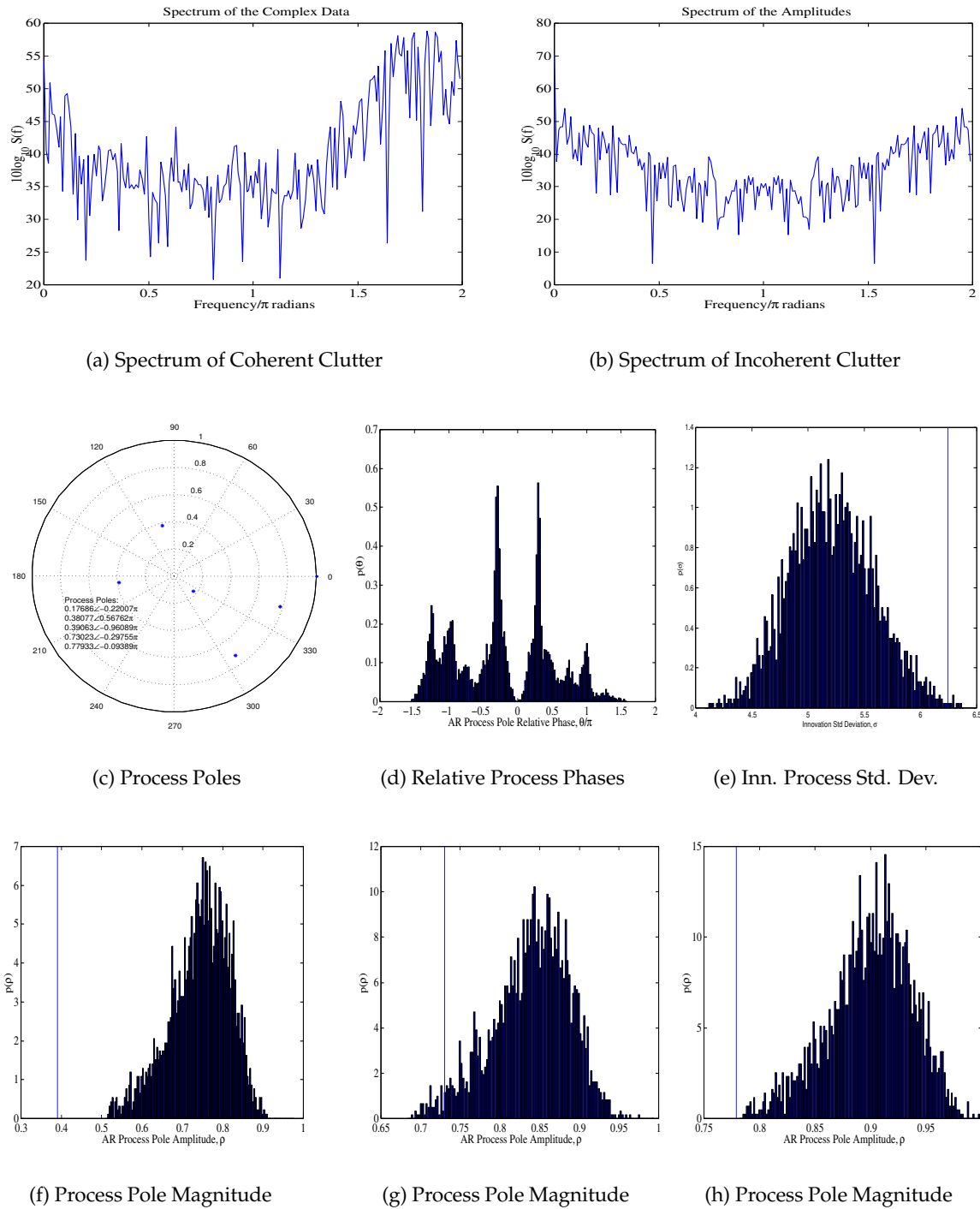


Figure 4.6: (a) Coherent and (b) Incoherent spectrum of 150 samples of high resolution radar clutter. (c) ML estimates of complex AR(5) process parameters obtained from the complex data, along with the samples of (d) relative phases, (e) innovation process standard deviation and (f)-(h) process pole magnitudes obtained using 10000 iterations of the Gibbs sampler with random initial conditions with 500 discarded burn-in samples. ML estimates indicated by vertical lines.

4.4.6 Synthetic $AR(1)$ target in the Presence of Incoherent Sea Clutter

High resolution radar returns data containing targets was not available for the purpose of this research. As a consequence, synthetic complex autoregressive target data was generated here, instead. The marginal statistics of a complex autoregressive model are Gaussian distributed, and hence the marginal statistics of the corresponding amplitudes are Rayleigh distributed. Therefore, the complex autoregressive model for the target is a special case of the Swerling I and II target models considered earlier.

However, it is very unlikely in practice that target returns would be well modelled as an autoregressive process. Consequently, autoregressive process-based characterisation of non-autoregressive process spectra in incoherent radar systems may need further consideration. Nevertheless, in the present case, 150 samples were generated from a complex $AR(1)$ process with the process pole at $0.99\angle 0.3\pi$. The synthetic $AR(1)$ target was then coherently (*i.e.* linearly) added to the vertically polarised high resolution sea clutter data. Two cases of the clutter plus target data are now treated in turn, with $SCR = 0dB$ and $SCR = 20dB$, respectively.

Sea Clutter plus Synthetic Target, $SCR = 0dB$

A complex $AR(6)$ model is presently used to model the 150 samples of clutter plus $AR(1)$ target data. The ML estimates of the $AR(6)$ process parameters are shown in table 4.5, along with the *relative* sum and difference components of the phases of the three dominant $AR(6)$ process poles. The ML complex $AR(6)$ process pole estimates obtained from the coherent data are also shown graphically in figure 4.7(c). It is interesting to note that the positions of the ML pole estimates of the clutter plus target data are different to the positions of the ML pole estimates of the individual clutter and target data. The reasons for this, along with the possible implications, are now briefly discussed.

It is relatively easy to show that a linear sum of autoregressive processes is itself an autoregressive-moving average (ARMA) process. Indeed, it is very unlikely that sea clutter data is a pure autoregressive process, as assumed in the analysis above. However, sampling for an ARMA process parameters is highly non-trivial [Wu, 1998] even when the full phase information is available. Therefore, modelling the incoherent clutter as an ARMA process-based state space model is virtually impossible. Instead, sea clutter and target characterisation based on a complex autoregressive process parameters is sought in this chapter.

One of the consequences of modelling a sum of two autoregressive processes as an autoregressive process is that the resulting process is no longer fully characterised by its poles. Therefore, the ML estimates of the poles of the combined process no longer coincide with the poles of the constituent processes. However, although not exact, it is well known that an autoregressive

process can approximate the ARMA process. Furthermore, it is still expected that an autoregressive process can largely characterise the spectral properties of an ARMA process. However, the issues associated with process identifiability in the case when only the amplitudes of the ARMA process are observed need further study.

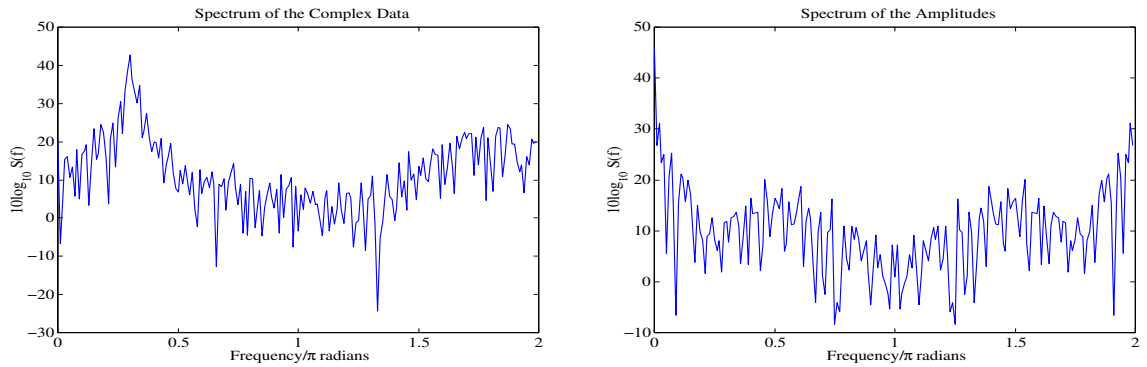
Process Poles		Relative Phases	
k	ML est. ($\rho_k \angle \phi_k$)	$[i, j]$	$ \phi_i \pm \phi_j $
1	$0.97 \angle 0.30\pi$	[1, 2]	0.15 π 0.45 π
2	$0.78 \angle -0.15\pi$	[1, 3]	0.04 π 0.64 π
3	$0.71 \angle -0.34\pi$	[2, 3]	0.19 π 0.49 π
4	$0.50 \angle 0.41\pi$		
5	$0.46 \angle 0.87\pi$		
6	$0.40 \angle -0.84\pi$		

Table 4.5: Complex $AR(6)$ process poles and their relative phases. Complex $AR(6)$ process parameters estimated by ML from 150 samples of vertically polarised high resolution sea clutter data with an artificial complex $AR(1)$ target ($SCR = 0dB$).

Figure 4.7(a) illustrates the presence of spectral peaks at frequencies around -0.15π and 0.3π in the log power spectrum of the coherent clutter plus target data. While the peak at -0.15π is associated with the background clutter, the large peak at 0.3π is associated with the $AR(1)$ target. Similarly, figure 4.7(b) illustrates a peak at 0.45π in log power spectrum of the incoherent clutter plus target data, corresponding to the relative frequency difference between the two peaks in coherent log power spectrum.

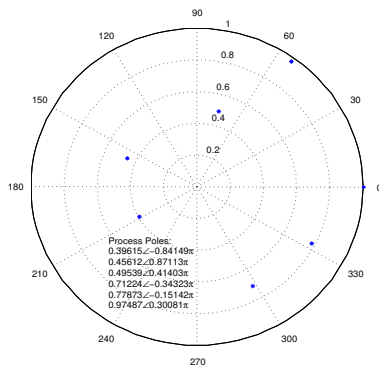
Figures 4.7(d)-4.7(h) show the results of 10000 iterations of a Gibbs sampler, with the first 500 samples discarded as 'burn-in'. As discussed in section 4.4.1, the Gibbs sampler was used to sample the complex $AR(6)$ model parameters, \mathbf{a} and σ^2 , conditional on the incoherent data. In particular, figure 4.7(d) shows a histogram of the relative phases of the sampled process poles. The information about the relative phases of the *dominant* process poles appears to be preserved and to be in line with the relative phases expected from the ML estimates obtained from the coherent data (see table 4.5). Figure 4.7(e) shows that the samples of the innovation process standard deviation are centered around 0.28, and thus overestimate the ML value of 0.23.

The histograms of the sampled magnitudes of three dominant poles of the $AR(6)$ model are shown in figures 4.7(f)-4.7(h), along with the ML estimates obtained from the coherent data. Figure 4.7(h) illustrates that the samples obtained for the magnitude of the largest dominant pole accurately reflect the ML pole magnitude estimate, while figures 4.7(f)-4.7(g) show that the samples obtained for the magnitudes of the other dominant poles overestimate the ML estimates.

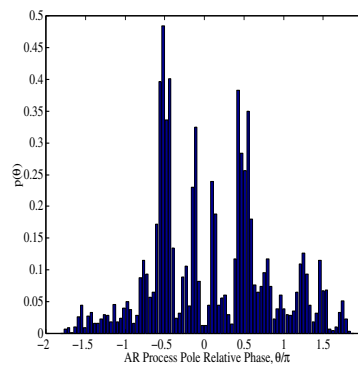


(a) Spectrum of Coherent Data

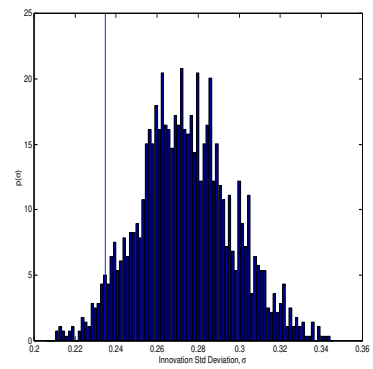
(b) Spectrum of Incoherent Data



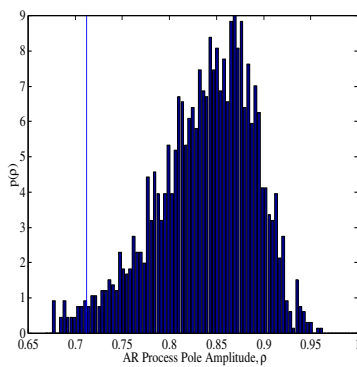
(c) Process Poles



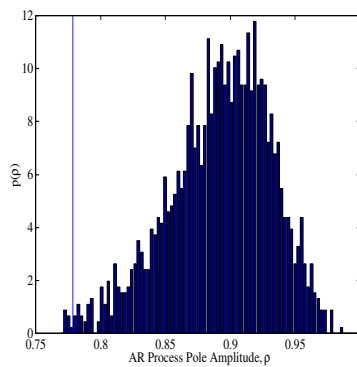
(d) Relative Process Phases



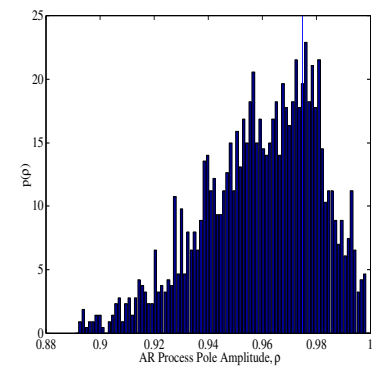
(e) Inn. Process Std. Dev.



(f) Process Pole Magnitude



(g) Process Pole Magnitude



(h) Process Pole Magnitude

Figure 4.7: (a) Coherent and (b) Incoherent spectrum of 150 samples of high resolution radar clutter with an artificial AR(1) target, with SCR = 0dB. (c) ML estimates of complex AR(6) process parameters obtained from the complex data, along with the samples of (d) relative phases, (e) innovation process standard deviation and (f)-(h) process pole magnitudes obtained using 10000 iterations of the Gibbs sampler with random initial conditions with 500 discarded burn-in samples. ML estimates indicated by vertical lines.

Sea Clutter plus Synthetic Target, $SCR = 20dB$

A complex $AR(6)$ model is presently used to model the 150 samples of clutter plus $AR(1)$ target data. The ML estimates of the $AR(6)$ process parameters are shown in table 4.6, along with the *relative* sum and difference components of the phases of three dominant $AR(6)$ process poles. The ML complex $AR(6)$ process pole estimates obtained from the coherent data are also shown graphically in figure 4.8(c). Once again, the positions of the ML pole estimates of the clutter plus target data are different to the positions of the ML pole estimates of the individual clutter and target data.

Process Poles		Relative Phases	
k	ML est. $(\rho_k \angle \phi_k)$	$[i, j]$	$ \phi_i \pm \phi_j $
1	$0.97 \angle 0.30\pi$	[1, 2]	0.24 π 0.84 π
2	$0.52 \angle -0.54\pi$	[1, 3]	0.04 π 0.56 π
3	$0.49 \angle 0.26\pi$	[2, 3]	0.28 π 0.80 π
4	$0.46 \angle -0.19\pi$		
5	$0.45 \angle 0.98\pi$		
6	$0.38 \angle 0.62\pi$		

Table 4.6: Complex $AR(6)$ process poles and their relative phases. Complex $AR(6)$ process parameters estimated by ML from 150 samples of vertically polarised high resolution sea clutter data with an artificial complex $AR(1)$ target ($SCR = 20dB$).

Figure 4.8(a) illustrates the presence of a spectral peak at the frequencies around 0.3π in the log power spectrum of the coherent clutter plus target data. Due to the relatively large $SCR = 20dB$, the target contributions associated with the synthetic $AR(1)$ target far outweigh any contributions from the background clutter. For this very reason, figure 4.8(b) illustrates that no spectral peaks are present in the log power spectrum of the incoherent clutter plus target data.

Figures 4.8(d)-4.8(h) show the results of 10000 iterations of a Gibbs sampler, with the first 500 samples discarded as 'burn-in'. As discussed in section 4.4.1, the Gibbs sampler was used to sample the complex $AR(6)$ model parameters, \mathbf{a} and σ^2 , conditional on the incoherent data. In particular, figure 4.8(d) shows a histogram of the relative phases of the sampled process poles. The information about the relative phases of the *dominant* process poles appears to be preserved and to be in line with the relative phases expected from the ML estimates obtained from the coherent data (see table 4.6). Figure 4.8(e) shows that the samples of the innovation process standard deviation are centered around 0.16, and thus slightly overestimate the ML value of 0.155.

The histograms of the sampled magnitudes of three dominant poles of the $AR(6)$ model are shown in figures 4.8(f)-4.8(h), along with the ML estimates obtained from the coherent data. Once again, figure 4.8(h) illustrates that the samples obtained for the magnitude of the largest

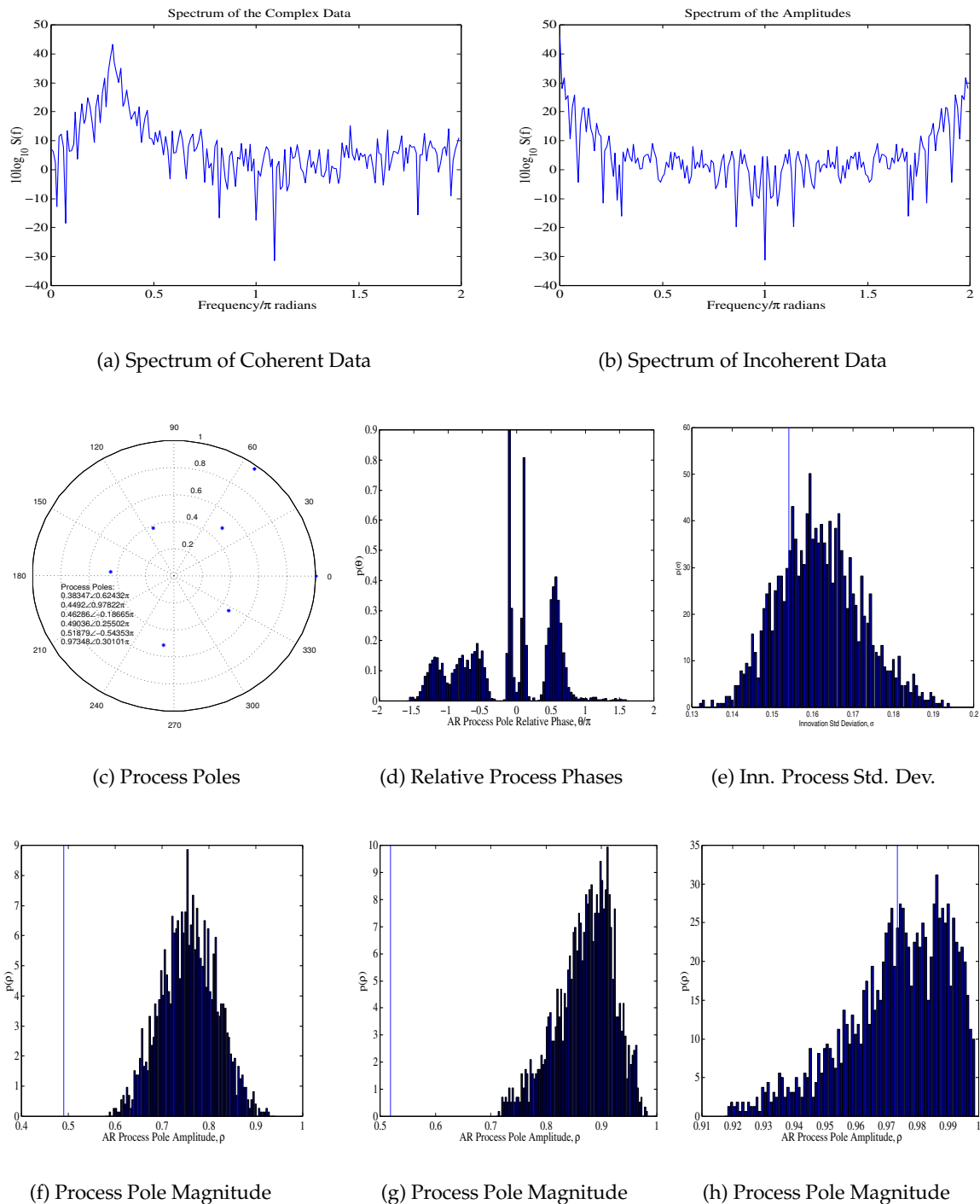


Figure 4.8: (a) Coherent and (b) Incoherent spectrum of 150 samples of high resolution radar clutter with an artificial AR(1) target, with $SCR = 20dB$. (c) ML estimates of complex AR(6) process parameters obtained from the complex data, along with the samples of (d) relative phases, (e) innovation process standard deviation and (f)-(h) process pole magnitudes obtained using 10000 iterations of the Gibbs sampler with random initial conditions with 500 discarded burn-in samples. ML estimates indicated by vertical lines.

dominant pole reflect the ML pole magnitude estimate well, while figures 4.8(f)-4.8(g) show that the samples obtained for the magnitudes of the other dominant poles overestimate the ML pole magnitude estimates.

Finally, it is interesting to compare the results presented in figures 4.7(f)-4.7(h) and 4.8(f)-4.8(h) to the results presented in figures 4.6(f)-4.6(h). In particular, the samples of the magnitudes of the autoregressive process poles estimated for the incoherent sea clutter are consistently larger than the corresponding ML estimates obtained from the coherent clutter. This behaviour is repeated for the samples of the poles corresponding to the background clutter, despite the presence of an $AR(1)$ target (see figures 4.7(f)-4.7(g) and 4.8(f)-4.8(g)). However, the samples of the magnitudes of the poles corresponding to the synthetic $AR(1)$ target accurately reflect the ML estimates, as evidenced by the results shown in figures 4.7(h) and 4.8(h).

There are two main differences between the background clutter and the synthetic $AR(1)$ target. Firstly, the magnitudes of the poles corresponding to the background clutter are much smaller than in the case of the synthetic target, indicating much larger spectral widths in the Doppler spectrum of the clutter. Secondly, the background clutter is most likely to be an ARMA process. However, the time series corresponding to the synthetic target in the presence of the background clutter is also an ARMA series. Nevertheless, the magnitude of the process pole is still accurately estimated from the process amplitudes. It is therefore interesting, as part of a future research effort, to investigate whether real targets with narrow spectral widths also lend themselves to correct estimation of the magnitude of the pole of the underlying complex process associated with the target when only the amplitudes of radar returns are observed.

4.5 DISCUSSION, CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

The principle difference between an incoherent and a coherent radar system lies in the availability of a stable phase reference. As a result, only the amplitudes of radar returns are observed in incoherent radar systems. For this reason, in the present chapter, the incoherent radar clutter was considered to correspond to the amplitudes of a complex Gaussian distributed process.

Being locally complex Gaussian distributed, the coherent radar clutter is completely characterised by its second order statistics (*i.e.* the spectrum). Furthermore, [Nohara and Haykin, 1993] states that the spectral width seems to be the only characteristic of the Doppler spectrum that uniquely distinguishes between clutter and growler spectra. For this reason, characterisation of coherent radar clutter spectrum based on the poles of a complex autoregressive process was considered in section 4.2.

Although the Doppler frequency information contained in the spectrum of the underlying complex Gaussian process is lost in the incoherent radar clutter spectrum, the information about the *relative* frequency contributions remains. Furthermore, based on a first order autoregressive model parametrisation of the underlying complex Gaussian process spectrum, it was argued that the information about the magnitude of the process pole is preserved in the spectrum of the incoherent radar clutter. A complex autoregressive process-based characterisation of high resolution sea clutter spectra was proposed in section 4.2 for incoherent radar systems. Unfortunately, due to the non-Gaussian nature of the incoherent clutter returns, it is not easy to establish the identifiability conditions for complex autoregressive processes of orders beyond unity. Instead, the Gibbs sampler was proposed to facilitate sampling from the autoregressive process parameter posterior distribution. To this end, the Hybrid Monte Carlo (HMC) algorithm was employed to conditionally sample for the missing phases.

The HMC algorithm was briefly reviewed in section 4.3. In particular, it was argued that the HMC algorithm is especially useful for sampling from high dimensional distributions whenever the first derivative of the distribution of interest is available. In addition to reviewing the basic HMC algorithm, a number of interesting insights were presented. Chiefly, the effects of a single guiding and acceptance Hamiltonian were discussed. To this end, it was argued that if the guiding and the acceptance Hamiltonians are identical, the only 'free parameter' is the step size, and that the resulting Markov Chain need not be irreducible. However, it was shown that by employing two different Hamiltonians, the 'momenta temperature' can improve the 'mobility' of the proposed state at the expense of the acceptance rate. As an extension of this argument, a symmetrical momentum-based tempering scheme was proposed for improved mixing of the HMC algorithm. The sampling performance of the proposed scheme was illustrated on a multi-modal example.

The research reported in this chapter was originally undertaken with a view of potential applications for target detection in incoherent radar systems. However, as was pointed out in section 4.4, each run of 10000 iterations of the sampling algorithm for 150 data samples takes in excess of 15 hours to complete on Pentium II based computers (*i.e.* about 10 minutes for 100 iterations!!!). In practice, each data block would have to be processed in real time for *each* range cell in a range profile. With 1kHz sampling frequency and up to 40000 range cells per range profile, it is very unlikely indeed that such an algorithm would prove to be of any practical value.

As the results in section 4.4 indicate, the relative phases and the magnitudes of complex autoregressive process poles appear to be largely preserved when only the process amplitudes are observed. However, as illustrated by the $AR(2)$ process example in section 4.4.3, ambiguities in the process pole magnitude *vs* the innovation process variance may arise. Therefore, other than in the case of a first order complex autoregressive process, the issue of process parameter

identifiability needs closer attention, as does the question of model order selection.

Finally, as was argued in section 4.4, autoregressive process-based characterisation of non-autoregressive process spectra (as is the case for some targets) in incoherent radar systems may need further consideration. Furthermore, since a sum of two autoregressive processes is an ARMA process, it is very likely that sea clutter may actually be better characterised using an ARMA process. However, in addition to exacerbating the problems of identifiability, the ARMA process would add another level of computational complexity to the already computationally intensive algorithm.

In short, autoregressive process-based characterisation of Doppler spectra of the underlying complex Gaussian processes associated with incoherent radar clutter is fraught with dangers. Instead, an approach based on the higher order spectra of the incoherent radar clutter may prove to be of more practical value (see [Nikias and Mendel, 1993],[Brillinger, 1992] and [Mendel, 1991] for a review of higher order statistics methods). However, the study of target detection techniques based on the higher order spectra of incoherent radar returns is left as an exercise for future research.

Modulating Component Modelling using Conditional Heteroscedastic Models

5

5.1 INTRODUCTION

Radar backscatter from an ocean surface consists of a fast oscillating speckle component modulated in power by a slowly varying modulating process. Being able to capture the statistics and the correlation structure of the clutter is of paramount importance for optimal target detection in the presence of sea clutter. While the modelling of the temporal statistics of the speckle component was addressed in chapter 4, the present chapter is concerned with the modelling of the spatial statistics of the modulating component of high resolution sea clutter.

The modulating component of sea clutter is widely agreed to be associated with the swell structure of the sea [Skolnik, 1980]. To the best of the author's knowledge, no satisfactory physical models for the scattering process exist. However, as [Jakeman and Tough, 1988] state: *as the only non-Rayleigh distribution based on tabulated functions for which [the distribution corresponding to the scattering from N scattering centres] could be evaluated analytically, while at the same time satisfying the criteria imposed by experimental data, the K -distribution was proposed as a semi-empirical model*. Subsequently, with the advent of the compound models for sea clutter, the modulating component of sea clutter is widely believed to be Gamma distributed [Ward et al., 1990].

In chapter 2, it was shown that the marginal distribution of the modulating component can range anywhere from Inverse Gamma, through log-Normal to Gamma distribution. In fact, target detection performance analysis based on the Gamma distribution assumption for the modulating component of high resolution sea clutter can lead to overly optimistic predictions of the probability of false alarm in compound clutter environment. For this reason, a more pragmatic view is adopted in this chapter. In particular, a flexible family of models is considered, such that the correlation properties and the marginal statistics of the modulating component are well (though *not* exactly) modelled, without the rather dogmatic assumptions being made about the univariate marginal of the process being Gamma distributed. To this end, as an extension of the arguments based on birth-death migration of a population of scattering centres on the sea surface, a family of conditionally heteroscedastic (*i.e.* non-constant prediction error variance) models is proposed for the modulating component of high resolution sea clutter.

The remainder of this chapter is structured as follows. In section 5.2, an argument based on a birth-death migration process for the evolution of a population of scattering centres on the sea surface, in conjunction with Ito's Lemma, is used to derive a form of heteroscedastic model in continuous time for the modulating process in the logarithm domain. Such a model, when discretised, takes on the form of a conditional heteroscedastic process. The properties of such models are briefly discussed in section 5.2. Sections 5.3 and 5.4 investigate the form of the heteroscedasticity present in the temporal dynamics of the modulating component, and high resolution sea clutter range profiles, respectively. Finally, conclusions and suggestions for further research are presented in section 5.5.

5.2 CONDITIONAL HETEROSCEDASTIC MODELS

Arguments based on random walk and birth-death migration processes are used by [Jakeman and Tough, 1988] to elucidate the significance of the K-distribution for modelling the statistics of scattered waves. In particular, based on a statistical-mechanical formulation of birth-death migration processes applied to a population of scattering centres, [Jakeman and Tough, 1988] develop a Stochastic Differential Equation (SDE) model for the modulating component of sea clutter, which corresponds to the normalised cross-section of scatterers in an illuminated area.

Although the SDE model is inadequate to capture the correlation structure of real sea clutter [Oliver and Tough, 1986], it does provide a convenient starting point for the development of more sophisticated, semi-phenomenological models for the modulating component of sea clutter.

In their paper, [Jakeman and Tough, 1988] use a statistical mechanical framework to derive a Fokker-Planck equation corresponding to the birth-death migration process, from which they derive the corresponding stochastic differential equation of the form

$$\begin{aligned} d\sigma &= (\alpha - \beta\sigma)dt + g^{1/2}dz \\ g &= 2\sigma \end{aligned} \tag{5.1}$$

where σ corresponds to the normalised cross-section of the scatterers in the illuminated area (*i.e.* the scattering power, or the *intensity* of the modulating process), and where dz is a zero mean, unity variance Wiener process. The parameters α and β determine the shape of the marginal distribution and the decay rate of the autocorrelation function, which for this SDE model happens to be exponential.

Such a model, in its present form, is of rather limited utility for the purpose of sea clutter modelling. This model is only capable of modelling exponentially decaying autocorrelations, and the marginal distribution of the process, σ , is $\Gamma(\alpha, \beta)$ distributed. Furthermore, when discretised, the resulting difference equation for the process assigns non-zero probabilities to negative intensities, which is clearly not physically possible.

It proves to be more convenient to transform the SDE given in equation 5.1 into the logarithm domain of the modulating process, *i.e.* let $y = \ln \sigma$. By applying Ito's Lemma [Hull, 1989], it can be shown that the SDE for y (corresponding to the SDE for σ in equation 5.1) is given by

$$\begin{aligned} dy &= \frac{1}{2} \{(\alpha - 1) \exp(-y) - \beta\} dt + h^{1/2} dz \\ h &= \exp(\ln 2 - y) \end{aligned} \quad (5.2)$$

where dz is the same Wiener process as that in equation 5.1. There are a number of reasons why modelling the modulating component of sea clutter in the logarithm, as opposed to the intensity, domain is preferable.

Firstly, equations 5.1 and 5.2 for σ and y , respectively, model the same stochastic process in two different domains, with $\sigma = \exp(y)$. However, unlike the intensity σ , its logarithm y , need not be positive. Therefore the discretisation of equation 5.2 does not result in any physically implausible models. Secondly, as was pointed out in chapter 2, depending on the viewing angle and environmental conditions, the single point statistics of the modulating component can range anywhere from Gamma, through log-Normal to Inverse-Gamma distributions. As, from the detection prediction point of view, the log-Normal distribution provides a good fit (though not necessarily the best one) to the data under most conditions, it would therefore seem that the logarithm domain is the natural domain for modelling the modulating component. Finally, as will be discussed in the following chapter, due to the multiplicative nature of sea clutter (*i.e.* the speckle and the modulating component), the process associated with the clutter becomes additive in the logarithm domain. State space methods can then in principle be employed to estimate the modulating component within a range cell, conditional on the neighbouring clutter samples within the given range profile.

The processes corresponding to equations 5.1 and 5.2 have state-dependent noise variances g and h , respectively. Equations of this type are frequently said to represent multiplicative noise processes, while the constant noise variance processes have been characterised as additive-noise processes. Using the statistical-mechanical framework [Jakeman and Tough, 1988] further demonstrate that processes with different marginal distributions (including Inverse Gamma marginal) can be obtained by merely changing the form of the variance in the SDE description of the process.

5.2.1 Discrete Time Conditional Heteroscedastic Models

The stochastic differential equation model presented in the previous section is an idealised *continuous time* model of the modulating component of sea clutter, based on the birth-death migration process model for a population of scattering centers on the sea surface. The SDE in equation 5.2 can be approximately discretised using a simple Euler integration algorithm [Hull, 1989] to give

$$\begin{aligned} y_t &= y_{t-1} + \frac{1}{2} \{(\alpha - 1) \exp(-y_{t-1}) - \beta\} \Delta t + h_t^{1/2} \epsilon_t & \epsilon_t &\sim N(0, 1) \\ h_t &= \exp(\ln(2\Delta t) - y_{t-1}) \end{aligned} \quad (5.3)$$

This equation is the discrete time analogue of the continuous Gamma process in the logarithm domain.

The discrete process described by equation 5.3 falls into a broader class of *conditional heteroscedastic* models which were first introduced by [Engle, 1982] in the context of modelling UK inflation. Such models are characterised by being conditionally Gaussian, and by non-constant and state-dependent process variance. In their most general form, the conditional heteroscedastic models (as introduced by [Engle, 1982]) for time series y_t can be written as

$$\begin{aligned} y_t &= \sum_{j=0}^q \beta_j x_j + e_t \\ e_t &= h_t^{1/2} \epsilon_t & \epsilon_t &\sim N(0, 1) \\ h_t &= h(e_{t-1}, e_{t-2}, \dots, y_{t-1}, y_{t-2}, \dots, h_{t-1}, h_{t-2}, \dots, \alpha) \end{aligned}$$

where α and β are the model parameter vectors. In this notation, e_t is a conditionally heteroscedastic (CH) process which is conditionally Gaussian, and for which the variance is a function of any *past* information (*e.g.* past observations y , past innovations e , or past variances h). The process y_t , on the other hand, is referred to as a CH regression, where the x_j 's refer to any exogenous or endogenous variables (such as the previous observations) on which y_t is being regressed.

As will subsequently become apparent (see section 5.3), it is the simple autoregressive processes driven by conditionally heteroscedastic errors (AR-CH) of the form

$$\begin{aligned} y_t &= \beta_0 + \sum_{j=1}^q \beta_j y_{t-j} + e_t \\ e_t &= h_t^{1/2} \epsilon_t & \epsilon_t &\sim N(0, 1) \end{aligned} \quad (5.4)$$

that are of interest in the context of modelling the modulating component of sea clutter. The form of the innovation process variance h_t will be discussed and studied subsequently.

It is interesting to note that the conditional mean of y_t is a function of the previous samples $\{y_{t-j}\}$, while the unconditional mean of y_t is constant ($\frac{\beta_0}{1-\sum_j \beta_j}$). The improvement in the forecast of the underlying modulating process y_t stems from the use of the conditional mean. However, ignoring the heteroscedastic effects may not greatly affect the point forecasts from the model, *i.e.* the point forecasts provided by the AR-CH models may not be too dissimilar to the point forecasts obtained using simple AR models. On the other hand, the AR-CH model may give more efficient estimates and information about the conditional variance as well as the conditional mean. This could prove to be particularly important in the case of CFAR detection, where the level of the underlying modulation process is estimated (*i.e. predicted*) from the neighbouring range cells. The knowledge of both the mean and the variance of the conditional forecasts can potentially provide means of obtaining better threshold estimates compared to the simple, cell averaging type CFAR detectors.

The Conditional Variance Model

As was mentioned in the previous section, the modulating component of sea clutter in the logarithm domain (*i.e.* $y = \ln \sigma$) could prove to be modelled well by an autoregressive process driven by conditionally heteroscedastic innovations (*i.e.* $e_t = h_t^{1/2} \epsilon_t$), as given by equation 5.4. However, the form of the conditional variance of e_t has not been yet defined. For the case of a Gamma distributed process corresponding to the birth-death migration process in the logarithm domain, equation 5.3 suggests the following form for the logarithm of the conditional variance

$$\text{AR-ECH} \quad \ln h_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i}$$

The resulting model is an autoregressive model driven by exponential conditional heteroscedastic process, and hence from now on, such a model will be referred to as an AR-ECH model.

The AR-ECH models provide only one family of conditional heteroscedastic models. In fact, in the original work of [Engle, 1982] and subsequently extended by [Weiss, 1984b], conditionally heteroscedastic models were proposed with the conditional variance of the form

$$\text{AR-ARCH} \quad h_t = \alpha_0 + \sum_{i=1}^p \alpha_i e_{t-i}^2 + \sum_{k=1}^r \alpha_{p+k} (y_{t-k} - \bar{y})^2$$

where $\bar{y} = E(y_t)$, *i.e.* the unconditional mean of the process y_t . The conditional variance for the AR-ARCH model depends on the past observations $\{y_{t-i}\}$ as well as the past prediction errors $\{e_{t-k}\}$.

As will subsequently be discussed, although the innovations e_t for AR-ARCH models are independent, the squared prediction errors e_t^2 exhibit autoregressive behaviour (hence the name ARCH). These models have been further extended by [Bollerslev, 1986] to include dependence of the conditional variance on previous variances $\{h_{t-j}\}$. The resulting models are referred to as generalised ARCH (GARCH) processes, and the autocorrelation function of e_t^2 for such models is analogous to that of an ARMA process [Bollerslev, 1986]. However, due to the associated complexity of such models, they will not be considered further for the purpose of sea clutter modelling. In fact, there is a myriad of other forms of CH models which can potentially be used to model heteroscedasticity in the data [Bollerslev et al., 1992]. Some models describe the variance itself as a stochastic process, and are referred to as the Stochastic Volatility (SV) models [Harvey and Shephard, 1996]. Unfortunately, analysis of SV models becomes prohibitively expensive and calls for sophisticated Monte Carlo sampling schemes. However, it is the purpose of this chapter to investigate *simple* models for the modulating component of sea clutter, which can be easily incorporated into a real time CFAR target detector.

The AR-ARCH model and its extensions such as the AR-GARCH model require the conditional variance to be positive, thus introducing constraints on the parameter space (all parameters in the conditional variance process must be non-negative). The positivity constraints on the parameters can be removed by modelling the logarithm of the variance as a regression on previous observations and prediction errors. The resulting models, analogous to AR-ARCH model in the logarithm domain, take on the form

$$\begin{aligned} \text{AR-EARCH} \quad \ln h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i \ln e_{t-i}^2 \\ \text{AR-ELCH} \quad \ln h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i \ln (y_{t-i} - \bar{y})^2 \end{aligned}$$

In particular, instead of the logarithm of the conditional variance being regressed on e_{t-i}^2 and $(y_{t-i} - \bar{y})^2$ as would be the case for the conditional variance of the AR-ARCH process, the logarithm of the conditional variances of AR-EARCH and AR-ELCH processes are regressed on $\ln e_{t-i}^2$ and $\ln (y_{t-i} - \bar{y})^2$, respectively. Although, as will be discussed in the following section, the stability conditions for $E(e^2)$ for AR-EARCH and AR-ELCH processes are not readily defined, stability conditions for $E(\ln e^2)$ are readily obtainable.

Furthermore, it is interesting to note that the conditional variance of AR-ARCH model is the average of previous squared prediction errors and observations, while for the AR-EARCH and AR-ELCH models, the conditional variance is effectively a geometric product of previous squared prediction errors and observations.

Properties of AR-CH Models

Regardless of the form of the conditional variance, it is true for all AR-CH models defined by equation 5.4 that $E(e_t|I_{t-1}) = 0$, where I_{t-1} is the information set consisting of all the observations y_t up to time $t - 1$. Furthermore, as is the case for ordinary autoregressive processes, the prediction errors for an AR-CH process are uncorrelated. It can also be readily seen from equation 5.4 that the conditional variance and the conditional moment of the logarithm of the innovation process e_t is given by

$$\begin{aligned} E(e_t^2|I_{t-1}) &= h_t \\ E(\ln e_t^2|I_{t-1}) &= E(\ln h_t|I_{t-1}) + E(\ln \epsilon_t^2|I_{t-1}) \\ &= \ln h_t + \{\psi^{(0)}(1/2) + \ln 2\} \end{aligned}$$

where $\psi^{(0)}(\cdot)$ is the digamma function. The unconditional variance and the unconditional moment of the logarithm of the prediction error are given by

$$E(e_t^2) = E(h_t) \quad (5.5)$$

$$E(\ln e_t^2) = E(\ln h_t) + \{\psi(1/2) + \ln 2\} \quad (5.6)$$

provided that the e_t 's are i.i.d. Gaussian.

In the context of AR-ARCH models, [Weiss, 1984b] argue that *the most basic assumptions on the process are those needed for the process to be well defined and for e_t and hence y_t to have finite and constant second moments (i.e. for the process to be wide sense stationary).*

One requirement regarding stationarity for all the above mentioned AR-CH processes, is that the roots of the autoregressive process polynomial (see equation 5.4)

$$a(B) = 1 - \beta_1 B - \dots - \beta_q B^q$$

lie outside the unit circle. Then, provided e_t is wide sense stationary, so is y_t .

By substituting the expression for h_t into equation 5.5 [Weiss, 1984b] show that, provided that $\alpha_0 > 0$ and $\alpha_i \geq 0$, the difference equation in $E(h_t)$ for AR-ARCH process is stable (i.e. $E(e_t^2) < \infty$) when

$$\sum_{i=1}^p \alpha_i + \sum_{k=1}^r \alpha_{p+k} \frac{\text{var}(y_t)}{\text{var}(e_t)} < 1$$

Furthermore, by studying the second and fourth moment structure of AR(0)-ARCH(q) processes (i.e. no autoregression present), [Engle, 1982] and [Bollerslev, 1986] show that AR(0)-ARCH(q) processes are leptokurtic (i.e. heavy tailed).

The AR-CH models with exponential conditional variances (*i.e.* AR-ECH, AR-ELCH and AR-EARCH) have no positivity constraints on the parameters. Unfortunately, the variance of e_t^2 , as well as the higher order moments, are not easily obtained for these models. Therefore, the conditions for wide sense stationarity are not easily derived. Instead, by substituting the expression for $\ln h_t$ into equation 5.6, conditions necessary to ensure that the difference equation in $E(\ln h_t)$ is stable (*i.e.* $E(\ln \epsilon^2) < \infty$) are straight-forwardly obtained as

$$\begin{aligned} \text{AR-ECH} & \quad \sum_i \alpha_i \frac{\beta_0}{1 - \sum_j \beta_j} < \infty \\ \text{AR-ELCH} & \quad \sum_i \alpha_i < 1 \\ \text{AR-EARCH} & \quad \sum_i \alpha_i \frac{\text{var}(y_t)}{\text{var}(e_t)} < 1 \end{aligned}$$

Parameter Estimation

Since ϵ_t in equation 5.4 is i.i.d. Gaussian, it follows that the likelihood for y_t is conditionally Gaussian. Hence, the log likelihood for T AR-CH process samples $\{y_t\}$, apart from some constants, is given by [Engle, 1982]

$$\begin{aligned} L(\Theta) &= \frac{1}{T} \sum_{t=1}^T l_t(\Theta) \\ l_t(\Theta) &= -\frac{1}{2} \ln h_t - \frac{1}{2} e_t^2 / h_t \end{aligned}$$

where $\Theta = \{\alpha, \beta\}$ is the set of model parameters. The score and the Hessian for the process follow automatically, as given in appendix C.1.

As discussed in [Weiss, 1984b], the Maximum Likelihood estimates of the parameters are consistent and asymptotically normal. Furthermore, as the off-diagonal block elements of the information matrix are zero, the estimation of the AR and CH parameters (β and α , respectively) can be undertaken separately without asymptotic loss of efficiency [Engle, 1982].

The parameter estimation procedure recommended by [Engle, 1982] is to initially estimate β by ordinary least squares (*i.e.* assume no heteroscedasticity present). Since the off-diagonal elements of the information matrix are exactly zero, efficient estimates of α and β can then be calculated by iterating between α and β using the scoring algorithm, given by

$$\phi^{i+1} = \phi^i + [\hat{\mathcal{I}}_{\phi\phi}^i]^{-1} \frac{\partial L^i}{\partial \phi}$$

where $\frac{\partial L^i}{\partial \phi}$ and $\hat{\mathcal{I}}_{\phi\phi}^i$ are the score and the information matrix at the i^{th} iteration. The interested reader is referred to appendix C.1.3 for the implementational details of the parameter estimation algorithm.

Identification and Testing

The final point associated with the AR-CH models concerns identification and testing. These issues have been discussed at length in [Engle, 1982] and [Weiss, 1984a]. In particular, the use of the correlation structure of e_t^2 is recommended at the identification stage, while more involved Lagrange Multiplier hypothesis tests are recommended for testing for presence of ARCH. Such tests, though interesting from a statistical point of view, will not be undertaken in this thesis. Instead, as discussed in section 5.3, a more pragmatic engineering approach based on the prediction and generalisation properties of the models on real data sets will be adopted herein.

[Weiss, 1984a] recommends firstly identifying and estimating the AR model in equation 5.4 in the usual way, with CH being treated as heteroscedasticity and requiring only the adjustment of the standard errors of the estimates. On the other hand, identification of the CH model can be performed informally using the correlation structure of e_t^2 and $\ln e_t^2$. This follows from results presented in [Bollerslev, 1986]. In particular, it can be shown that equivalent representations for AR-CH models discussed above are given by

$$\begin{aligned} \text{AR-ARCH} \quad e_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i e_{t-i}^2 + \sum_{k=1}^r \alpha_{p+k} (y_{t-k} - \bar{y})^2 + \nu_t \\ \text{AR-ECH} \quad \ln e_t^2 &= \alpha_0 + c + \sum_{i=1}^p \alpha_i y_{t-i} + \eta_t \end{aligned} \quad (5.7)$$

$$\begin{aligned} \text{AR-ELCH} \quad \ln e_t^2 &= \alpha_0 + c + \sum_{i=1}^p \alpha_i \ln(y_{t-i} - \bar{y})^2 + \eta_t \\ \text{AR-EARCH} \quad \ln e_t^2 &= \alpha_0 + c + \sum_{i=1}^p \alpha_i \ln e_{t-i}^2 + \eta_t \end{aligned}$$

where

$$\begin{aligned} c &= \psi^{(0)}(1/2) + \ln 2 \\ \nu_t &= e_t^2 - h_t = (\epsilon_t^2 - 1)h_t \\ \eta_t &= \ln(e_t^2/h_t) - c = \ln \epsilon_t^2 - c \end{aligned}$$

with $\epsilon_t \sim N(0, 1)$, and hence, by definition, ν_t and η_t being serially uncorrelated and having a zero mean. These equivalent representations, though in themselves not as useful as the origi-

nal models, can be used as an informal tool for identification of conditional heteroscedasticity based on the correlation structure of e_t^2 and $\ln e_t^2$.

5.3 HETEROSCEDASTICITY IN SEA CLUTTER

As was pointed out in the discussion above, based on the arguments on birth-death migration for a population of scattering centres in the logarithm domain, non-constant prediction error variance is expected for the modulating component of sea clutter. It is the aim of this section to illustrate that there indeed may be some evidence to support the premise of the presence of heteroscedasticity in sea clutter.

The modulating component data used in this section has been obtained by averaging blocks of 250 pulse by pulse radar return samples from a single range cell. The data was obtained using a high resolution radar at look angle of 215.1° into the wind at 7.5km range. The resolution of each range cell is effectively 4m in range and 160m in azimuth (see chapter 2 for more details).

As the birth-death migration process arguments are formulated in terms of the temporal evolution of a population of scattering centres, the present discussion is concerned with the study of temporal evolution of the modulating component within a single range cell. The analysis is extended to the more interesting case of range profiles in section 5.4.

Figures 5.1(a) and 5.1(b) show 400 time samples corresponding to 200 sec of temporal evolution of the modulating component within a single range cell, while figures 5.1(c) and 5.1(d) show the corresponding histograms of the amplitudes of horizontally and vertically polarised clutter, respectively. The order parameter, α , for the horizontally and the vertically polarised data was found to be 0.61 and 1.28, respectively, while the corresponding correlation length ρ (*i.e.* the rate of decay of the autocorrelation function) was found to be 0.61 and 0.84, respectively.

The log likelihood corresponding to the Inverse Gamma, the log-Normal and the Gamma distribution, based on the Maximum Likelihood parameter estimates for the modulating component in the intensity domain, is given in figure 5.1. The likelihood for the data indicates that the intensity of the horizontally polarised clutter is best modelled by the Inverse Gamma distribution, while the intensity of the vertically polarised data is best modelled by the log-Normal distribution. However, the likelihood does not provide a particularly solid foundation for model comparison [Wu, 1998] and one has to be very wary of drawing any conclusions on the basis of likelihood alone. In fact, figure 5.1 shows that while the Gamma distribution provides a rather poor fit to the data in either polarisation, the log-Normal distribution appears to fit the single point statistics well. This point was discussed in more detail in chapter 3.

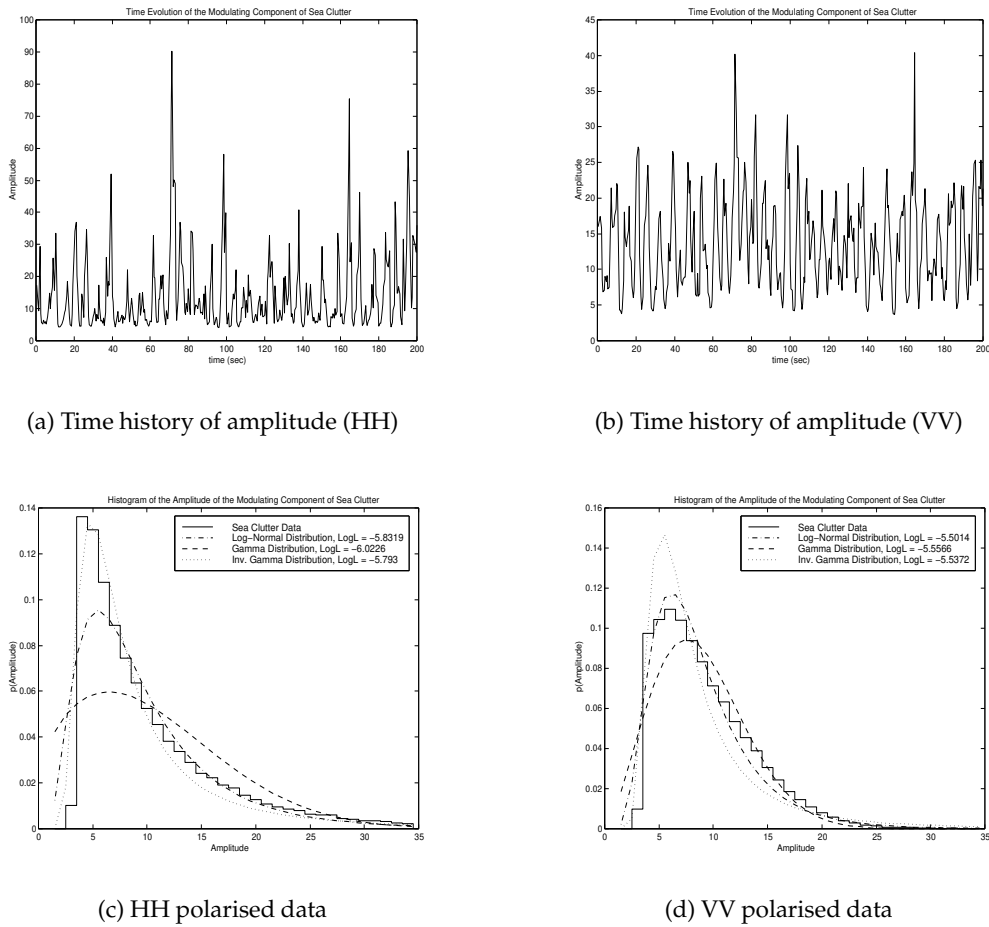


Figure 5.1: The time series and the histograms of the amplitude of the modulating component of horizontally and vertically polarised sea clutter. The log likelihoods for the data, based on the log-Normal, Gamma and Inverse Gamma distributions are shown in the legends.

5.3.1 The Form of Regression

The discretised model for the birth-death migration process (equation 5.3) suggests that the modulating component in the logarithm domain, $y_t = \ln \sigma_t$, evolves as a regression on $\{y_{t-j}\}$ and $\{\exp(-b_j y_{t-j})\}$. As the model is not exact, and only has a semi-phenomenological basis, the form of regression stipulated by the model can only be taken as a guide and it is therefore of interest to investigate the potential presence of any non-linearity in real clutter data.

The autocorrelation of the time series corresponding to the modulating component in the logarithm domain is shown in figure 5.2. The results in the figure illustrate that there is a clear correlation structure in the modulating component of sea clutter in the logarithm domain in either polarisation. These *linear* correlations can be removed by regressing y_t on the lagged observations $\{y_{t-j}\}$. The resulting prediction errors (innovations), $e_{AR}(t)$, corresponding to the

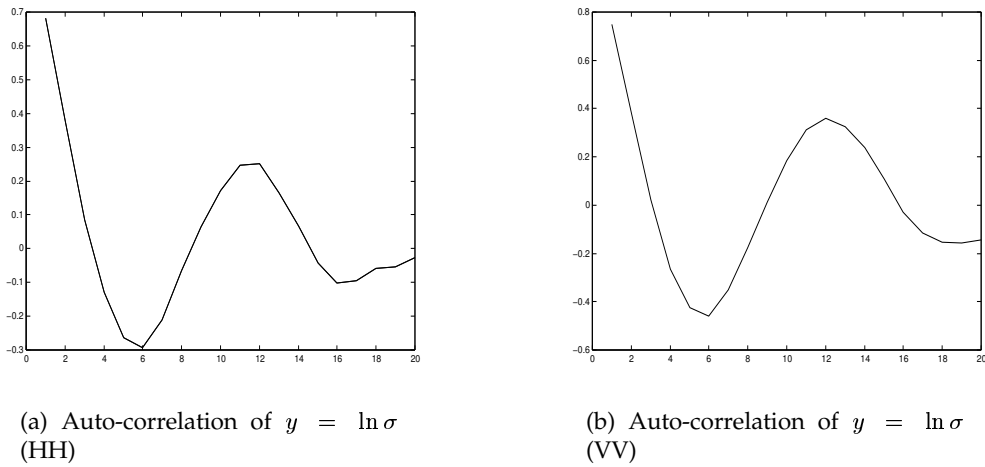


Figure 5.2: Autocorrelation of the modulating component in the logarithm domain, $y = \ln \sigma$, for (a) horizontally and (b) vertically polarised sea clutter.

linear autoregressive process are given by

$$e_{AR}(t) = y_t - \left(\beta_0 + \sum_{j=1}^Q \beta_j y_{t-j} \right)$$

If the process associated with the modulating component in the logarithm domain were purely linear and Gaussian, then assuming stationarity and no Moving Average component in the innovations (which would show up in the autocorrelation structure of e_{AR}), the innovations corresponding to the autoregressive model, e_{AR} , would be i.i.d. Gaussian distributed, and independent of any *past* observations. Figure 5.3 shows that the innovations, e_{AR} , resulting from fitting the data with an AR(12) model are indeed uncorrelated up to lag 12.

The model of equation 5.3 suggests the presence of a non-linearity of the form e^{-by} (where equation 5.3 stipulates $b = 1$). The exponent b , minimising the mean squared prediction error, $E(e^2)$ (with $e_n = e_{AR}(n) - \sum_k \gamma_k e^{-by_{n-k}}$), was found to be 0.05 and 8.5 for horizontally and vertically polarised clutter, respectively. Figure 5.3 also shows the cross correlation between the autoregressive innovations, e_{AR} , and the non-linear component e^{-by} , where b is set to 1, as well as to the least mean squared error value specified above. From visual inspection of the figure, it appears that the non-linear component is largely decorrelated with e_{AR} in the case of the horizontally polarised clutter, while correlations appear to be present in the case of the vertically polarised clutter.

To investigate the potential utility of including the non-linearity in the regression model for the modulating component of sea clutter, the prediction errors of the autoregressive model with

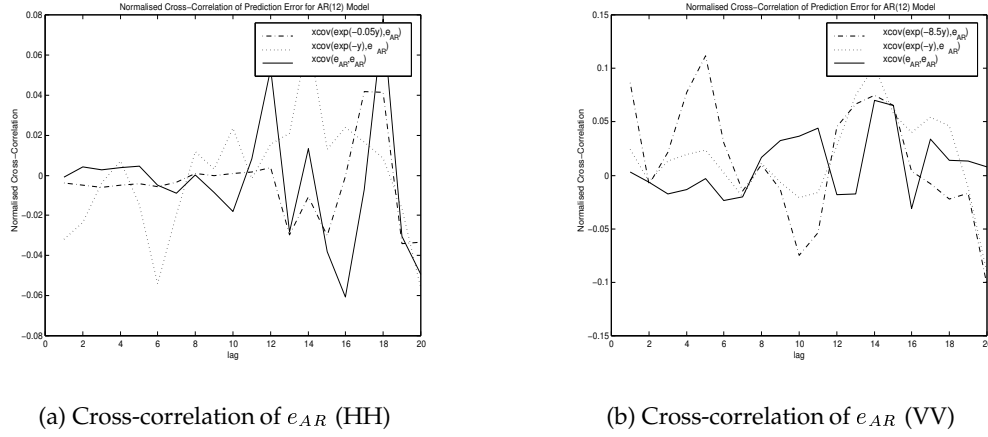


Figure 5.3: Cross-correlation of e_{AR} and the non-linear regression on the observations, given by e^{-by} , where b is set to 1, as well as to the value minimising the mean squared prediction error, $E(e^2)$. Results shown for (a) horizontally and (b) vertically polarised sea clutter.

non-linear terms were evaluated for a training and a test sequence. The training set consists of 200 sequential samples from a single range cell, while the test sequence consists of 200 sequential samples corresponding to the same time period, but from an adjacent range cell. In this way, it is hoped that any non-stationarity effects are avoided, while the two data sets are sufficiently independent to serve as a suitable validation set. The choice of the neighbouring range cell for the test sequence is also based on the observation that in radar systems, the threshold and clutter parameters are derived from the statistics in the neighbouring range cells.

Model	Prediction Error (dB) Training Set		Prediction Error (dB) Test Set	
	HH	VV	HH	VV
AR(12)	3.26	5.00	4.28	5.74
AR(12)+NL(5,b=1)	3.28	5.09	4.29	5.82
AR(12)+NL(5,b)	3.29 (b=0.05)	5.19 (b=8.5)	4.30 (b=0.05)	5.92 (b=8.5)

Table 5.1: Improvement in the prediction error in dB, for 200 samples of the modulating component of sea clutter, using an AR(12), and an AR(12)-NL(5,b) models.

The results for the prediction error improvement for the different models are presented in table 5.1, for both horizontally and vertically polarised clutter. The prediction error improvement is given in dB, with reference to the original signal power $10 \log_{10} E[(\sigma - E[\sigma])^2]$. The models entertained include an AR(12) model (order 12 chosen on the basis of the autocorrelation function), as well as an AR(12) model augmented with non-linear terms, denoted by AR(12)-NL(K,b), of the form

$$e_{ARNL}(t) = y_t - (\beta_0 + \sum_{j=1}^{12} \beta_j y_{t-j}) - \sum_k^K \gamma_k e^{-by_{t-k}}$$

The improvement in the prediction error given in table 5.1 is given in terms of the power (in dB) of the prediction errors, $10 \log_{10} E[e_{AR}^2]$ and $10 \log_{10} E[e_{ARNL}^2]$ for the AR(12) and AR(12)-NL(K,b) models, respectively.

The results in table 5.1 show that the additional improvement resulting from the inclusion of the non-linearity, relative to the improvement obtained using the simple linear autoregressive model, is very modest (fraction of a dB) for either polarisation. It is clearly of interest for future research to investigate whether other, more general, non-linear models are capable of *substantially* and *consistently* improving on the prediction performance of simple linear regression models.

5.3.2 Presence of Heteroscedasticity

In the previous section, the form of regression in the temporal evolution of the modulating component of sea clutter in the logarithm domain was investigated. In particular, table 5.1 shows that compared to linear regression, the non-linear component stipulated by equation 5.3 provides only an insignificant additional improvement in the mean squared prediction error.

It is interesting to note that once the linearity has been accounted for in the form of an autoregression, any remaining non-linearities become part of the innovation process, e_{AR} . These can usually be studied by analysing the prediction errors for correlations, non-linearities, non-Gaussianity, *etc.* However, it may be that the prediction errors themselves are uncorrelated with each other, as well as with the observed process. As was pointed out in section 5.2, for conditionally heteroscedastic models, although the prediction errors, e , may be uncorrelated and independent of the previous observations, the squared prediction errors, e^2 (or $\ln e^2$), may themselves be correlated or depend in some way on the previous observations.

The simple birth-death migration argument for a Gamma distributed process, leads to the possibility of heteroscedasticity in the modulating component of sea clutter. In order to investigate the form of the heteroscedasticity, figure 5.4 illustrates the cross correlation of the log of the squared prediction error, $\ln e^2$, with itself, with the past observations, y , and with the normalised log squared observations, $\ln(y - E[y])^2$.

The prediction errors were obtained by using the AR(12) regression model, as well as by using the AR(12)-NL(5,b=1) model. In particular, figure 5.4(a) and 5.4(b) show the correlations in the AR(12) innovations for horizontally and vertically polarised clutter, respectively, while figure 5.4(c) and 5.4(d) show the correlations in the innovations obtained from an AR(12)-NL(5,b=1) model.

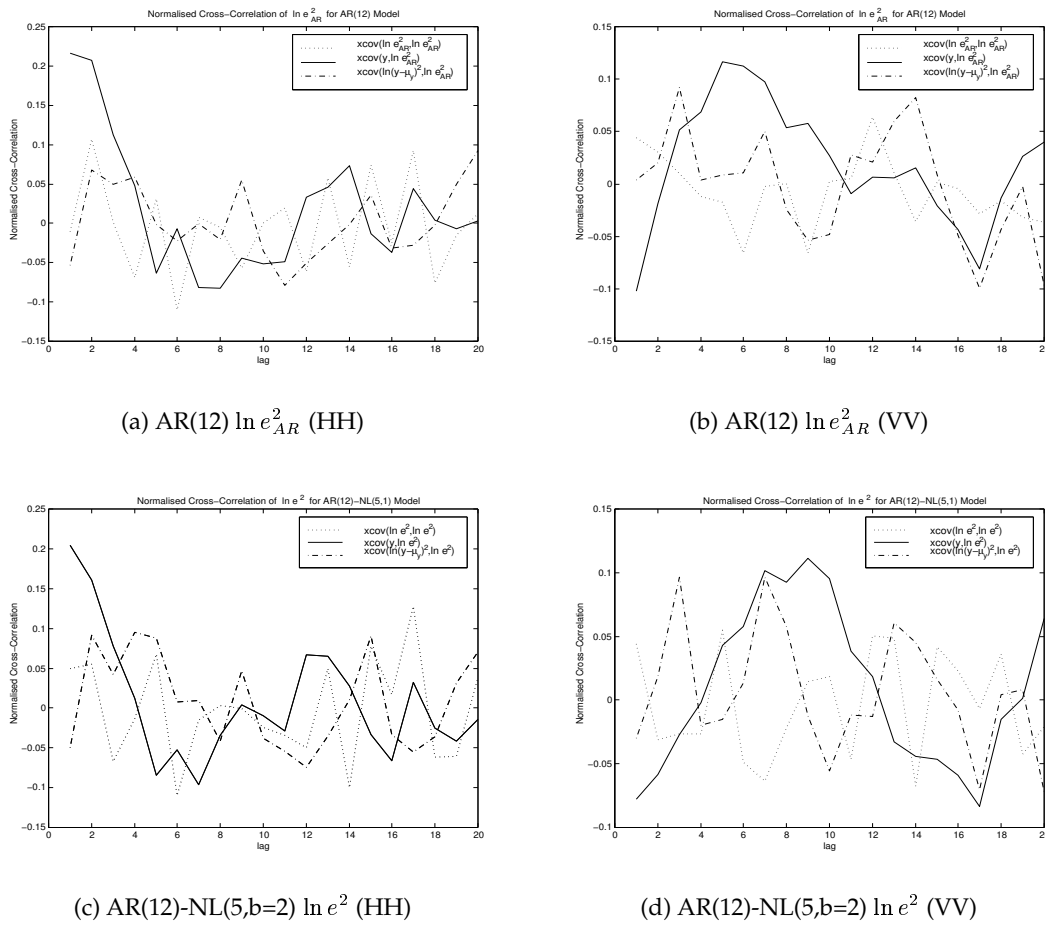


Figure 5.4: Cross-correlation of the log squared prediction errors, $\ln e^2$, for AR(12) and AR(12)-NL(5,b=1) models. The results are shown for horizontally (see (a) and (c)) and vertically (see (b) and (d)) polarised clutter.

Figure 5.4 shows that, particularly in the case of horizontally polarised clutter, there is a strong correlation between $\ln e_{AR}^2$ and the previous observations y . The correlation does not seem to be significantly affected by the inclusion of the non-linearity (see figure 5.4(a) and 5.4(c)). Similarly, for vertically polarised clutter, clear correlation is visible between $\ln e_{AR}^2$ and the observations y , while the form of the correlation appears to be affected by the non-linearity (see figures 5.4(b) and 5.4(d)).

The results in figure 5.4 show that there indeed may be a case for further investigation of heteroscedasticity in the modulating component of sea clutter. The present section has illustrated some of the issues associated with an informal diagnosis of heteroscedasticity in real data. Admittedly, only a single time series from a single range cell was used for illustrative purposes. However, modelling the time series associated with the modulating component of sea clutter may have limited practical utility, as this component decorrelates on the scale of a few seconds. It is of more practical interest to model the correlations of the modulating compo-

ment within a range profile. A more thorough study of heteroscedasticity will now be undertaken in the following section for a number of range profiles taken from a sea clutter database [Hirst and Baker, 1987].

5.4 EXTENSIONS TO RANGE PROFILES

The previous section has illustrated that, based on an informal inspection of the correlation structure of log squared prediction errors, there may be some evidence for the existence of heteroscedasticity in the time series corresponding to the modulating component of sea clutter in the logarithm domain. However, the utility of a time series model is rather limited in the context of target detection. Based on the discussion presented in chapter 3, a model for spatial correlations within a range profile is of more interest for CFAR target detection.

Intuitively, there is no apparent reason why the arguments which have led to the conditional heteroscedastic model for the modulating component of sea clutter, could not be extended to range profile modelling. In particular, whether the time series is obtained from a single range cell, or a range profile taken at one particular instant of time, the fluctuations observed in the data correspond to the fluctuations in the effective number of scattering centres on the sea surface. Whether evolving in range, or in time, the single point statistics of the process is still the same and corresponds to the statistics of the same underlying process (the swell), while the correlations in the modulating component arise from the evolution of the population of scattering centres within a resolution cell.

In principle, it should be possible to model the modulating component within a range profile in the same way as the time series obtained from within a single range cell. However, a question of process symmetry arises in the context of range profile modelling. Unlike the time series, which has a preferred direction of evolution (*i.e.* forward in time), the range profile (obtained as an instantaneous snapshot in time) may not. Furthermore, while the resolution of a single range cell remains constant, the resolution of the different range cells within a single range profile can vary widely, due to the range-dependent azimuth resolution of the radar. This can lead to spatially inhomogeneous statistics of the data. However, for the purposes of this chapter, such complications are not being considered and the study of their potential effects is left as an exercise for future research.

In the present section, the utility of modelling the evolution of the modulating component in range (in *either* direction) as a heteroscedastic process is investigated. The single sided (*i.e.* Markov) models are then extended in section 5.4.3 to symmetric Markov models, whereby the modulating component in a given range cell is modelled as being dependent on the modulat-

ing component in the neighbouring range cells, on *either* side of the current range cell, within a range profile. The proposed models are tested on a large set of range profiles obtained using a high resolution radar [Hirst and Baker, 1987]. The statistics of the subset consisting of 440 averaged range profiles used in the ensuing analysis are discussed at length in section 3.5.1.

5.4.1 Non-linearity in Range Profiles

The form of the regression for a time series corresponding to the modulating component of sea clutter in the logarithm domain was briefly investigated in section 5.3. In particular, it was found that a 3-4dB improvement in the prediction error is obtained by simply modelling the time series as a linear autoregressive process and a negligible additional improvement in the prediction performance is obtained by including the non-linearity stipulated by equation 5.3. However, the results presented in section 5.3 were only illustrative and were only obtained for a single realisation of the time series associated with the modulating component.

The analysis is now extended to the database of 440 range profiles. The range profiles are effectively treated as a time series. The improvements in the prediction errors obtained by using a linear autoregressive model are shown in figures 5.5(a) and 5.5(b) for the training and the test set, respectively. Similarly, the improvements in the prediction errors obtained using a linear autoregressive model with non-linear terms are shown in figures 5.5(c) and 5.5(d) for the training and the test set, respectively. In all cases, the model order is chosen such that the prediction error on the test set is minimised.

Figure 5.5 shows that about 2 to 12dB improvement in prediction error can be expected from a simple linear regression (*i.e.* from an autoregressive model), while an additional improvement in prediction error of up to 0.5dB can be expected by including the non-linearities of the form discussed in section 5.3. Although this additional improvement does appear to generalise to the test set in most cases, it is insignificant compared to the improvement obtained from the simple linear autoregression.

5.4.2 Heteroscedasticity in Range Profiles

In addition to the possible presence of non-linearities of the form described in equation 5.3, the preliminary analysis presented in section 5.3 indicated a possible presence of heteroscedasticity (*i.e.* non-constant prediction error variance).

The choice of a measure that best illustrates the presence of heteroscedasticity proves to be a rather daunting task. Improvement in the model likelihood may constitute one such measure, though relying on the likelihood alone is fraught with dangers. Alternatively, it is the observa-

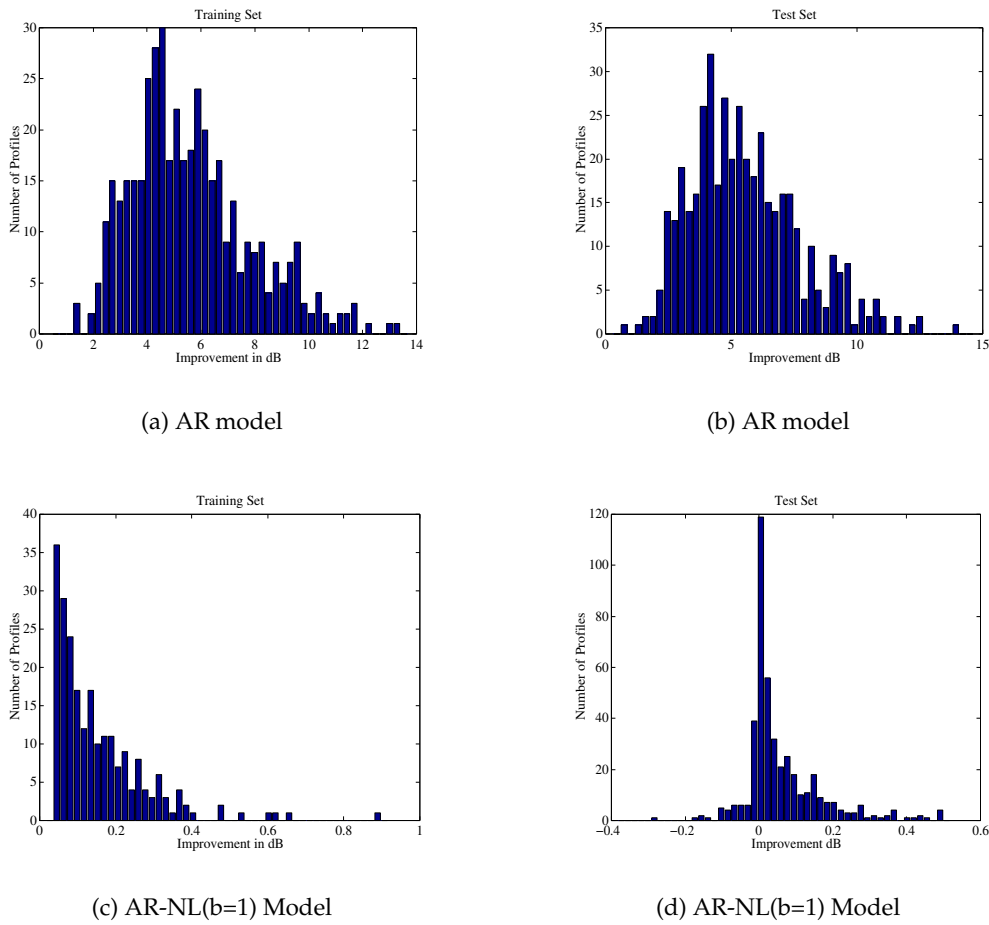


Figure 5.5: Histogram of the prediction error improvement in dB obtained using a linear AR model (see (a) and (b)) and an additional prediction error improvement in dB obtained by including non-linearities (see (c) and (d)), i.e. using an AR-NL($b=1$) model. Results are shown for training (see (a) and (c)) and test (see (b) and (d)) sets.

tion that heteroscedasticity is manifested through the correlations in the log squared prediction errors, that can form the basis for a rather informal investigation. In fact it is the latter approach, based on the correlations in the log squared prediction errors, that is adopted in the present section. An illustration of how the heteroscedasticity translates into target detection performance enhancement in compound clutter environment is presented in section 5.4.4.

A measure that quantifies the degree of correlation between the log squared prediction error and any other variables of interest is therefore required. As was discussed in section 5.2, the log squared prediction error, $\ln e^2$, is related to the log variance, $\ln h$, through

$$\begin{aligned}\ln e_t^2 &= \ln h_t + c + \eta_t \\ \eta_t &= \ln e_t^2 - c\end{aligned}$$

with $c = \psi^{(0)}(1/2) + \ln 2$, $\epsilon_t \sim N(0, 1)$, and hence, by definition, η_t being serially uncorrelated with zero mean. The form of $\ln h_t$ depends on the type of heteroscedasticity present. Therefore, to assess the degree of correlation between $\ln e_t^2$ and $\ln h_t$ (which effectively 'predicts' $\ln e^2$), the improvement in the 'prediction' of log squared prediction error, defined as

$$\text{Improvement (dB)} = 10 \log_{10} \frac{E[(\ln e^2 - \ln h - c)^2]}{E[(\ln e^2 - E[\ln e^2])^2]}$$

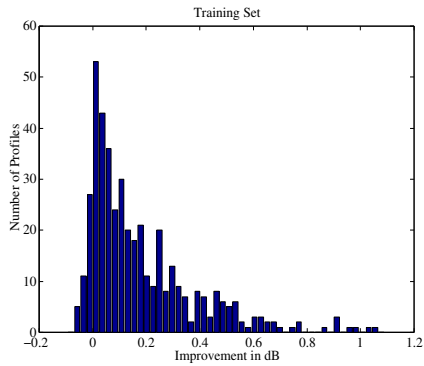
is calculated, where e and h are obtained by Maximum Likelihood (see section 5.2.1). This measure is designed to provide a quantitative measure of the degree of correlation between $\ln e^2$ and $\ln h$.

The prediction errors for the range profiles of the modulating component were obtained using a linear autoregressive model and are denoted by e_{AR} . The order of the linear autoregressive process was chosen such that the prediction error is minimised on the test set. For all the models considered here (*i.e.* AR-ECH, AR-ELCH, AR-EARCH), the order of the regression model for the variance was fixed at 2, as it was found that the models seldom generalise well in the likelihood sense [Noga and Fitzgerald, 1998a] for larger model orders. This result is also borne out in the analysis based on the correlations in the log squared prediction errors. The reader is referred to section 5.2.1 for the definitions of the regression models for the variance.

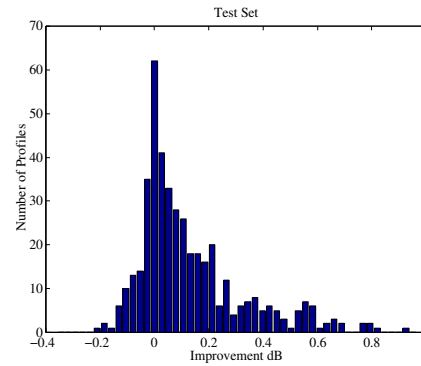
Figure 5.6¹ illustrates the results for the prediction errors obtained from the linear autoregressive model. The improvement in the 'prediction' of $\ln e_{AR}^2$ is shown for AR-ECH (see (a) and (b)), AR-ELCH (see (c) and (d)) and AR-EARCH (see (e) and (f)) models, for both the training and the test set data. The model parameters were estimated by Maximum Likelihood on the training set. The figure shows that the largest improvement is obtained using an AR-ECH model. This improvement also generalises well to the test set. The AR-ELCH and AR-EARCH models, although providing an improvement on the training set (not as large as the improvement obtained for AR-ECH model), do not generalise well to the test set. The results in figure 5.6 suggest that the log squared prediction error for a linear autoregressive model, $\ln e_{AR}^2$, may be correlated with the observations, y .

Finally, figure 5.7 illustrates the improvement in 'prediction' of the log squared prediction error as a function of the deviation of the range profile marginal statistics from log-Normality, in the log-likelihood sense for both the training and the test sets. The results in the figure suggest that the correlations in the log squared prediction error increase as the range profile data deviates from the log-Normal distribution in the log-likelihood sense. Whether this is due to some non-linearities being unaccounted for in the non-Gaussian (in the logarithm domain) data, or

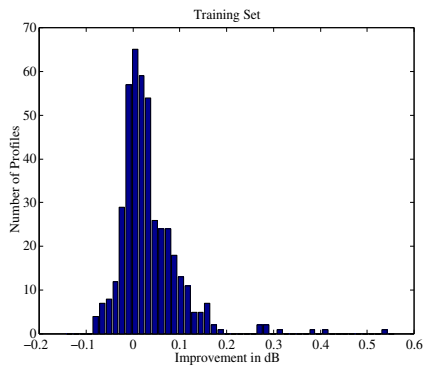
¹In some cases, a net degradation in the prediction of log squared prediction error can be observed in the figure, for both the training and the test sets. The reason for the apparent degradation is that the model parameters are estimated by Maximum Likelihood, which does not necessarily maximise the correlation between the log squared prediction error and the regressed variables.



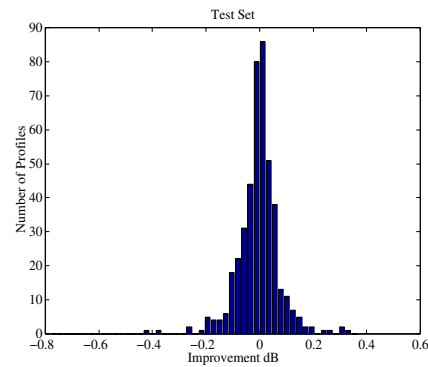
(a) AR-ECH



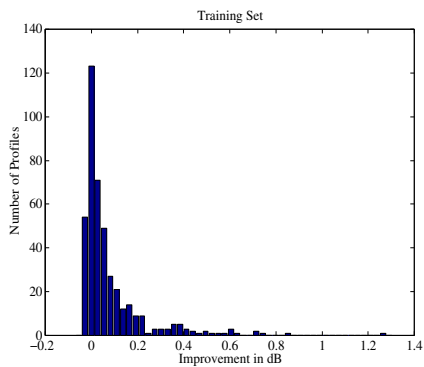
(b) AR-ECH



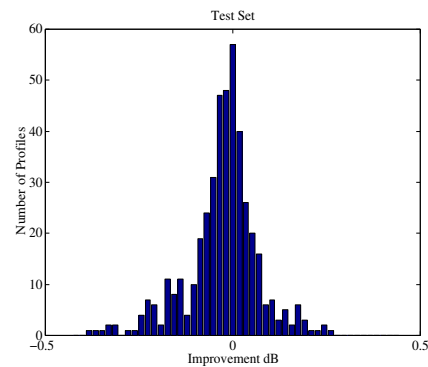
(c) AR-ELCH



(d) AR-ELCH



(e) AR-EARCH



(f) AR-EARCH

Figure 5.6: Histograms of the improvement in the 'prediction' of log squared prediction error for AR-ECH (see (a) and (b)), AR-ELCH (see (c) and (d)) and AR-EARCH (see (e) and (f)) models. Results shown for the training (see (a), (c) and (e)) and test (see (b), (d) and (f)) sets.

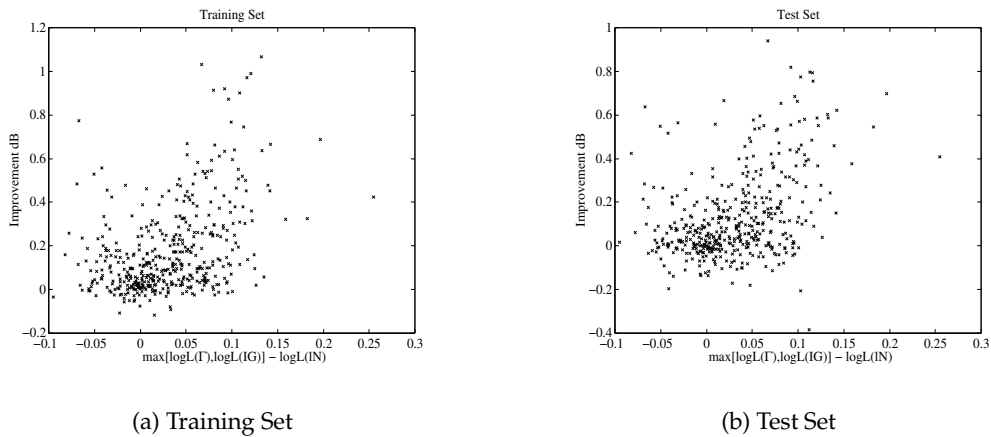


Figure 5.7: Improvement in the ‘prediction’ of log squared prediction error for AR-ECH model, as a function of the deviation of the range profile marginal statistics from the log-Normal distribution in the log-likelihood sense. Results shown for (a) training and (b) test sets.

simply due to the inherent heteroscedasticity in the process associated with the modulating component of sea clutter, requires further study.

5.4.3 Symmetric Models

In the previous section, a Markov model, inspired by the stochastic differential equation associated with the birth-death migration process, was used to model the evolution of the modulating component of sea clutter. The model was applied to range profile data which, for the purposes of the analysis, was treated as a time series. However, the range profile is not a time series, but rather a snapshot in time of the radar returns from the different ranges. As was discussed in chapter 3, from the constant false alarm rate (CFAR) detection point of view, it is of interest to obtain an estimate of the local clutter power (*i.e.* the modulating component) in the current range cell, based on the radar returns in the neighbouring range cells on *either* side of the range cell under test.

The estimate for the modulating component is normally obtained by some form of averaging of the samples in the neighbouring range cells. Cell Averaging (CA) CFAR detector achieves this by averaging the samples in either the amplitude, or the intensity domain. Order Statistics CFAR detectors obtain an estimate by ordering the samples according to the intensity and obtaining an estimate as some sort of a function (*e.g.* k^{th} largest sample) of the ordered samples. These techniques are inherently suboptimal for correlated clutter, as they do not account for the correlation structure of the data.

The LMAP detector [Bucciarelli et al., 1996] obtains an estimate of the local clutter power in the current range cell, by obtaining the weighted average of the radar returns in the surrounding range cell, in the logarithm domain (which is also the domain that linearises the compound clutter). The weights attached to the returns in the different range cells are evaluated, based on the correlation structure of the data. In fact, for a jointly Gaussian distributed process, the prediction equation for the modulating component in the n^{th} range cell can be written as

$$\begin{aligned} y_n &= \beta_0 + \sum_{j=-Q, j \neq 0}^Q \beta_j y_{n+j} + e_n \\ e_n &= h_n^{1/2} \epsilon_n \quad \epsilon_n \sim N(0, 1) \end{aligned} \quad (5.8)$$

The weighted linear average of the modulating component samples in the neighbouring range cells is the optimal (in the *maximum a posteriori* sense) estimator of the modulating component in the given range cell. The model in equation 5.8 is no longer Markov, but instead is symmetrically Markov (*i.e.* the current sample depends on Q samples on its either side). Such linear symmetric autoregressive models of order Q , shall be denoted herein as SAR(Q).

For non-Gaussian data, it is most likely that the simple weighted *linear* estimator is no longer optimal, and it may be the case that a non-linear estimator should be sought. A study of such estimators, and the associated improvement in the detection performance, is left as a potentially fruitful avenue for future research. Alternatively, as was the case in the linear Markov model for the modulating component of sea clutter, the non-Gaussian nature of the data may be manifested in non-constant prediction error variance. Such an observation could potentially have a profound effect on adaptive threshold estimation in CFAR detection.

Following from the Markov models of the previous section, the simple extension to the symmetric weighted linear estimator, is to model the prediction error variance as being state dependent, and of the form

$$\ln h_n = \alpha_0 + \sum_{i=-P, i \neq 0}^P \alpha_i y_{n+i} \quad (5.9)$$

The resulting symmetric autoregressive model with the state dependent prediction error variance is conditionally heteroscedastic, and is denoted as SAR(Q)-SCH(P). Whether or not such state dependence in the prediction variance is present in the modulating component of sea clutter is the subject of the present investigation. The issues associated with parameter estimation by Maximum Likelihood for symmetric AR-CH models are dealt with in appendix C.2. However, due to the complexity associated with the evaluation of the Jacobian term in the parameter likelihood for the symmetric Markov models, the model parameters are obtained herein by Pseudo Maximum Likelihood, whereby the parameter-dependent Jacobian term in the likelihood is neglected for the purpose of parameter estimation.

Figure 5.8 shows the improvement (in dB) for the prediction error obtained using the symmetric autoregressive (SAR) model, for which the model order is chosen to minimise the prediction error on the test set. Figures 5.5 and 5.8 show that the improvement obtained using the symmetric models ranges from about 5 to 20dB, compared to the 2 to 12dB for the single-sided Markov models. Clearly, the prediction power of the symmetric models exceeds the prediction power of the single Markov models by a significant factor, far exceeding any improvements obtained by including the non-linearities of the kind considered in section 5.4.1.

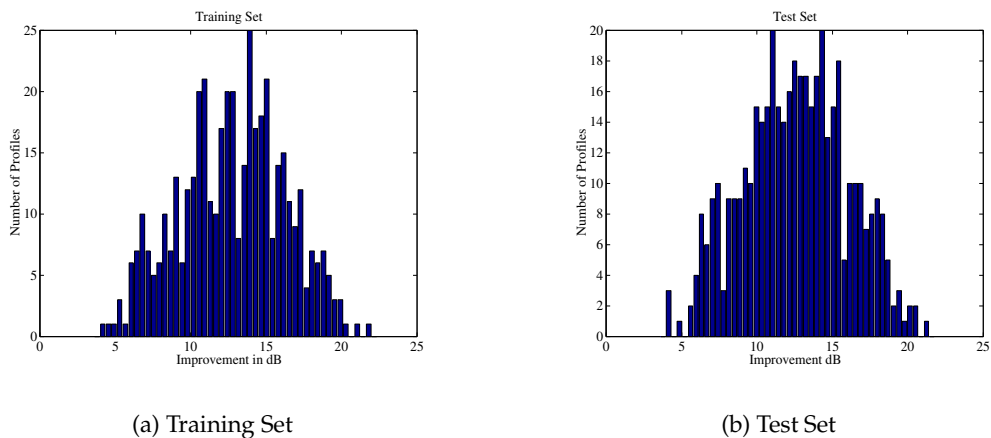


Figure 5.8: Histogram of the prediction error improvement in dB obtained using a symmetric autoregressive (SAR) model. Results shown for (a) training and (b) test sets.

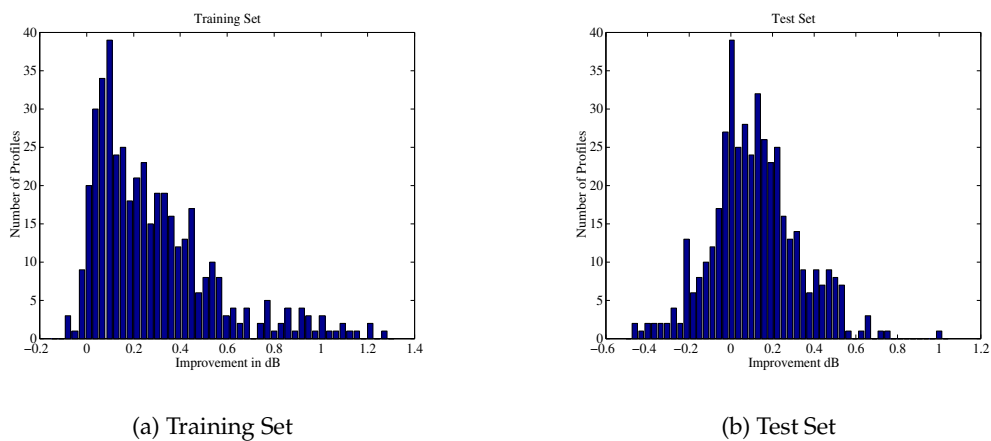


Figure 5.9: Histogram of the improvement in the 'prediction' of log squared prediction error for SAR-SCH model. Results shown for (a) training and (b) test sets.

Figure 5.9 shows the improvement in the ‘prediction’ of the log squared prediction errors obtained using the SAR model, when the variance is modelled as a SCH(2) process. The figure shows some correlation between $\ln e^2$ and y is still evident, although the correlations do not appear to generalise as well as in the case of single sided Markov models. Also, as was observed for the single sided Markov AR-CH models, the order of the heteroscedastic process is very small, with higher order heteroscedastic models not generalising well to the test set data.

5.4.4 Implications for CFAR Detection

The analysis presented thus far has attempted to informally diagnose heteroscedasticity in the data, by employing a very simple measure of the correlation between the log squared prediction errors and the variables on which the prediction variance is being regressed. The foregoing analysis, based on a large database of averaged range clutter profiles, has shown that there is some evidence for heteroscedasticity in the modulating component of sea clutter. However, the above analysis does not give any tangible feel for the practical significance of the finding. To this end, the present section will illustrate the effects of heteroscedasticity on target detection in the presence of sea clutter.

Figure 5.10 shows the relative performance (in terms of the detection loss) of the CA CFAR and LMAP CFAR detectors in an ideal (*i.e.* speckle free) environment, for the case when $P_d = 0.5$ and $P_{fa} = 10^{-4}$. In particular, the detection loss of a CA CFAR detector relative to a fixed threshold detector is shown in figure 5.10(a). Figure 5.10(b) illustrates the increase in the spatial variance of the probability of false alarm for the CA CFAR detector, relative to the fixed threshold detector, operating in ideal clutter conditions. Similarly, the additional detection loss incurred by an LMAP CFAR detector, relative to the CA CFAR detector is shown in 5.10(c). Figure 5.10(d) illustrates the increase in the spatial variance of the probability of false alarm for the LMAP CFAR detector, relative to the CA CFAR detector. In short, the results in figure 5.10 show that the LMAP CFAR detection scheme can potentially provide up to a 12dB improvement in detectability, in addition to a 3dB improvement in the spatial variation of the probability of false alarm, relative to the CA CFAR detection scheme.

In the case of conditionally heteroscedastic models based on the prediction equation 5.8, not only is the prediction mean, $\mu_y = E[y_n | \mathbf{y}_{-n}]$, available, but also the prediction variance $\sigma_y^2 = \text{var}[y_n | \mathbf{y}_{-n}]$, as given by equation 5.9. Therefore, to increase detectability, the CFAR detection threshold can be set as

$$T_{CHLMAP} = G \exp \frac{1}{2} (\mu_y + K \sigma_y^2)$$

The resulting conditional heteroscedastic LMAP CFAR detector is from now on denoted as CHLMAP CFAR. The analysis of the performance of such detectors is undertaken in more

detail in the following chapter. For the purpose of the present exposition, the idealised case of speckle free environment in the reference window, is once again assumed.

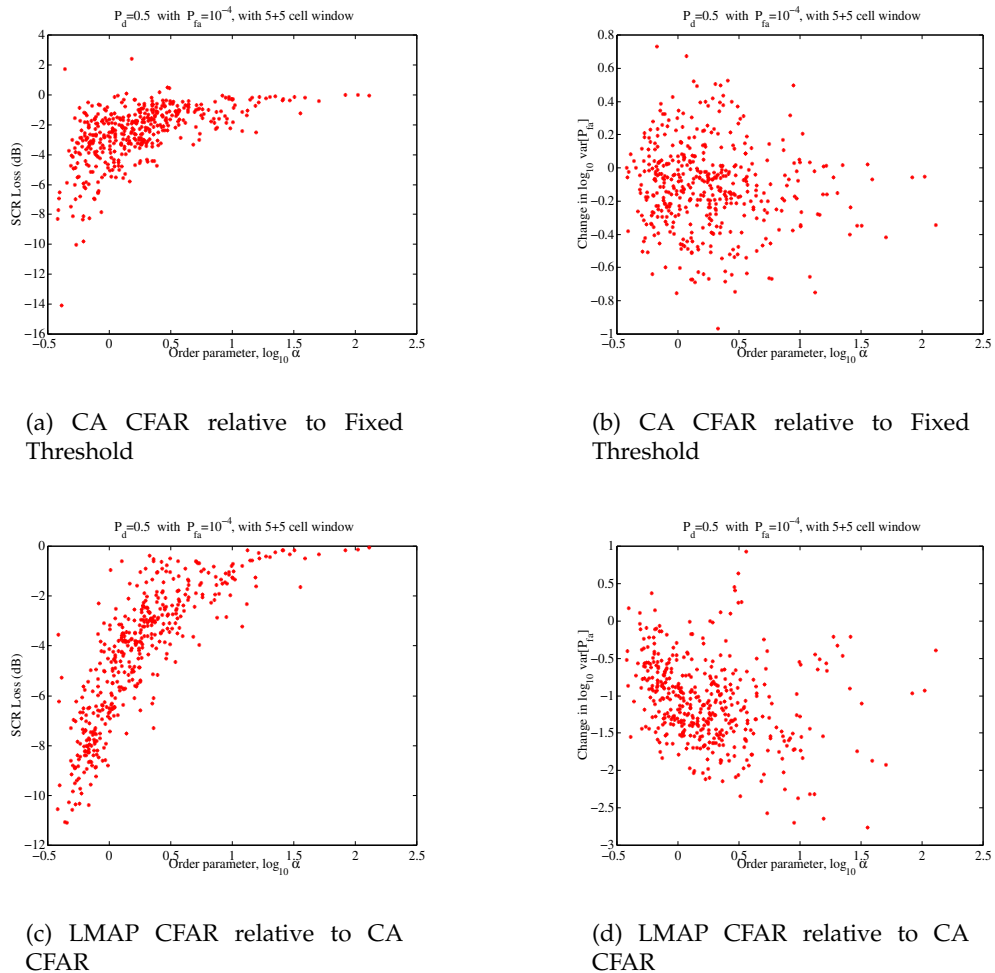


Figure 5.10: (a) Detection loss and (b) the change in the logarithm of the variance of P_{fa} , for a CA CFAR detector, relative to a Fixed Threshold detector. Also shown, (c) an additional detection loss and (d) the change in the logarithm of the variance of P_{fa} , for an LMAP CFAR detector, relative to a CA CFAR detector. $P_d = 0.5$ and $P_{fa} = 10^{-4}$.

The prediction variance gain, K , is presently set to maximise detectability, while the threshold gain, G , is adjusted for the correct probability of false alarm. Figure 5.11 illustrates the additional increase in the detection loss and the increase in the spatial probability of false alarm incurred by including the prediction error variance in the setting of the CFAR detection threshold, relative to the LMAP CFAR detector.

Figure 5.11 shows that an additional improvement of up to 1dB in detectability can be obtained by including the prediction error variance in setting the threshold. The increase is particularly prominent for spiky clutter (*i.e.* small α). Furthermore, the increased detectability (see figure

5.11(a)) appears to be accompanied by the corresponding decrease in the spatial variation of probability of false alarm (see figure 5.11(b)). Whether such improvements can be sustained in non-ideal conditions, *i.e.* when the local clutter power is estimated from neighbouring samples contaminated by the speckle noise, is the subject of the investigation presented in the following chapter.

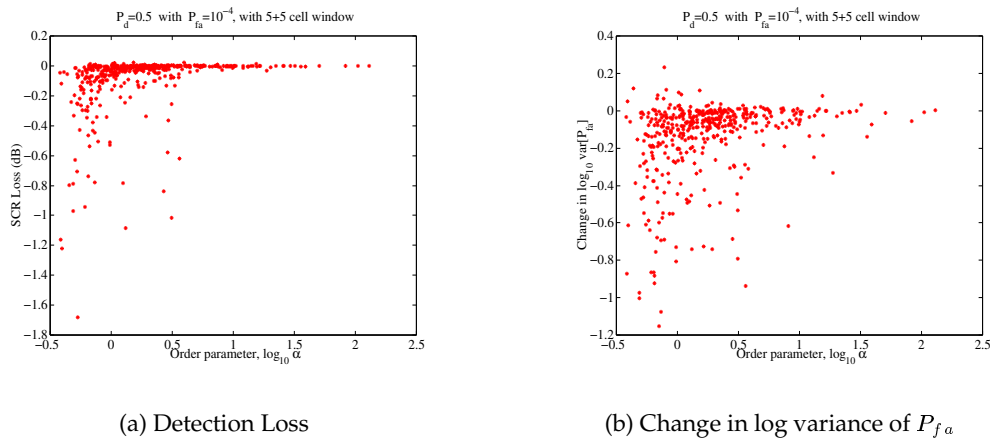


Figure 5.11: (a) Detection loss and (b) the change in the logarithm of the variance of P_{fa} , for a CHLMAP CFAR detector, relative to an LMAP CFAR detector. $P_d = 0.5$ and $P_{fa} = 10^{-4}$.

5.5 DISCUSSION, CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

This chapter was concerned with the modelling of the modulating component of high resolution sea clutter. In section 5.2, as an extension of birth-death migration arguments for a population of scattering centres on the sea surface, a flexible family of conditional heteroscedastic models was proposed for the modulating component of sea clutter. The proposed models are a generalisation of linear Gaussian processes, and have a number of attractive properties, which were reviewed in section 5.2.

As was discussed in section 5.2, the conditionally heteroscedastic models are characterised by the correlation structure of (the logarithm of) the squared prediction errors. This observation forms the basis for an illustrative study undertaken in section 5.3 for a time series corresponding to the modulating component extracted from a single cell of a range profile. The purpose of the study was to illustrate a number of issues associated with modelling of the modulating component of sea clutter. There may be some evidence for heteroscedasticity in the modulating component of sea clutter, as manifested through the linear and non-linear model prediction errors. In particular, the log squared prediction errors and the past observations appear to be correlated in the absence of any correlations between the past observations and the model prediction errors.

The results in section 5.3 were only obtained for a single time series from a single range cell. Furthermore, the possible presence of heteroscedasticity was diagnosed simply by visual inspection of the correlation function of the log squared prediction errors and the past observations. Such an informal inspection does not provide any insight into the potential significance of heteroscedasticity to CFAR target detection in the presence of sea clutter. For this reason, results presented in section 5.4 were extended to a large database of high resolution sea clutter range profiles. Also, instead of modelling the time series of the modulating component from a single range cell, the results presented in section 5.4 were obtained for the heteroscedastic models applied to the range profiles (*i.e.* spatially distributed snapshots in time).

In order to quantify the heteroscedasticity present in the modulating component of sea clutter, a measure of the amount of correlation present in the log squared prediction errors was defined in section 5.4. Evidence for heteroscedasticity, as manifested by the correlations between the log squared prediction errors and the process observations, was found for the modulating component in the logarithm domain. However, the order of heteroscedasticity (*i.e.* the number of lagged terms in the variance) was found to be very low, with the heteroscedasticity not generalising well to the test set for larger model orders. Furthermore, it was found that the correlations in the log squared prediction errors are strongest when the modulating component deviates from the log-Normal distribution in the log-likelihood sense.

The potential implications of heteroscedasticity in the modulating component to CFAR target detection performance in compound clutter environment were studied in section 5.4.4. It was found that a small improvement in detectability (as measured by the decrease in the SCR required for the given probability of detection) was obtained when the conditional prediction error variance was included in the CFAR detection threshold. Also, the inclusion of the conditional variance has resulted in a decrease in spatial variation in the probability of false alarm, as measured by its variance. However, even in the ideal, speckle free conditions considered in this chapter, any additional improvements obtained by including the heteroscedasticity were insignificant compared to the improvements obtained by simply accounting for the correlation structure of the modulating component.

It was shown in chapter 3 that, contrary to the widely held beliefs, rather than being Gamma distributed, the modulating component of sea clutter may, in fact, be better modelled as being log-Normal distributed. This finding has profound consequences on the results introduced in the present chapter, as the conditional heteroscedastic models have their roots in the birth-death migration processes, with Gamma distributed marginals. In particular, as was shown in the present chapter, the evidence for heteroscedasticity in the modulating component of sea clutter is not very compelling, and any improvements in detectability in speckle-free environment appear to be insignificant. The issue of whether the heteroscedasticity provides any improvements in target detection performance in noisy clutter (*i.e.* when the local clutter power

is predicted from the returns including the speckle component) is addressed in the following chapter.

Another question still remains to be answered. In particular, does the (very limited) heteroscedasticity observed arise due to its inherent presence in the clutter, or whether it is a manifestation of some other, unaccounted for property of sea clutter, such as a non-linearity, or possibly non-Gaussian distributed prediction errors. The utility of studying such non-linear and non-Gaussian models and any potential benefits that can be derived from them, in the context of target detection in compound clutter environment, is left as a potentially fruitful avenue for future research. Such complicated models are purposefully avoided here and, instead, a study of the potential benefits that can be derived from the *simple* extensions to the linear models remains the focus of the research presented in the remainder of this thesis.

CFAR Target Detection in Spatially Correlated Sea Clutter

6.1 INTRODUCTION

Sea clutter is a compound process consisting of Rayleigh distributed speckle modulated in power by the underlying modulating process. Unlike the speckle component, which is largely decorrelated spatially and in range, the modulating component has a definite spatial correlation structure.

As was discussed in chapter 3, the spatial correlations of the clutter have a profound influence on the target detection performance in compound clutter environment. While in spatially uncorrelated clutter, fixed threshold detection is the optimal detection strategy, in spatially correlated clutter, significant improvements in detectability can be obtained by adaptive threshold techniques. The largest improvements can be achieved in spiky and highly correlated clutter, where the ideal CFAR detector (for which the local clutter power level is assumed to be known exactly) represents the maximum attainable improvement in detectability relative to the fixed threshold detection in compound clutter environment.

[Bucciarelli et al., 1996], [Watts, 1996] and [Armstrong and Griffiths, 1991a] study the effects of the correlations in the modulating component on CFAR detection performance. As was borne out by an idealised analysis presented in chapter 3, a simple cell averaging CFAR detector is inherently incapable of approaching the ideal CFAR detection performance. However, ideal CFAR performance, at least in an idealised speckle free environment, can be approached by adaptively setting the threshold based on a *weighted* average of the raw clutter samples in the neighbouring range cells in the logarithm domain, where the weighting coefficients are obtained from the correlation structure of the modulating component.

The remainder of this chapter is structured as follows. A linearised, conditionally Gaussian state space model for high resolution sea clutter is developed in section 6.2. In section 6.3, it is shown that a detection scheme proposed by [Bucciarelli et al., 1996] suffers from a major shortcoming, in that the local clutter power estimate is not obtained from the true predictive conditional. Instead, based on the state space representation, an 'optimal' (in the minimum mean squared linear sense) estimator for the local clutter power level is derived in section 6.3. Furthermore, an estimate of the uncertainty associated with the local clutter power prediction

is also obtained. The estimate of the uncertainty is utilised in an adaptive threshold detection scheme discussed in section 6.4. An analytical approximation to the probability of false alarm for the proposed scheme is also presented in this section. The detection performance of the proposed scheme is compared in section 6.5 to other detection techniques, including the fixed threshold, ideal CFAR, CA CFAR and LMAP CFAR detectors. A brief summary of the work presented in this chapter is presented in section 6.6, along with suggestions for future research.

6.2 COMPOUND SEA CLUTTER MODEL

Throughout this dissertation, one topic is more recurrent than any other. From the study of incoherent speckle in chapter 4 and the modulating component in chapter 5, to the present chapter dealing with target detection in the presence of sea clutter, it is the compound nature of the clutter, which lies at the heart of this work.

In chapter 2, it was argued that the (in)coherent sea clutter return in range cell n at time t (denoted by x_n^t) can be written as

$$x_n^t = v_n^t s_n^t \quad (6.1)$$

where v_n^t is the level of the underlying modulating process (sometimes referred to as the *texture* component) in range cell n at time t , and s_n^t refers to the corresponding speckle. The above equation holds true for the case of incoherent, as well as of coherent, radar returns. In the former case s_n^t is Rayleigh distributed, while in the latter case, s_n^t is a complex Gaussian variate.

The disparate correlation properties of the speckle and the modulating components allow the overall detection problem to be effectively split in two, and dealt with in a piecemeal fashion. In particular, the long term temporal correlations of the modulating component renders it effectively constant over the typical detector integration periods. Furthermore, unlike the speckle component, the modulating component is correlated in range.

Consider T consecutive clutter returns in range cell n , denoted by $\mathbf{x}_n = [x_n^t, \dots, x_n^{t+T-1}]'$ (note that the superscript t has been dropped for clarity, and will be implicitly assumed, unless stated otherwise). Provided T corresponds to a time period much shorter than the decorrelation time of the modulating component of sea clutter, the local power of the returns, $\sigma_n = 2v_n^2$, can be effectively considered to be constant. In the case of incoherent clutter, \mathbf{x}_n corresponds to amplitudes of the returns, while in the case of coherent clutter, \mathbf{x}_n corresponds to the complex in-phase and quadrature components of clutter returns. The two cases shall now be treated in turn.

Coherent Clutter Returns

The joint distribution for the T complex clutter samples in range cell n , as observed using a coherent radar, is given by

$$p(\mathbf{x}_n|\sigma_n, \Sigma) = \frac{1}{(\pi\sigma_n)^T|\Sigma|} \exp\left(-\frac{1}{\sigma_n}\mathbf{x}'_n\Sigma^{-1}\mathbf{x}_n\right) \quad (6.2)$$

where Σ is the temporal covariance matrix associated with the speckle s_n , and where \mathbf{x}_n is a vector of T complex samples $x_n^t = x_{In}^t + ix_{Qn}^t$.

Incoherent Clutter Returns

As was discussed in chapters 2 and 4, there is no simple analytical expression for the joint distribution for correlated Rayleigh distributed clutter resulting from the complex Gaussian process. Instead, a simplifying assumption of independent clutter samples can be made, resulting in the joint density for the Rayleigh distributed amplitudes given by

$$p(\mathbf{x}_n|\sigma_n) = \left(\prod_{\tau=0}^{T-1} \frac{2x_n^{t+\tau}}{\sigma_n}\right) \exp\left(-\frac{1}{\sigma_n}\mathbf{x}'_n\mathbf{x}_n\right) \quad (6.3)$$

This assumption is not unreasonable in practice, as frequency agility can be used to ensure temporal decorrelation of the speckle.

6.2.1 The Speckle Component

The multiplicative noise process of equation 6.1 can be generalised to T samples as

$$\mathbf{x}_n = v_n \mathbf{s}_n \quad (6.4)$$

where the speckle component, \mathbf{s}_n , is correlated, with the covariance matrix, Σ , as given in equation 6.2.

Equation 6.4 for the compound clutter process is in a multiplicative form for v_n (and hence for $\sigma_n = 2v_n^2$). However, it is possible to obtain an approximate additive Gaussian equation for $\ln \sigma_n$ in terms of the observed clutter \mathbf{x}_n . There are a couple of alternative approaches which can be followed to this end. The first one, discussed presently, is based on approximating the posterior density of $\ln \mathbf{x}_n^2$ by a Gaussian distribution, while an alternative approach, discussed subsequently, is based on a Gaussian approximation to the posterior distribution for $\ln \sigma_n$.

In order to simplify the problem for the case of correlated coherent clutter, frequency agility can be employed in order to remove the temporal correlations. Alternatively, the processes \mathbf{x}_n and \mathbf{s}_n can be pre-whitened by pre-multiplying by the square root of the covariance matrix Σ^{-1} , resulting in uncorrelated processes $\tilde{\mathbf{x}}_n$ and $\tilde{\mathbf{s}}_n$, respectively, such that

$$\begin{aligned}\tilde{\mathbf{x}}_n &= v_n \tilde{\mathbf{s}}_n \\ \tilde{\mathbf{x}}_n &= \Lambda \mathbf{x}_n \quad \tilde{\mathbf{s}}_n = \Lambda \mathbf{s}_n \\ \Sigma^{-1} &= \Lambda' \Lambda\end{aligned}\tag{6.5}$$

For the purpose of clarity of exposition, it is from now on assumed in this section that the T consecutive clutter samples are decorrelated either by the use of frequency agility, or by pre-whitening, as outlined above.

Once decorrelated, the joint distribution for $\tilde{\mathbf{x}}_n$ factorises into a product of marginal distributions for individual samples \tilde{x}_n^t . Therefore, using the arguments presented in appendix E, the components of equation 6.5 can individually and independently be linearised to yield an equivalent linear Gaussian form, *i.e.*

$$\begin{aligned}\ln(\tilde{x}_n^t)^2 - \psi^{(0)}(\alpha) &= \ln \sigma_n + \left(\psi^{(1)}(\alpha)\right)^{1/2} \nu_n^t \quad \nu_n^t \sim N(0, 1) \\ &\vdots \\ \ln(\tilde{x}_n^{t+T-1})^2 - \psi^{(0)}(\alpha) &= \ln \sigma_n + \left(\psi^{(1)}(\alpha)\right)^{1/2} \nu_n^{t+T-1} \quad \nu_n^{t+T-1} \sim N(0, 1)\end{aligned}\tag{6.6}$$

where $\psi^{(0)}(\cdot)$ is the digamma function, and $\psi^{(1)}(\cdot)$ is the trigamma function. For Rayleigh distributed speckle $\alpha = 1$, while for the in-phase and quadrature components of complex Gaussian distributed speckle, $\alpha = \frac{1}{2}$ (see appendix E).

An Alternative Approach

In the previous section, it was the distribution of the pre-whitened clutter, $\tilde{\mathbf{x}}_n$, which was approximated by a Gaussian distribution in the logarithm domain. Presently, a different approach is proposed, where the distribution of the modulating component, σ_n , is being approximated by a Gaussian distribution in the domain of the logarithm. This approach is particularly useful in the case of correlated speckle, in that it does not require the clutter to be explicitly pre-whitened.

In particular, following the arguments in appendix E, and noting that the likelihood for $\sigma_n = 2v_n^2$ is given by equations 6.2 and 6.3 for coherent and incoherent clutter, respectively, equation 6.4 can be cast into a linear Gaussian form as

$$\ln(\mathbf{x}_n' \Sigma^{-1} \mathbf{x}_n) - \psi^{(0)}(T) = \ln \sigma_n + \left(\psi^{(1)}(T)\right)^{1/2} \nu \quad \nu \sim N(0, 1)\tag{6.7}$$

where $\Sigma = \mathbf{I}_T$ for the incoherent clutter.

It is interesting to compare equations 6.6 and 6.7. Both equations are linear and Gaussian in $\ln \sigma_n$, and are identical for the univariate case (*i.e.* when $T = 1$ both equations simplify to simple linearisation of equation 6.1). However, in the case when $T > 1$, equation 6.6 constitutes a linear set of equations for $\ln \sigma_n$, and hence, an estimate of σ_n is equivalent to a geometric mean of the observations in the intensity domain. Equation 6.7, on the other hand, results in the estimate of σ_n being an arithmetic average of the observations in the intensity domain.

6.2.2 The Modulating Component

The previous section has dealt with the speckle component and how the multiplicative clutter can be approximately transformed into a linear Gaussian form, by approximating the likelihood for $\ln \sigma_n$ by a Normal distribution. The resulting linear Gaussian regression equation for $\ln \sigma_n$ can form the basis for the observation equation of a *quasi optimal* (*i.e.* optimal in the minimum mean square *linear* sense [Harvey and Shephard, 1996]) state space estimator for the modulating component, to be discussed in section 6.3.

In the discussion presented in the previous section, it was assumed that the integration time corresponding to T is much shorter than the decorrelation time of the underlying modulating component, and hence, σ_n was considered to be effectively constant over the integration period. This is a very convenient result arising from the very disparate correlation properties of the two components of sea clutter.

Furthermore, while the speckle component can be considered to be fully decorrelated in range (rendering the observation noise process spatially independent), the modulating component has a definite spatial correlation structure, dependent on environmental conditions and the viewing angle. In chapter 5, it was argued that the modulating component is well modelled in the logarithm domain, and that the conditional heteroscedastic regression models provide a good model for this process. In particular, for range cell n in a given range profile, the modulating process in the logarithm domain, $y_n = \ln \sigma_n$, can be modelled as a symmetric regression on its neighbours with conditional heteroscedastic errors, *i.e.*

$$\begin{aligned}
 y_n &= \beta_0 + \sum_{j=-q, j \neq 0}^{j=q} \beta_j y_{n+j} + h_n^{1/2} \epsilon_n & \epsilon_n &\sim N(0, 1) \\
 \ln h_n &= \alpha_0 + \sum_{i=-p, i \neq 0}^{i=p} \alpha_i y_{n+i}
 \end{aligned} \tag{6.8}$$

Equations 6.7 and 6.8 constitute a linearised, conditionally Gaussian state space model for high resolution sea clutter. The issues associated with estimation of local clutter power level from a reference window of raw clutter samples are dealt with in the following section.

6.3 STATE SPACE ESTIMATION OF LOCAL CLUTTER POWER

In the previous section, the multiplicative form of the compound model for sea clutter has been approximately transformed into a linear, conditionally Gaussian state space model, manifested by equations 6.7 and 6.8. The aforementioned state space model can be written down more succinctly as

$$\mathbf{w}_{-n} = \mathbf{y}_{-n} + \sigma_{w|y} \nu_{-n} \quad \nu_{-n} \sim N(0, \mathbf{I}_{-n}) \quad (6.9)$$

$$y_n = \mu_n + h_n^{1/2} \epsilon_n \quad \epsilon_n \sim N(0, 1) \quad (6.10)$$

where $\sigma_{w|y}^2 = \psi^{(1)}(T)$ is the observation noise variance, resulting from observing the modulating component in speckle. The *observations*, $\mathbf{w}_n = \ln(\mathbf{x}_n' \Sigma^{-1} \mathbf{x}_n) - \psi^{(0)}(T)$, correspond to the (integrated) radar returns in the logarithm domain, while the *states* (i.e. directly unobserved variables), y_n , correspond to the logarithm of the local clutter power, $\ln \sigma_n$. Furthermore, the mean and the variance of the state transition equation for the state space model are, respectively, given by

$$\mu_n = \beta_0 + \sum_{j \neq 0} \beta_j y_{n+j} \quad (6.11)$$

$$\ln h_n = \alpha_0 + \sum_{i \neq 0} \alpha_i y_{n+i} \quad (6.12)$$

where β and α are the regression parameters for the state prediction mean and variance, respectively.

As was discussed in chapter 3 in the context of adaptive threshold detection, an estimate of the local clutter power level for the range cell under test, σ_n , based on the raw clutter samples in the reference window, \mathbf{x}_{-n} , is required. Alternatively, in the logarithm domain, an estimate of the log of the local clutter power, y_n , conditional on the logarithm of the observations in the reference window, \mathbf{w}_{-n} , is sought. An 'optimal' local clutter power estimate can be obtained from the predictive distribution, $p(y_n | \mathbf{w}_{-n})$, which can, in principle, be obtained from the state space model outlined above, by noting that

$$p(y_n | \mathbf{w}_{-n}) \propto \int p(\mathbf{w}_{-n} | \mathbf{y}_{-n}) p(y_n, \mathbf{y}_{-n}) d\mathbf{y}_{-n} \quad (6.13)$$

where $p(\mathbf{w}_{-n}|\mathbf{y}_{-n})$ corresponds to the observation equation 6.9, while $p(y_n, \mathbf{y}_{-n})$ corresponds to the state transition equation 6.10, which defines the joint prior distribution for the states.

6.3.1 LMAP Estimator for Local Clutter Power

[Bucciarelli et al., 1996] propose a scheme, which they denote as LMAP, whereby the local clutter power level in the logarithm domain is obtained as a MAP estimate from a joint conditional posterior for the state vector, $\mathbf{y} = \{y_n, \mathbf{y}_{-n}\}$. The joint conditional posterior is obtained based on the assumption of the states being jointly Gaussian distributed, with the covariance matrix for \mathbf{y} denoted as Σ_y .

From equation 6.9, the joint likelihood for the logarithm of the states, \mathbf{y} , given the observations, \mathbf{w}_{-n} , in the reference window, is readily shown to be of the form

$$p(\mathbf{w}_{-n}|\mathbf{y}) = (2\pi\sigma_{w|y}^2)^{-L} \exp\left(-\frac{1}{2\sigma_{w|y}^2}(\mathbf{w} - \mathbf{y})'\mathbf{I}_{-(L+1)}(\mathbf{w} - \mathbf{y})\right)$$

where $\mathbf{I}_{-(L+1)}$ is an identity matrix with its $(L + 1)^{th}$ diagonal element nulled. The prior distribution for the states is assumed to be jointly Gaussian, with the covariance matrix estimated from the clutter map (*i.e.* the preceding range profile). This prior distribution basically contains the information about the correlation structure of the texture, which can, in principle, be used to obtain a better estimate of the local clutter power in the range cell under test, from the returns in the reference window. In particular, the posterior distribution for the clutter texture (*i.e.* the modulating component), conditional on the returns in the reference window, is given by

$$\begin{aligned} p(\mathbf{y}|\mathbf{w}_{-n}) &\propto p(\mathbf{w}_{-n}|\mathbf{y})p(\mathbf{y}) \\ &= \frac{1}{(2\pi\sigma_{w|y}^2)^L} \exp\left(-\frac{1}{2\sigma_{w|y}^2}(\mathbf{w} - \mathbf{y})'\mathbf{I}_{-(L+1)}(\mathbf{w} - \mathbf{y})\right) \\ &\quad \times \frac{1}{(2\pi)^{-(2L+1)/2}|\Sigma_y|} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{m}_y)'\Sigma_y^{-1}(\mathbf{y} - \mathbf{m}_y)\right) \end{aligned}$$

where \mathbf{m}_y and Σ_y are the *prior* estimates of the modulating component mean and covariance, respectively, as obtained from the clutter map. Maximising $p(\mathbf{y}|\mathbf{w}_{-n})$ jointly with respect to \mathbf{y} results in an LMAP estimate (*i.e.* MAP estimate in the logarithm domain) of \mathbf{y} given by

$$\mathbf{y}^{MAP} = (\mathbf{I}_{-(L+1)} + \sigma_{w|y}^2 \Sigma_y^{-1})^{-1} \left[\mathbf{w} \mathbf{I}_{-(L+1)} + \sigma_{w|y}^2 \mathbf{m}_y \Sigma_y^{-1} \right] \quad (6.14)$$

The $(L + 1)^{th}$ element of \mathbf{y}^{MAP} , $\hat{y}_n = \mathbf{y}^{MAP}(L + 1)$, can then be used as an estimate of the local clutter power in the n^{th} range cell, based on which the CFAR threshold can be set as $T = G \exp \hat{y}_n$.

The LMAP estimator, as given by equation 6.14, suffers from a major shortcoming, in that when the observation noise variance, $\sigma_w|y$, tends to zero (*i.e.* in speckle free case), the LMAP estimate of the local clutter power for the n^{th} range cell, y_n , is ill-defined, since the matrix $\mathbf{I}_{-(L+1)}$ is not invertible. This is the consequence of obtaining an estimate from the joint conditional for \mathbf{y} , rather than integrating out the remaining states in the reference window, as is done in equation 6.13. To overcome this problem, a full state space treatment is presented in the following section, whereby the marginal posterior for the local clutter power in the range cell under test is obtained, conditional on the radar returns in the reference window.

6.3.2 Predictive LMAP Estimator for Local Clutter Power

As was pointed out in the previous section, the LMAP estimation technique advocated by [Bucciarelli et al., 1996] suffers from a shortcoming, in that as the observation noise is reduced to zero, the local power estimate becomes ill-defined. Ideally, to overcome this problem, an estimate maximising the predictive distribution, $p(y_n|\mathbf{w}_{-n})$, should be sought. However, due to the presence of conditional heteroscedasticity, obtaining the marginal predictive distribution for the state proves to be no easy task. The reason for this is that, although conditionally Gaussian, the joint prior distribution for the states is not easily obtained.

Instead, it is the state prediction mean and the state prediction error variance, conditional on the noisy observations, *i.e.*

$$\mu_{n|w} = E[\mu_n|\mathbf{w}_{-n}] \quad (6.15)$$

$$\ln h_{n|w} = E[\ln h_n|\mathbf{w}_{-n}] \quad (6.16)$$

that are derived presently. To this end, a couple of simplifying assumptions are made. In particular:

- Assume that \mathbf{y} is *a priori* jointly Gaussian with mean \mathbf{m}_y and the covariance matrix Σ_y , estimates of which can be obtained from the clutter map (*i.e.* time history of clutter returns). This assumption greatly simplifies the analysis, while retaining the most important ingredient in the analysis, the covariance structure of the modulating component.
- Assume that \mathbf{y} and $\ln \mathbf{e}^2$ (where $e_k = y_k - \mu_k$) are *a priori* jointly Gaussian, with the mean and the covariance matrix that can be obtained from the clutter map. This assumption also greatly simplifies the analysis, while retaining the important information about the correlations between the prediction errors and the states, which lie at the heart of the conditional heteroscedastic formulation.

With these simplifying, though very reasonable assumptions, consistent estimates of $\mu_{n|w}$ and $\ln h_{n|w}$ can be obtained as follows.

Estimation of State Prediction Mean

Given $p(\mathbf{w}_{-n}|\mathbf{y}_{-n})$ and $p(\mathbf{y}) = p(\mathbf{y}_{-n}|y_n)p(y_n)$, the following identity holds true

$$p(y_n|\mathbf{w}_{-n}) \propto p(y_n) \int p(\mathbf{w}_{-n}|\mathbf{y}_{-n})p(\mathbf{y}_{-n}|y_n)d\mathbf{y}_{-n}$$

Denoting by

$$\begin{aligned} p(y_n|\mathbf{w}_{-n}) &= N(\mu_n|w, \sigma_n^2|w) \\ p(y_n) &= N(\mu_n, \sigma_n^2) \\ p(\mathbf{w}_{-n}|\mathbf{y}_{-n}) &= N(\mathbf{y}_{-n}, \sigma_w^2|y\mathbf{I}) \\ p(\mathbf{y}_{-n}|y_n) &= N(\mu_{y|n}, \Sigma_{y|n}) \end{aligned}$$

and using some standard Gaussian identities [Anderson and Moore, 1979], it can be shown that

$$\mu_n|w = \mu_n - \mathbf{M}_{ny}\mu_y + \mathbf{M}_{nw}\mathbf{w}_{-n} \quad (6.17)$$

where

$$\begin{aligned} \mathbf{A} &= \frac{1}{\sigma_n^2}\Sigma_{yn} \\ \mathbf{S}_{yw}^{-1} &= \Sigma_{y|n}^{-1} \left(\mathbf{I} - \sigma_w^2|y(\Sigma_{y|n} + \sigma_w^2|y\mathbf{I})^{-1} \right) \\ \mathbf{M}_{ny} &= (\mathbf{A}'\mathbf{S}_{yw}^{-1}\mathbf{A} + \frac{1}{\sigma_n^2})^{-1}\mathbf{A}'\mathbf{S}_{yw}^{-1} \\ \mathbf{M}_{nw} &= (\mathbf{A}'\mathbf{S}_{yw}^{-1}\mathbf{A} + \frac{1}{\sigma_n^2})^{-1}\mathbf{A}'(\Sigma_{y|n} + \sigma_w^2|y\mathbf{I})^{-1} \end{aligned}$$

Estimation of State Prediction Error Variance

Given $p(\mathbf{w}_{-n}|\mathbf{y}_{-n})$ and $p(\ln e_n^2, \mathbf{y}_{-n})$, the following identity holds true

$$p(\ln e_n^2|\mathbf{w}_{-n}) \propto \int p(\mathbf{w}_{-n}|\mathbf{y}_{-n})p(\ln e_n^2|\mathbf{y}_{-n})p(\mathbf{y}_{-n})d\mathbf{y}_{-n}$$

Denoting by

$$\begin{aligned} p(\ln e_n^2|\mathbf{w}_{-n}) &= N(\mu_e|w, \sigma_e^2|w) \\ p(\mathbf{w}_{-n}|\mathbf{y}_{-n}) &= N(\mathbf{y}_{-n}, \sigma_w^2|y\mathbf{I}) \\ p(\ln e_n^2|\mathbf{y}_{-n}) &= N(\mu_e|y, \Sigma_e|y) \\ p(\mathbf{y}_{-n}) &= N(\mu_y, \Sigma_y) \end{aligned}$$

and using some standard Gaussian identities [Anderson and Moore, 1979], it can be shown that

$$\mu_{e|w} = \mu_e - \mathbf{M}_{ey}\mu_y + \mathbf{M}_{ew}\mathbf{w}_{-n}$$

where

$$\begin{aligned} \mathbf{A} &= \Sigma_{ey}\Sigma_{yy}^{-1} \\ \mathbf{S}_{ye}^{-1} &= \mathbf{I} + \sigma_{w|y}^2 \left(\Sigma_{yy}^{-1} + \mathbf{A}'\Sigma_{e|y}^{-1}\mathbf{A} \right) \\ \mathbf{M}_{ey} &= \mathbf{A} - \sigma_{w|y}^2 \left(\mathbf{I} - \sigma_{w|y}^2 \mathbf{A}\mathbf{S}_{ye}\mathbf{A}'\Sigma_{e|y}^{-1} \right)^{-1} \mathbf{A}\mathbf{S}_{ye}\Sigma_{yy}^{-1} \\ \mathbf{M}_{ew} &= \left(\mathbf{I} - \sigma_{w|y}^2 \mathbf{A}\mathbf{S}_{ye}\mathbf{A}'\Sigma_{e|y}^{-1} \right)^{-1} \mathbf{A}\mathbf{S}_{ye} \end{aligned}$$

The auxiliary equation for a conditionally heteroscedastic process (equation 5.7) provides a Gaussian approximation to $p(\ln h_n | \ln e_n^2)$. Therefore, using

$$p(\ln h_n | \mathbf{w}_{-n}) = \int p(\ln h_n | \ln e_n^2) p(\ln e_n^2 | \mathbf{w}_{-n}) d(\ln e_n^2)$$

with the distributions given by

$$\begin{aligned} p(\ln h_n | \mathbf{w}_{-n}) &= N(\ln h_n | w, \sigma_{h|w}^2) \\ p(\ln h_n | \ln e_n^2) &= N(\ln h_n - c, \sigma_{h|e}^2) \\ p(\ln e_n^2 | \mathbf{w}_{-n}) &= N(\mu_{e|w}, \sigma_{e|w}^2) \end{aligned}$$

where $c = \psi^{(0)}(1/2) + \ln 2$ and $\sigma_{h|e}^2 = \psi^{(1)}(1/2)$, it can be shown that

$$\ln h_n | w = \mu_{e|w} - c \quad (6.18)$$

6.3.3 Extensions to non-Homogeneous Clutter Environment

In the previous section, based on a couple of simplifying assumptions, the expectation of the state prediction mean and the state prediction error variance, conditional on the returns in the reference window, were derived in equations 6.17 and 6.18, respectively. In the case of the modulating component being Gaussian distributed in the logarithm domain, equation 6.17 constitutes a *quasi optimal*, consistent estimator of the local clutter power. The expected prediction error variance provides a measure of uncertainty associated with this estimate and is constant in the case of Gaussian distributed modulating component.

The above analysis was concerned with the local clutter prediction, based on the observations in the reference window, immediately surrounding the range cell under test. In particular, no guard cells were taken into account, and the returns within the reference window were assumed to be target free and to consist of homogeneous background clutter.

However, as was discussed in chapter 3, multiple target and inhomogeneous clutter scenarios can adversely affect the local clutter power prediction performance of cell averaging-type detectors. A simple way to account for multiple targets and clutter edges, would be to apply some kind of censoring to the reference window (not unlike the case of outlier detection). In fact, the study of censoring schemes for the different radar return scenarios could provide an interesting direction of future research.

Once censored, the range cells (including the guard cells) could be tagged as being potentially 'contaminated' and hence discarded for the purpose of local clutter power prediction for the range cell under test. In particular, the observation vector, \mathbf{w}_{-n} , could be truncated to exclude any guard cells, or any 'contaminated' range cells, and the local clutter power could then be estimated as before using equations 6.17 and 6.18, with the covariance matrix for \mathbf{y}_{-n} being trivially collapsed to exclude the censored range cells.

6.4 CHLMAP CFAR DETECTION PERFORMANCE ANALYSIS

A brief review of target detection techniques in compound clutter environment was presented in chapter 3. As was discussed therein, fixed threshold detection is optimal in homogeneous, uncorrelated clutter environment. Any adaptive scheme, where the threshold is adjusted based on a few radar return samples within a reference window, suffers an inherent detection loss, as compared to fixed threshold detection in uncorrelated clutter.

In partially correlated clutter, particularly when the clutter is spiky (*i.e.* its single point statistics deviates from the Rayleigh distribution), a detection gain, with respect to the fixed threshold detection, can be obtained by locally adjusting the detection threshold according to the local clutter power. The maximum gain is determined by the ideal CFAR limit, whereby the local clutter power level is known exactly.

The cell averaging CFAR detector is optimal in homogeneous, uncorrelated exponential clutter environment, where the sample mean is the optimal estimate of the clutter power. The

threshold for a CA CFAR detector is set as

$$V_T(n) = G \sum_{k \neq 0} r_{n+k} \quad (6.19)$$

where r_n is the amplitude of the radar return in the n^{th} range cell. As was shown in chapter 3, when spatial correlations are present in the modulating component, simple cell averaging does not fare particularly well. In fact, even under ideal, speckle free conditions, the detection performance of a cell averaging CFAR detector in compound clutter environment is usually close to the worst case, determined by the fixed threshold detection.

[Bucciarelli et al., 1996] propose to use a weighted linear estimator in the logarithm domain to obtain an estimate of the local clutter power, $\hat{y}_n = \mu_{n|w} = \sum_{i \neq 0} \alpha_i w_{n+i}$. Based on this estimate, the logarithm of the threshold for an LMAP CFAR detector is set as

$$\ln V_T(n) = \ln G + \frac{1}{2} \mu_{n|w} \quad (6.20)$$

The parameters for the local clutter power estimator in the logarithm domain, $\mu_{n|w}$, can be obtained from the covariance matrix of \mathbf{y} , as given by equation 6.17. Furthermore, a potentially spatially more uniform false alarm rate can be obtained by including the local clutter power prediction error variance, $\sigma_{n|w}^2$, in the threshold setting for the range cell under test (this is also known as Class 3 detection [Eaves and Reedy, 1987]). This can be justified intuitively by noting that if one is uncertain as to the local clutter power, then the threshold should be increased for that range cell, in order to ensure that the local probability of false alarm is not too great. The threshold for the resulting detector, denoted as CHLMAP CFAR, is given by

$$\ln V_T(n) = \ln G + \frac{1}{2} \left(\mu_{n|w} + K \sigma_{n|w}^2 \right) \quad (6.21)$$

where G ensures the overall probability of false alarm is within the specified level, while K is adjusted to minimise the SCR required for the given probability of detection (*i.e.* maximise the detectability) of the targets of interest. The expressions for the local clutter power estimate, $\mu_{n|w}$, and the corresponding prediction error variance, $\sigma_{n|w}^2$, conditional on the radar returns in the reference window, \mathbf{w}_{-n} , are given in equations 6.17 and 6.18, respectively.

6.4.1 Distribution of the Threshold

As $\ln V_T(n)$ is a sum of a number of (approximately Gaussian) variates, by the Central Limit Theorem, the distribution of the threshold in the logarithm domain is expected to be well approximated as being Gaussian distributed. In fact, it has been found in practice that the Gaussian distribution provides a good approximation for as few as ten range cells (*i.e.* five plus five, either side of the range cell under test) in the reference window.

For the (CH)LMAP CFAR detector, the logarithm of the normalised threshold, τ_n , is given by

$$\begin{aligned}\tau_n &= \ln(V_T(n)/G)^2 = \mu_{n|w} + K\sigma_{n|w}^2 & (6.22) \\ \mu_{n|w} &= \beta_0 + \sum_{j \neq 0} \beta_j w_{n+j} \\ \ln \sigma_{n|w}^2 &= \alpha_0 + \sum_{i \neq 0} \alpha_i w_{n+i}\end{aligned}$$

where $\mathbf{w}_{-n} = \{w_{n+k} : k \neq 0\}$ are the radar returns in the reference window, in the logarithm domain. As the returns are assumed to be obtained in open ocean conditions, with no interfering targets, they consist of compound clutter samples, *i.e.* speckle locally modulated in power by \mathbf{y}_{-n} . Therefore, using equation 6.9, the mean and the variance of the logarithm of the threshold are given by

$$\mu_\tau = E[\tau_n | \mathbf{y}_{-n}] = m_X + K \exp(m_Y + \frac{1}{2}s_Y^2) \quad (6.23)$$

$$\begin{aligned}\sigma_\tau^2 = \text{var}[\tau_n | \mathbf{y}_{-n}] &= s_X^2 + K^2 \exp(2m_Y + s_Y^2) (\exp(s_Y^2) - 1) & (6.24) \\ &+ 2K \exp(m_Y + \frac{1}{2}s_Y^2) \sigma_{w|y}^2 \sum_{k \neq 0} \alpha_k \beta_k\end{aligned}$$

where, for notational simplicity, the following have been defined as

$$\begin{aligned}m_X = E[\mu_{n|w}] &= \beta_0 + \sum_{j \neq 0} \beta_j y_{n+j} \\ s_X = \text{var}[\mu_{n|w}] &= \sum_{j \neq 0} \beta_j^2 \sigma_{w|y}^2 \\ m_Y = E[\ln \sigma_{n|w}^2] &= \alpha_0 + \sum_{i \neq 0} \alpha_i y_{n+i} \\ s_Y = \text{var}[\ln \sigma_{n|w}^2] &= \sum_{i \neq 0} \alpha_i^2 \sigma_{w|y}^2\end{aligned}$$

with $\sigma_{w|y}^2$ corresponding to the observation noise, associated with the speckle, as given by the linearised equation 6.9. It is interesting to note that the above equations apply to CHLMAP CFAR detector, and simplify trivially to LMAP CFAR detector by setting $\{\alpha_i = 0 : i \neq 0\}$.

6.4.2 Target Detection Performance Analysis

The normalised detection threshold in the logarithm domain for the CHLMAP CFAR detector is given in equation 6.22. As was argued in the previous section, this threshold is approximately Gaussian distributed, with the mean and the variance given by equations 6.23 and 6.24, respectively.

As was pointed out in chapter 3, in order to estimate the *average* probability of detection (and hence the average P_{fa}) for an adaptive threshold detector, the probability of detection in range cell under test, conditional on the threshold for the given range cell, needs to be averaged over all the possible values of the threshold. More concisely, following from equation 3.9,

$$\hat{P}_d(n|\sigma, \bar{A}) = \int_{V_T(n)} P_d(n|V_T(n), \sigma_n, \bar{A}) p(V_T(n)|\sigma_{-n}) dV_T(n) \quad (6.25)$$

where $\sigma_n = \exp y_n$ is the local clutter power.

As was previously mentioned in chapter 3, for reasons of analytical tractability, the analysis presented in this thesis is limited to single pulse detection of Swerling I and II targets in compound clutter environment. The probability of detection, based on a single detection opportunity, for a fluctuating target in locally Rayleigh distributed clutter is given by equation 3.6, *i.e.*

$$P_d(V_T(n)|\sigma_n, \bar{A}) = \exp \frac{-G^2 V_T^2(n)}{\sigma_n + 2\bar{A}^2}$$

Unfortunately, since τ_n is Gaussian distributed (*i.e.* $\tau_n \sim N(\mu_\tau, \sigma_\tau)$), the normalised detection threshold, $V_T(n) = \exp \tau_n/2$, is log-Normal distributed, and therefore, the average probability of detection, as given by equation 6.25, is not easy to evaluate analytically.

However, by approximating the log-Normal distribution for $T(n) = V_T^2(n) = \exp \tau_n$ by a Gamma distribution (*i.e.* $T(n) \sim \Gamma(\alpha_T, \beta_T)$), the average probability for single pulse detection of a fluctuating target, can be shown to be

$$\hat{P}_d(n|\sigma, \bar{A}) = \left[\frac{\beta_T}{\beta_T + G^2/(\sigma_n + 2\bar{A}^2)} \right]^{\alpha_T} \quad (6.26)$$

where the Gamma distribution parameters for $T(n)$ are evaluated, such that the Kullback-Leibler (KL) divergence between the distributions for $\tau_n = \ln T(n)$ and $T(n)$ is minimised. In particular, given that $\tau_n \sim N(\mu_\tau, \sigma_\tau)$ and that $T(n) \sim \Gamma(\alpha_T, \beta_T)$, it can be shown that KL measure is minimised when the parameters α_T and β_T satisfy

$$\begin{aligned} \ln \alpha_T - \psi^{(0)}(\alpha_T) &= \sigma_\tau^2/2 \\ \beta_T &= \alpha_T \exp[-\mu_\tau - \sigma_\tau^2/2] \end{aligned} \quad (6.27)$$

where μ_τ and σ_τ depend on the local clutter power in the reference window, $\sigma_{-n} = \exp \mathbf{y}_{-n}$, as given by equations 6.23 and 6.24, respectively. The corresponding KL measure between the two distributions is given by

$$KL_1 = \alpha_T \ln \beta_T - \ln \Gamma(\alpha_T) + \alpha_T \mu_\tau - \beta_T \exp(\mu_\tau + \sigma_\tau^2/2) \quad (6.28)$$

Through extensive Monte Carlo simulations it has been found that, for most range profiles encountered in practice, the Gamma distribution-based estimate for the probability of false alarm and the probability of detection provides a very accurate prediction for the detection performance. However, it was also found that, in some cases, a prediction based on a Gaussian approximation to $V_T(n)$ provides a more accurate estimate of the radar detection performance. Specifically, approximating the log Normal distribution for $V_T(n) = \exp \tau_n/2$ by a Gaussian distribution (*i.e.* $V_T(n) \sim N(\mu_V, \sigma_V)$), the average probability of detection is given as in equation 3.10, *i.e.*

$$\begin{aligned} \hat{P}_d(n|\sigma, \bar{A}) &= \frac{1}{2\sqrt{\alpha}} \exp \left[\frac{m_V^2}{2\sigma_V^2} \left(\frac{1}{\alpha} - 1 \right) \right] \left(1 + \operatorname{erf} \left[\frac{m_V}{\sqrt{2\alpha\sigma_V^2}} \right] \right) \\ \alpha &= 1 + \frac{2G^2\sigma_V^2}{2A^2 + \sigma_n^2} \end{aligned} \quad (6.29)$$

where the approximating Gaussian distribution parameters for $V_T(n)$ are evaluated, such that the KL divergence between the distributions for τ_n and $V_T(n)$ is minimised. In particular, given that $\tau_n \sim N(\mu_\tau, \sigma_\tau)$ and that $V_T(n) \sim N(\mu_V, \sigma_V)$, it can be shown that KL measure is minimised when the parameters μ_V and σ_V satisfy

$$\begin{aligned} \mu_V &= \exp \left[\mu_\tau/2 + \sigma_\tau^2/8 \right] \\ \sigma_V^2 &= \exp \left[\mu_\tau + \sigma_\tau^2/4 \right] (\exp(\sigma_\tau^2/4) - 1) \end{aligned} \quad (6.30)$$

with the corresponding KL measure given by

$$KL_2 = -\frac{1}{2} [\ln(8\pi\sigma_V^2) - \mu_\tau + 1] \quad (6.31)$$

In summary, in order to analytically estimate the average probability of detection (and hence P_{fa} , by setting $\bar{A} = 0$) for the adaptive threshold detectors, whereby the local clutter power estimate is obtained in the logarithm domain, the following procedure has been adopted throughout this chapter:

- Obtain the parameter estimates for the approximating distributions for the threshold using equations 6.27 and 6.30.
- Evaluate the KL divergence measure for the approximating distributions using equations 6.28 and 6.31.
- Using the approximating distribution with the smaller KL measure, evaluate the probability of detection, based either on equation 6.26 or equation 6.29.

Throughout this chapter, in order to ensure that the analytical expressions provide a good approximation to the actual probabilities of false alarm and detection, the estimates obtained

using the above procedure were compared to the estimates obtained using Monte Carlo methods, whereby the integral equation 6.25 is evaluated, based on samples obtained directly from $p(V_T(n)|\sigma_{-n})$.

6.5 DETECTION PERFORMANCE IN SPATIALLY CORRELATED CLUTTER

In the preceding sections of this chapter, a linearised, conditionally Gaussian state space model for high resolution sea clutter was developed. Based on the state space form of the background clutter, estimators of the local clutter power and the associated expected prediction error variance were proposed in section 6.3 and utilised in an adaptive threshold detection framework proposed in section 6.4. Also in section 6.4, analytical expressions for the detection performance of the proposed detector were derived for the case of single pulse Swerling I and II target detection in homogeneous background clutter environment.

The subsequent analysis in the present section is based on single pulse detection of fluctuating (Swerling I and II) targets, in compound clutter environment. For the purpose of the exposition, the speckle is assumed to be spatially uncorrelated and Rayleigh distributed. This assumption is expected to hold true in most of high resolution radar systems, though it could break down in very high resolution radars, where the returns are effectively oversampled in range. The database of 440 averaged range profiles is used for the analysis of the detection performance of a number of different detection schemes, including fixed threshold, ideal CFAR, CA CFAR, as well as the *predictive* LMAP and CHLMAP CFAR detectors (note that in the present section, the results are presented for the predictive LMAP detector, and not the LMAP detector as originally proposed by [Bucciarelli et al., 1996]).

The following analysis is meant to serve as an illustrative comparison of the different detection schemes in compound clutter environment. The purpose of the study is to reveal the inherent limitations and the potential gains of the different detection techniques, while further analysis, including incoherent integration and censoring, is left as an exercise for future research.

6.5.1 Fixed Threshold and Ideal CFAR Detection Performance Limits

The detection performance measures introduced in section 3.5 will presently be adopted to illustrate the target detection performance of the aforementioned detection schemes in compound clutter environment. In particular, *target detectability* refers to the Signal-to-Clutter Ratio (SCR) required to achieve a given probability of target detection, P_d , for the specified *average* probability of false alarm, P_{fa} . Similarly, the *spatial variation* in the probability of false alarm

refers to the logarithm of the spatial variance of the average probability of false alarm across a range profile.

As was discussed in section 3.5, both in terms of target detectability and the spatial variation of the probability of false alarm, the target detection performance of both the ideal CFAR and the fixed threshold detectors is identical in Rayleigh distributed clutter. In non-Rayleigh clutter, the ideal CFAR detector achieves a spatially constant level of P_{fa} and also determines the maximum achievable level of target detectability in compound clutter environment. However, it was shown in section 3.5 that in non-Rayleigh clutter, the fixed threshold detector is no longer optimal and suffers a detectability loss with respect to the ideal CFAR detector. The detectability loss increases with the spikiness of the background clutter. Adaptive threshold detection techniques can potentially outperform the fixed threshold detector in spiky clutter and approach the ideal CFAR detectability limit in highly correlated background clutter environment.

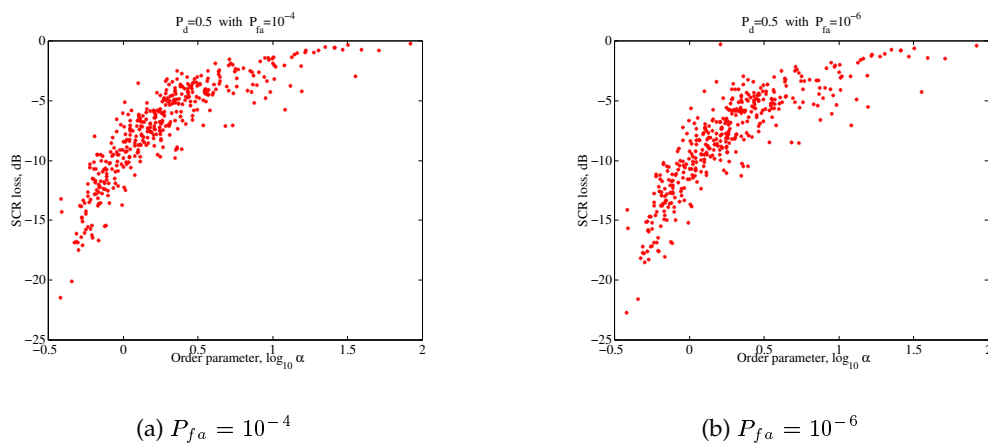


Figure 6.1: Detection loss for the ideal CFAR detector, relative to the Fixed Threshold detector. $P_d = 0.5$ with (a) $P_{fa} = 10^{-4}$ and (b) $P_{fa} = 10^{-6}$. Results shown for database of 440 averaged range profiles.

Figure 6.1 shows the single pulse detection loss for the ideal CFAR detector relative to the fixed threshold detector for a database of 440 averaged range profiles, with the probability of false alarm set at $P_{fa} = 10^{-4}$ and 10^{-6} , and with the probability of detection $P_d = 0.5$. The detectability results are based on the assumption of the speckle being locally Rayleigh distributed and spatially uncorrelated. No assumptions are made about the distribution of the modulating component. Instead, the actual averaged range profiles were used as the basis for the results presented. Although not assumed to be Gamma distributed, in order to give an indication of the 'spikiness' of the data (and to be consistent with the bulk of the radar detection literature), the detection results presented in this chapter are plotted as a function of the Gamma distribution order parameter for the modulating component.

For large values of the Gamma distribution order parameter (*i.e.* for near Rayleigh distributed clutter), figure 6.1 shows that the ideal CFAR detector achieves only a slight gain in target detectability relative to the fixed threshold detector. However, for smaller values of the order parameter (*i.e.* more spiky clutter), the gain in detectability of the ideal CFAR detector relative to the fixed threshold detector can exceed 20dB. Furthermore, the relative gain in detectability is marginally larger for smaller values of the probability of false alarm (*i.e.* $P_{fa} = 10^{-6}$ vs $P_{fa} = 10^{-4}$). Under all conditions, for all the range profiles in the database, the ideal CFAR detector outperforms the fixed threshold detector in terms of single pulse Swerling I and II target detectability in compound clutter environment. Whether or not the adaptive threshold detectors proposed in section 6.4 can achieve a near ideal CFAR detection performance is the subject of the investigation reported in the remaining sections of the present chapter.

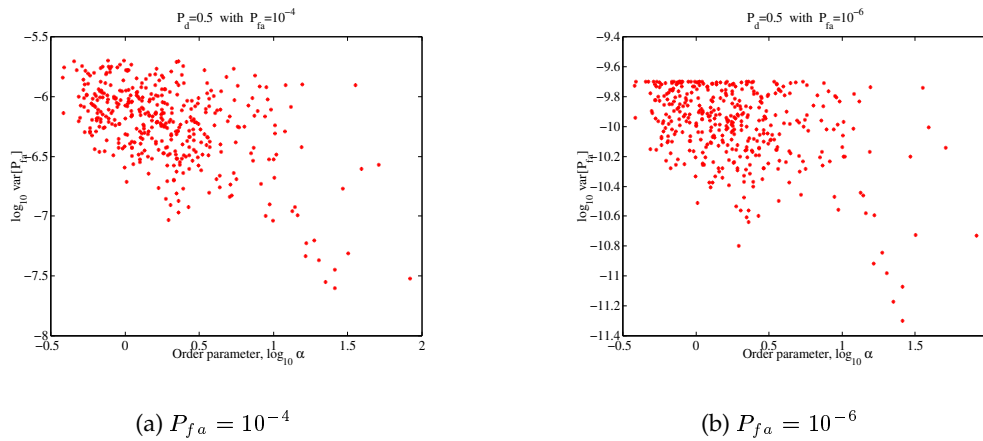


Figure 6.2: Logarithm of the variance of P_{fa} for the Fixed Threshold detector. $P_d = 0.5$ with (a) $P_{fa} = 10^{-4}$ and (b) $P_{fa} = 10^{-6}$. Results shown for database of 440 averaged range profiles.

As a result of the exact knowledge of the local clutter power, the ideal CFAR detector is characterised by a spatially uniform probability of false alarm. This is not necessarily the case for a fixed threshold detector, except for the case of spatially uncorrelated clutter, or in the case of Rayleigh distributed clutter. This is illustrated in figure 6.2, which shows the spatial variation in the probability of false alarm obtained using a fixed threshold detector for the database of 440 range profiles. In particular, the figure shows that the smallest spatial variations are observed for large values of the order parameter (*i.e.* near Rayleigh distributed clutter), with substantial reduction in the spatial variations observed for a smaller average probability of false alarm ¹.

¹The smaller *absolute* value of the spatial variance of P_{fa} is simply due to the smaller *absolute* average value of the probability of false alarm.

6.5.2 CA CFAR Detector Performance

Cell averaging (CA) CFAR detector was briefly discussed in section 3.5. CA CFAR is an adaptive threshold technique, whereby the local clutter power level in the cell under test is obtained by simply averaging the radar return samples contained in the reference window. In addition to any potential targets, the returns in the reference window consist of the speckle locally modulated in power by the modulating component. It was shown in chapter 3 that, even in ideal speckle-free conditions, the detection performance of a CA CFAR detector in compound clutter environment is inherently close to the worst case, as determined by the fixed threshold detection. The reason for this is that the simple cell averaging does not constitute a good local clutter power estimator.

The results presented in chapter 3 corresponds to an idealised, speckle free clutter samples in the reference window. However, even this simplified analysis demonstrates the inherent limitations of the CA CFAR detection scheme. This is also borne out by the findings reported by [Armstrong and Griffiths, 1991a], where the high resolution sea clutter is modelled as being K-distributed with a correlated Gamma distributed modulating component. Presently, the analysis is extended to the database of 440 averaged range profiles, and the effect of speckle in the reference window is investigated for single pulse CA CFAR detection of Swerling I and II targets in compound clutter environment.

Single pulse Swerling I and II target detectability achieved by the CA CFAR detector, relative to the fixed threshold detector, for the database of 440 range profiles is shown in figure 6.3. The results are shown for the average probability of detection $P_d = 0.5$ with the average probability of false alarm $P_{fa} = 10^{-4}$ and 10^{-6} . The reference window size was adjusted to 5+5, 10+10 and 20+20 reference cells, with no guard cells inserted. Similarly, under the same conditions, the corresponding results for the spatial variation of the probability of false alarm for the database of the averaged range profiles is shown in figure 6.4. These results can now be directly compared to the limiting case of ideal CFAR detection, discussed in the previous section.

Figures 6.3 and 6.4 demonstrate that the reference window size has a profound effect on the target detection performance of a CA CFAR detector. To better illustrate this point, a brief discussion of the results obtained with the average probability of false alarm set to $P_{fa} = 10^{-4}$ will now be presented. To this end, figures 6.3(a)-6.3(c) show the detectability loss of the CA CFAR detector, relative to the fixed threshold detector, for reference window size of 5+5, 10+10 and 20+20 cells, respectively. Similarly, figures 6.4(a)-6.4(c) show the increase in the spatial variation of probability of false alarm for the CA CFAR detector, relative to the fixed threshold detector, for reference window size of 5+5, 10+10 and 20+20 cells, respectively.

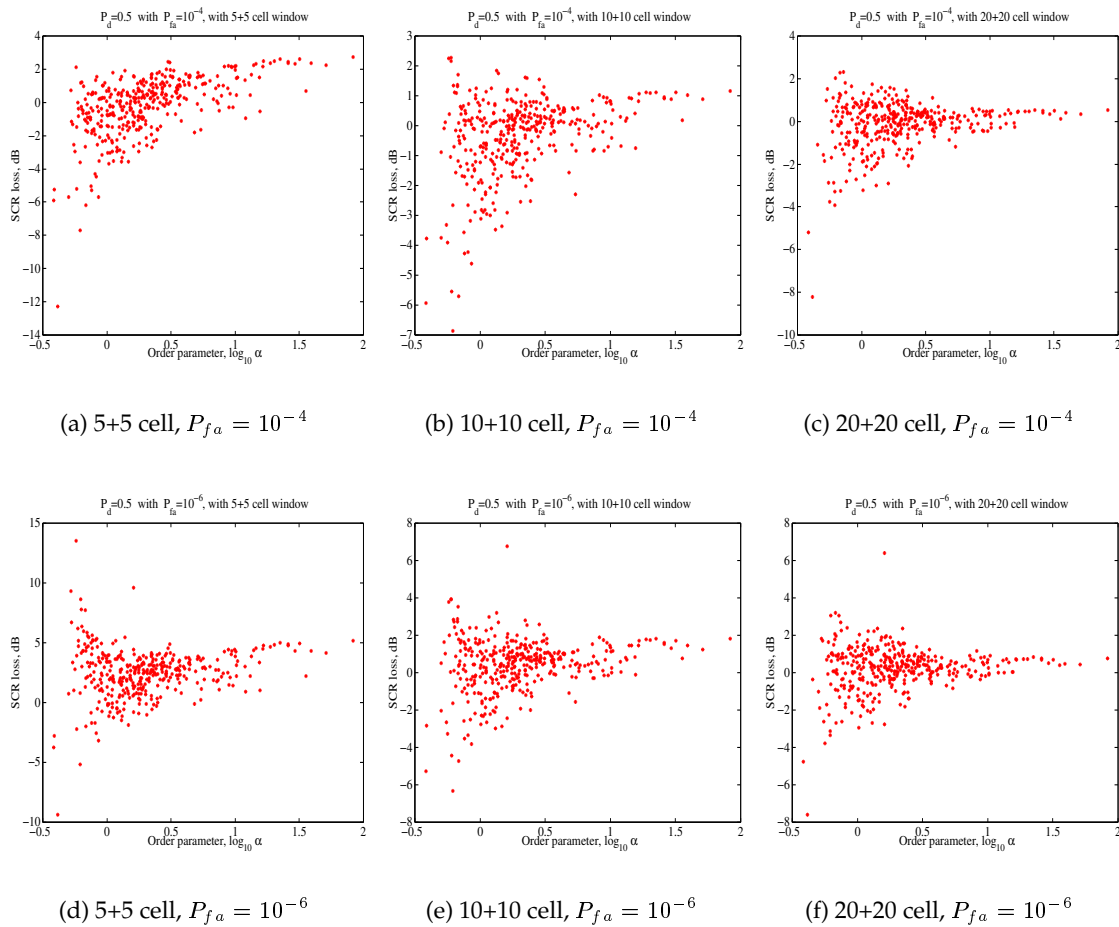


Figure 6.3: Detection loss, as a function of reference window size, for a CA CFAR detector, relative to a Fixed Threshold detector. $P_d = 0.5$ with $P_{fa} = 10^{-4}$ (see (a), (b) and (c)) and $P_{fa} = 10^{-6}$. (see (d), (e) and (f)). Reference window size 5+5 (see (a) and (d)), 10+10 (see (b) and (e)) and 20+20 (see (c) and (f)) range cells with no guard cells inserted. Results shown for database of 440 averaged range profiles.

As shown in figures 6.3(a)-6.3(c), the CA CFAR detector suffers from a characteristic loss in target detectability, relative to the fixed threshold detector, at large values of the order parameter. This loss in near Rayleigh distributed clutter arises from the random variations in the threshold, which itself is estimated from a finite size window. The loss in detectability decreases with increasing size of the reference window. In particular, while the detection loss is about 3dB for CA CFAR with 5+5 cell reference window, it is only 1dB for 10+10 cell reference window, and 0.5dB when the reference window has 20+20 cells.

As was shown in the previous section, the ideal CFAR detector exhibits a gain in target detectability for small values of the order parameter. In particular, for $P_{fa} = 10^{-4}$, a gain of up to 20dB relative to fixed threshold detection was observed for spiky clutter. However, as figures 6.3(a)-6.3(c) show, a single pulse CA CFAR detector can suffer a net detection loss (as opposed

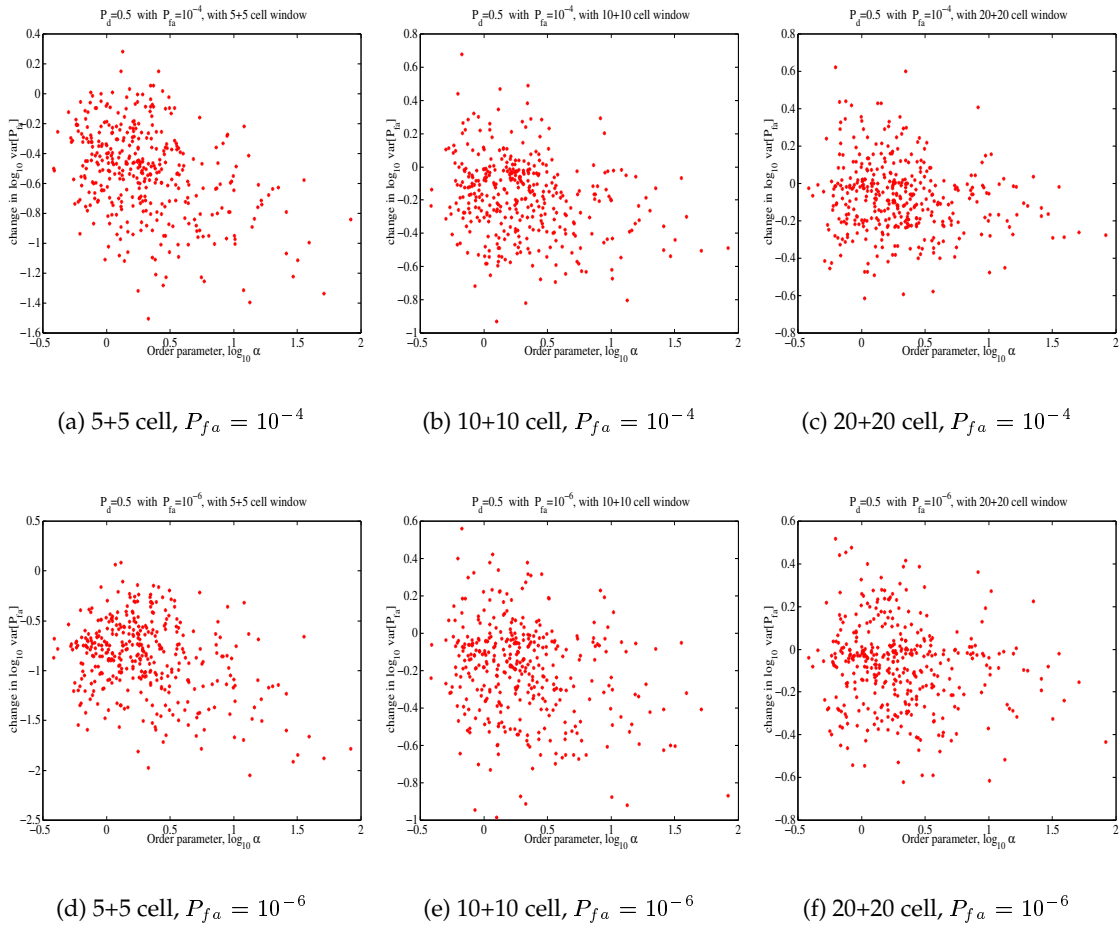


Figure 6.4: Change in the logarithm of the variance of P_{fa} , as a function of reference window size, for a CA CFAR detector, relative to a Fixed Threshold detector. $P_d = 0.5$ with $P_{fa} = 10^{-4}$ (see (a), (b) and (c)) and $P_{fa} = 10^{-6}$. (see (d), (e) and (f)). Reference window size 5+5 (see (a) and (d)), 10+10 (see (b) and (e)) and 20+20 (see (c) and (f)) range cells with no guard cells inserted. Results shown for database of 440 averaged range profiles.

to the ideal CFAR gain) at small values of order parameter (*i.e.* spiky clutter). While a 2dB loss is observed for some of the range profiles, CA CFAR detection exhibits a gain in detectability for some other profiles in the database. The maximum observed gains in detectability range from 12dB for CA CFAR detector with 5+5 cell reference window, down to an 8dB gain for a CA CFAR detector with 20+20 cell reference window. The reduction in the gain observed with the increasing reference window size is not unexpected, since as the reference window size is increased, fixed threshold detection is effectively reached in the limit.

As figures 6.4(a)-6.4(c) show, no significant improvement in the spatial variation of false alarm for a CA CFAR detector, relative to the fixed threshold detector, is observed. For a small reference window size (5+5 cell), a net improvement of up to 1.5dB is observed over the fixed threshold detector. This could mainly be due to the 'randomness' of the CA CFAR threshold

estimated from a few noisy samples in the reference window. As expected, the spatial variation of the probability of false alarm of CA CFAR detector approaches the fixed threshold limit, as the reference window size is increased.

The target detectability and spatial variation of the probability of false alarm for the database of the averaged range profiles obtained using the CA CFAR detector with the average $P_{fa} = 10^{-6}$ are shown in figures 6.3(d)-6.3(f) and 6.4(d)-6.4(f), respectively. With a few notable exceptions, the results obtained for $P_{fa} = 10^{-6}$ correspond closely with the results obtained for $P_{fa} = 10^{-4}$. In particular, as a result of the increased sensitivity of the threshold at lower average probability of false alarm ², a significantly larger loss in detectability is observed in figures 6.3(d)-6.3(f) compared to the loss observed in figures 6.3(a)-6.3(c). This is particularly pronounced for small reference window size, where a loss of up to 14dB is observed for small values of the order parameter (*i.e.* spiky clutter). However, as figures 6.4(d)-6.4(f) show, due to the rather random nature of the threshold, the effects of the reduced level of average probability of false alarm on the spatial variance of P_{fa} is not nearly as pronounced.

In short, based on the results presented in this section, the following observations have been made about single pulse CA CFAR detection of Swerling I and II targets in compound clutter environment. Firstly, the CA CFAR detector target detectability performance, on average, is closer to the fixed threshold than to the ideal CFAR detector performance. Additionally, for some range profiles, the CA CFAR detector exhibits an additional detectability loss, relative to the fixed threshold detector, which is particularly pronounced for small reference window sizes. As the reference window size is increased, the detectability performance of a CA CFAR detector approaches that of the fixed threshold detector. The abovementioned effects are particularly pronounced for smaller values of the probability of false alarm.

Finally, in terms of the spatial probability of false alarm, no significant improvements were observed for CA CFAR detector, relative to the fixed threshold detector. In the case of small reference window size, some improvement was observed due to the random nature of the CA CFAR detection threshold setting. However, such improvements in spatial variations are achieved at the expense of detectability. Clearly, ideal CFAR performance has not been achieved and, as will be discussed in the following section, other detection schemes may need to be entertained.

6.5.3 (CH)LMAP CFAR Detector Performance

In the previous section, single pulse target detection performance of CA CFAR detector was investigated for Swerling I and II targets in spatially correlated compound clutter environment.

²At smaller P_{fa} , the detection threshold is more sensitive to the *tails* of the clutter distribution.

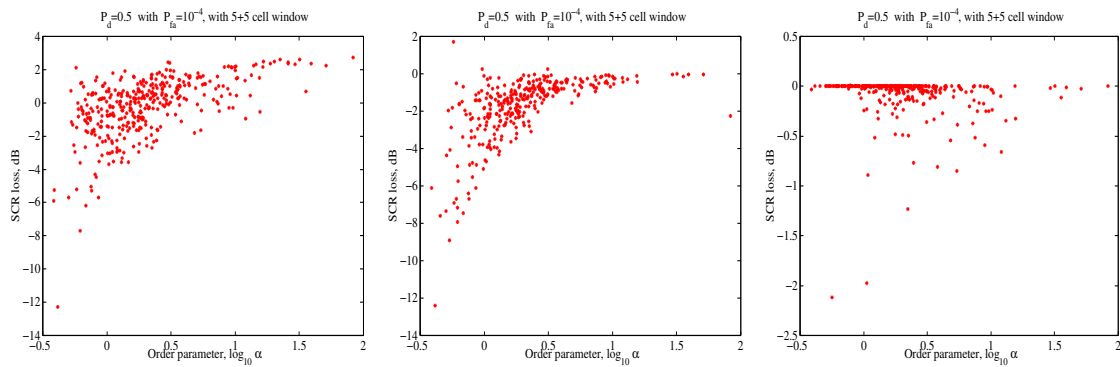
In particular, it was found that the detection performance (as measured by target detectability) of the CA CFAR is always closer to the fixed threshold detector, than to the ideal CFAR detector performance. Furthermore, no significant improvement in spatial variation of the probability of false alarm was observed for the CA CFAR compared to the fixed threshold detector, except for the case of CA CFAR detector with a small reference window size.

The effects of the CA CFAR detector reference window size were also investigated in the previous section. In this regard, it was found that the performance of the CA CFAR detector approaches the fixed threshold detector performance as the reference window size is increased. It was also found that the average probability of false alarm does not have a great influence (in qualitative terms) on the overall target detection behaviour of the CA CFAR detector.

In the present section, the effects of incorporating the spatial correlation structure of the background clutter into detection threshold setting are being investigated. To this end, the models introduced and discussed in sections 6.3 and 6.4 are being studied. Unlike the case of CA CFAR detection, it was found that the reference window size does not play a major factor in the detection performance when the spatial correlation structure is being incorporated in the threshold setting. The reason for this is that predominantly, most of the spatial correlation information is contained in the few immediate range cells around the cell under test. Furthermore, a large reference window can have detrimental effects on the detector performance, resulting from weighting and inclusion of noisy radar returns samples, which are not greatly correlated with the sample in the reference cell under test.

Figures 6.5 and 6.6 show the target detection results obtained using a CA CFAR, LMAP CFAR and CHLMAP CFAR detectors for the previously discussed database of 440 averaged range profiles. The average probability of false alarm for each profile was set to $P_{fa} = 10^{-4}$, with the average probability of detection for Swerling I and II targets set to $P_d = 0.5$. The reference window size was fixed at 5+5 range cells (with no guard cells, and 2+2 range cells for the prediction error variance estimate of the CHLMAP CFAR detector), while the number of pulses integrated in the reference window (but not the range cell under test) was changed from $T = 1$ to $T = 5$ and to $T = 20$ pulses. The present analysis is motivated by the idealised results reported in chapter 3, where the LMAP CFAR detector was shown to achieve a near ideal CFAR detection performance for speckle free radar returns in the reference window (*i.e.* the limiting case of $T \rightarrow \infty$).

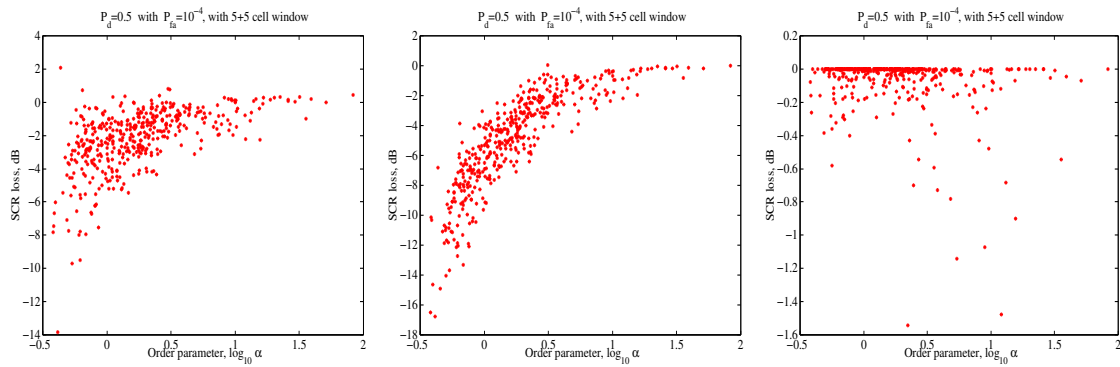
Figures 6.5(a)-6.5(c) and 6.6(a)-6.6(c) illustrate the results obtained for single pulse detection, with no integration in the reference window (*i.e.* $T = 1$). In particular, figure 6.5(a) illustrates the detectability loss of a CA CFAR detector, relative to the fixed threshold detector. Figure 6.5(b) shows the detectability loss of an LMAP CFAR detector, relative to the fixed threshold detector, while figure 6.5(c) shows an additional loss incurred by the CH LMAP CFAR detector,



(a) CA CFAR relative to Fixed Threshold, $T = 1$

(b) LMAP CFAR relative to Fixed Threshold, $T = 1$

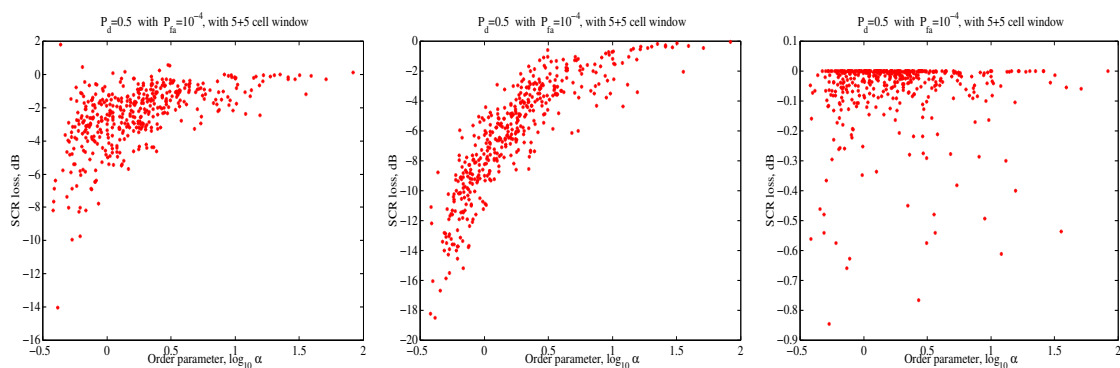
(c) CHLMAP CFAR relative to LMAP CFAR, $T = 1$



(d) CA CFAR relative to Fixed Threshold, $T = 5$

(e) LMAP CFAR relative to Fixed Threshold, $T = 5$

(f) CHLMAP CFAR relative to LMAP CFAR, $T = 5$



(g) CA CFAR relative to Fixed Threshold, $T = 20$

(h) LMAP CFAR relative to Fixed Threshold, $T = 20$

(i) CHLMAP CFAR relative to LMAP CFAR, $T = 20$

Figure 6.5: Effect of the number of integrated pulses, T , on the detection loss. Reference window size fixed at 5+5 range cells, with no guard cells. $P_d = 0.5$ with $P_{fa} = 10^{-4}$. Results shown for database of 440 averaged range profiles.

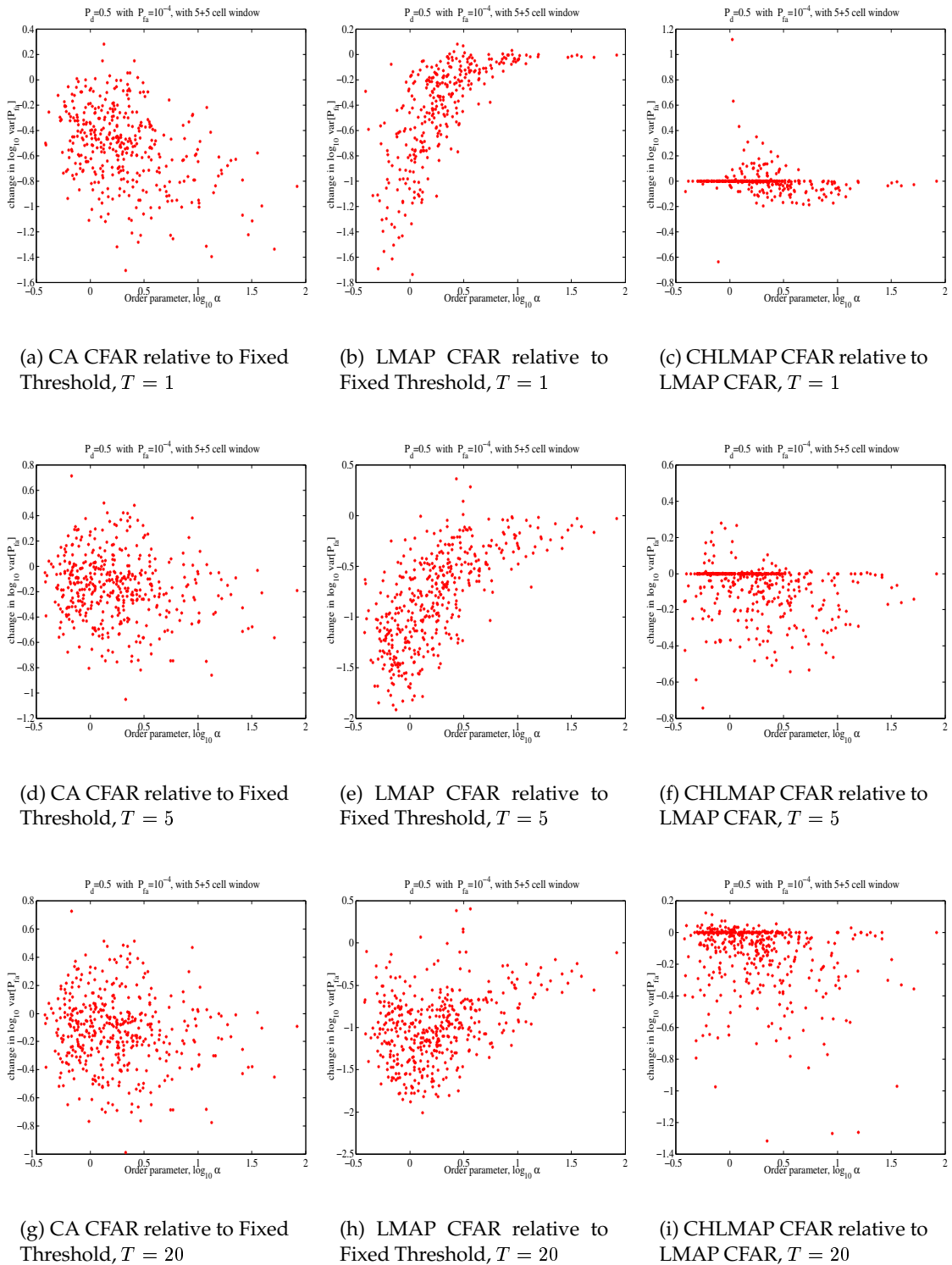


Figure 6.6: Effect of the number of integrated pulses, T , on the spatial variation of P_{fa} . Reference window size fixed at 5+5 range cells, with no guard cells. $P_d = 0.5$ with $P_{fa} = 10^{-4}$. Results shown for database of 440 averaged range profiles.

relative to the LMAP CFAR detector for the database of averaged range profiles.

A number of notable results become apparent by looking at figures 6.5(a)-6.5(c). In particular, the detection loss associated with the CA CFAR detector, relative to the fixed threshold detector, for large values of the order parameter (*i.e.* near Rayleigh clutter) is no longer present in the case of the LMAP CFAR detector. This could potentially be explained by the fact that the threshold estimate of the LMAP CFAR detector is better 'tuned in' to the correlation structure of the modulating component, and thus does not suffer from the bias incurred by the CA CFAR detector, which equally weighs noisy clutter samples in the reference window. However, other than the reduction in the detection loss relative to the fixed threshold detector (associated with the small reference window size), the LMAP CFAR detector does not show any significant improvement in terms of target detectability over the CA CFAR detector. Furthermore, the improvements obtained for CHLMAP CFAR, relative to the LMAP CFAR detector (figure 6.5(c)) are rather sporadic and insignificant.

Figure 6.6(a) illustrates the change in spatial variation of the probability of false alarm obtained using a CA CFAR detector, relative to the fixed threshold detector. Figure 6.6(b) shows the change in the spatial variation in the probability of false alarm for an LMAP CFAR detector, relative to the fixed threshold detector, while figure 6.6(c) shows the additional increase in the spatial variation of the probability of false alarm incurred by the CHLMAP CFAR detector, relative to the LMAP CFAR detector. In particular, while the CA CFAR detector exhibits some sporadic improvement in the spatial variation of P_{fa} , the LMAP CFAR detector exhibits a marked and consistent improvement, relative to the fixed threshold detector, particularly at small values of the order parameter (*i.e.* spiky clutter). The CHLMAP CFAR detector performance is largely the same as that of the LMAP CFAR detector, with no consistent pattern of any improvement.

The above analysis was also extended to the case of pulse integration in the reference window. The motivation for introducing the pulse integration in the reference window, rather than in the reference cell under test, stems from the results presented in chapter 3. In particular, it was shown above that for the case of single pulse detection, the LMAP CFAR detector shows no significant improvement over the simple CA CFAR detector. However, as was demonstrated in section 3.5, the detection performance (as measured by Swerling I and II target detectability) of LMAP CFAR detector approaches the performance of ideal CFAR detector, under the idealised conditions of no speckle present in the reference window (*e.g.* through pulse integration techniques).

Another reason for introducing pulse integration in reference window stems from the ease of analysis of the modified detector performance. In particular, because of the spatial and temporal independence of the speckle in the profiles considered, and since no targets are deemed

to be present in the reference window, the effects of pulse integration on the distribution of the detection threshold are trivially incorporated in the model outlined in section 6.4 above. However, as was discussed in chapter 3, due to the complications associated with the presence of fluctuating targets, pulse integration in the cell under test is not easily treated in adaptive threshold detection framework in compound clutter environment. In fact the effects of the presence of extraneous targets and clutter edges, as well as the effects of pulse integration in the cell under test, are left as a rather cumbersome and difficult exercise for future research.

Results for the detection loss and the increase in the spatial variation of the probability of false alarm for the CA CFAR, LMAP CFAR and CHLMAP CFAR detection schemes are shown, respectively, in figures 6.5(d)-6.5(f) and 6.6(d)-6.6(f) for $T = 5$ integrated pulses in the reference window, and figures 6.5(g)-6.5(i) and 6.6(g)-6.6(i) for $T = 20$ integrated pulses in the reference window. The reference window size for all detectors was maintained at 5+5 neighbouring range cells (2+2 for the prediction error variance estimate for the CHLMAP CFAR detector), with no guard cells inserted. A number of interesting results follow from these figures, and are now briefly discussed.

Figures 6.5(d) and 6.5(g) show that the pulse integration in the reference window does not have a significant effect on the detection performance (in terms of target detectability) of the CA CFAR detector. Similarly, figures 6.6(d) and 6.6(g) demonstrate that as the number of integrated pulses is increased, the improvement in the spatial variation of the probability of false alarm of the CA CFAR detector, relative to the fixed threshold detector, decreases. The effective increase in the variance of P_{fa} with increasing T is qualitatively similar to the increase observed when the reference window size of the CA CFAR detector was increased (as discussed in the previous section). The reasons for both effects is similar, in that as the number of samples averaged increases, either through pulse integration or increased reference window size, the spatial variability of the threshold decreases. This effectively increases the *average* spatial variance of the probability of false alarm, since the cell average does not constitute a particularly good local clutter power estimate.

Figures 6.5(e) and 6.5(h) show analogous results for the detection loss of the LMAP CFAR detector, relative to the fixed threshold detector, with $T = 5$ and $T = 20$ integrated pulses, respectively. Unlike the case of the CA CFAR detector, even for a modest number of 5 integrated pulses, the LMAP CFAR detection scheme demonstrates a near ideal CFAR target detection performance in compound clutter environment. Similarly, figures 6.6(e) and 6.6(h) also show a slight improvement in the spatial variance of the probability of false alarm of the LMAP CFAR detector, relative to the fixed threshold detector, as the number of integrated pulses, T , is increased.

Figures 6.5(f) and 6.5(i) illustrate that the inclusion of the prediction error variance in the de-

tection threshold setting (*i.e.* CHLMAP CFAR detector) provides only a negligible additional improvement in target detectability over the LMAP CFAR detector. This point is further borne out in figures 6.6(f) and 6.6(i) which similarly demonstrate that only a small and sporadic improvement in terms of the spatial variance of the probability of false alarm is obtained for CHLMAP CFAR detector, relative to the LMAP CFAR detector.

In summary, accounting for the prediction error variance (CHLMAP CFAR) provides an insignificant benefit compared to simply accounting for the spatial correlation structure of the clutter (LMAP CFAR). The possible reasons for this observation can be found in the discussion presented in the previous chapters. In particular, the non-constant prediction error variance was based on the premise of the modulating component of sea clutter being Gamma distributed. But, as was argued in chapters 3 and 5, this may not always be the case.

Finally, a brief comment on the parameter estimation schemes is in order. Although the weighting coefficients for all the detection schemes presented herein were obtained from the clutter map (*i.e.* an immediately preceding *raw*, *i.e.* non-averaged, range profile), a further sensitivity analysis may need to be undertaken, particularly for cases when thermal noise, clutter edges or other extraneous interference is present in the radar returns. Based on the above experiments, the LMAP CFAR detector coefficients for local clutter power estimation appear to be robust for the datasets considered here. More attention needs to be paid to the CHLMAP CFAR detector prediction error variance coefficients. The reason for this is that the CHLMAP CFAR detection threshold is currently set by adjusting the gain G in equation 6.21 to ensure the correct average probability of false alarm, while the gain K is adjusted such that the target detectability is maximised. If the gain K is large, any bias in the estimate of the prediction error variance will be accordingly amplified, resulting in a possible degradation of the CHLMAP detector performance. This effect may require further study. However, unlike the LMAP CFAR detector, the inclusion of heteroscedastic effects in the adaptive threshold framework provides only a small and insignificant improvement, rendering the utility of future studies of the CHLMAP CFAR detection scheme highly questionable.

6.6 DISCUSSION, CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

This chapter has dealt with adaptive threshold detection in the presence of spatially correlated sea clutter. Based on the compound form of the clutter, a linearised, conditionally Gaussian state space model for sea clutter in the logarithm domain was developed in section 6.2. In particular, spatially uncorrelated, Rayleigh distributed speckle provides an approximately linear and Gaussian observation equation, while the conditional heteroscedastic model studied in chapter 5 provides a state transition equation for the unobserved modulating component, in the logarithm domain.

In section 6.3, the issue of local clutter power estimation from the returns in the reference window was considered. The returns in the reference window were assumed to consist of target free, homogeneous background clutter (*i.e.* no clutter edges). The LMAP scheme for estimating the local clutter power from the returns in the reference window, as originally proposed by [Bucciarelli et al., 1996], was shown to be ill-conditioned when the observation noise variance tends to zero (*e.g.* through the use of pulse integration). Furthermore, due to the conditional heteroscedastic nature of the state space model for sea clutter in the logarithm domain, the estimation of the predictive distribution for the local clutter power, conditional on the radar return samples in the reference window, proves to be no easy task.

Based on a number of simplifying assumptions, conditional estimates of the prediction mean and the prediction error variance for the local clutter power were obtained in section 6.3. These estimates are *quasi optimal* (in the minimum mean squared linear sense) for the case of log-Normal distributed modulating component. As was shown in chapter 3, this is often the case. In the case of non-Gaussian modulating component in the logarithm domain, the expected prediction mean provides an informed estimate of the local clutter power, while the expected prediction error variance provides an estimate of the uncertainty associated with the prediction of the local clutter power. These two estimates are then used together in an adaptive detection scheme, denoted as CHLMAP CFAR, considered in section 6.4.

Although the discussion in section 6.3 was concerned with local clutter power estimation in target-free, homogeneous background clutter, the proposed framework can be readily extended to multiple target, inhomogeneous clutter scenario. In particular, by using some form of censoring scheme, some of the range cells in the reference window are tagged as being 'contaminated' (*i.e.* contain clutter edges, interfering targets, *etc.*) and hence discarded for the purpose of local power estimation. The study of the censoring techniques for different radar return scenarios provides scope for future research.

An adaptive threshold detector utilising both the prediction mean and the prediction error variance of local clutter power was introduced in section 6.4. An analytical method was developed for assessing the target detection performance of the proposed CHLMAP CFAR detector in compound clutter environment. The analytical approximation to the probability of detection, accuracy of which was monitored by Monte Carlo simulation, was utilised in section 6.5 to assess the target detection performance of the proposed detection scheme.

In section 6.5, the target detection performance of LMAP CFAR and CHLMAP CFAR detectors was compared to the detection performance of CA CFAR, ideal CFAR and fixed threshold detectors, in compound clutter environment. To this end, a large database of averaged high resolution sea clutter range profiles was used for the purpose of target detection performance analysis. The comparative study was undertaken under a number of simplifying assumptions.

Following [Armstrong and Griffiths, 1991a], the study was limited to the detection of Swerling I and II fluctuating targets, based on a single detection opportunity. Furthermore, the speckle was assumed to be spatially uncorrelated, and no thermal noise present (*i.e.* thermal noise being negligible compared to the background sea clutter). Under these conditions it was found that CHLMAP CFAR detector provides only a negligible improvement in detectability relative to LMAP CFAR detector.

In the case of single pulse detection with the threshold evaluated from a reference window containing a single pulse range profile, both LMAP CFAR and CA CFAR detectors perform close to the worst case scenario, as determined by the fixed threshold detector. For small reference window size, the CA CFAR detector suffers an additional detection loss, relative to LMAP CFAR and fixed threshold detectors, as it does not take the prior knowledge of the clutter correlation structure into account, and the threshold is based on the local clutter power estimate obtained from the reference window consisting of only a few range cells. Furthermore, while the CA CFAR detection performance does not improve greatly when the threshold is estimated from an integrated range profile, the LMAP CFAR detection performance approaches that of ideal CFAR, even for as few as five integrated pulses in the reference window.

In short, it was found that while the CHLMAP CFAR detector ostensibly provides negligible improvement relative to the LMAP CFAR detector under most conditions, the LMAP CFAR detector can potentially result in the detection performance close to that of the ideal CFAR detector in compound clutter environment. Although the improvement for LMAP CFAR relative to CA CFAR detector is negligible for single pulse detection, the improvement obtained can be very large when the reference window contains a small number of integrated pulses. It is therefore interesting for future research to compare the performance of CA CFAR and LMAP CFAR detectors employing incoherent integration techniques. Unfortunately, to the best of the author's knowledge, this is not easily analytically tractable and would require extensive Monte Carlo simulations.

Finally, in the results presented in this chapter, no significant deterioration in the detection performance was noted when the modulating process parameters were obtained from the clutter map. However, the robustness of the proposed scheme to the model parameter estimates deserves a more thorough study, particularly with respect to what effects the thermal noise and the spatial correlations in the speckle might have on the local clutter power estimates. Other issues worthy of future considerations include the study of the benefits of retaining the spatial correlation information in multiple target and inhomogeneous clutter scenarios, when censoring techniques are employed to suppress the contaminated range cells within a reference window. Any potential improvements are clearly limited by the results obtained in the present chapter, wherein the radar returns within the whole reference window were used for the purpose of local clutter power estimation and hence setting of the adaptive threshold.

Summary, Conclusions and Suggestions for Further Research

The contributions of the work reported in this thesis are manifold. In addition to providing an extensive review on the subjects of sea clutter modelling and target detection in compound clutter environment, a number of interesting insights were also presented. These will now be summarised briefly, along with a number of suggestions for future research.

Throughout this thesis a statistical paradigm was adopted within a compound, multiplicative process framework. In particular, high resolution sea clutter was modelled as a stochastic process, consisting of a fast oscillating speckle component, modulated in power by a slowly varying modulating component, associated with the swell.

Two major issues associated with the modelling of the compound high resolution sea clutter were identified in chapter 2 of this thesis. Firstly, while the temporal correlations of sea clutter in coherent radar systems are modelled readily, this is not necessarily the case for sea clutter in incoherent radar systems. Secondly, due to the non-Gaussian nature of its statistics, no simple models have been presented to date that capture both the statistics and the spatial correlation structure of the modulating component of high resolution sea clutter.

The significance of the aforementioned issues in the context of target detection in compound clutter environment was discussed in chapter 3. In particular, it was shown that a significant improvement in target detectability can be obtained by accounting for the spatial correlation structure of the modulating component. It was also argued that, particularly in the case of small moving targets, the Doppler information contained in the temporal correlation structure of the radar returns provides a powerful means for target detection in the presence of sea clutter.

As an extension of the previously reported work on target detection in coherent radar systems, a complex autoregressive process was proposed in chapter 4 as the basis for characterisation of high resolution sea clutter spectra in incoherent radar systems. As no phase information is available in incoherent radar returns, the Gibbs sampler was used to facilitate sampling from the autoregressive process parameter posterior distribution, conditional on the observed amplitudes. To this end, the Hybrid Monte Carlo (HMC) algorithm was employed to conditionally sample for the missing phases.

In addition to the computational expense associated with sampling the process parameters, autoregressive process-based characterisation of sea clutter spectra in incoherent radar systems was found to be fraught with dangers. In particular, it was found that sea clutter may actually be better characterised as an autoregressive moving-average (ARMA) process, thus exacerbating the identifiability concerns. Also, ARMA modelling would add another level of computational complexity to the already computationally intensive algorithm. Instead, an approach based on higher order spectra characterisation may better be considered as part of a future research effort.

Chapter 4 has also dealt with the HMC algorithm. In particular, it was argued that the basic HMC algorithm suffers from a number of potential shortcomings. To this end, a symmetrical momentum-based tempering scheme was proposed for improved mixing of the HMC algorithm. The sampling performance of the proposed scheme was illustrated on a multi-modal distribution.

Modelling of the spatial correlation structure of high resolution sea clutter was dealt with in chapter 5. An argument based on birth-death migration for a population of scattering centres on the sea surface was used to propose a family of conditional heteroscedastic processes for the modulating component of sea clutter in the logarithm domain. The attractiveness of such models stems from their linear, conditionally Gaussian structure, which can in principle be utilised for improved adaptive threshold detection in compound clutter environment.

As was discussed in chapter 5, heteroscedasticity manifests itself through correlations in log squared prediction errors of the conditionally heteroscedastic model. Through an informal inspection, a possible presence of heteroscedasticity was found in the time series corresponding to the modulating component of sea clutter. Upon closer inspection of a large database of averaged range profiles, the evidence for heteroscedasticity in the modulating component of high resolution sea clutter in the logarithm domain was shown to be not very compelling.

As a consequence of the research undertaken in chapter 5 (with the findings reported in chapter 3) it was found that, contrary to the widely held belief, the modulating component of high resolution sea clutter, rather than being Gamma distributed, may better be modelled as being log-Normal distributed. The implications of this finding are twofold.

The first consequence is that, as was shown in chapter 3, the Gamma distribution assumption for the marginal statistics of the modulating component of sea clutter can result in an overly optimistic assessment of the target detection performance in compound clutter environment with log-Normal distributed modulating component. Therefore, as part of a future research effort, it may prove fruitful to more thoroughly assess the implications of the Gamma distribution assumption on target detection performance prediction in compound clutter environment.

The second consequence of the above observation is that the marginal statistics of the modulating component of high resolution sea clutter in the logarithm domain is better modelled as being Gaussian distributed. This seems to corroborate the findings presented in this dissertation, in that no strong evidence for heteroscedasticity (which was expected for Gamma distributed modulating component) was found in the modulating component of high resolution sea clutter.

Although the evidence for heteroscedasticity in the modulating component of sea clutter is not very strong, one question still remains to be answered. In particular, does the (very limited) heteroscedasticity observed arise due to its inherent presence in the clutter, or whether it is a manifestation of some other, unaccounted for property of sea clutter, such as a non-linearity, or possibly non-Gaussian distributed prediction errors. Although not pursued further at present, this question provides ample scope for future research.

Finally, a Gaussian state space model for high resolution sea clutter was presented in chapter 6. To this end, the compound multiplicative form of the clutter was linearised by modelling the clutter in the logarithm domain, where the modulating component also happens to be Gaussian distributed. In order to quantify the effects of heteroscedasticity on target detection performance in compound clutter environment, conditional heteroscedasticity was introduced for the modulating component in the state space model for sea clutter in logarithm domain.

It was shown in chapter 6 that the LMAP scheme proposed by [Bucciarelli et al., 1996] for local clutter power estimation is ill-conditioned when the observation noise variance tends to zero (*e.g.* through the use of pulse integration). Based on a number of simplifying assumptions, estimators for the prediction mean and the prediction error variance of the local clutter power, conditional on the returns in the reference window, were derived instead. The local clutter power estimate and its uncertainty were then utilised in an adaptive threshold detector.

The results presented in chapters 3 and 6 reveal a number of interesting insights. In particular, it was shown that the target detection performance of a simple cell averaging CFAR detector in spatially correlated clutter is almost always close to the detection performance of the fixed threshold detector. The reason for this is that the cell averaging CFAR detector does not account particularly well for the spatial correlation structure of the underlying modulating component.

It was also shown in chapter 6 that the target detection performance of a CFAR detector can be greatly enhanced by adjusting the weighting coefficients to take account of the spatial correlation structure of the modulating component of sea clutter in the logarithm domain. However, no significant improvement was obtained by accounting for the heteroscedastic effects in the modulating component of sea clutter, thus further supporting the notion that there is no strong evidence for heteroscedasticity in high resolution sea clutter.

In short, the greatest improvement in target detection performance of adaptive threshold detectors was observed when the weighting coefficients were adjusted to account for the correlation structure of high resolution sea clutter in the logarithm domain. Although no significant improvement is obtained in the case of single pulse detection, it was shown that close to ideal CFAR performance can be achieved by using pulse integration techniques in conjunction with adaptive threshold detection utilising the spatial correlation structure.

Although the results presented in chapter 6 are extensive, there remains ample scope for future research. In particular, the effects of incorporating pulse integration techniques in adaptive threshold detection framework deserve further study, as does the study of robustness of the proposed schemes to the model parameter estimates (particularly with respect to the effects of thermal noise and the spatial correlations in the speckle). Finally, the study of the proposed detection schemes in multiple target and inhomogeneous clutter scenarios may prove to be worthy of future consideration. To this end, censoring techniques can be employed in order to suppress the contaminated range cells within a reference window.

Correlations in Rayleigh Clutter

A

Beyond two dimensions, there is no simple analytical form for a joint distribution of Rayleigh distributed amplitudes corresponding to a complex Gaussian process. Furthermore, multivariate Rayleigh distribution is not uniquely defined by its univariate marginal and autocorrelation function.

In this appendix, based on the results of [Szajnowski, 1977] and [Li and Yu, 1989], the relationship between the autocorrelations of the Rayleigh distributed amplitudes and the autocorrelations of the corresponding *complex* Gaussian process is derived (equation A.3).

Consider a zero-mean, N -dimensional, *complex* Gaussian vector

$$\mathbf{x} = [x_{I1} + jx_{Q1}, \dots, x_{IN} + jx_{QN}]'$$

with the joint distribution given by

$$p(\mathbf{x}|\Sigma) = \frac{1}{(2\pi)^N |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}' \Sigma^{-1} \mathbf{x}\right) \quad (\text{A.1})$$

where the covariance matrix

$$\Sigma_{ij} = \begin{cases} 2\sigma^2 & \text{if } i = j \\ R_{\mathbf{x}}(i - j) & \text{otherwise} \end{cases}$$

and where the auto-covariance of \mathbf{x} is defined as

$$\begin{aligned} R_{\mathbf{x}}(k) &= E[(x_{Im} + jx_{Qm})(x_{Im+k} - jx_{Qm+k})] \\ &= R_{x_I x_I}(k) + R_{x_Q x_Q}(k) + j(R_{x_Q x_I}(k) - R_{x_I x_Q}(k)) \end{aligned}$$

With \mathbf{x} being wide-sense stationary narrowband process, and with

$$\begin{aligned} R_{x_I x_I}(k) &= R_{x_Q x_Q}(k) \\ R_{x_Q x_I}(k) &= -R_{x_I x_Q}(k) \end{aligned}$$

the normalised autocorrelation function for \mathbf{x} is defined by

$$\begin{aligned}\rho_{\mathbf{x}}(k) &= \frac{R_{\mathbf{x}}(k)}{2\sigma^2} \\ &= \rho_{x_I x_I}(k) - j\rho_{x_I x_Q}(k) \\ \rho_{x_I x_I}(k) &= \frac{R_{x_I x_I}(k)}{\sigma^2} \\ \rho_{x_I x_Q}(k) &= \frac{R_{x_I x_Q}(k)}{\sigma^2}\end{aligned}$$

Upon marginalising the distribution in equation A.1 over the remaining $N - 2$ complex variables, followed by a change of variables to polar coordinates, *i.e.*

$$\begin{aligned}r_k &= \sqrt{x_{Ik}^2 + x_{Qk}^2} \\ \phi_k &= \tan^{-1} \frac{x_{Qk}}{x_{Ik}}\end{aligned}$$

the distribution for (r_l, ϕ_l) and (r_m, ϕ_m) can be shown to be [Li and Yu, 1989]

$$\begin{aligned}p(r_l, \phi_l, r_m, \phi_m) &= \frac{r_l r_m}{4\pi^2 |\Lambda|^{1/2}} \exp\left(-\frac{r_l^2 + r_m^2 - \Xi}{2\sigma^2 \Delta}\right) \\ \Xi &= 2\rho_{x_I x_I} r_l r_m \cos(\phi_m - \phi_l) + 2\rho_{x_I x_Q} r_l r_m \sin(\phi_m - \phi_l) \\ |\Lambda|^{1/2} &= \sigma^4 \Delta \\ \Delta &= |1 - \rho_{x_I x_I}^2 - \rho_{x_Q x_Q}^2|\end{aligned}$$

The autocorrelation for the Rayleigh distributed amplitudes can now be obtained from

$$\begin{aligned}R_r(k) &= \int \int \int \int r_l r_m p(r_l, r_m, \phi_l, \phi_m) dr_l dr_m d\phi_l d\phi_m \\ &= \int_0^\infty \int_0^\infty \frac{r_l^2 r_m^2}{4\pi^2 |\Lambda|^{1/2}} \exp\left(-\frac{r_l^2 + r_m^2}{2\sigma^2 \Delta}\right) I(r_l, r_m) dr_l dr_m\end{aligned}\quad (\text{A.2})$$

where $m = l + k$, and also

$$\begin{aligned}I(r_l, r_m) &= \int_0^{2\pi} \int_0^{2\pi} \exp\left(\frac{\Xi}{2\sigma^2 \Delta}\right) d\phi_l d\phi_m \\ &= \int_0^{2\pi} \int_0^{2\pi} \exp(A \cos(\phi_m - \phi_l - \beta)) d\phi_l d\phi_m \\ &= 4\pi^2 I_0(A)\end{aligned}$$

where $I_0(\cdot)$ is a modified Bessel's function [Abramowitz and Stegun, 1970], while

$$\begin{aligned}A &= \frac{r_l r_m}{\sigma^2 \Delta} \sqrt{\rho_{x_I x_I}^2 + \rho_{x_I x_Q}^2} \\ \beta &= \tan^{-1} \frac{\rho_{x_I x_Q}}{\rho_{x_I x_I}}\end{aligned}$$

By using the following identities [Abramowitz and Stegun, 1970]

$$\begin{aligned}
 I_0(z) &= \sum_{k=0}^{\infty} \frac{(z^2/4)^k}{k! \Gamma(k+1)} \\
 \int_0^{\infty} x^b \exp(-ax^2) dx &= \frac{1}{2a^{(b+1)/2}} \Gamma((b+1)/2) \\
 {}_2F_1(a, b, c, z) &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b+k)z^k}{\Gamma(c+k)k!}
 \end{aligned}$$

where ${}_2F_1(\cdot)$ is the Gaussian hypergeometric function, and by substituting for $I(r_l, r_m)$ in equation A.2, it can be shown that the k^{th} lag autocorrelation of the Rayleigh amplitudes of the complex Gaussian process is given by

$$R_r(k) = \frac{\pi}{2} \sigma^2 \Delta^2 {}_2F_1\left(3/2, 3/2, 1, \rho_{x_I x_I}^2 + \rho_{x_I x_Q}^2\right)$$

Finally, noting that the mean and the variance of a Rayleigh process is given by

$$\begin{aligned}
 E[r] &= \sqrt{\pi/2} \sigma \\
 \text{var}[r] &= \frac{\sigma^2}{2} (4 - \pi)
 \end{aligned}$$

and using the following identity [Abramowitz and Stegun, 1970]

$$F(a, b, c, z) = (1-z)^{c-a-b} F(c-a, c-b, c, z)$$

it can be shown that the normalised covariance of the Rayleigh distributed amplitudes

$$\rho_r(k) = \frac{R_r(k) - E[r]^2}{\text{var}[r]}$$

is related to the complex autocorrelation of the complex Gaussian process through the following identity (*c.f.* the result in [Szajnowski, 1977])

$$\rho_r(k) = \frac{\pi}{4 - \pi} \left({}_2F_1(-1/2, -1/2, 1, \rho_{x_I x_I}^2 + \rho_{x_I x_Q}^2) - 1 \right) \quad (\text{A.3})$$

The relationship between $\rho_r(k)$ and $\rho_{x_I x_I}^2(k) + \rho_{x_I x_Q}^2(k)$ in equation A.3 is valid only over the range $|\rho_{x_I x_I}^2 + \rho_{x_I x_Q}^2| < 1$. Fortunately, as [Li and Yu, 1989] point out, it is easy to prove that the covariance sequence of the narrowband process generally meets this requirement. There are also other constraints on $\{\rho_r(k)\}$, the most important one being the non-negative definite property.

The mapping from $\rho_{x_I x_I}^2(k) + \rho_{x_I x_Q}^2(k)$ to $\rho_r(k)$ (equation A.3) is shown graphically in figure A.1. It can be seen that the relationship is almost linear, and certainly can be assumed to

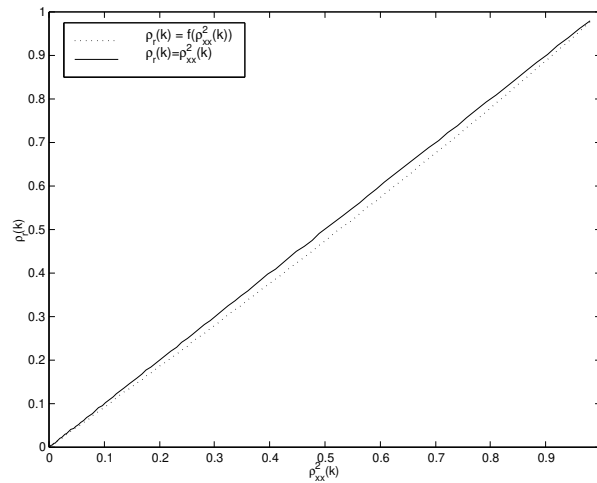


Figure A.1: Relationship given by equation A.3 between the autocorrelation of a complex Gaussian process $\rho_{x_I x_I}(k)^2 + \rho_{x_I x_Q}(k)^2$, and the covariance of the corresponding Rayleigh distributed amplitudes $\rho_r(k)$.

be so over the range of interest ($0 \leq \rho_r(k) \leq 1$) to within the sample estimation errors of the correlations. Therefore, for all practical cases of interest, the normalised covariance of the Rayleigh distributed amplitudes is related to the normalised autocorrelation of the corresponding complex Gaussian process by

$$\rho_r(k) \approx \rho_{x_I x_I}(k)^2 + \rho_{x_I x_Q}(k)^2 \quad |\rho_r(k)| \leq 1$$

Review of Markov Chain Theory

B

Consider a sequence of E -valued random samples $\{\mathbf{x}_t; t = 0, 1, 2, \dots\}$ on state space E . The elements of E can be thought of as the possible states of a system, with \mathbf{x}_t representing the state at time t . Such sequence constitutes a Markov chain if

$$p(\mathbf{x}_{t+1}|\mathbf{x}_t, \dots, \mathbf{x}_0) = p(\mathbf{x}_{t+1}|\mathbf{x}_t)$$

This is often referred to as the Markov condition, and its significance is that any future sample depends solely on the present.

A Markov chain is defined by two components: the initial distribution $p(\mathbf{x}_0)$ and the transition kernel $T(\mathbf{x}'|\mathbf{x})$. It is the transition kernel that largely determines the *dynamics* of the Markov chain. The transition kernel specifies the probability of the next state being \mathbf{x}' given that the current state is \mathbf{x} . For efficient sampling, transition kernels which produce as uncorrelated samples as possible, with good mixing properties need to be sought.

The transition kernel needs to satisfy a number of conditions in order to ensure that the resulting Markov chain has the desired invariant distribution, which is unique and reachable from any starting point of the chain. To this end, consider the following definitions:

Definition B.1 Invariant Distribution. *A distribution π is an invariant or stationary distribution of a Markov chain, if the transition kernel T is such that*

$$\pi(\mathbf{x}') = \int_{\mathbf{x}} \pi(\mathbf{x})T(\mathbf{x}'|\mathbf{x})d\mathbf{x}$$

In other words, once a Markov chain has reached the stationary distribution, all subsequent samples will also be from that distribution. It is important to note that the invariant distribution need not be unique, and that the Markov chain is not guaranteed to be able to reach it from every starting point.

Definition B.2 Irreducibility. A Markov chain is irreducible if $\forall \mathbf{x}' \in E, \forall \mathbf{x}$ with $\pi(\mathbf{x}) > 0, \exists t > 0$ such that $T^t(\mathbf{x}'|\mathbf{x}) > 0$.

where $T^t(\mathbf{x}'|\mathbf{x})$ denotes the transition kernel corresponding to t iterations of the Markov chain, i.e. $T^t(\mathbf{x}'|\mathbf{x}) = \int_{\mathbf{z}} T(\mathbf{x}'|\mathbf{z})T^{t-1}(\mathbf{z}|\mathbf{x})d\mathbf{z}$. In other words, a Markov chain is irreducible, if it is possible to eventually get to any other state, from any state to which π assigns positive probability measure.

Definition B.3 Recurrence. An irreducible Markov chain is recurrent if for any set $A \subset E$ with $\pi(A) > 0$ the conditions

1. $P\{\mathbf{x} \in A \text{ infinitely often}\} > 0$ for all \mathbf{x}
2. $P\{\mathbf{x} \in A \text{ infinitely often}\} = 1$ for π -almost all \mathbf{x}

are both satisfied. An irreducible recurrent chain is positive recurrent if it has an invariant probability distribution. Otherwise it is null recurrent.

Definition B.4 Harris Recurrence. An irreducible Markov chain is Harris recurrent if for any $A \subset E$ with $\pi(A) > 0$ the following condition

$$P\{\mathbf{x} \in A \text{ infinitely often}\} = 1 \quad \forall \mathbf{x} \in E$$

is satisfied.

The concept of recurrence is crucial in the study of convergence of Markov Chains. While irreducibility guarantees that any interesting state can be reached, a recurrent Markov chain is guaranteed to revisit any state infinitely often, at least from almost all starting points. Harris recurrence ensures that each state is revisited infinitely often, from *any* starting point.

Definition B.5 Aperiodicity. An irreducible Markov chain is aperiodic if for some (and hence for all) $\mathbf{x} \in E, \gcd\{t > 0 : T^t(\mathbf{x}|\mathbf{x})\} = 1$.

Bringing all the definitions together now, it can be seen that any irreducible, positive recurrent, aperiodic Markov chain has the correct invariant distribution, and can reach any state infinitely often from almost any starting point. Furthermore, the invariant distribution of an aperiodic positive recurrent Markov chain is also the unique limiting distribution of successive samples drawn from the chain. This is true regardless of the choice for the initial state distribution $p_0(\mathbf{x})$. This is more concisely stated in the following theorem

Theorem B.1 *If a Markov chain is aperiodic and positive recurrent, then its invariant distribution $\pi(\cdot)$ is the unique probability distribution satisfying $\pi(\mathbf{x}') = \int_{\mathbf{x}} \pi(\mathbf{x})T_t(\mathbf{x}'|\mathbf{x})d\mathbf{x} \forall \mathbf{x}' \in E$ and $t \geq 0$. The chain is then said to be ergodic and*

$$\lim_{t \rightarrow \infty} p_t(\mathbf{x}) = \pi(\mathbf{x}) \quad \forall \mathbf{x} \in E$$

Most Markov chains used in practice satisfy a more restrictive condition of being *time reversible*. In particular

Definition B.6 Reversibility. *A Markov chain is time reversible with respect to distribution π if and only if $\forall B, C \in E$, the detailed balance condition*

$$\int_B \int_C \pi(\mathbf{x})T(\mathbf{x}'|\mathbf{x})d\mathbf{x}'d\mathbf{x} = \int_C \int_B \pi(\mathbf{x}')T(\mathbf{x}|\mathbf{x}')d\mathbf{x}d\mathbf{x}'$$

is satisfied. Equivalently, detailed balance holds if and only if

$$\pi(d\mathbf{x})T(\mathbf{x}, d\mathbf{x}') = \pi(d\mathbf{x}')T(\mathbf{x}', d\mathbf{x}) \quad \forall \mathbf{x}, \mathbf{x}' \in E$$

Detailed balance is a sufficient (but *not* a necessary) condition for having an invariant distribution.

ML Parameter Estimation for AR-CH Models

C

This appendix deals with parameter estimation for AR-CH models. In particular, section C.1 deals with parameter estimation for the different kinds of uni-directional (*i.e.* Markov) AR-CH models, while section C.2 deals with parameter estimation for symmetric AR-CH models.

C.1 AR-CH MODELS

The autoregressive models driven by conditionally heteroscedastic process (AR-CH models) have the form

$$\begin{aligned}y_t &= \beta_0 + \sum_{j=1}^q \beta_j y_{t-j} + e_t \\e_t &= h_t^{1/2} \epsilon_t \quad \epsilon_t \sim N(0, 1)\end{aligned}$$

with the form of the conditional variance h_t specified in section 5.2.1.

The likelihood for y_t is conditionally Gaussian, and hence it can be shown that the joint log likelihood for T samples, to within an additive constant, is given by

$$\begin{aligned}L(\Theta) &= \frac{1}{T} \sum_{t=1}^T l_t(\Theta) \\l_t(\Theta) &= -\frac{1}{2} \ln h_t - \frac{1}{2} e_t^2 / h_t\end{aligned}$$

where $\Theta = \{\alpha, \beta\}$ is the set of the model parameters, and

$$e_t = y_t - \mathbf{y}'_{t-1} \beta$$

with $\mathbf{y}_{t-1} = [1, y_{t-1}, \dots, y_{t-p}]'$ and $\beta = [\beta_0, \dots, \beta_p]'$. To estimate the unknown parameters α and β , this likelihood function can be maximised. According to [Weiss, 1984b], the resulting parameter estimates will be consistent and asymptotically normal.

It can be shown that the first derivative of the log likelihood for the autoregressive parameters β is given by

$$\frac{\partial l_t}{\partial \beta} = \frac{e_t}{h_t} \mathbf{y}_{t-1} + \frac{1}{2} \left[\frac{e_t^2}{h_t} - 1 \right] \frac{\partial \ln h_t}{\partial \beta}$$

with the corresponding Hessian of the form

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \beta \partial \beta'} &= \frac{1}{h_t} \mathbf{y}_{t-1} \frac{\partial e_t}{\partial \beta'} - \frac{1}{2} \frac{e_t^2}{h_t} \frac{\partial \ln h_t}{\partial \beta} \frac{\partial \ln h_t}{\partial \beta'} + \\ &\frac{e_t}{h_t} \left(\mathbf{y}_{t-1} \frac{\partial \ln h_t}{\partial \beta'} - \frac{\partial \ln h_t}{\partial \beta} \frac{\partial e_t}{\partial \beta'} \right) + \frac{1}{2} \left(\frac{e_t^2}{h_t} - 1 \right) \frac{\partial^2 \ln h_t}{\partial \beta \partial \beta'} \end{aligned}$$

The conditional expectation of the last two terms, given I_{t-1} , is zero. Hence, the information matrix, which is simply the negative expectation of the Hessian averaged over all observations becomes

$$\begin{aligned} \mathcal{I}_{\beta\beta} &= -\frac{1}{T} \sum_t E \left[E \left(\frac{\partial^2 l_t}{\partial \beta \partial \beta'} \middle| \psi_{t-1} \right) \right] \\ &= \frac{1}{T} \sum_t E \left[\frac{\mathbf{y}_{t-1} \mathbf{y}_{t-1}'}{h_t} + \frac{1}{2} \frac{\partial \ln h_t}{\partial \beta} \frac{\partial \ln h_t}{\partial \beta'} \right] \end{aligned}$$

Similarly, for the conditional heteroscedastic parameters α , the first derivative of the log likelihood is given by

$$\frac{\partial l_t}{\partial \alpha} = \frac{1}{2} \left(\frac{e_t^2}{h_t} - 1 \right) \frac{\partial \ln h_t}{\partial \alpha}$$

with the associated information matrix for the CH parameters α

$$\mathcal{I}_{\alpha\alpha} = \frac{1}{2T} \sum_t E \left[\frac{\partial \ln h_t}{\partial \alpha} \frac{\partial \ln h_t}{\partial \alpha'} \right]$$

Finally, it can be shown that the off-diagonal elements of the information matrix are given by

$$\mathcal{I}_{\alpha\beta} = \frac{1}{2T} \sum_t E \left[\frac{\partial \ln h_t}{\partial \alpha} \frac{\partial \ln h_t}{\partial \beta'} \right]$$

These elements are exactly zero for all of the models of interests. In particular, arguments based on symmetry and regularity of h_t [Engle, 1982] can be used to prove that this is true for AR-ARCH and AR-EARCH models, while $\frac{\partial \ln h_t}{\partial \beta} = 0$ for AR-ECH and AR-ELCH models. This result has profound influence on parameter estimation algorithm, in that the estimation of α and β can be undertaken separately without asymptotic loss of efficiency.

The parameter estimation procedure recommended by [Engle, 1982] is to initially estimate β by ordinary least squares (*i.e.* assume no heteroscedasticity present). Since the off-diagonal elements of the information matrix are exactly zero, efficient estimates of α and β can then be calculated by iterating between α and β using the scoring algorithm, given by

$$\phi^{i+1} = \phi^i + [\hat{\mathcal{I}}_{\phi\phi}^i]^{-1} \frac{\partial L^i}{\partial \phi}$$

where $\frac{\partial L^i}{\partial \phi}$ and $\hat{\mathcal{I}}_{\phi\phi}^i$ are the score and the information matrix at the i^{th} iteration. The forms of the score vector and information matrix for the different models are now derived, while the actual parameter estimation algorithm is discussed in section C.1.3.

C.1.1 Estimation of α

The derivative of the log likelihood with respect to the CH parameter vector α can be shown to be given by

$$\frac{\partial l_t}{\partial \alpha} = \frac{1}{2} \left(\frac{e_t^2}{h_t} - 1 \right) \mathbf{z}_t$$

where the form of the vector \mathbf{z}_t depends on the form of the conditional variance. In particular

$$\begin{aligned} \text{AR-ARCH} \quad \mathbf{z}'_t &= [1, e_{t-1}^2, \dots, e_{t-p}^2, (y_{t-1} - \bar{y})^2, \dots, (y_{t-r} - \bar{y})^2] / h_t \\ \text{AR-ECH} \quad \mathbf{z}'_t &= [1, y_{t-1}, \dots, y_{t-p}] \\ \text{AR-ELCH} \quad \mathbf{z}'_t &= [1, \ln(y_{t-1} - \bar{y})^2, \dots, \ln(y_{t-p} - \bar{y})^2] \\ \text{AR-EARCH} \quad \mathbf{z}'_t &= [1, \ln e_{t-1}^2, \dots, \ln e_{t-p}^2] \end{aligned}$$

Similarly, the information matrix for the CH parameters can be shown to be

$$\mathcal{I}_{\alpha\alpha} = \frac{1}{2T} \sum_t \mathbf{z}_t \mathbf{z}'_t$$

Therefore, using the above expressions for the score vector and the information matrix, and by defining

$$\begin{aligned} \mathbf{Z} &= [\mathbf{z}_1, \dots, \mathbf{z}_T] \\ f_t &= \frac{e_t^2}{h_t} - 1 \\ f &= [f_1, \dots, f_T]' \end{aligned}$$

it can be shown that the i^{th} iteration of the scoring algorithm for the CH parameters α is given by

$$\alpha^{i+1} = \alpha^i + (\mathbf{Z}\mathbf{Z}')^{-1}\mathbf{Z}\mathbf{f} \quad (\text{C.1})$$

C.1.2 Estimation of β

Before obtaining expressions for the score vector and the associated information matrix for the AR parameter vector β , a few preliminary results need to be established. In particular, for AR-ECH and AR-ELCH models

$$\frac{\partial \ln h_t}{\partial \beta} = 0$$

For the AR-EARCH model, on the other hand, it can be shown that

$$\frac{\partial \ln h_t}{\partial \beta} = -2 \sum_{i=1}^p \frac{\alpha_i}{e_{t-i}} \mathbf{y}_{t-i-1}$$

while for the case of the AR-ARCH model refer to [Engle, 1982] for the appropriate expressions.

Using the above results for $\frac{\partial \ln h_t}{\partial \beta}$, it can be shown that (in the limit of large T) the score vector and the information matrix are given by

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= \frac{1}{T} \sum_t \mathbf{y}_{t-1} s_t \\ \mathcal{I}_{\beta\beta} &= \frac{1}{T} \sum_t \mathbf{y}_{t-1} \mathbf{y}'_{t-1} r_t \end{aligned}$$

where r_t and s_t are of the following form

$$\begin{aligned} \text{AR-ARCH} \quad r_t &= \left[\frac{1}{h_t} + 2e_t^2 \sum_{i=1}^p \alpha_i^2 \frac{1}{h_{t+i}^2} \right] \\ s_t &= e_t \left[\frac{1}{h_t} - \sum_{i=1}^p \alpha_i \frac{1}{h_{t+i}} \left(\frac{e_{t+i}^2}{h_{t+i}} - 1 \right) \right] \\ \text{AR-ECH \& AR-ELCH} \quad r_t &= 1/h_t \\ s_t &= e_t/h_t \\ \text{AR-EARCH} \quad r_t &= \frac{1}{e_t^2} \left[\frac{e_t^2}{h_t} + 2 \sum_{i=1}^p \alpha_i^2 \right] \\ s_t &= \frac{1}{e_t} \left[\frac{e_t^2}{h_t} - \sum_{i=1}^p \alpha_i \left(\frac{e_{t+i}^2}{h_{t+i}} - 1 \right) \right] \end{aligned}$$

Therefore, using the above expressions for the score vector and the information matrix, and by defining

$$\begin{aligned}\mathbf{R} &= \text{diag}[r_1, \dots, r_T] \\ \mathbf{s} &= [s_1, \dots, s_T]' \\ \mathbf{y}_{t-1} &= [1, y_{t-1}, \dots, y_{t-p}]' \\ \mathbf{Y} &= [\mathbf{y}_0, \dots, \mathbf{y}_{T-1}]\end{aligned}$$

it can be shown that the i^{th} iteration of the scoring algorithm for the AR parameters β is given by

$$\beta^{i+1} = \beta^i + (\mathbf{YRY}')^{-1} \mathbf{Ys} \quad (\text{C.2})$$

C.1.3 Parameter Estimation Algorithm

Following from the discussion in the preceding sections, the parameter estimation algorithm for Markov AR-CH models can be summarised thus. Given a data vector \mathbf{y}

1. Obtain an initial estimate of β using ordinary least squares

$$\beta^0 = (\mathbf{Y}\mathbf{Y}')^{-1} \mathbf{Y}\mathbf{y}$$

2. Based on the initial estimate of β^0 , calculate the prediction errors $\mathbf{e} = \mathbf{y} - \mathbf{Y}'\beta^0$
3. Set the initial estimate $\alpha_i^0 = 0 \ \forall i \geq 1$ and set α_0^0 such that $h_t = \text{var}(\mathbf{e})$
4. Starting with the initial estimates α^0 and β^0 , iterate the scoring algorithm for parameters α and β in equations C.1 and C.2 until convergence, *i.e.*

$$\begin{aligned}\alpha^{i+1} &= \alpha^i + (\mathbf{Z}\mathbf{Z}')^{-1} \mathbf{Z}\mathbf{f} \\ \beta^{i+1} &= \beta^i + (\mathbf{YRY}')^{-1} \mathbf{Ys}\end{aligned}$$

where matrices \mathbf{Z} and \mathbf{R} , and vectors \mathbf{f} and \mathbf{s} are re-evaluated at every iteration

The convergence of the scoring algorithm can be formulated in many ways. [Engle, 1982] recommend a simple criterion based on the gradient around the inverse Hessian. Thus, for a parameter vector ϕ the convergence criterion is given by

$$\Theta = \frac{\partial L'}{\partial \phi} \left(\frac{\partial^2 L'}{\partial \phi \partial \phi'} \right)^{-1} \frac{\partial L}{\partial \phi}$$

C.2 SYMMETRIC AR-CH MODELS

Symmetric autoregressive models with conditionally heteroscedastic errors (SAR-CH models) have the form

$$\begin{aligned} y_n &= \beta_0 + \sum_{j=-q, j \neq 0}^q \beta_j y_{n+j} + e_n \\ e_n &= h_n^{1/2} \epsilon_n \quad \epsilon_n \sim N(0, 1) \end{aligned}$$

with the form of the conditional variance is given by

$$\begin{aligned} \text{SAR-CH} \quad \ln h_n &= \alpha_0 + \sum_{i=-p, i \neq 0}^p \alpha_i y_{n+i} \\ \text{SAR-ELCH} \quad \ln h_n &= \alpha_0 + \sum_{i=-p, i \neq 0}^p \alpha_i \ln(y_{n+i} - \bar{y})^2 \end{aligned}$$

Since ϵ_n are i.i.d. Normal, the likelihood for ϵ is given by $p(\epsilon) = \prod_n p(\epsilon_n)$. The likelihood for the observations \mathbf{y} can be obtained from the likelihood for ϵ by change of variables, namely

$$\begin{aligned} p(\mathbf{y}|\alpha, \beta) &= p(\epsilon)|\mathbf{J}| \\ \mathbf{J} &= \left| \frac{\partial \epsilon}{\partial \mathbf{y}} \right| \end{aligned}$$

where \mathbf{J} is the Jacobian associated with the change of variables. In the case of Markov AR-CH models of section C.1, the Jacobian is upper triangular with ones on the main diagonal, and hence the determinant of the Jacobian reduces to unity. Unfortunately in the case of symmetric AR-CH models, the transformation from ϵ to \mathbf{y} is two sided, and hence the determinant of the Jacobian will have a number of terms non-linearly dependent on α and β , which make Maximum Likelihood parameter estimation for SAR-CH models very laborious.

Denoting by $L = \ln p_{\mathbf{y}}(\mathbf{y}|\alpha, \beta)$, $L_J = \ln |\mathbf{J}|$ and $L_P = \sum_n \ln p_{\epsilon}(y_n|\mathbf{y}_{-n}, \alpha, \beta)$, it is readily seen that the log likelihood L consists of two parts, *i.e.*

$$L = L_P + L_J$$

where L_P is the pseudo log likelihood (*i.e.* log likelihood akin to that for a Markov process), and L_J is the Jacobian contribution to the likelihood. The SAR-CH model parameters α and β can therefore be estimated using a gradient based optimisation algorithm, using the gradients of the log likelihood L derived in sections C.2.2 and C.2.3 below.

C.2.1 The Jacobian

The $(n, n+k)^{th}$ element of the Jacobian matrix is given by

$$\begin{aligned} \mathbf{J}_{n,n+k} &= \frac{\partial \epsilon_n}{\partial y_{n+k}} \\ &= \begin{cases} h_n^{-1/2} & \text{if } k = 0 \\ -h_n^{-1/2} (\beta_k + \frac{1}{2} \alpha_k e_n z_{n+k}) & \text{otherwise} \end{cases} \end{aligned}$$

where z_{n+k} is given by

$$\begin{aligned} \text{SAR-ECH} \quad z_{n+k} &= 1 \\ \text{SAR-ELCH} \quad z_{n+k} &= \frac{2}{y_{n+k} - \bar{y}} \end{aligned}$$

It can further be shown that

$$\frac{\partial \ln |\mathbf{J}|}{\partial \phi} = \sum_n \sum_k \mathbf{J}_{n+k,n}^{-1} \frac{\partial \mathbf{J}_{n,n+k}}{\partial \phi}$$

and since $\frac{\partial \mathbf{J}_{n,n}}{\partial \beta} = 0$ for SAR-ECH and SAR-ELCH models, it follows that

$$\begin{aligned} \frac{\partial \ln |\mathbf{J}|}{\partial \alpha_l} &= \begin{cases} \frac{1}{2} \sum_n h_n^{-1/2} \Delta_{\alpha n} & \text{if } l = 0 \\ \frac{1}{2} \sum_n h_n^{-1/2} \left[\Delta_{\alpha n} x_{n+l} - \mathbf{J}_{n+l,n}^{-1} e_n z_{n+l} \right] & \text{otherwise} \end{cases} \\ \Delta_{\alpha n} &= -\mathbf{J}_{n,n}^{-1} + \sum_{k \neq 0} \mathbf{J}_{n+k,n}^{-1} (\beta_k + \frac{1}{2} \alpha_k e_n z_{n+k}) \end{aligned}$$

$$\begin{aligned} \frac{\partial \ln |\mathbf{J}|}{\partial \beta_l} &= \begin{cases} \frac{1}{2} \sum_n h_n^{-1/2} \Delta_{\beta n} & \text{if } l = 0 \\ \frac{1}{2} \sum_n h_n^{-1/2} \left[\Delta_{\beta n} y_{n+l} - 2 \mathbf{J}_{n+l,n}^{-1} \right] & \text{otherwise} \end{cases} \\ \Delta_{\beta n} &= \sum_{k \neq 0} \alpha_k \mathbf{J}_{n+k,n}^{-1} z_{n+k} \end{aligned}$$

where $x_n = y_n$ for SAR-ECH and $x_n = \ln(y_n - \bar{y})^2$ for SAR-ELCH models.

C.2.2 Estimation of α

The derivatives of L_P with respect to α can be shown to be

$$\begin{aligned} \frac{\partial L_P}{\partial \alpha} &= \frac{1}{2N} \sum_n \frac{e_n^2}{h_n} \mathbf{z}_n \\ \frac{\partial^2 L_P}{\partial \alpha \partial \alpha'} &= -\frac{1}{2N} \sum_n \frac{e_n^2}{h_n} \mathbf{z}_n \mathbf{z}_n' \end{aligned}$$

where the form of the vector \mathbf{z}_n depends on the form of the conditional variance, *i.e.*

$$\begin{aligned} \text{SAR-ECH} \quad \mathbf{z}_n &= [y_{n-p}, \dots, y_{n-1}, 1, y_{n+1}, \dots, y_{n+p}]' \\ \text{SAR-ELCH} \quad \mathbf{z}_n &= [\ln(y_{n-p} - \bar{y})^2, \dots, \ln(y_{n-1} - \bar{y})^2, 1, \\ &\quad \ln(y_{n+1} - \bar{y})^2, \dots, \ln(y_{n+p} - \bar{y})^2]' \end{aligned}$$

It therefore follows that the derivatives of the log likelihood L are given by

$$\begin{aligned} \frac{\partial L}{\partial \alpha} &= \frac{1}{2N} \sum_n \frac{e_n^2}{h_n} \mathbf{z}_n + \frac{\partial \ln |\mathbf{J}|}{\partial \alpha} \\ \mathcal{I}_{\alpha\alpha} &= \frac{1}{2N} \sum_n \mathbf{z}_n \mathbf{z}_n' - E \left[\frac{\partial^2 \ln |\mathbf{J}|}{\partial \alpha \partial \alpha'} \right] \end{aligned}$$

where $\frac{\partial \ln |\mathbf{J}|}{\partial \alpha}$ is given in section C.2.1. Although the first derivative is fairly straight forward to determine, the contributions to the information matrix resulting from the Jacobian are not as easily obtained (though it is possible). Instead, a simple optimisation algorithm based on gradients can be employed to estimate α .

C.2.3 Estimation of β

The derivatives of L_P with respect to β can be shown to be

$$\begin{aligned} \frac{\partial L_P}{\partial \beta} &= \frac{1}{N} \sum_n \frac{e_n}{h_n} \mathbf{y}_n \\ \frac{\partial^2 L_P}{\partial \beta \partial \beta'} &= -\frac{1}{N} \sum_n \frac{1}{h_n} \mathbf{y}_n \mathbf{y}_n' \end{aligned}$$

with \mathbf{y}_n of the form

$$\mathbf{y}_n = [y_{n-q}, \dots, y_{n-1}, 1, y_{n+1}, \dots, y_{n+q}]'$$

It therefore follows that the derivatives of the log likelihood L are given by

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= \frac{1}{N} \sum_n \frac{e_n}{h_n} \mathbf{y}_n + \frac{\partial \ln |\mathbf{J}|}{\partial \beta} \\ \mathcal{I}_{\beta\beta} &= \frac{1}{N} \sum_n \frac{1}{h_n} \mathbf{y}_n \mathbf{y}_n' - E \left[\frac{\partial^2 \ln |\mathbf{J}|}{\partial \beta \partial \beta'} \right] \end{aligned}$$

where $\frac{\partial \ln |\mathbf{J}|}{\partial \beta}$ is given in section C.2.1. Although the first derivative is fairly straight forward to determine, the contributions to the information matrix resulting from the Jacobian are not as easily obtained (though it is possible). Instead, a simple optimisation algorithm based on gradients can be employed to estimate β .

Detection Threshold Setting

D

D.1 INTRODUCTION

The probability of false alarm for a detector with the single pulse detection threshold level set to T is given by

$$P_{fa} = \int_T^\infty p(r) dr$$

where $p(r)$ is the probability distribution of the clutter envelope in the current range cell under test.

Consider a compound model with Rayleigh distributed speckle, and the modulating component v distributed as $p(v^2)$, *i.e.*

$$\begin{aligned} p(r|v^2) &= \frac{r}{v^2} \exp\left(-\frac{1}{2} \frac{r^2}{v^2}\right) \\ p(r) &= \int_0^\infty p(r|v^2) p(v^2) dv^2 \end{aligned}$$

For such models the probability of false alarm simplifies to

$$P_{fa} = \int_0^\infty \exp\left(-\frac{1}{2} \frac{T^2}{v^2}\right) p(v^2) dv^2 \quad (\text{D.1})$$

Unfortunately, for most distributions of v^2 of interest, either the above integral does not have a simple analytical form, or the relationship between P_{fa} and the threshold T is highly non-linear (*e.g.* for K-distributions).

D.2 APPROXIMATE RELATIONSHIP OF T AND P_{fa}

Throughout this thesis it has been found that the useful forms for $p(v^2)$ include the two-parameter Gamma, Inverse Gamma and the log-Normal distributions. It is of immense practical utility to find a relationship between the required threshold level and the distribution parameters, such that the required probability of false alarm is achieved.

As the simple analytical relationship between the threshold and the probability of false alarm does not always exist, simulation methods are here employed, and approximate functional relationships are proposed.

D.2.1 The Inverse Gamma(α, β) Distribution

The Inverse Gamma distribution for v^2 is one of the very few distributions for which a simple analytical result exists. In particular, if

$$p(v^2|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} (v^2)^{-(\alpha+1)} \exp(-\beta/v^2)$$

then it can be shown that

$$P_{fa} = \left(\frac{\beta}{T^2/2 + \beta} \right)^\alpha$$

and hence

$$T^2 = 2\beta \left(P_{fa}^{-1/\alpha} - 1 \right)$$

D.2.2 The Gamma(α, β) Distribution

In the case when v^2 is Gamma distributed, *i.e.*

$$p(v^2|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} (v^2)^{\alpha-1} \exp(-\beta v^2)$$

the probability of false alarm is given by

$$P_{fa} = \frac{2}{\Gamma(\alpha)} \left(\frac{1}{2} \beta T^2 \right)^{\alpha/2} K_\alpha \left(\sqrt{2\beta T^2} \right)$$

where $K_\alpha(\cdot)$ is the Modified Bessel function [Abramowitz and Stegun, 1970].

Clearly, it is not possible to analytically invert the above equation to obtain an estimate of the threshold T corresponding to the required probability of false alarm P_{fa} and parameters α and β .

Instead, an approximate method for estimating the threshold is proposed in this section. In a typical radar system, the required probability of false alarm P_{fa} is specified as a design parameter of the system. It is typically specified to be 10^i , where i is a non-positive integer. The

problem therefore reduces to estimating T as a function of distribution parameters α and β for a given value of P_{fa} .

To further reduce the degrees of freedom, one of the distribution parameters, namely β is concentrated out of the above equation by redefining the threshold as $\tau = \sqrt{2\beta T^2}$, from which it follows that in terms of the new threshold

$$P_{fa} = \frac{2}{\Gamma(\alpha)} (\tau/2)^\alpha K_\alpha(\tau) \quad (\text{D.2})$$

This equation can be solved iteratively to obtain a plot of $\log_{10} \tau$ as a function of $\log_{10} \alpha$. It can be seen from figure D.1(a) that for $10^{-3} \leq \alpha \leq 10^2$, and for $\log_{10} P_{fa} \in \{-12, \dots, -1\}$, the relationship between $\log_{10} \tau$ and $\log_{10} \alpha$ is well approximated polynomially, *i.e.*

$$\log_{10} \tau = a_N (\log_{10} \alpha)^N + \dots + a_1 \log_{10} \alpha + a_0 \quad (\text{D.3})$$

where the polynomial coefficients depend on P_{fa} (see figure D.2).

Therefore, in order to estimate the detection threshold based on a Gamma distribution with parameters α and β for a given probability of false alarm, the following steps need to be followed

1. estimate τ from equation D.3 for given α and P_{fa}
2. set the threshold $T = \frac{\tau}{\sqrt{2\beta}}$

D.2.3 The log-Normal(μ, σ) Distribution

In the case when v^2 is log-Normally distributed (*i.e.* $p_y(y_n = \log v_n) = N(\mu, \sigma)$), *i.e.*

$$p(v^2 | \mu, \sigma) = \frac{1}{2v^2 \sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2} \left(\frac{1}{2} \ln v^2 - \mu\right)^2\right)$$

the integral in equation D.1 cannot be performed analytically. It is therefore not possible to determine the detection threshold T directly.

Approximation scheme, such as the one used in the previous section for the Gamma distributed modulating component, therefore needs to be employed. As was done above to reduce the degrees of freedom, the threshold is redefined as $\tau = T^2 \exp(-2\mu)$ in order to concentrate out the parameter μ . In terms of the new threshold, the probability of false alarm is given by

$$P_{fa} = E_{y \sim N(0, \sigma)} \left(\exp\left(-\frac{\tau}{2} \exp(-2y)\right) \right) \quad (\text{D.4})$$

This expectation can be evaluated using Monte Carlo methods in order to obtain plots of $\log_{10} \tau$ as a function of σ for a given probability of false alarm.

Figure D.1(b) shows that over the range $0 < \sigma \leq 1$, the relationship between $\log_{10} \tau$ and σ is well approximated polynomially, *i.e.*

$$\log_{10} \tau = a_N \sigma^N + \dots + a_1 \sigma + a_0 \quad (\text{D.5})$$

where the polynomial coefficients depend on P_{fa} (see figure D.3).

Therefore, in order to estimate the detection threshold based on a log-Normal distribution with parameters μ and σ for a given probability of false alarm, the following steps need to be followed

1. estimate τ from equation D.5 for given σ and P_{fa}
2. set the threshold $T^2 = \tau \exp(2\mu)$

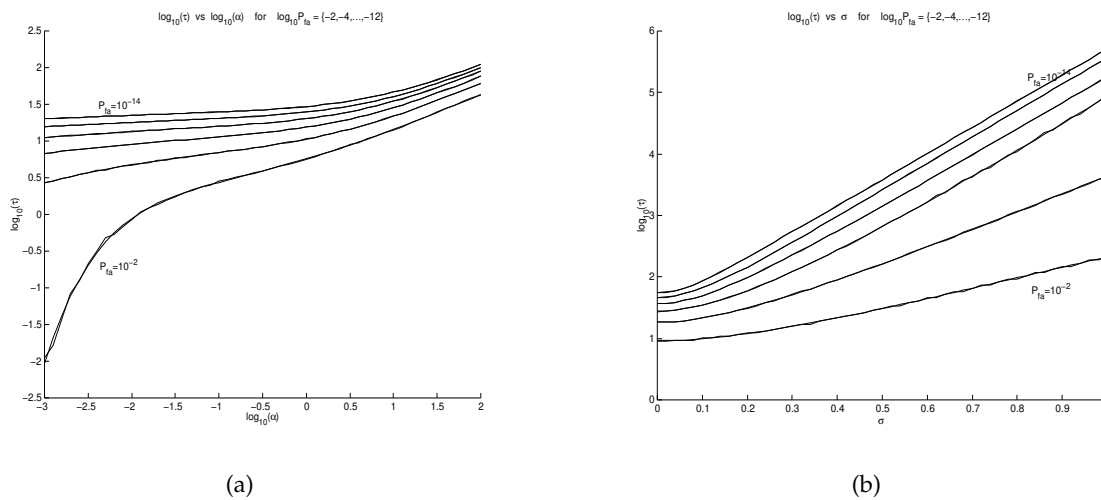


Figure D.1: (a) The true solution and 7th order polynomial approximation to the true solution to equation D.2 for $10^{-3} \leq \alpha \leq 10^2$ and (b) the true solution and 7th order polynomial approximation to the true solution to equation D.4 for $0 < \sigma \leq 1$

$$\begin{aligned}
P_{fa} = 10^{-1} \quad \log_{10}(\tau) &= 0.50858 + 0.522(\log_{10} \alpha)^1 + -0.012(\log_{10} \alpha)^2 \\
&\quad + -0.0725(\log_{10} \alpha)^3 + -0.0207(\log_{10} \alpha)^4 + 0.0673(\log_{10} \alpha)^5 \\
&\quad + -0.00915(\log_{10} \alpha)^6 + -0.00539(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-2} \quad \log_{10}(\tau) &= 0.75987 + 0.348(\log_{10} \alpha)^1 + 0.0353(\log_{10} \alpha)^2 \\
&\quad + 0.00624(\log_{10} \alpha)^3 + 0.00206(\log_{10} \alpha)^4 + 0.00395(\log_{10} \alpha)^5 \\
&\quad + -0.00229(\log_{10} \alpha)^6 + -0.000258(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-3} \quad \log_{10}(\tau) &= 0.91403 + 0.265(\log_{10} \alpha)^1 + 0.0702(\log_{10} \alpha)^2 \\
&\quad + 0.028(\log_{10} \alpha)^3 + -0.0112(\log_{10} \alpha)^4 + -0.00538(\log_{10} \alpha)^5 \\
&\quad + 0.000968(\log_{10} \alpha)^6 + 0.00053(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-4} \quad \log_{10}(\tau) &= 1.0268 + 0.234(\log_{10} \alpha)^1 + 0.0699(\log_{10} \alpha)^2 \\
&\quad + 0.016(\log_{10} \alpha)^3 + -0.00513(\log_{10} \alpha)^4 + -0.00149(\log_{10} \alpha)^5 \\
&\quad + 0.000136(\log_{10} \alpha)^6 + 6.45e - 05(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-5} \quad \log_{10}(\tau) &= 1.1151 + 0.204(\log_{10} \alpha)^1 + 0.0737(\log_{10} \alpha)^2 \\
&\quad + 0.0186(\log_{10} \alpha)^3 + -0.00586(\log_{10} \alpha)^4 + -0.00212(\log_{10} \alpha)^5 \\
&\quad + 0.00033(\log_{10} \alpha)^6 + 0.000123(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-6} \quad \log_{10}(\tau) &= 1.1893 + 0.18(\log_{10} \alpha)^1 + 0.0691(\log_{10} \alpha)^2 \\
&\quad + 0.02(\log_{10} \alpha)^3 + -0.00262(\log_{10} \alpha)^4 + -0.00195(\log_{10} \alpha)^5 \\
&\quad + -9.73e - 05(\log_{10} \alpha)^6 + 4.19e - 05(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-7} \quad \log_{10}(\tau) &= 1.2504 + 0.165(\log_{10} \alpha)^1 + 0.0695(\log_{10} \alpha)^2 \\
&\quad + 0.0196(\log_{10} \alpha)^3 + -0.0025(\log_{10} \alpha)^4 + -0.00176(\log_{10} \alpha)^5 \\
&\quad + -7.28e - 05(\log_{10} \alpha)^6 + 3.24e - 05(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-8} \quad \log_{10}(\tau) &= 1.3044 + 0.15(\log_{10} \alpha)^1 + 0.0689(\log_{10} \alpha)^2 \\
&\quad + 0.0206(\log_{10} \alpha)^3 + -0.0024(\log_{10} \alpha)^4 + -0.00193(\log_{10} \alpha)^5 \\
&\quad + -4.15e - 05(\log_{10} \alpha)^6 + 5.06e - 05(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-9} \quad \log_{10}(\tau) &= 1.3521 + 0.14(\log_{10} \alpha)^1 + 0.0669(\log_{10} \alpha)^2 \\
&\quad + 0.0215(\log_{10} \alpha)^3 + -0.00196(\log_{10} \alpha)^4 + -0.00218(\log_{10} \alpha)^5 \\
&\quad + -5.15e - 05(\log_{10} \alpha)^6 + 7.31e - 05(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-10} \quad \log_{10}(\tau) &= 1.3954 + 0.13(\log_{10} \alpha)^1 + 0.0639(\log_{10} \alpha)^2 \\
&\quad + 0.0208(\log_{10} \alpha)^3 + -0.000482(\log_{10} \alpha)^4 + -0.00168(\log_{10} \alpha)^5 \\
&\quad + -0.000241(\log_{10} \alpha)^6 + 7.04e - 06(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-11} \quad \log_{10}(\tau) &= 1.4346 + 0.122(\log_{10} \alpha)^1 + 0.0627(\log_{10} \alpha)^2 \\
&\quad + 0.0208(\log_{10} \alpha)^3 + -0.000477(\log_{10} \alpha)^4 + -0.00164(\log_{10} \alpha)^5 \\
&\quad + -0.000209(\log_{10} \alpha)^6 + 1.04e - 05(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-12} \quad \log_{10}(\tau) &= 1.4701 + 0.115(\log_{10} \alpha)^1 + 0.0612(\log_{10} \alpha)^2 \\
&\quad + 0.0205(\log_{10} \alpha)^3 + -0.00018(\log_{10} \alpha)^4 + -0.0015(\log_{10} \alpha)^5 \\
&\quad + -0.000219(\log_{10} \alpha)^6 + 2.82e - 06(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-13} \quad \log_{10}(\tau) &= 1.5033 + 0.109(\log_{10} \alpha)^1 + 0.0595(\log_{10} \alpha)^2 \\
&\quad + 0.021(\log_{10} \alpha)^3 + 4.74e - 05(\log_{10} \alpha)^4 + -0.00157(\log_{10} \alpha)^5 \\
&\quad + -0.000216(\log_{10} \alpha)^6 + 1.06e - 05(\log_{10} \alpha)^7 \\
P_{fa} = 10^{-14} \quad \log_{10}(\tau) &= 1.534 + 0.103(\log_{10} \alpha)^1 + 0.0575(\log_{10} \alpha)^2 \\
&\quad + 0.0212(\log_{10} \alpha)^3 + 0.000629(\log_{10} \alpha)^4 + -0.00153(\log_{10} \alpha)^5 \\
&\quad + -0.00028(\log_{10} \alpha)^6 + -9.91e - 07(\log_{10} \alpha)^7
\end{aligned}$$

Figure D.2: Approximate polynomial relationship between $\log_{10} \tau$ and $\log_{10} \alpha$ (where $\tau = \sqrt{2\beta T^2}$) for Gamma distributed clutter with $10^{-3} < \alpha < 10^2$

$$\begin{aligned}
P_{fa} = 10^{-1} \quad \log_{10}(\tau) &= 0.66414 + -0.338\sigma^1 + 5.02\sigma^2 \\
&\quad + -18.9\sigma^3 + 46.7\sigma^4 + -66\sigma^5 \\
&\quad + 48.2\sigma^6 + -14.1\sigma^7 \\
P_{fa} = 10^{-2} \quad \log_{10}(\tau) &= 0.96094 + -0.0818\sigma^1 + 4.61\sigma^2 \\
&\quad + -7.87\sigma^3 + 8.53\sigma^4 + -4.63\sigma^5 \\
&\quad + 0.517\sigma^6 + 0.29\sigma^7 \\
P_{fa} = 10^{-3} \quad \log_{10}(\tau) &= 1.1363 + 0.0698\sigma^1 + 4.35\sigma^2 \\
&\quad + 1.42\sigma^3 + -23.1\sigma^4 + 43.9\sigma^5 \\
&\quad + -35.5\sigma^6 + 10.7\sigma^7 \\
P_{fa} = 10^{-4} \quad \log_{10}(\tau) &= 1.266 + -0.206\sigma^1 + 11\sigma^2 \\
&\quad + -30.8\sigma^3 + 60.5\sigma^4 + -73.3\sigma^5 \\
&\quad + 48.2\sigma^6 + -13.1\sigma^7 \\
P_{fa} = 10^{-5} \quad \log_{10}(\tau) &= 1.3621 + -0.14\sigma^1 + 12.8\sigma^2 \\
&\quad + -35.2\sigma^3 + 65.4\sigma^4 + -72\sigma^5 \\
&\quad + 41.4\sigma^6 + -9.48\sigma^7 \\
P_{fa} = 10^{-6} \quad \log_{10}(\tau) &= 1.4385 + 0.0939\sigma^1 + 11.4\sigma^2 \\
&\quad + -22.8\sigma^3 + 33.2\sigma^4 + -30.6\sigma^5 \\
&\quad + 15.1\sigma^6 + -2.82\sigma^7 \\
P_{fa} = 10^{-7} \quad \log_{10}(\tau) &= 1.5083 + -0.0187\sigma^1 + 15.4\sigma^2 \\
&\quad + -41.6\sigma^3 + 74.5\sigma^4 + -80.9\sigma^5 \\
&\quad + 47.7\sigma^6 + -11.6\sigma^7 \\
P_{fa} = 10^{-8} \quad \log_{10}(\tau) &= 1.5643 + -0.0308\sigma^1 + 17.4\sigma^2 \\
&\quad + -44.2\sigma^3 + 68.8\sigma^4 + -62.7\sigma^5 \\
&\quad + 30.3\sigma^6 + -5.9\sigma^7 \\
P_{fa} = 10^{-9} \quad \log_{10}(\tau) &= 1.616 + 0.00294\sigma^1 + 19.2\sigma^2 \\
&\quad + -53\sigma^3 + 92.7\sigma^4 + -98.5\sigma^5 \\
&\quad + 57.4\sigma^6 + -14\sigma^7 \\
P_{fa} = 10^{-10} \quad \log_{10}(\tau) &= 1.6598 + 0.198\sigma^1 + 18.3\sigma^2 \\
&\quad + -45.5\sigma^3 + 68.4\sigma^4 + -62.2\sigma^5 \\
&\quad + 31.8\sigma^6 + -7.06\sigma^7 \\
P_{fa} = 10^{-11} \quad \log_{10}(\tau) &= 1.7015 + 0.112\sigma^1 + 23\sigma^2 \\
&\quad + -75.9\sigma^3 + 147\sigma^4 + -164\sigma^5 \\
&\quad + 97.2\sigma^6 + -23.8\sigma^7 \\
P_{fa} = 10^{-12} \quad \log_{10}(\tau) &= 1.7412 + -0.0598\sigma^1 + 28.1\sigma^2 \\
&\quad + -97.7\sigma^3 + 192\sigma^4 + -213\sigma^5 \\
&\quad + 125\sigma^6 + -29.8\sigma^7 \\
P_{fa} = 10^{-13} \quad \log_{10}(\tau) &= 1.7747 + 0.245\sigma^1 + 23.6\sigma^2 \\
&\quad + -93.5\sigma^3 + 206\sigma^4 + -254\sigma^5 \\
&\quad + 165\sigma^6 + -43.6\sigma^7 \\
P_{fa} = 10^{-14} \quad \log_{10}(\tau) &= 1.8035 + 0.545\sigma^1 + 22.6\sigma^2 \\
&\quad + -88.4\sigma^3 + 191\sigma^4 + -232\sigma^5 \\
&\quad + 146\sigma^6 + -37.5\sigma^7
\end{aligned}$$

Figure D.3: Approximate polynomial relationship between $\log_{10} \tau$ and σ (where $\tau = T^2 \exp(-2\mu)$) for log-Normally distributed clutter with $0 < \sigma < 1$

(Inverse) Gamma-log Normal Distribution Matching using Moments

E

E.1 INTRODUCTION

As in the case of Stochastic Volatility models [Harvey and Shephard, 1996], it is often of interest to transform a multiplicative type model of the form

$$x = h\epsilon \quad \epsilon \sim N(0, 1) \quad (\text{E.1})$$

into a non-Gaussian, but linear model. This can be achieved in the logarithm domain, resulting in the following linear process¹

$$\ln 2x^2 = \ln 2h^2 + \ln \epsilon^2 \quad \epsilon \sim N(0, 1)$$

Furthermore, in order to make such a transformation useful for filtering purposes, it proves to be beneficial to approximate the resulting error distribution $\ln \epsilon^2$, by a Gaussian distribution, so as to make the above equation for $\ln 2x^2$ and $\ln 2h^2$ linear and Gaussian. This can, in principle, be achieved in one of the two ways.

Either the posterior for $\ln 2x^2$ is approximated by a Gaussian distribution, *i.e.*

$$\begin{aligned} p(\ln 2x^2 | 2h^2) &= N(\mu, \sigma) \\ \Rightarrow \ln 2x^2 &= \mu + \sigma\nu \quad \nu \sim N(0, 1) \end{aligned}$$

where μ is a function of $\ln 2h^2$. Alternatively, the posterior distribution for $\ln 2h^2$ is approximated by a Gaussian distribution, in which case

$$\begin{aligned} p(\ln 2h^2 | 2x^2) &= N(\mu, \sigma) \\ \Rightarrow \ln 2h^2 &= \mu + \sigma\nu \quad \nu \sim N(0, 1) \end{aligned}$$

where μ is a function of $\ln 2x^2$.

¹Note that the linearisation is obtained in terms of $\ln 2h^2$, *i.e.* the logarithm of the mean power, which is equal to $2h^2$ for Rayleigh distributed clutter.

Before progressing any further, it is useful to inspect the distributions associated with the above variables. In particular, two cases need be considered.

Rayleigh Distributed ϵ

Consider Rayleigh distributed ϵ , as is the case in incoherent clutter, *i.e.*

$$p(\epsilon) = \epsilon \exp\left(-\frac{1}{2}\epsilon^2\right)$$

Using equation E.1 it can be shown that the posterior distribution for $x^2 = h^2\epsilon^2$ is Gamma distributed, *i.e.*

$$\begin{aligned} p(x^2|2h^2) &= \frac{1}{h^2} \exp\left(-\frac{1}{2h^2}x^2\right) \\ &\sim \Gamma\left(1, \frac{1}{2h^2}\right) \end{aligned}$$

Furthermore, using a non-informative prior for $2h^2$ (*i.e.* $p(2h^2) \propto 1/h^2$), the posterior for $2h^2$ can be shown to be Inverse Gamma distributed, *i.e.*

$$\begin{aligned} p(2h^2|x^2) &\propto \frac{1}{(2h^2)^2} \exp\left(-\frac{x^2}{2h^2}\right) \\ &\sim IG(1, x^2) \end{aligned}$$

Gaussian Distributed ϵ

Consider Gaussian distributed ϵ , as is the case in coherent clutter, *i.e.*

$$p(\epsilon) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\epsilon^2\right)$$

Using equation E.1 it can be shown that the posterior distribution for $x^2 = h^2\epsilon^2$ is Gamma distributed, *i.e.*

$$\begin{aligned} p(x^2|2h^2) &\propto \frac{1}{xh} \exp\left(-\frac{1}{2h^2}x^2\right) \\ &\sim \Gamma\left(\frac{1}{2}, \frac{1}{2h^2}\right) \end{aligned}$$

Furthermore, using a non-informative prior for $2h^2$ (i.e. $p(2h^2) \propto 1/h^2$), the posterior for $2h^2$ can be shown to be Inverse Gamma distributed, i.e.

$$\begin{aligned} p(2h^2|x^2) &\propto \frac{1}{(2h^2)^{3/2}} \exp\left(-\frac{x^2}{2h^2}\right) \\ &\sim IG\left(\frac{1}{2}, x^2\right) \end{aligned}$$

E.2 GAMMA-LOG NORMAL DISTRIBUTION MATCHING

Consider a Gamma distributed variable x^2 (this is true when x is either Gaussian or Rayleigh distributed), i.e.

$$p(x^2) = \frac{\beta^\alpha}{\Gamma(\alpha)} (x^2)^{\alpha-1} \exp(-\beta x^2)$$

The problem is then to approximate $p(\ln x^2)$ by a Gaussian distribution $N(\mu, \sigma)$. To this end, the first two moments of $\ln x^2$ under $p(x^2)$ (i.e. the mean and variance of $\ln x^2$) are given by

$$E_{p(x^2)}(\ln x^2) = \psi^{(0)}(\alpha) - \ln \beta \quad (\text{E.2})$$

$$\text{var}_{p(x^2)}(\ln x^2) = \psi^{(1)}(\alpha) \quad (\text{E.3})$$

where $\psi^{(0)}(\cdot)$ is the digamma function [Abramowitz and Stegun, 1970], while $\psi^{(1)}(\cdot)$ is the trigamma function. It is interesting to note that α solely determines the variance of the distribution in the logarithm domain.

Therefore, in the logarithm domain, with $x^2 \sim \Gamma(\alpha, \beta)$

$$\begin{aligned} p(\ln x^2) &= N(\mu, \sigma) \\ \mu &= \psi^{(0)}(\alpha) - \ln \beta \\ \sigma^2 &= \psi^{(1)}(\alpha) \end{aligned}$$

E.3 INVERSE GAMMA-LOG NORMAL DISTRIBUTION MATCHING

Consider an Inverse Gamma distributed variable h^2

$$p(h^2) = \frac{\beta^\alpha}{\Gamma(\alpha)} (h^2)^{-(\alpha+1)} \exp(-\beta/h^2)$$

The problem is then to approximate $p(\ln h^2)$ by a Gaussian distribution $N(\mu, \sigma)$. To this end, the first two moments of $\ln h^2$ under $p(h^2)$ (i.e. the mean and variance of $\ln h^2$) are given by

$$E_{p(h^2)}(\ln h^2) = \ln \beta - \psi^{(0)}(\alpha) \quad (\text{E.4})$$

$$\text{var}_{p(h^2)}(\ln h^2) = \psi^{(1)}(\alpha) \quad (\text{E.5})$$

where $\psi^{(0)}(\cdot)$ is the digamma function [Abramowitz and Stegun, 1970], while $\psi^{(1)}(\cdot)$ is the trigamma function. It is interesting to note that α solely determines the variance of the distribution in the logarithm domain.

Therefore, in the logarithm domain, with $h^2 \sim IG(\alpha, \beta)$

$$\begin{aligned} p(\ln h^2) &= N(\mu, \sigma) \\ \mu &= \ln \beta - \psi^{(0)}(\alpha) \\ \sigma^2 &= \psi^{(1)}(\alpha) \end{aligned}$$

Bibliography

- [Abramowitz and Stegun, 1970] Abramowitz, M. and Stegun, I. (1970). *Handbook of Mathematical Functions*. Dover Publications.
- [Anderson and Moore, 1979] Anderson, B. and Moore, J. (1979). *Optimal Filtering*. Prentice-Hall.
- [Armstrong and Griffiths, 1991a] Armstrong, B. and Griffiths, H. (1991a). CFAR detection of fluctuating targets in spatially correlated K-distributed clutter. *IEE Proc.-F*, 138(2):139–152.
- [Armstrong and Griffiths, 1991b] Armstrong, B. and Griffiths, H. (1991b). Modelling spatially correlated K-distributed clutter. *Electronics Letters*, 27(15):1355–56.
- [Azzarelli, 1995] Azzarelli, T. (1995). General class of non-Gaussian coherent clutter models. *IEE Proc.-Radar, Sonar, Navig.*, 142(2):61–70.
- [Bernardo and Smith, 1994] Bernardo, J. and Smith, A. (1994). *Bayesian Theory*. John Wiley & Sons.
- [Blacknell, 1994a] Blacknell, D. (1994a). Comparison of parameter estimators for K-distribution. *IEE Proc. - Radar, Sonar, Navig.*, 141(1):45–52.
- [Blacknell, 1994b] Blacknell, D. (1994b). New method for the simulation of correlated K-distributed clutter. *IEE Proc.-Radar, Sonar, Navig.*, 141(1).
- [Bollerslev, 1986] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, pages 307–327.
- [Bollerslev et al., 1992] Bollerslev, T., Chou, R., and Kroner, K. (1992). ARCH modelling in finance. *Journal of Econometrics*, 52:5–59.
- [Branson, 1998] Branson, J. (1998). Tropospheric radar propagation and applications to atmospheric inversion. In *RSSP98*, Peebles, Scotland.
- [Brillinger, 1992] Brillinger, D. (1992). *Basic aspects of Higher-Order Spectra and some of their uses*. Elsevier Science Publishers.
- [Brown, 1967] Brown, M. (1967). Synthetic Aperture Radar. *IEEE Trans. AES*, AES-3(2):217–229.
- [Brown and Porcello, 1969] Brown, W. and Porcello, L. (1969). An introduction to Synthetic Aperture Radar. *IEEE Spectrum*, pages 52–62.
- [Bucciarelli et al., 1996] Bucciarelli, T., Lombardo, P., and Tamburrini, S. (1996). Optimum CFAR detection against compound Gaussian clutter with partially correlated texture. *IEE Proc.-Radar, Sonar Navig.*, 143(2):95–112.
- [Cole, 1992] Cole, H. (1992). *Understanding Radar*. Blackwell Scientific Publications.

- [Conte and Longo, 1987a] Conte, E. and Longo, M. (1987a). Characterisation of radar clutter as a Spherically Invariant Random Process. *IEE Proceedings F*, 134(2):191–97.
- [Conte and Longo, 1987b] Conte, E. and Longo, M. (1987b). On a coherent model for log-Normal clutter. *IEE Proceedings-F*, 134(2):198–201.
- [Conte et al., 1991a] Conte, E., Longo, M., and Lops, M. (1991a). Modelling and simulation of non-Rayleigh radar clutter. *IEE Proceedings F*, 138(2):121–30.
- [Conte et al., 1991b] Conte, E., Longo, M., Lops, M., and Ullo, S. (1991b). Radar detection of signals with unknown parameters in K-distributed clutter. *IEE Proceedings-F*, 138(2):131–38.
- [Conte et al., 1994] Conte, E., Lops, M., and Ricci, G. (1994). Radar detection in k-distributed clutter. *IEE Proc.-Radar, Sonar Navig.*, 141(2):116–18.
- [Davies, 1998] Davies, M. (1998). Looking for non-linearities in sea clutter. In *RSSP98*, Peebles, Scotland.
- [Duane et al., 1987] Duane, S., Kennedy, A., Pendleton, B., and Roweth, D. (1987). Hybrid monte carlo. *Physics Letters B*, 195(2):216–22.
- [Eaves and Reedy, 1987] Eaves, J. and Reedy, E. (1987). *Principles of Modern Radar*. Van Nostrand Reinhold.
- [Engle, 1982] Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with estimates of the variance of UK inflation. *Econometrica*, 50(4):987–1007.
- [Farina and Lombardo, 1994] Farina, A. and Lombardo, P. (1994). Modelling of a mixture of K-distributed and Gaussian clutter for coherent radar detection. *Electronics Letters*, 30(6):520–521.
- [Gandhi and Kassam, 1988] Gandhi, P. and Kassam, S. (1988). Analysis of CFAR processors in nonhomogeneous background. *IEEE Trans. AES*, 24(4):427–445.
- [Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions, PAMI-6(6)*:721–741.
- [Gilks et al., 1996] Gilks, W., Richardson, S., and Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall.
- [Gini et al., 1995] Gini, F., Greco, M., and Verrazzani, L. (1995). Detection problem in mixed clutter environment as a gaussian problem by adaptive processing. *Electronics Letters*, 31(14):1189–90.
- [Harvey and Shephard, 1996] Harvey, A. and Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14(4):429–434.
- [Hastings, 1970] Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):437–443.
- [Haykin, 1996] Haykin, S. (1996). Neural networks expand SP's horizons. *IEEE Signal Processing Magazine*, pages 24–49.
- [Haykin and Li, 1995] Haykin, S. and Li, X. (1995). Detection of signals in chaos. *Proceedings of the IEEE*, 83(1):95–122.
- [He, 1994] He, Y. (1994). Performace of some generalised modified order statistics cfar detectors with automatic censoring technique in multiple target situations. *IEE Proc.-Radar, Sonar Navig.*, 141(4):205–212.
- [Hirst and Baker, 1987] Hirst, P. and Baker, C. (1987). Sea clutter data base. *RSRE Divisional Memo BS2*.

- [Hull, 1989] Hull, J. (1989). *Options, futures, and other derivative securities*. Prentice Hall.
- [Jakeman and Tough, 1988] Jakeman, E. and Tough, R. (1988). Non-Gaussian models for the statistics of scattered waves. *Advances in Physics*, 37(5):471–529.
- [Kanter, 1986] Kanter, I. (1986). Exact detection probability for partially correlated rayleigh targets. *IEEE Trans. AES*, AES-22(2):184–195.
- [Kazakos and Papantoni-Kazakos, 1990] Kazakos, D. and Papantoni-Kazakos, P. (1990). *Detection and Estimation*. Computer Science Press.
- [Kingsley and Quegan, 1992] Kingsley, S. and Quegan, S. (1992). *Understanding Radar Systems*. McGraw-Hill.
- [Leung, 1992] Leung, H. (1992). Experimental modeling of electromagnetic wave scattering from an ocean surface based on chaotic theory. *Chaos, Solitons and Fractals*, 2(1):25–43.
- [Leung, 1995] Leung, H. (1995). Applying chaos to radar detection in an ocean environment: An experimental study. *IEEE Journal of Oceanic Engineering*, 20(1):56–64.
- [Leung and Haykin, 1990] Leung, H. and Haykin, S. (1990). Is there a radar clutter attractor. *Appl. Phys. Lett.*, 6(5):593–595.
- [Leung and Lo, 1993] Leung, H. and Lo, T. (1993). Chaotic radar signal processing over the sea. *IEEE Journal of Oceanic Engineering*, 18(3):287–295.
- [Levanon, 1988] Levanon, N. (1988). *Radar Principles*. John Wiley & Sons.
- [Lewis et al., 1989] Lewis, P., McKenzie, E., and Hugus, D. (1989). Gamma processes. *Commun. Statist. - Stochastic Models*, 5(1):1–30.
- [Li and Yu, 1989] Li, G. and Yu, K. (1989). Modelling and simulation of coherent Weibull clutter. *IEE Proc. F*, 136(1):2–12.
- [Lombardo and Farina, 1996] Lombardo, P. and Farina, A. (1996). Coherent radar detection against k-distributed clutter with partially correlated texture. *Signal Processing*, (48):1–15.
- [Lombardo and Oliver, 1994] Lombardo, P. and Oliver, C. (1994). Estimation of texture parameters in K-distributed clutter. *IEE Proc. - Radar, Sonar, Navig.*, 141(4):197–204.
- [Lombardo and Oliver, 1995] Lombardo, P. and Oliver, C. (1995). Estimating the correlation properties of K-distributed SAR clutter. *IEE Proc. Radar, Sonar, Navig.*, 142(4).
- [Lombardo et al., 1995] Lombardo, P., Oliver, C., and Tough, R. (1995). Effect of noise on order parameter estimation for K-distributed clutter. *IEE Proc.-Radar, Sonar, Navig.*, 142(1):33–40.
- [Lops and Willett, 1994] Lops, M. and Willett, P. (1994). LI-CFAR: a flexible and robust alternative. *IEEE Trans. AES*, 30(1):41–53.
- [Marier, 1995] Marier, L. (1995). Correlated K-distributed clutter generation for radar detection and track. *IEEE Transactions on Aerospace and Electronic Systems*, 31(2):568–580.
- [Marshall and Olkin, 1967] Marshall, A. and Olkin, I. (1967). A multivariate exponential distribution. *American Statistical Association Journal*, pages 30–44.
- [Mendel, 1991] Mendel, J. (1991). Tutorial on HOS (Spectra) in Signal Processing and System Theory: theory and applications. *Proceedings of IEEE*, 79(3).

- [Metropolis et al., 1953] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1091.
- [Neal, 1992] Neal, R. (1992). Bayesian training of backpropagation networks by the hybrid monte carlo method. Technical Report CRG-TR-92-1, University of Toronto.
- [Neal, 1994] Neal, R. (1994). Sampling from multimodal distributions using tempered transitions. Technical Report 9421, University of Toronto.
- [Nikias and Mendel, 1993] Nikias, C. and Mendel, J. (1993). Signal processing with higher order spectra. *IEEE Signal Processing Magazine*.
- [Noga and Fitzgerald, 1997] Noga, J. and Fitzgerald, W. (1997). Modelling correlated Rayleigh distributed clutter using the Gibbs sampler. *American Statistical Association: Proceedings of the Section on Bayesian Statistics*, pages 234–239.
- [Noga and Fitzgerald, 1998a] Noga, J. and Fitzgerald, W. (1998a). Modelling sea clutter using conditional heteroscedastic models. In *EUSIPCO98*, Rhodes, Greece.
- [Noga and Fitzgerald, 1998b] Noga, J. and Fitzgerald, W. (1998b). Radar detection in spatially correlated sea clutter using Kalman Filter based CFAR detector. In *RSSP98*, Peebles, Scotland.
- [Nohara and Haykin, 1993] Nohara, T. and Haykin, S. (1993). AR-based growler detection in sea clutter. *IEEE Transactions on Signal Processing*, 41(3):1259–1271.
- [Ó Ruanaidh and Fitzgerald, 1996] Ó Ruanaidh, J. and Fitzgerald, W. (1996). *Numerical Bayesian Methods Applied to Signal Processing*. Springer-Verlag.
- [Oliver, 1993] Oliver, C. (1993). Optimum texture estimators for SAR clutter. *J. Phys. D: Appl. Phys.*, 26:1824–1835.
- [Oliver and Tough, 1986] Oliver, C. and Tough, R. (1986). On the simulation of correlated K-distributed random clutter. *Optica Acta*, 33(3):223–250.
- [Parker and Chua, 1987] Parker, T. and Chua, L. (1987). Chaos: A tutorial for engineers. *Proceedings of the IEEE*, 75(8):982–1007.
- [Pentini et al., 1992] Pentini, F., Farina, A., and Zirilli, F. (1992). Radar detection of targets located in a coherent k distributed clutter background. *IEE Proceedings F*, 139(3):239–245.
- [Raghavan, 1991a] Raghavan, R. (1991a). A method for estimating parameters of K-distributed clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 27(2):238–247.
- [Raghavan, 1991b] Raghavan, R. (1991b). A model for spatially correlated radar clutter. *IEEE Trans. on Aerospace and Electronic Systems*, 27(2):268–275.
- [Rangaswamy et al., 1995] Rangaswamy, M., Weiner, D., and Ozturk, A. (1995). Computer generation of correlated non-Gaussian radar clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 31(1):106–115.
- [Rangaswamy et al., 1993] Rangaswamy, M., Wiener, D., and Ozturk, A. (1993). Non-Gaussian random vector identification using Spherically Invariant Random Processes. *IEEE Transactions on Aerospace and Electronic Systems*, 29(1):111–23.
- [Rifkin, 1994] Rifkin, R. (1994). Analysis of cfar performance in weibull clutter. *IEEE Trans. on aerospace and electronic systems*, 30(2):315–328.

- [Roberts and Sahu, 1997] Roberts, G. and Sahu, S. (1997). Updating schemes, correlation structure, blocking and parametrization for the gibbs sampler. *J. Royal Stat. Soc. B*, 59(2):291–317.
- [Shephard, 1994] Shephard, N. (1994). Local scale models: state space alternative to integrated GARCH processes. *Journal of Econometrics*, (60):181–202.
- [Shnidman, 1995] Shnidman, D. (1995). Radar detection probabilities and their calculation. *IEEE Trans. on aerospace and electronic systems*, 31(3):928–950.
- [Singpurwalla and Youngren, 1993] Singpurwalla, N. and Youngren, M. (1993). Multivariate distributions induced by dynamic environments. *Scand J Statist*, (20):251–261.
- [Skolnik, 1980] Skolnik, M. (1980). *Introduction to Radar Systems*. McGraw Hill.
- [Szajnowski, 1976] Szajnowski, W. (1976). Generation of correlated log-Normal clutter samples. *Electronics Letters*, 12(19):497–98.
- [Szajnowski, 1977] Szajnowski, W. (1977). The generation of correlated Weibull clutter for signal detection problems. *IEEE Trans. on Aerospace and Electronic Systems*, AES-13(5):536–541.
- [Ward et al., 1990] Ward, K., Baker, C., and Watts, S. (1990). Maritime surveillance radar part 1: Radar scattering from the ocean surface. *IEE Proceedings F*, 137(2):51–62.
- [Watts, 1985] Watts, S. (1985). Radar detection prediction in sea clutter using the compound k-distribution model. *IEE Proceedings F*, 132(7):613–620.
- [Watts, 1987] Watts, S. (1987). Radar detection prediction in k-distributed sea clutter and thermal noise. *IEEE Trans. on aerospace and electronic systems*, AES-23(1):40–45.
- [Watts, 1996] Watts, S. (1996). Cell-averaging CFAR gain in spatially correlated k-distributed clutter. *IEE Proc.-Radar, Sonar Navig.*, 143(5):321–327.
- [Watts et al., 1990] Watts, S., Baker, C., and Ward, K. (1990). Maritime surveillance radar part 2: Detection performance prediction in sea clutter. *IEE Proceedings F*, 137(2):63–72.
- [Watts and Ward, 1987] Watts, S. and Ward, K. (1987). Spatial correlation in K-distributed sea clutter. *IEE Proc. F*, 134(6):526–532.
- [Weiss, 1984a] Weiss, A. (1984a). ARMA models with ARCH errors. *Journal of time series analysis*, 5(2):129–143.
- [Weiss, 1984b] Weiss, A. (1984b). Asymptotic theory for ARCH models: Estimation and testing. *Econometric Theory*, (2):107–131.
- [Wu, 1998] Wu, M. (1998). *Markov Chain Monte Carlo Methods Applied to Bayesian Data Analysis*. PhD thesis, University of Cambridge.

