Australian Government
Department of Defence
Defence Science and
Technology Organisation

Analysis of High Grazing Angle Sea-clutter with the KK-Distribution

Luke Rosenberg, David J. Crisp and Nick J. Stacy

National Security and ISR Division

Defence Science and Technology Organisation

DSTO–TR–2915

ABSTRACT

An estimated probability distribution of the backscatter is commonly used to determine the threshold for distinguishing targets from clutter at a given false alarm rate. Data collected at high grazing angles ($15^\circ - 45^\circ$) by the Defence Science Technology Organisation's Ingara fully polarimetric X-band radar demonstrates that the commonly used K-distribution is not always adequate for modelling the probability distribution. This is especially the case for the horizontal polarisation when sea-spikes can cause high false alarm rates. An alternative proposed as a more accurate model is known as the KK-distribution. The analysis presented in this report describes this model with the addition of multiple looks and a thermal noise component to produce an estimate of the underlying mean and shape. This then enables the KK-distribution to be used as a proxy for data in radar detection performance studies. The threshold required to achieve a constant false alarm rate is then studied and compared with that obtained from the K-distribution.

APPROVED FOR PUBLIC RELEASE

Published by

DSTO Defence Science and Technology Organisation
PO Box 1500
Edinburgh, South Australia 5111, Australia

Telephone: 1300 333 362
Facsimile: (08) 7389 6567

AR No. 015-796
November 2013

APPROVED FOR PUBLIC RELEASE

Analysis of High Grazing Angle Sea-clutter with the KK-Distribution

Executive Summary

Traditionally, maritime surveillance of small targets is conducted from low altitude platforms due to the relatively low clutter backscatter. This surveillance scenario has been well studied and relevant models have been developed. In the future however, high altitude airborne platforms will offer improved area coverage at the expense of increased radar backscatter. This geometry results in sea-clutter with a higher grazing angle where there is very little data available for analysis. In August 2004 and July 2006, the Defence Science and Technology Organisation's (DSTO) Ingara X-band airborne radar collected fine resolution fully polarimetric data in the high grazing angle region $15^\circ - 45^\circ$. The data was collected from the ocean off the coasts of Port Lincoln and Darwin respectively. This report builds on work undertaken at the DSTO in characterising the maritime environment from high altitude airborne platforms.

The focus of this report is to characterise sea-clutter probability distribution functions using the collected data with the goal of more accurately detecting targets in radar backscatter. At finer radar resolution, the effects from both discrete and persistent sea-spikes will have a big impact on the radar backscatter. The commonly used K-distribution is shown to not always be accurate enough for modelling the probability distribution, especially in the horizonal polarisation channels and at lower grazing angles ($\sim 15^\circ$). Instead, a recently proposed model known as the KK-distribution is presented as an alternative to model the backscatter in this region. Understanding and modelling these sea-spike events are important for the prediction of radar performance, especially when the goal is to detect small targets present within the sea-clutter. The results in this report show that incorrectly calculating the detection threshold can result in under-prediction of the required signal to interference ratio by up to 6 dB. In an operational radar, an error of this magnitude will undoubtedly result in missed detections of small targets.

The main contributions in this report include a description of the KK-distribution model with extensions to include multiple looks and thermal noise, thereby allowing an estimate of the underlying mean and shape of the distribution. Rather than fitting all five parameters, constraints are introduced to overcome difficulties in fitting the full PDF and the data is fitted in the $\log_{10}$ domain in order to focus on the tail of the distribution. A sea-clutter simulation is used to verify the fitting procedure and show that the K-distribution is not able to estimate the underlying shape of the distribution with both thermal noise and sea-spikes present, while the modified KK-distribution can.

Both single and multi-look data are analysed, with a poor fit shown for the K-distribution in the horizontal and cross polarised channels. The fit is substantially improved when the modified KK-distribution is used. The dynamic range of the mean is greatly increased and there is now more detail in the shape estimates, particularly where the clutter to noise ratio is low. The KK parameter which measures the extent of the tail has its highest values in the low grazing angle regions of the horizontal polarisation channel.

To assess target detection performance, the false alarm rate is calculated from the distributions down to 10^{-6}. To determine the model suitability, the 'threshold error' is measured as the difference between the data and model at a specific false alarm rate. These results show a good match between the data and the KK-distribution fit, while the K-distribution fit for the horizontal and cross polarised channels mismatched the data from approximately 10^{-2} and below. The

mismatch is in the lower grazing region and reduces linearly as the grazing angle increases. The final result shows a comparison between the original formulation of the KK-distribution without considering noise and the modified version from this report. There is only a small improvement in the threshold error, thus demonstrating that the parameters estimated with and without thermal noise are both able to construct the distribution accurately enough for target detection purposes at the CNRs consistent with false alarm rates lower than 10^{-4}. However, estimates of the underlying mean and shape are important as it will enable the KK-distribution to be used as a proxy for data in radar detection performance studies.

Authors

Luke Rosenberg
National Security and ISR Division

Luke Rosenberg received his BE (Elec.) with honours from Adelaide University in 1999 and joined the DSTO in January 2000. Since this time he has completed both a masters degree in signal and information processing and a PhD in multichannel synthetic aperture radar through Adelaide University. He has worked at the DSTO as an RF engineer in the missile simulation centre, as a research scientist in the imaging radar systems group and now in the maritime radar group. Current research interests include radar and clutter modelling, radar imaging and detection theory.

David J. Crisp
National Security and ISR Division

David Crisp graduated from the University of Adelaide in 1987 with a B.Sc. (Hons) in Mathematics and completed his Ph.D. at the same institution in 1993. For the three years following that he held a post-doctoral research position in Mathematics at the Flinders University of South Australia. In 1997 he commenced employment as a Post-doctoral Research Fellow with the Cooperative Research Centre for Sensor, Signal and Information Processing where he worked in the Pattern Recognition Group on the application of machine learning techniques to real world problems. In September 1999 he joined the Australian DSTO. He spent several years working as a Research Scientist in the Imagery Analysis and Exploitation Group before transferring to the Imaging Radar Systems Group - both within National Security and Intelligence, Surveillance and Reconnaissance Division. His current research is focused on the detection of targets in synthetic aperture radar imagery of the ocean.

Nick Stacy
National Security and ISR Division

Nick Stacy received the B.E. (Hons), M.S. and Ph.D. degrees in electrical engineering from the University of Adelaide in 1984, Stanford University in 1985 and Cornell University in 1993 respectively. He worked at the National Astronomy and Ionosphere Center at Arecibo Observatory in Puerto Rico from 1985 to 1986 on Earth-based planetary radar imaging of Venus and the Moon. From 1987 to 1989 he was with British Aerospace Australia where he worked on the AETHERS-1 system for the Australian Centre for Remote Sensing to process SAR data from the ERS-1 satellite. He joined DSTO in 1993 where he has worked in the field of imaging radar systems, phenomenology, image formation and analysis primarily using the Ingara airborne radar system. He was the Australian sensor lead for the Global Hawk deployment to Australia in 2001 and was awarded a Defence Science Fellowship to support collaborative R&D in the US for 18 months from 2005 to 2007. He led the Imaging Radar Systems group from 2001 to 2008 and is currently the Research Leader for the Imagery Systems Branch in NSI Division, DSTO.

Contents

Figures

Tables

THIS PAGE IS INTENTIONALLY BLANK

Glossary

BAM:	Breaking Area Model
CDF:	Cumulative Distribution Function
CFAR:	Constant False Alarm Rate
CNR:	Clutter to Noise Ratio
DSTO:	Defence Science and Technology Organisation
FFT:	Fast Fourier Transform
HH:	Horizontal transmit and Horizontal receive polarisation
HV:	Horizontal transmit and Vertical receive polarisation
MAST06:	Maritime Surveillance Trial 2006
NRMS:	Normalised Root Mean Square
PDF:	Probability Distribution Function
PRF:	Pulse Repetition Frequency
RCS:	Radar Cross Section
SAR:	Synthetic Aperture Radar
SCT04:	Sea Clutter Trial 2004
VH:	Vertical transmit and Horizontal receive polarisation
VV:	Vertical transmit and Vertical receive polarisation

THIS PAGE IS INTENTIONALLY BLANK

1 Introduction

An active area of research at the DSTO is to understand the characteristics of sea-clutter as seen by an X-band radar, specifically the mean backscatter, amplitude statistics and the Doppler spectrum, [Rosenberg 2012, Crisp et al. 2008, Dong 2006, Rosenberg, Crisp & Stacy 2010, Rosenberg & Stacy 2008, Rosenberg, Crisp & Stacy 2008]. This report is focussed on the amplitude statistics and aims to get a more accurate estimate of the underlying sea-clutter parameters by accounting for the high magnitude sea-spikes and the thermal noise which is present in the radar.

Sea-clutter probability distribution functions (PDFs) have been studied for decades, primarily on coarse resolution data collected at low grazing angles. However, at finer range resolution, the effects from discrete sea-spike events will have a bigger impact on the radar backscatter. Understanding and modelling these events are important for the prediction of radar performance, especially when the goal is to discriminate targets from the sea-clutter. Typically the Weibull, log-normal or K-distributions are used to model radar backscatter from the sea, but none of these perfectly describes the sea under all conditions, [Ward, Tough & Watts 2006]. Also, thermal noise from the radar can effect the total received signal and change the expected distribution. A number of newer distributions have since been developed to account for the effects of sea-spikes and/or the thermal noise found in the radar. These include the KA-distribution, [Middleton 1999], the KK-distribution, [Dong 2006] and more recently the Pareto distribution [Farshchian & Posner 2010, Weinberg 2011, Rosenberg & Bocquet 2013].

The KK-distribution was developed to capture the non Rayleigh sea-spike events, but does not account for thermal noise. The work presented in this report extends the KK-distribution model to include thermal noise and multiple looks. The modified model is then applied to both simulated and real data to determine the regions where sea-spikes are present and to extract the underlying shape parameter of the sea-clutter PDF with no thermal noise present. Data taken from the DSTO Ingara radar is used to characterise results over all polarisations, azimuth directions and a range of grazing angles from 15° to 45°. This knowledge will enable the KK-distribution to be used as a proxy for data in radar detection performance studies. Furthermore, the cumulative distribution function (CDF) can be formed from the PDF and is related to the probability of false alarm by $P_{\text{fa}} = 1-\text{CDF}$. This enables detection thresholds to be determined for a desired CFAR. By carefully combining the radar backscatter over azimuth and grazing angles to increase the sample size, detection thresholds can then be determined for P_{fa}'s down to 10^{-6}. A discussion of the impact of this averaging is presented in Section 5.1.3. However, the exact effect of this windowing on the interpreted phenomenology is unknown and cannot be resolved except with repeated experimentation.

This report contains five main sections. Section 2 contains background to the distributions mentioned above, the DSTO's Ingara radar and the trial data. Section 3 develops the modified KK-distribution and describes how the important parameters are estimated from the data. Section 4 then applies the estimation method to simulated data to demonstrate its accuracy and measure the effect of sea-spike events and thermal noise on the PDF. Finally, Section 5 looks at real data collected from the DSTO trials, with results shown that demonstrate the suitability of the KK-distribution when compared to the commonly used K-distribution.

2 Background

This section provides relevant background material to the latter sections of this report. The first Section 2.1 contains a brief literature review of the commonly used PDFs for sea-clutter and the newer ones developed to account for sea-spikes and/or thermal noise in the radar. A brief description of the Ingara radar is presented in Section 2.2, while background to the two sea-clutter trials is presented in Section 2.3.

2.1 Sea-clutter probability distributions

Sea-clutter distributions were originally developed to look separately at the temporal and spatial returns from the sea surface. The radar systems used were typically stationary with low grazing angles and coarse resolution. Commonly used PDFs include the Rayleigh, log-normal and Weibull, with the latter two used when longer tails were observed in the radar backscatter, [Long 2001]. A more useful model however, includes both components in a 'compound' representation. The most widely used is known as the K-distribution which is described in terms of Rayleigh speckle fluctuations modulated by Gamma underlying radar cross section (RCS) variations.

The K-distribution was first applied to modelling probability distributions of sea-clutter by Jakeman & Pusey [1976]. It was later put into a Bayesian or compound formulation by Ward [1981] which allows a more meaningful understanding of the two main components. These being the temporal or fast varying component with a correlation time on the order of 10 ms and the spatial varying component with a correlation length on the order of metres. The fast component is commonly known as speckle and is the result of constructive and destructive interference effects between multiple scatterers. It is typically associated with the small local wind-driven ripples (capillary waves) on the ocean surface. The spatial component represents changes in the large and medium scale waves which modulate the speckle. In terms of scattering processes, reflections off the wind-driven ripples are known as Bragg (or resonant) scattering, while the non-Bragg scattering can be described as a combination of whitecaps and discrete sea-spikes, [Ward, Tough & Watts 2006]. Whitecaps are the result of wave crest breaking and occur for seconds with a Rayleigh fluctuation and decorrelation time on the order of milliseconds, while sea-spikes are produced by specular-type reflections which are spatially more isolated and can last for a number of seconds. It is these sea-spikes that cause the K-distribution model to break down, particularly at finer resolutions where they have higher levels of backscatter relative to adjacent cells and extend the tail of the distribution. A second source of mismatch can also arise when there is a low clutter to noise ratio (CNR) and the distribution becomes more Rayleigh. Consequently, both sea-spikes and thermal noise should be accounted for to improve the accuracy of the distribution model and allow an estimate of the underlying mean and shape.

In the literature, there have been a few methods developed to modify the K-distribution to deal with non-Rayleigh scattering. The first is the KA-distribution developed by Middleton [1999] and applied by Ward & Tough [2002]. The model includes components for the thermal noise and speckle with the discrete spikes modelled with a Poisson distribution. This model is then incorporated with the Gamma model for the underlying waves to give the form of the KA-distribution. A similar model known as the breaking area model (BAM) has been developed by Clements and Yurtsever[1] and also uses a Poisson model to characterise the sea-spikes. It differs from the KA-

[1]The original reference for this work is no longer available.

distribution in that the former is used to add the sea-spike component coherently, while the BAM is used to model the mean power and the sea-spike component is added incoherently, [Ward, Tough & Watts 2006]. More recently, the Pareto distribution has been applied to model the non-Rayleigh scattering [Farshchian & Posner 2010, Weinberg 2011, Rosenberg & Bocquet 2013]. It was found that this distribution can equally fit the long tails present in the data with the advantage of only requiring two parameters to be estimated.

An alternative model for capturing the non-Rayleigh sea-spike components has been presented by Dong [2006]. His solution uses a mixture model for the distribution consisting of the sum of two K components, one associated with the Bragg components and one for the sea-spikes. It should be noted that the choice of the second K component is arbitrary and has no direct physical basis. However, when compared to the KA-distribution, the KK is equally able to model the tail region with the benefit of better matching in the region where the tail extends from the bulk of the distribution. Dong's implementation of the KK-distribution doesn't account for thermal noise however, which means that the closed analytic form of the K-distribution can be used. This provides a large computational benefit over the KA-distribution making it more suitable for real-time use in a constant false alarm rate (CFAR) target detection scheme. However, if the thermal noise in the radar is stronger than the received backscatter, there will be significant errors in the threshold estimate.

The effect of thermal noise on the K-distribution has been studied by a number of authors. Watts [1987] first looked at this effect from a target detection point of view. He formulated a modified distribution and showed that estimating the K-distribution parameters with thermal noise present leads to an 'effective' shape, which can still be used to achieve acceptable target detection results. Further studies by Gini et al. [1998] has extended the K-distribution model to include correlation and detailed analysis showed better target detection results when thermal noise is included in the model, at the expense of greater computation. A slightly different focus by Lombardo, Oliver & Tough [1995] looked at the effect of thermal noise on the parameter estimation for K-distributed clutter and derived a different shape estimator which is less sensitive to thermal noise.

2.2 Radar description

The DSTO Ingara system is an airborne multi-mode X-band imaging radar system. It operates with a centre frequency of 10.1 GHz and supports a 600 MHz bandwidth for fine resolution in a spotlight mode. The sea-clutter trials however used a bandwidth of 200 MHz to achieve a larger swath width. The radar is fully polarimetric and utilises a dual linear polarised antenna developed by the Australian CSIRO for both transmitting and receiving [Parfitt & Nikolic 2001]. In fully polarimetric collections, the system is operated at double the normal pulse repetition frequency (PRF) with the polarisation switch being used to alternate the transmit polarisation between horizontal and vertical polarisations while receiving horizontal and vertical polarisations simultaneously. A more detailed description of the system may be found in [Stacy et al. 2003] and the references therein cite earlier descriptions. The standard radar operating parameters used during the sea-clutter collections are shown in Table 1. Also, to simplify the analysis in this report, the backscatter measurements are based on the range-compressed real beam data without any Doppler processing.

Table 1: *Standard radar operating parameters for ocean backscatter collections.*

Parameter	Value
Frequency	10.1 GHz
Transmitted bandwidth	200 MHz
Pulse length	20 μs
Peak transmit power	1 kW
3 dB beamwidth - azimuth / elevation	$1^\circ/13^\circ$

2.3 Trials background

The trial data was obtained with Ingara on two separate occasions and two distinctly different regions. The first 'sea-clutter' trial was conducted in 2004 (SCT04) in the southern ocean approximately 100 km south of Port Lincoln, South Australia [Crisp, Stacy & Goh 2006]. The site chosen was at the edge of the South Australian continental shelf where there was little chance of shallow water affecting the wave field. During the trial, ocean backscatter was collected for a range of different geometries on eight separate days with different ocean conditions. The second 'maritime surveillance' trial was conducted in 2006 (MAST06) in littoral and open sea environments near Darwin in the Northern Territory. Again, data was collected for a range of different geometries and ocean conditions. In this trial, a total of four days data were collected: two of the days were in the littoral zone approximately 25 km north of Darwin and the other two days were in the open ocean approximately 200 km west of Darwin. These trials have been reported in detail in [Crisp, Stacy & Goh 2006] and [Crisp, Preiss & Goh 2008].

During the ocean backscatter collections, Ingara was operated in the circular spotlight-mode. Figure 1 shows this mode, where the aircraft flies a circular orbit in an anti-clockwise direction (as seen from above) around a nominated point of interest, while the radar beam is continuously directed toward this point. Radar echo data is continuously collected during the full 360° orbit, with the instantaneous PRF appropriately adjusted to maintain a constant spatial separation between pulse transmission positions. Once collected, the echo data may be processed either immediately (in real-time) or subsequently (off-line) to produce either range-compressed profiles or spotlight synthetic aperture radar (SAR) images of the scene at various azimuth angles. Further, owing to the continuous nature of the data collection, the images can be formed at any desired azimuth look direction. Each collection of data in this mode is referred to as a 'run' and there may be several complete orbits in a single run.

In order to examine the effect of grazing angle on ocean backscatter, runs were made with different grazing angles. In circular spotlight mode, this is done by controlling the aircraft altitude and orbit radius. For both the SCT04 and MAST06 trials, data was collected at the centre of the spotlight for the nominal grazing angles of 15° to 40° in 5° increments. Owing to the wide elevation beam width of the radar, its footprint on the ocean surface has a significant range extent. This means that the grazing angle varies across the footprint. It follows that, with appropriate range compression and data processing, the variation in backscatter with grazing angle across the range extent of the radar beam footprint can be measured. In this way, backscatter measurements for most grazing angles in the range from 10° to 45° could be extracted from the data. Note that the aircraft speed was approximately 100 m/s and so a 1.5 NM radius orbit took approximately 3 minutes while a 1.9 NM radius orbit took 3.5 to 4 minutes. The total collection across all grazing angles took approximately 90 minutes. It is reasonable to assume that over such short

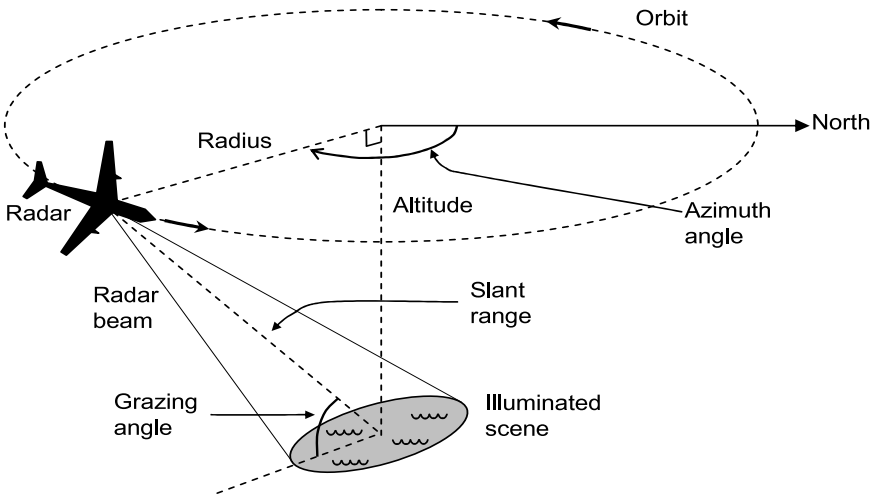

Figure 1: *Circular spotlight mode collection geometry.*

time intervals, the ocean surface conditions are relatively unchanged and that mean backscatter variations are mostly due to the changing imaging geometry rather changing ocean conditions. Nevertheless, it is possible that wind gusts and changes of wind strength and direction may have effected the measurements. Table 2 shows the wind and wave ground truth for the two days analysed in this report.

Table 2: *Wind and wave ground truth. Directions are "from" not "to".*

Trial	**Flight**	**Date**	**Wind**		**Wave**	
			Speed (m/s)	**Direction (°)**	**Height (m)**	**Direction (°)**
SCT04	F35	11/8/04	10.3	315	2.6	210
MAST06	F9	24/5/06	10.2	124	1.21	128

Finally, using the geometry in Figure 1, the azimuth resolution can be calculated approximately by $R_s \Delta\psi$, where R_s is the slant range and $\Delta\psi$ is the measured two-way azimuth antenna 3 dB beamwidth equal to 1.02° for the horizontal transmit, horizontal receive (HH) channel and 0.99° for the vertical transmit, vertical receive (VV) channel. Table 3 shows the geometry for collections at various grazing angles and slant-ranges for a beamwidth of 1°. Over these nominal parameters, the average azimuth resolution is 62.7 m. However since each collection spans a range of grazing angles, the actual azimuth resolution will always differ slightly.

Table 3: *Nominal geometric parameters for circular spotlight-mode collections.*

Grazing angle (°)	Altitude (m)	Radius (NM / m)	Slant range (m)	Azimuth res. (m)
15	932	1.9 / 3519	3643	63.54
20	1353	1.8 / 3334	3548	62.80
25	1522	1.8 / 3334	3679	63.97
30	1711	1.6 / 2963	3421	59.72
35	2073	1.6 / 2963	3617	63.11
40	2314	1.5 / 2778	3626	63.10

3 Distribution functions and parameter estimation

This section describes the development of the KK-distribution model with thermal noise and techniques for estimating its descriptive parameters. This however requires descriptions of both the K and KK-distributions with and without thermal noise, which are presented in Sections 3.1 to 3.4 respectively. The final Section 3.5 then looks at radar performance modelling by introducing the relationship between the cumulative distribution function and the probability of false alarm. This enables the threshold error to be determined at a given P_{fa} and provides a means of comparing the performance of each distribution.

3.1 K-distribution

The K-distribution is a well established model used to describe the amplitude or intensity of ocean backscatter. Among its advantages over other models is its theoretical justification with an underlying physical model. It has also proved to be realistic and is commonly used to model sea-clutter, [Ward, Tough & Watts 2006]. It is usually presented in terms of an intensity product model combining an underlying RCS component, x, with an uncorrelated speckle component, z. Assuming analysis of baseband radar data in complex format (in-phase and quadrature), the magnitude PDF becomes Rayleigh and the power (or intensity) PDF is modelled as an exponential or gamma distribution with unity shape and mean power x,

$$P_{z|x}(z|x) = \frac{1}{x}\exp\left[-\frac{z}{x}\right], \quad z \geq 0 \tag{1}$$

For a multi-look radar however, the average of M independent looks (assuming that the speckle is random and the underlying RCS is constant) is determined by,

$$\sigma = \frac{1}{M}\sum_{m=1}^{M} z_m \tag{2}$$

and the clutter power σ is now described by a gamma distribution with shape M,

$$P_{\sigma|x}(\sigma|x) = \frac{M^M \sigma^{M-1}}{x^M \Gamma(M)}\exp\left[-\frac{M\sigma}{x}\right] \tag{3}$$

where $\Gamma(\cdot)$ is the gamma function. In the K-distribution model, the underlying RCS, x is also a random variable and can be modelled as a Gamma distribution,

$$P_x(x) = \frac{b^\nu}{\Gamma(\nu)}x^{\nu-1}\exp\left[-bx\right], \quad 0 \leq x \leq \infty \tag{4}$$

where ν is the shape, $b = \nu/\mu$ is the scale and μ is the mean. The compound formulation for the K-distribution is then obtained by integrating the speckle PDF over the PDF of the underlying RCS,

$$P(\sigma) = \int_0^\infty P_{\sigma|x}(\sigma|x)P_x(x)dx \tag{5}$$

$$= \frac{2}{\sigma}\left(Mb\sigma\right)^{\frac{M+\nu}{2}}\frac{1}{\Gamma(M)\Gamma(\nu)}K_{\nu-M}\left(2\sqrt{Mb\sigma}\right) \tag{6}$$

where $K_{\nu-M}(\cdot)$ is the modified Bessel function (of the second kind) with order $\nu - M$.

The order parameter ν controls the shape of the PDF and in particular it specifies the 'length' of the tail. At one extreme, when $\nu = \infty$ the underlying RCS is effectively constant and the K-distribution reduces to a gamma distribution (or the negative exponential distribution for single-look data). At the other extreme, as ν decreases to 0, the tail of the K-distribution grows longer due to large amplitude bursts (sea-spikes) in the radar backscatter. In practice [Ward, Tough & Watts 2006, Antipov 1998], ν lies in the range $0.1 < \nu \leq \infty$. When $\nu > 25$, the distribution is generally considered to be the same as if $\nu = \infty$ in which case the K-distribution degenerates to a gamma distribution.

There are three parameters required to estimate a fit for the K-distribution: the number of looks, the mean and the shape. However the number of looks, M is known a-priori in the radar processor and the mean estimate can be calculated directly from the data, $\hat{\mu} = \langle\sigma\rangle$, where $\langle\cdot\rangle$ represents the averaging. That leaves the shape parameter ν, which can be estimated a number of different ways, [Ward, Tough & Watts 2006]. Blacknell & Tough [2001] note that the maximum likelihood estimate of ν for the K-distribution is analytically intractable and that computationally expensive iterative methods must be used. Their X-statistic formulation has been widely accepted as a useful approximation for estimating the shape. It is based on the formula,

$$\frac{1}{\hat{\nu}} + \frac{1}{M} = \frac{\langle\sigma \log\sigma\rangle}{\langle\sigma\rangle} - \langle\log\sigma\rangle \tag{7}$$

where $\hat{\nu}$ is the estimate of ν. Blacknell and Tough examine the accuracy of this estimator very closely and compare it with the other moment based methods. They conclude that this estimator provides the best performance. However, Ward, Tough & Watts [2006] warn that in practice, the total received signal will contain thermal noise and that Equation 7 performs poorly when the CNR is low.

3.2 K-distribution with thermal noise

The total received data collected with a real radar contains a thermal noise component. For many radar systems, the thermal noise floor will be well below the typical clutter backscatter response and detection performance is usually clutter limited. However, there are situations where the clutter response is low and detection performance becomes noise limited. This is often the case for long range, low grazing angle or cross-polarised ocean backscatter. The importance of taking account of thermal noise to correctly model the ocean backscatter is noted by [Ward, Tough & Watts 2006].

The thermal noise in-phase and quadrature components are modelled as zero-mean Gaussian random variables with variance $p_n/2$ in the same way as the speckle component. Using the principle of superposition, the combined variance of the speckle and thermal noise is $x/2 + p_n/2$ and the modified multi-look power distribution is,

$$\tilde{P}_{\sigma|x}(\sigma|x, p_n) = \frac{M^M \sigma^{M-1}}{(x+p_n)^M \Gamma(M)} \exp\left[-\frac{M\sigma}{x+p_n}\right]. \tag{8}$$

With this modification however, the compound PDF has no closed form solution and must be

evaluated numerically,

$$\tilde{P}(\sigma|p_n) = \int_0^\infty \tilde{P}_{\sigma|x}(\sigma|x, p_n) P_x(x) dx. \tag{9}$$

Assuming the thermal noise level is known either through analysis of the radar hardware or by measuring the received signal in the absence of a transmit pulse, the true sea-clutter mean can be found by subtracting the mean of the thermal noise from the clutter plus noise mean, $\hat{\mu} = \langle\sigma\rangle - p_n$. To estimate the underlying shape when noise is present, a fit using Equation 9 could be implemented, but here an identity based on calculating moments of the distributions is used, [Watts 1987]. It relates the shape $\tilde{\nu}$ of the underlying noise free K-distribution to the shape ν through,

$$\tilde{\nu} = \nu(1 + 1/CNR)^{-2}. \tag{10}$$

Note that this relationship has also been verified by the authors to hold for multi-look data. While this approach improves the estimate of the mean and underlying shape when the CNR is low, it does not address the extended tail region when sea-spikes are present and consequently, a second modification must be included in the distribution model.

3.3 KK-distribution

When there are sea-spikes present, Dong [2006], has proposed using two K-distributions to model both the Bragg and sea-spike components, with the overall clutter distribution being the sum / mixture of the two components. Although, the choice of the second component is arbitrary and has no direct physical basis, by fitting the KK-distribution in the $\log_{10}$ domain, an improved fit in the tail region can be achieved. If the two components are represented by P_1 and P_2,

$$Q(\sigma) = (1 - k)P_1(\sigma|\nu_1, \mu_1) + kP_2(\sigma|\nu_2, \mu_2) \tag{11}$$

where the individual mean and shapes have been included explicitly. If the ratio, k between the two components reduces to 0, $Q(\sigma) = P_1(\sigma|\nu_1, \mu_1)$ and a single K-distribution remains without the sea-spike component.

With this distribution, there are now 5 parameters which must be estimated. Dong [2006] has reduced this number to 2 however, by first calculating the probability of false alarm, $P_{\text{fa}} = 1 - CDF$ and observing that the deviation in the tail typically occurs at 10^{-3} or higher. Due to the small number of samples which fall in the tail region, he has set the mean and underlying shape to be the same as that estimated by Blacknell's method in Equation 7, $\hat{\mu}_1 = \hat{\mu}, \hat{\nu}_1 = \hat{\nu}$. He has also found empirically that both the underlying shape parameters can be set the same, $\hat{\nu}_2 = \hat{\nu}$, leaving the second mean or ratio of means, $\hat{\rho} = \hat{\mu}_2/\hat{\mu}_1$ and the ratio between components, $\hat{k}$ to be determined. He has observed that the ratio of means ρ mainly influences the degree of separation between the K and KK-distributions, while the ratio k affects both the separation and the departure point of the tail from the bulk of the distribution. Analysis of the separation between the K-distribution fit and the data led Dong to a fitting procedure based on a lookup table where the ratio between components, $\hat{k} \equiv 0.01$ and the false alarm rate differences between the fitted K-distribution and data at 10^{-3}, 10^{-4} and 10^{-5} can be used to determine the ratio of means estimate, $\hat{\rho}$. Note, that these differences are known as the threshold error and will be discussed further in the following sections. Compared to the estimation method in Section 2.1, Dong's technique requires only a single extra parameter to be estimated and is yet able to model the tail of the distribution to a good degree of accuracy.

3.4 KK-distribution with thermal noise

To increase the usefulness of the KK-distribution, thermal noise is now included by modifying the KK model in Equation 11. This is achieved by substituting the integral expressions from Equation 9 for $P_1(\cdot)$ and $P_2(\cdot)$,

$$\begin{aligned}\tilde{Q}(\sigma|p_n) &= (1-k)\tilde{P}_1(\sigma|\nu_1,\mu_1,p_n) + k\tilde{P}_2(\sigma|\nu_2,\mu_2,p_n) \\ &\approx (1-k)\tilde{P}_1(\sigma|\tilde{\nu}_1,\mu_1+p_n,0) + k\tilde{P}_2(\sigma|\tilde{\nu}_2,\mu_2+p_n,0) \end{aligned} \quad (12)$$

where the approximation from Equation 10 can be used for each component. The model is then fitted using similar assumptions to Dong's method in the previous section. The underlying shape is assumed to be the same for both components, $\hat{\nu}_1 = \hat{\nu}_2 = \hat{\nu}$, and is estimated within the realistic range described earlier, $0.1 \leq \hat{\nu} \leq 25$. As in the previous section, the thermal noise power p_n, is assumed to be known, thereby allowing the mean to be determined by subtracting the noise power from the clutter plus noise mean, $\hat{\mu} = \langle\sigma\rangle - p_n$. The ratio between components is fixed at $\hat{k} \equiv 0.01$ which simplifies the estimation procedure and lets the mean ratio, $\hat{\rho}$ be the sole measure of separation between the KK components. Its valid range was determined by analysis of the data and was found to lie within $1 \leq \hat{\rho} \leq 15$, where an estimate close to 1 indicates that the distribution is very close to a K-distribution, while a higher value indicates a larger deviation in the tail. This choice provides a generous upper bound for the mean ratio as the majority of fits are typically much closer to 1. While constraining the ratio between components does not have any physical significance or provide the most intuitive solution, it does guarantee a consistent and smoother estimate for the mean ratio over a variety of different geometries.

Using these constraints, the model in Equation 12 is fitted to the histogram lineshape using the 'Levenberg-Marquardt' algorithm. Random initial values for each parameter are chosen from within their valid ranges. To obtain a more balanced estimate over the distribution and due to the large dynamic range, the model is also modified to have the clutter power probability distribution represented in dBs,

$$\tilde{Q}(\sigma_{\text{dB}}|p_n) = \frac{\sigma\tilde{Q}(\sigma|p_n)}{10\log_{10}e}, \quad (13)$$

and the cost function is formed by taking the difference between the logarithm of this model and the data, histogrammed in the dB domain, $D(\sigma_{\text{dB}})$. This cost function is not designed to find a solution meeting the standard chi-square criteria, but rather to more accurately match the tail region of the distribution. The root mean square of this difference is then calculated over L points evenly spaced in the range, $\sigma_{\text{dB}} \in [\sigma_{\min},\ldots,\sigma_{\max}]$ to give the error, E for the optimisation algorithm,

$$E = \sqrt{\frac{1}{L}\sum_{\sigma_{\text{dB}}}\Big[\log_{10}(\tilde{Q}(\sigma_{\text{dB}},p_n)) - \log_{10}(D(\sigma_{\text{dB}}))\Big]^2}. \quad (14)$$

3.5 Radar performance modelling

The end goal of an accurate PDF model is to enable improved target detection. That is achieved by setting the most accurate threshold to distinguish between the target plus interference (clutter and noise) and the interference alone at a defined false alarm rate. Classical hypothesis target detection involves comparing PDF models of these two alternatives and then choosing an appropriate threshold to classify the incoming signal, [Ward, Tough & Watts 2006]. Consider Figure 2 where two distributions are presented representing the interference, $f_I(x)$ and the target plus interference, $f_{TI}(x)$. To classify the radar backscatter into either one of these distributions, a threshold, τ must be determined which maximises the probability of detection, P_{d} while minimising the probability of false alarm, P_{fa} and the miss probability P_{m}. If hypothesis H_0 is for interference and H_1 is for target plus interference, then the following definitions can be defined for a radar return of magnitude X:

$$P_{\text{d}} = P(X > \tau | H_1) = \int_{\tau}^{\infty} f_{TI}(x)dx \tag{15}$$

$$P_{\text{fa}} = P(X > \tau | H_0) = \int_{\tau}^{\infty} f_I(x)dx \tag{16}$$

For the work in this report, only the probability of false alarm is of interest as it can be directly related to the CDF of the clutter plus noise signal by,

$$P_{\text{fa}} = 1 - P(X \leq \tau | H_0) = 1 - \int_{-\infty}^{\tau} f_I(x)dx \tag{17}$$

Using this relationship, different distribution models can then be compared against the true distribution of the data and a threshold error determined. This error must be defined at a given P_{fa} as is shown in Figure 3. The magnitude of this error will be used extensively in the remainder of the report to measure the accuracy of different PDF models.

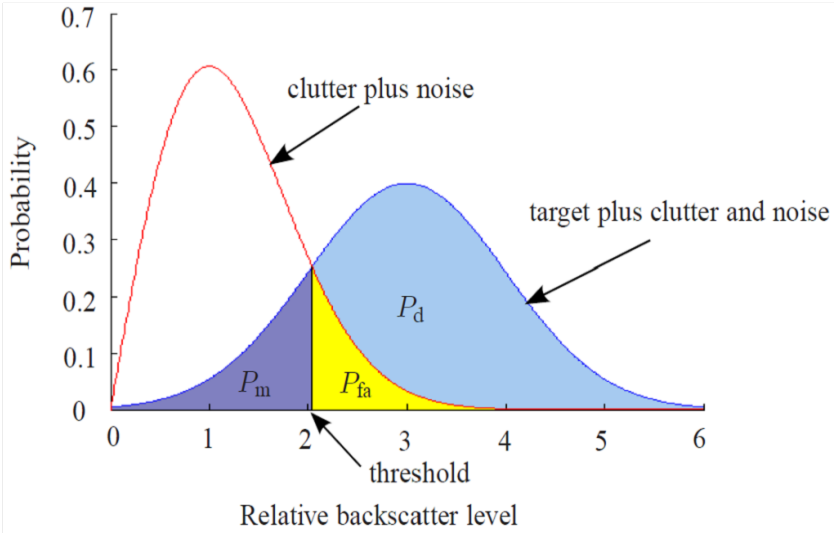

Figure 2: *Target detection regions for two distributions representing the interference (clutter plus noise) and the target plus interference.*

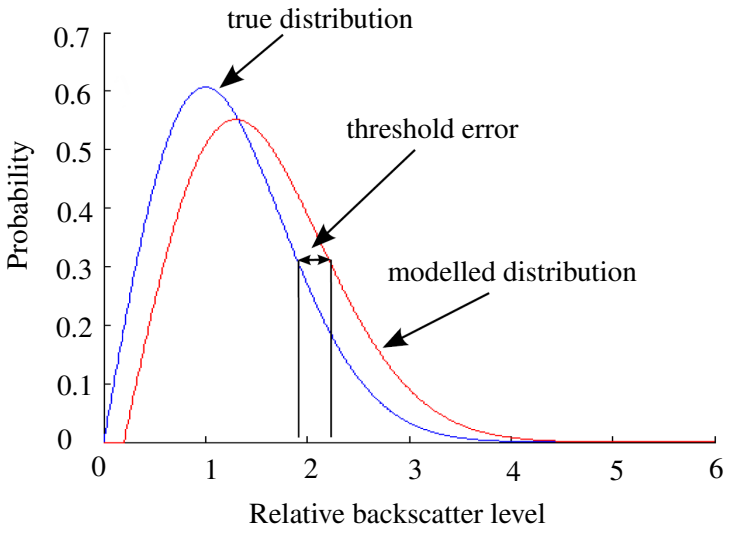

Figure 3: *Threshold error example. The error is determined by the backscatter difference between the true and modelled distributions.*

4 Analysis using simulated data

To effectively quantify the effects of thermal noise and verify the fitting procedures in the previous section, a sea-clutter simulation is now presented based on the single-look KK model presented in Section 3.4. The first Section 4.1 describes the simulation method with a number of examples to show how the different components combine together to form the final result. Section 4.2 then looks at the accuracy of the KK-distribution fitting algorithm described in the previous section. The final Section 4.3 then looks at the K-distribution model and how its accuracy degrades when sea-spikes and thermal noise are present.

4.1 Simulation method

The goal of the sea-clutter simulation is to form a KK-distributed PDF with thermal noise and random speckle. The KK model from Equation 12 is formed by creating two K-distributions where the number of samples for each is specified by the desired ratio between components, k. Also, to correctly model the multi-look formulation in Equation 6, the speckle component is randomly generated for each range bin and pulse, while the underlying RCS is kept constant for each pulse.

Figure 4 shows the simulation procedure as it is implemented in MATLAB. Due to memory constraints, if the total number of points, N is greater than $\tilde{N} = 10^6$, the simulation is repeated ceil$(N/\tilde{N})$ times where ceil$(\cdot)$ is the ceiling function. After the first iteration, the histogram bins, σ_l are defined by the maximum value of clutter power, σ_{max}, the histogram range, σ_{range} and the bin spacing, σ_Δ,

$$\sigma_l = \sigma_{\text{max}} - \sigma_{\text{range}} + (l-1)\sigma_\Delta, \qquad l = 1 \ldots L \tag{18}$$

with the total number of bins, $L = \text{ceil}(\sigma_{\text{range}}/\sigma_\Delta) + 1$. Each iteration of the simulation must therefore align its histogram so the bins line up according to this definition. Any bins which fall outside of the range are truncated.

To correctly simulate the two-dimensional (2D) range/azimuth backscatter matrix, the desired number of azimuth rows, $\tilde{N}_{\rm a}$ and range lines, $N_{\rm r}$ need to be specified to match with the desired ratio, k between the two instances of the K-distribution. If the number of range lines is fixed at $N_{\rm r} = 1024$, then the number of rows for each instance is determined by $N_{\rm a1} = \text{round}\left[(1-k)\tilde{N}_a/N_{\rm r}\right]$ and $N_{\rm a2} = \text{round}\left[k\tilde{N}_a/N_{\rm r}\right]$, with the total number of rows given by $N_{\rm a} = N_{\rm a1} + N_{\rm a2}$. The rounding will slightly change the total number of samples and alter the desired ratio k, but the effect is minor and will not adversely influence the simulation results.

For each azimuth row, a Rayleigh random vector, $\mathbf{R}$ with mode $1/\sqrt{2}$ and a Gamma random variable, G with scale b and shape ν is created. The K-distribution (magnitude) is then formed by the product of $\mathbf{R}$ with the square root of G, while the backscatter phase is modelled as a uniform random vector with elements $\phi \in [0, 2\pi]$ and is included with a dot product. Consider the two KK components, $\mathbf{P}_1$ and $\mathbf{P}_2$ where the r^{th} row can be written as,

$$\begin{aligned}\mathbf{P}_{1,r}\left(\nu_1, b_1\right) &= \mathbf{R}_1\left(1/\sqrt{2}\right)\sqrt{G_1\left(\nu_1, b_1\right)} \odot \exp\left[j\boldsymbol{\phi}_1\right] \in \mathcal{C}^{N_r \times 1},\\ \mathbf{P}_{2,r}\left(\nu_2, b_2\right) &= \mathbf{R}_2\left(1/\sqrt{2}\right)\sqrt{G_2\left(\nu_2, b_2\right)} \odot \exp\left[j\boldsymbol{\phi}_2\right] \in \mathcal{C}^{N_r \times 1}.\end{aligned} \tag{19}$$

The 2D representation is then formed by combining these vectors,

$$\begin{aligned}\mathbf{P}_1\left(\nu_1, b_1\right) &= \left[\mathbf{P}_{1,1}\left(\cdot\right), \mathbf{P}_{1,2}\left(\cdot\right) \ldots \mathbf{P}_{1,N_{\rm a1}}\left(\cdot\right)\right] \in \mathcal{C}^{N_r \times N_{a1}},\\ \mathbf{P}_2\left(\nu_2, b_2\right) &= \left[\mathbf{P}_{2,1}\left(\cdot\right), \mathbf{P}_{2,2}\left(\cdot\right) \ldots \mathbf{P}_{2,N_{\rm a2}}\left(\cdot\right)\right] \in \mathcal{C}^{N_r \times N_{a2}}.\end{aligned} \tag{20}$$

The KK-distributed vector of size $N_r \times N_a$ is then formed by appending the two vectors, $\mathbf{Q} = [\mathbf{P}_1, \mathbf{P}_2]$. The next step is to include the thermal noise, $\mathbf{T}$ which is created by simulating two Gaussian random vectors, $\mathbf{T}_1$ and $\mathbf{T}_2$ with zero mean and variance determined by the square root of half the desired noise power, p_n,

$$\mathbf{T}(p_n) = \left[\mathbf{T}_1\left(0, \sqrt{\frac{p_n}{2}}\right) + j\mathbf{T}_2\left(0, \sqrt{\frac{p_n}{2}}\right)\right] \in \mathcal{C}^{N_{\rm r} \times N_{\rm a}}. \tag{21}$$

The clutter plus noise vector, $\tilde{\mathbf{Q}}$ is then formed by summing the two vectors,

$$\tilde{\mathbf{Q}}\left(\nu_1, \nu_2, b_1, b_2, p_n\right) = \mathbf{Q}\left(\nu_1, b_1\right) + \mathbf{T}(p_n) \in \mathcal{C}^{N_{\rm r} \times N_{\rm a}} \tag{22}$$

Multi-looking using a spectral method with M looks can now be performed by first taking the inverse fast Fourier transform (FFT) along the range dimension of the data. The spectrum is then split into M equal parts and the FFT taken for each part. They are then converted to the power domain by taking their absolute square and averaging to produce the multi-look result. As a result of this processing however, the range resolution will also be reduced by a factor M. Finally, due to the large dynamic range of the backscatter powers, the simulated data is converted to a dB representation before the PDF is formed.

Figures 5-8 show examples of the simulation procedure using the parameters in Table 4 and with all histograms presented in the dB domain. Figure 5 shows a comparison of simulated and theoretical histograms for the two simulated K-distributed components, $\mathbf{P}_{1,\rm dB}$, $\mathbf{P}_{2,\rm dB}$ and the final KK-distributed component, $\mathbf{Q}_{\rm dB}$. The histogram of the first component has its mean at -20 dB as is expected, while the second is shifted by 7 dB to be located at -13 dB. The histogram of the final component clearly models the KK-distribution with a large deviation on the right hand

side. The second Figure 6 shows histograms of the sea-clutter, $\mathbf{Q}_{\text{dB}}$, thermal noise, $\mathbf{T}_{\text{dB}}$, and the sea-clutter plus thermal noise, $\tilde{\mathbf{Q}}_{\text{dB}}$. The effect of the thermal noise is to shift the distribution slightly to the right, according to the sum of means (in the linear domain) of the two distributions, $10\log_{10}\left[10^{-40/10}+10^{-30/10}\right] = -29.59$ dB. The third Figure 7 shows the effect as the noise power is increased and the CNR decreases from 10 dB to -10 dB. The distribution shifts further to the right, becomes more Rayleigh and the effect of the deviation in the tail becomes less pronounced. The fourth Figure 8 shows the simulated PDF for 1, 2 and 4-looks. As the number of looks increases, the PDF compresses in its dynamic range from 70 dB, down to 40 dB for 2-looks and 30 dB for 4-looks. The deviation in the tail also becomes more pronounced as the number of looks increases. Note that the result in Figure 28 later in the report shows that the ratio of means actually reduces due to the multi-looking and this pronounced deviation is not present.

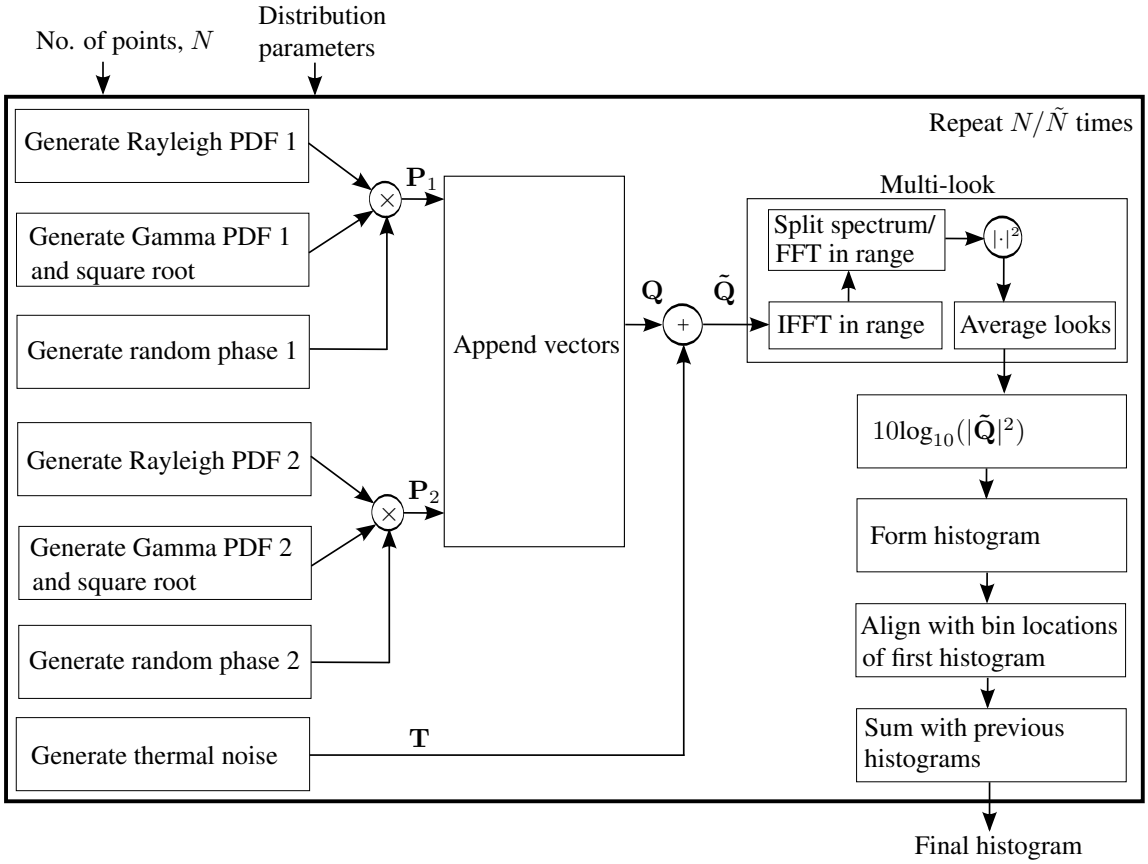

Figure 4: *Simulation block diagram.*

Table 4: *Simulation parameters.*

Parameter	Value
Mean, μ_1	-20 dB
Shape, ν_1, ν_2	5
Ratio between components, k	0.01
Ratio of means, ρ	5
Noise power, p_n	-30 dB
Number of samples, N	10^7
Dynamic range of histogram	70 dB
Histogram bin spacing, σ_Δ	0.1 dB

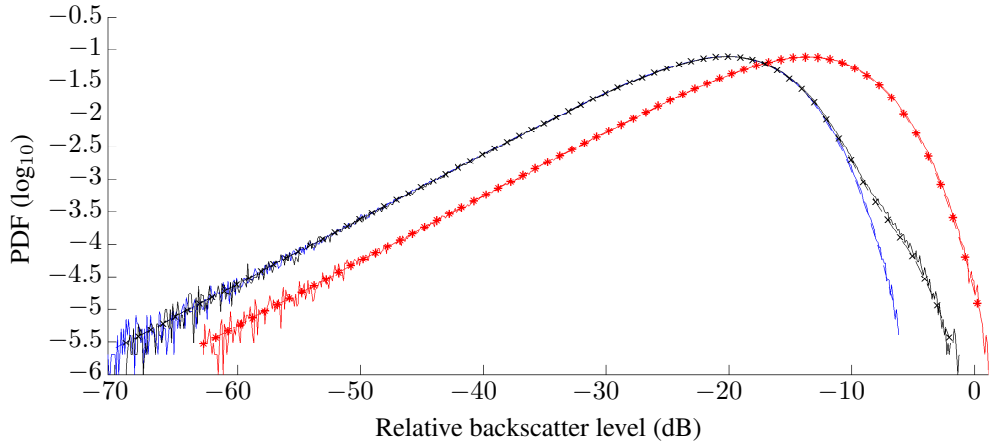

Figure 5: *Simulated and theoretical single-look PDFs (log_{10}).: (—) $\mathbf{P}_{1,dB}$, (-*-) $\mathbf{P}_{2,dB}$, (-×-) $\mathbf{Q}_{dB}$. Dashed lines show the theoretical PDFs.*

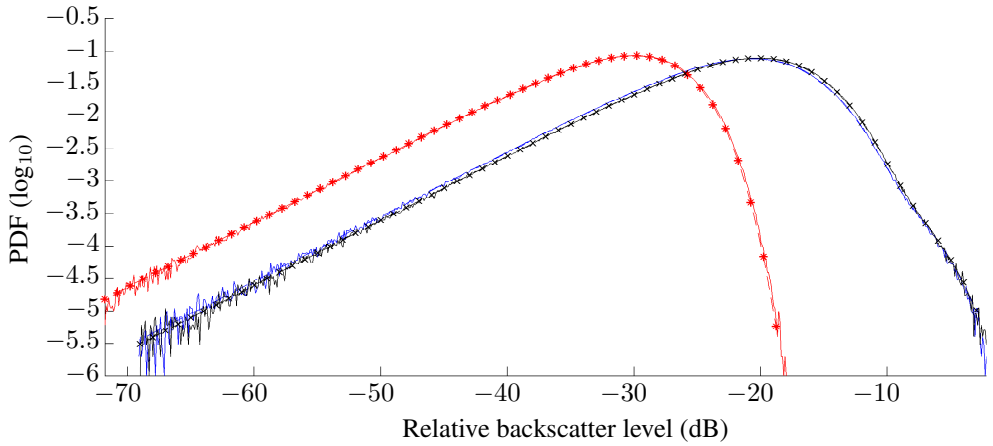

Figure 6: *Simulated and theoretical single-look PDFs (log_{10}): (—) $\mathbf{Q}_{dB}$, (-*-) $\mathbf{T}_{dB}$, (-×-) $\tilde{\mathbf{Q}}_{dB}$. Dashed lines show the theoretical PDFs.*

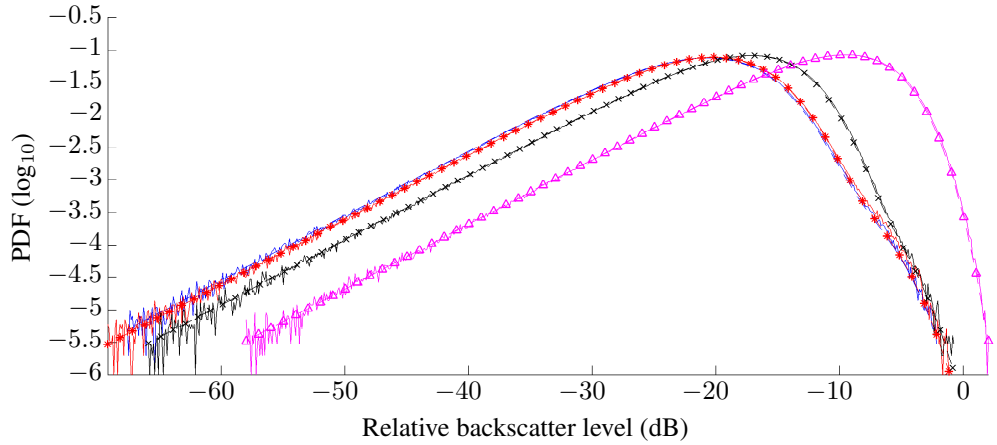

Figure 7: *Simulated and theoretical single-look PDFs (log_{10}) with CNR (dB): (—) ∞, (-*-) 10, (-×-) 0, (-Δ-) −10. Dashed lines show the theoretical PDFs.*

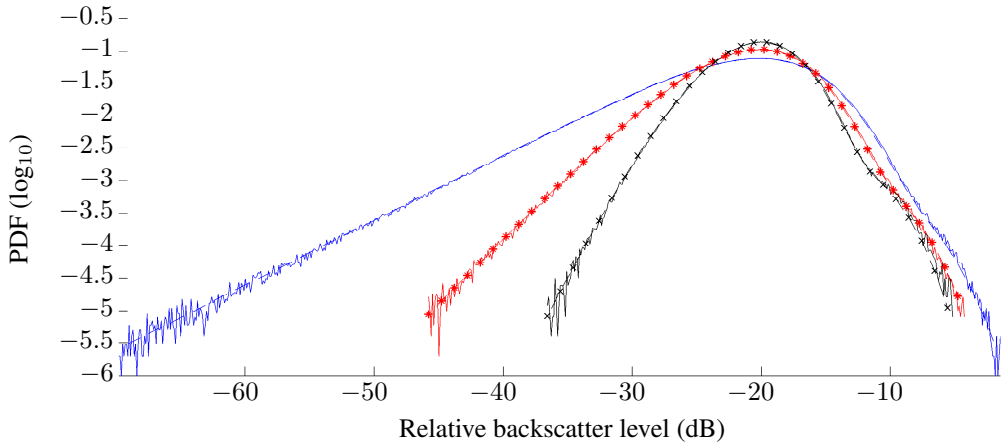

Figure 8: *Simulated and theoretical multi-look PDFs (log_{10}), $\mathbf{H}_{dB}$: (—) 1 look, (-*-) 2 looks, (-×-) 4 looks. Dashed lines show the theoretical PDFs.*

4.2 Fitting accuracy

To verify the fitting accuracy of the KK plus thermal noise fitting algorithm described in Section 3.4, a series of Monte Carlo simulations were run and the input parameters compared to the estimated ones. The range of valid simulated values which are randomly selected for each iteration are shown in Table 5, with the clutter and noise power fixed to provide a constant CNR. A number of statistical goodness of fit measures were considered such as the chi-squared and Kolmogorov-Smirnov tests. However, due to the focus on the tail of the distribution, the data is fitted in the log domain and these tests, which are applied in the linear domain and are not suitable. As a result, the normalised root mean square (NRMS) error was calculated between the model fit, $\tilde{\mathbf{Q}}'$ and the data, $\tilde{\mathbf{Q}}$ in the log domain. The metric g_{NRMS} varies between zero and one, where closer to zero the better the fit. If there are N points,

$$g_{\text{NRMS}} = \frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\tilde{\mathbf{Q}}'_n - \tilde{\mathbf{Q}}_n\right)^2}}{\max\left\{\tilde{\mathbf{Q}}'\right\} - \min\left\{\tilde{\mathbf{Q}}'\right\}}. \tag{23}$$

Two examples are compared, the first with no thermal noise present and the second with a noise power of -25 dB giving a CNR of 5 dB. After 1000 iterations the no-noise example has 99.9% of the fits with an NRMS error of less than 0.02, while 100% had an error of less than 0.05. With thermal noise present, the former reduces slightly to 99.2%. This demonstrates that the procedure is working very well at fitting the model to the data.

The second goodness of fit measure is based on measuring how accurately the estimated parameters match the desired input. The scatter plots in Figure 9 show the results of 1000 iterations where the estimated parameters which match the input parameters lie close to the diagonal line. The plots on the left show results with no thermal noise present, while those on the right show results with thermal noise. From top to bottom, the scatter plots show the result from the shape, ratio of means and the inverse of the shape.

Along the first row, the shape estimate is a good match for small values while higher values show some variation. With no thermal noise present, the spread shown on the left starts deviating from around 8 and is slightly biased upwards, with many hitting the ceiling of 25. With thermal noise present, the spread on the right starts deviating around 5 with a slightly wider spread. However, since the model still fitted the data with a high level of accuracy, this result also indicates that due to the thermal noise in the data, the overall distribution is less sensitive to higher shapes. In addition, the correlation coefficient has been calculated between the input and output estimates. The correlation value for this example is 0.98 for the no-noise case and 0.95 for the noisy case. As both values are close to 1, this result indicates that the correlation between the input and estimated output is high for both cases.

The second row shows the ratio of means for each case with a good match for nearly all points. The results both show a slight positive bias with values above 5 but still demonstrates a good match. The correlation values are 0.99 for both cases indicating a high correlation.

The third row then shows a comparison of the inverse of the shape, which has been proposed as a more physical quantity related to the square of the underlying cross section, [Lombardo, Oliver & Tough 1995]. Both scatter plots now have a much better match for nearly all points with high correlation values of 0.99 for both cases.

Table 5: *Simulation parameter range for testing fit accuracy.*

Parameter	Value
Mean, μ_1	-20 dB
Shape, ν_1, ν_2	0.1 to 25
Ratio between components, k	0.01
Ratio of means, ρ	1 to 15
Noise power, p_n	-∞ or -30 dB
Number of samples, N	10^7
Dynamic range of histogram	70 dB
Histogram bin spacing, σ_Δ	0.1 dB

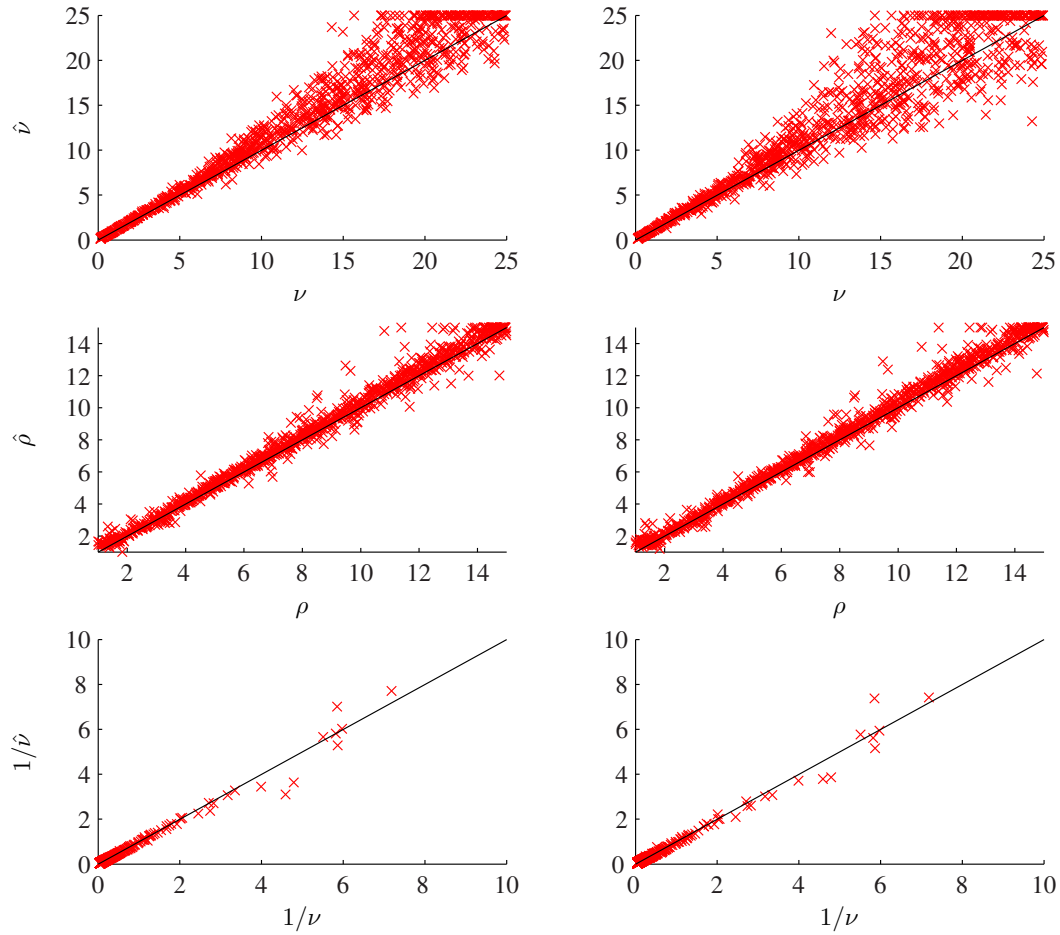

Figure 9: *Goodness of fit scatter plot for 1000 iterations where the best estimates lie close to the diagonal line. From top to bottom: shape, ratio of means, and 1/shape. Left: no thermal noise, right: with thermal noise.*

4.3 The effect of thermal noise and sea-spikes on the K-distribution

If the radar backscatter is assumed to be K-distributed and there are both sea-spikes present in the sea-clutter and thermal noise present in the radar, then the threshold required for a constant P_{fa} will be inaccurate. To measure this difference, the KK-distribution with thermal noise is assumed to be the true representation of the radar backscatter and the threshold error is measured between the data and a K-distribution fit with a P_{fa} of 10^{-5}. For the first part of this section, the KK-distributed shape is fixed at $\nu_1 = \nu_2 = 5$ and the thermal noise and ratio of means are varied.

Figure 10 shows the threshold error in dBs with no error on the far left where the ratio of means is 1. It then increases up to 8.5 dB as the ratio of means increases to 15. The variation in threshold error is small with changes in CNR in the clutter dominant region, CNR $>$ 0 dB. When the ratio of means is 15 however, the threshold error varies from 8.5 dB when the CNR is 30 dB down to 7 dB when the CNR is 0 dB. Then in the noise dominant region, the error drops significantly to 1 dB at CNR = -10 dB.

The next result looks at the estimation accuracy of the underlying shape when thermal noise is present and there is a large deviation in the tail, $\rho = 5$. A number of simulations were run

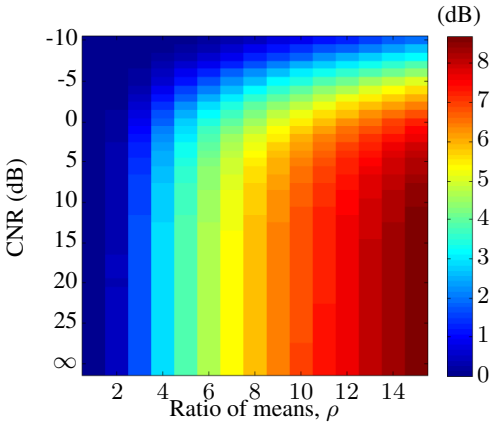

Figure 10: *Threshold error with shape = 5, varying CNR and the ratio of means.*

with variations in both the underlying shape, $\tilde{\nu}$ and the CNR. Both the X-statistic and the KK-distribution fit were then applied to compare estimates of the underlying shape with the simulated values. The first plot in Figure 11 shows the result from using the X-statistic to estimate the effective shape and then modified using the relationship from Equation 10 to determine the underlying shape. It is clear that the estimate of the underlying shape is consistently less than the simulated value, particularly in the region where the underlying shape is greater than 5. Also, in the noise dominant region where CNR $<$ 0 dB, the underlying shape approaches 0 and the result becomes meaningless. The former errors are not unexpected however, as this relationship was derived assuming a K-distribution and does not consider the case when sea-spikes extend the tail of the distribution. The second plot then shows the estimated underlying shape from fitting the KK parameters using the new method described in Section 3.4. The result is now a reasonably good match to the underlying trend. The estimate starts to fall down however as the noise starts to dominate in the region where CNR $<$ -3 dB.

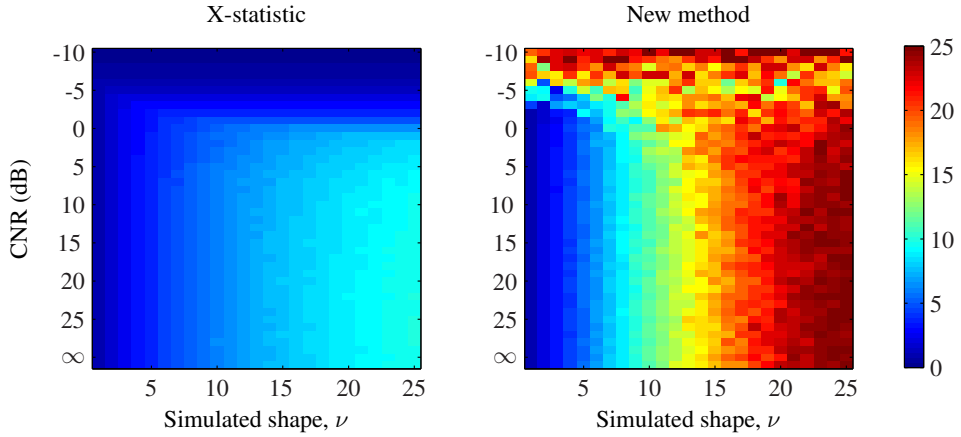

Figure 11: *Underlying shape comparison with ratio of means = 5, varying CNR and underlying shape. Left plot shows results based on X-statistic estimate. Right plot shows shows the results with the KK parameter estimate.*

4.4 KK-distribution sensitivity analysis with thermal noise

To measure the impact of errors in the thermal noise estimate, a sensitivity analysis was performed on the shape and ratio of means. A number of simulations were run with a shape of 5, ratio of means of 5 and three CNRs: 0, 5 and 10 dB. The noise power used in the KK estimation algorithm was then varied away from the true value and the parameters estimated. Figure 12 shows the mean difference between the shape and ratio of mean estimates.

For the shape variation, $\Delta\nu$, the deviation with a CNR of 0 dB has a significant mean error of more than -2 when the noise deviation is 0.5 dB and a slight error/bias of 0.5 when the noise deviation is not present. However, when the CNR is 5 dB, the maximum mean error reduces to 0.7, and for the higher CNR of 10 dB, shows only a very small mean error. This result demonstrates that shape estimates in the low CNR regions are quite sensitive to incorrect noise values. The ratio of means deviation, $\Delta\rho$ shows only a small mean error for each CNR, even when the noise deviation is 0.5 dB which demonstrates that the ratio of means is quite robust. In the data sets analysed in this report, the relative noise estimate error is less than 0.1 dB, so the effect of thermal noise should not significantly bias the parameter estimates of the shape and ratio of means.

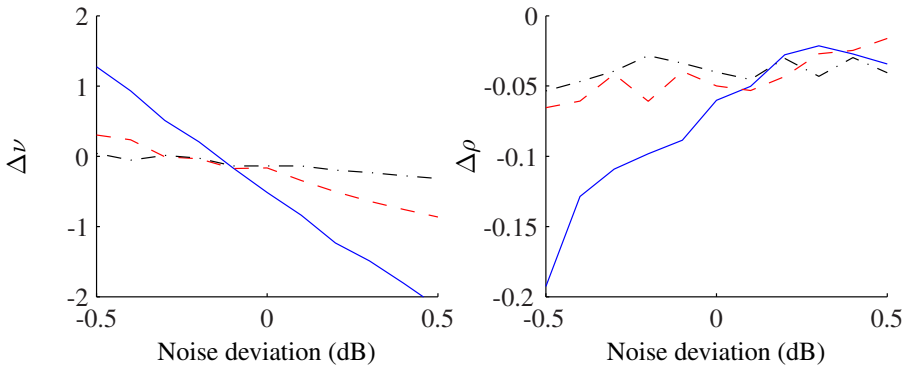

Figure 12: *Noise sensitivity analysis shows the mean difference between the shape (left) and the ratio of means (right), CNR = (—) 0 dB, (-.-) 5 dB, (- -) 10 dB.*

5 Analysis using Ingara backscatter data

In this section, data collected by the DSTO's Ingara radar is now analysed to verify the suitability of the KK-distribution with thermal noise. The first Section 5.1 describes the pre-processing required to get the data into the correct form for analysis. This includes the pre-processing of the trial data, formulation of the noise only data and details on the averaging used to achieve PDFs with the required number of samples. The data analysis in Section 5.2 then looks at a single representative day from each of the two sea-clutter trials. The different parameters of interest are studied over the three polarisations, multiple looks and a range of grazing and azimuth angles. The final Section 5.3 looks at the threshold error associated with the data, a typical K-distribution and then the KK-distribution with noise.

5.1 Ingara data pre-processing

Before the data is analysed, a number of pre-processing steps must be applied. Firstly, to avoid biasing the statistics, the data set is first scanned to identify and remove any 'bad' regions. These include some transmit-off regions which were sampled and any other artifacts due to the radar collection. This is done visually and less than 1% of data is removed in this way. The raw data then goes through the processing chain shown in Figure 13. Range processing occurs in hardware as a stretch process. The sampled signal is then processed to first remove bandpass filter modulations and adjust for motion compensation using both the inertial navigation unit and the global positioning system onboard the radar platform. The next steps include a radiometric correction due to changes in grazing angle and removal of the elevation beampattern. Using observations from data of a land calibration site, the data is then calibrated using the procedure described in [Quegan 1994]. During the calibration procedure, reciprocity is assumed to hold and the two cross-polarised channels are averaged giving a single cross-polarised channel.

5.1.1 Thermal noise

The thermal noise power is found by processing part of the collection where the RF transmitter was turned off. In these regions, the sampled signal (or raw data) then represents the thermal noise. By assuming the sea-clutter and thermal noise are independent throughout the processing chain, a reproduction of the processing chain which was applied to the clutter plus noise signal can be used to form an estimate of the noise power. This is complicated however as the transmit off data was collected once and then assumed constant over the entire day. Therefore a different procedure is required to calculate the thermal noise data for each possible collection geometry.

Figure 14 shows this procedure split into two stages. The first stage involves processing the data from regions when the transmitter is off. The data follows the processing chain in Figure 13 with the exception of the geometric dependent terms. After the noise data has been calibrated, it is then averaged over all angles and over each run which has transmit off sections to form a calibrated 'pseudo' mean noise power.

The second processing stage then uses this value to create a simulated noise signal in the backscatter coefficient domain with the same number of pulses and range bins as the clutter plus noise signal. This is achieved using the noise form in Equation 21 with the 'pseudo' mean noise

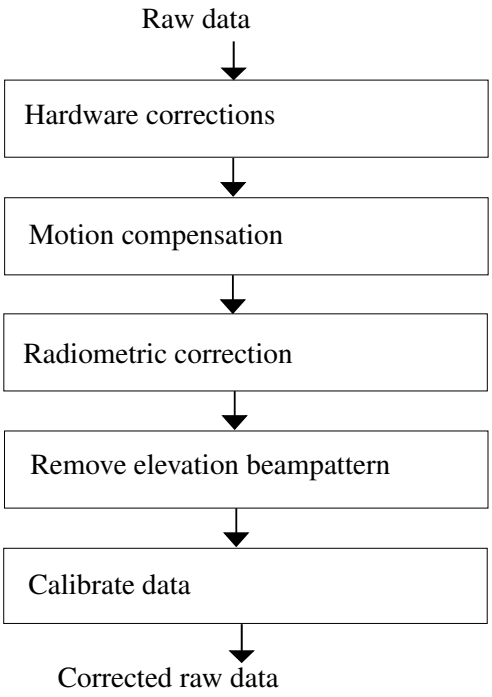

Figure 13: *Ingara pre-processing diagram.*

power equal to p_n. The simulated noise data then has the same geometric compensations (radiometric correction and elevation beam pattern compensation) applied as the clutter plus noise signal in Figure 13. To maintain this relationship with the clutter plus noise signal, any further processing due to multi-looking and data pooling (described below) is also applied to the noise only signal.

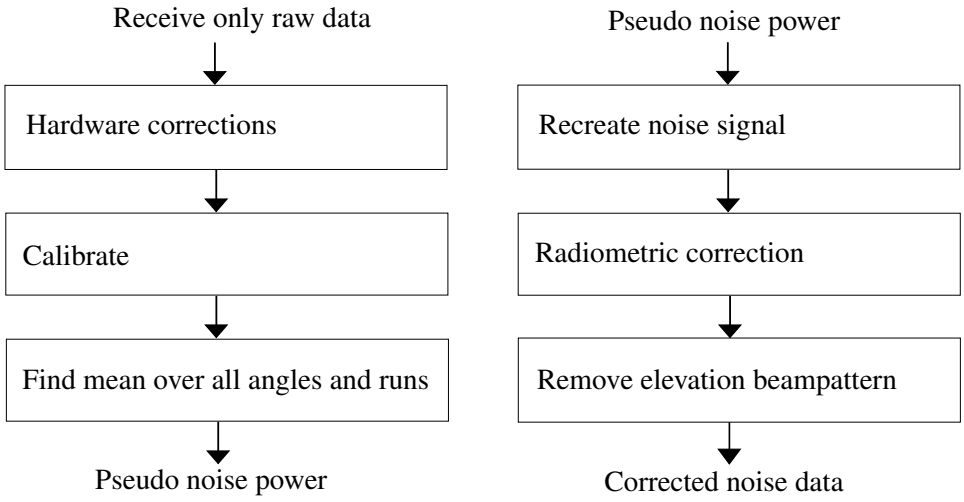

Figure 14: *Thermal noise pre-processing diagram, left: stage 1, right: stage 2.*

5.1.2 Resolution reduction and multi-looking

Multi-looking is a technique used in radar signal processing to reduce the fluctuation of the radar backscatter in order to better detect targets of interest. For the case of sea-clutter, this means reducing the magnitude of both the speckle and the sea-spikes which will in turn alter the observed PDF of the backscatter. Typically in SAR, multi-looking is implemented in the azimuth frequency domain, but it can also be done in the range frequency domain. This is implemented by taking an inverse FFT in range, splitting the data up into M blocks and then forming the range FFT of each block. The average of the absolute square of the data blocks then contains less variation in the radar backscatter than the equivalent single-look data. The effect of this multi-looking implementation will coarsen the range resolution instead of azimuth, which is desirable in a real-beam radar where the azimuth resolution is much coarser than range.

5.1.3 Forming histograms

In order to form PDFs with a tail that reaches down to 10^{-6}, a large number of data samples must be used. For example, to ensure more than 10 samples lie above the threshold of CFAR detector with a false alarm rate of 10^{-6}, at least 10^7 samples are needed in total. The first stage of processing is to pool all the (clutter plus noise and recreated noise) data samples which lie in a 1° grazing by 5° azimuth region. Since each data block spans a different physical area in range and azimuth, a weighted average was determined by the amount of the data block which falls into the defined grid. However, the sample sizes produced by this step are not large enough and therefore a second processing stage has been implemented to pool data from neighbouring viewing geometry bins.

The issue with pooling data is that the neighbouring distributions are all slightly different. In particular, assuming the K-distribution is a good model, the neighbouring means and shape parameters will vary slightly. The differing means are easily dealt with since the distributions can be shifted around without altering their inherent shapes. Pooling data with differing shapes is somewhat more problematic since the distribution of the pooled data will no longer be K-distributed.

To reduce the impact of the mixing data from different distributions, an elliptical window can be used to pool the data rather than a square one. If the viewing geometry bins are arranged in an array with horizontal dimension corresponding to azimuth angle and vertical dimension to grazing angle, then the data is pooled by running an elliptical window over the array. To meet the desired P_{fa} levels in this report, an elliptical window is used which covers 15° in grazing and 36° in azimuth. Figure 15 shows a comparison between the shape of the KK distribution for a $1^\circ \times 5^\circ$ square window and a $15^\circ \times 36^\circ$ elliptical window. The larger window has smoothed the spatial variations in the data and produced a larger valid area. Unfortunately with this approach, any data outliers have the potential to significantly alter the final distribution of the pooled data. An example of this can be seen in the downwind direction for the HH channel, where the outlier has skewed the average and the shape of the elliptical window is visible. The exact effect of this windowing on the interpreted phenomenology is also an unknown which can only be resolved by repeating the analysis with more real data and comparing our results with those from smaller bin sizes.

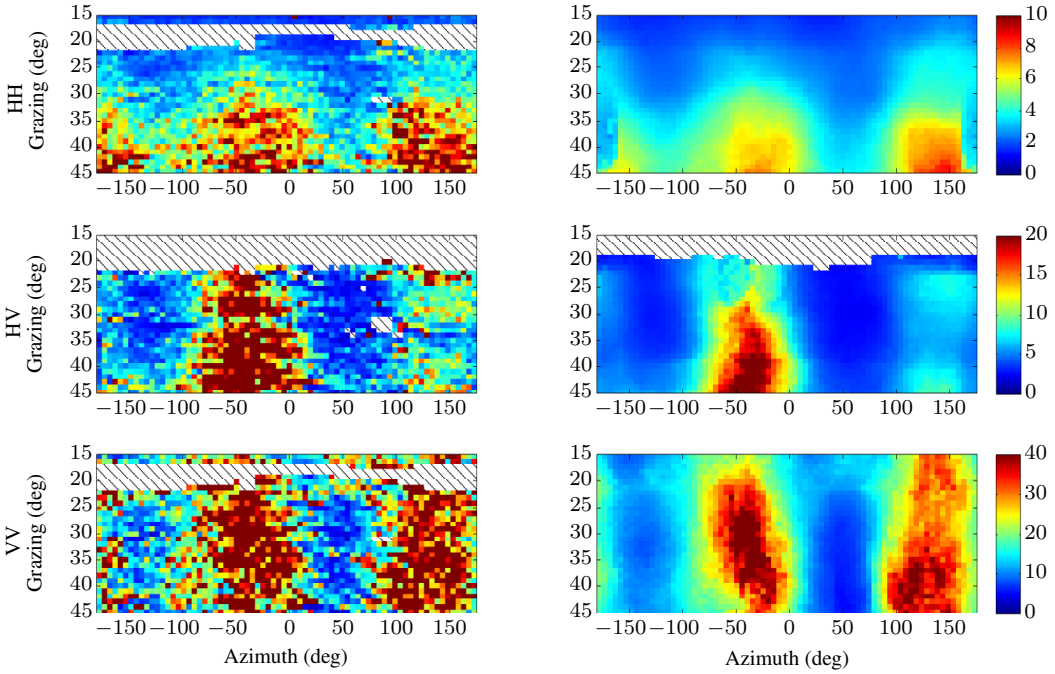

Figure 15: *F35 KK shape comparison with different windows, left:* $1^\circ \times 5^\circ$ *square window, right:* $15^\circ \times 36^\circ$ *elliptical window. Hashed areas indicate regions with poor or missing data.*

5.2 Sea-clutter analysis

To demonstrate the suitability of the KK-distribution fit for sea-clutter analysis, a single representative day is chosen from each of the two sea-clutter trials described in Section 2.3. Both correspond to a Douglas sea-state between 3 and 4 with upwind directions (wind towards the radar) for each day recorded as 115° and 315° respectively. The following results which show grazing against azimuth angle have however been rotated so the upwind direction is centred at 0° with the downwind direction at -180° and the cross wind directions at $\pm 90^\circ$. Any regions with poor or missing data are shown with 'hashed' diagonal stripes.

To limit the number of incorrect estimates due to a noise dominated distribution, only results with a CNR>-3 dB are used in the following analysis. Many of the figures in the following sections have a common scale. If this is located on the right, then the horizontal plots have the same scale and if the scale is on the bottom, then the vertical plots have the same scale. Also, unless otherwise stated, the comparisons below are for the K and KK-distributions with thermal noise.

5.2.1 Single-look

The first group of results focus on the data processed with a single-look, which corresponds to a range resolution of 0.75 m. Figure 16 shows example PDFs from each polarisation with the entire distribution on the left and a close up of the tail on the right. The data is from the upwind direction with a 30° grazing angle and fits are shown for both the K and KK-distributions with thermal noise. The K-distribution fit is found by estimating the mean from the data and the shape using the X-statistic. The thermal noise is then accounted for by offsetting the mean by p_n and

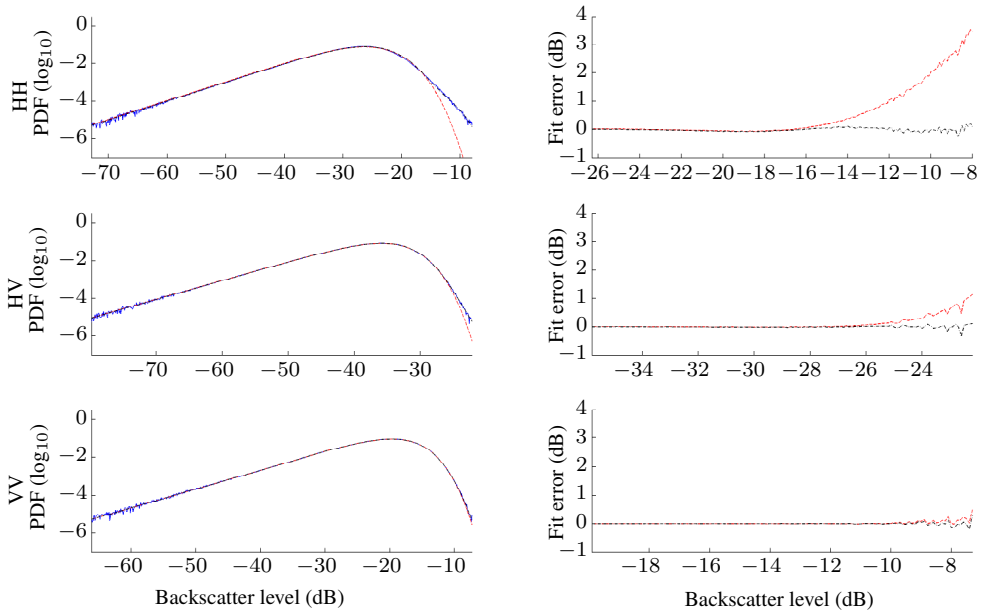

Figure 16: *F35 probability distributions (log_{10}), 30° grazing, upwind, (—) data, (- -) K, (-.-) KK. Left: model fit, right: fit error between data and model fit.*

modifying the shape using the relationship from Equation 10 to determine the underlying shape. The KK-distribution fit is determined using the estimation method outlined in Section 2.4 and uses an optimisation algorithm to find the best fit to the data. It is clear that the K-distribution does not completely match any of the polarisations, with the mismatched tail on the right hand side slightly larger for HH than the other two channels.

The following results are from the F35 data set and includes the calculated noise power, mean, shape and ratio of means for each polarisation. Figure 17 shows the calculated noise power, p_n in dB on the left and the CNR in dB on the right. For the noise power, the HH shows the largest power level at approximately -38 dB, while the VV channel and HV channels are -40 and -41 dB respectively. The CNR on the right shows a high level for both the HH and VV channels with average levels around 5 and 20 dB. There are also noticeable peaks in the upwind direction and to a lesser extent in the downwind direction. For the HV channel, there are a number of regions where the noise is dominant, i.e. CNR < 0 dB. The second figure 18 shows a comparison of the mean estimate with and without accounting for thermal noise. When noise is subtracted from the mean, the dynamic range shown over the collected geometry has extended to lower power levels for both the HH and HV channels and reveals a more complete picture of the mean backscatter.

Figure 19 shows a comparison of both the effective shape estimate for the K-distribution, ν_{eff} and the underlying shape estimates, ν for the K and the KK-distribution fits. These plots clearly show the improvements gained in the shape estimate when first thermal noise is included in the data model and secondly when the tail region has been modelled accurately. The inclusion of thermal noise makes a significant difference in the shape, particularly for the HV channel which has now been reduced in value by approximately 5 units. A sinusoidal trend across azimuth with peaks in the upwind and downwind directions can also be seen for each of the channels. When the second K component is included in the model, the level has now increased closer to the original values where noise was not considered.

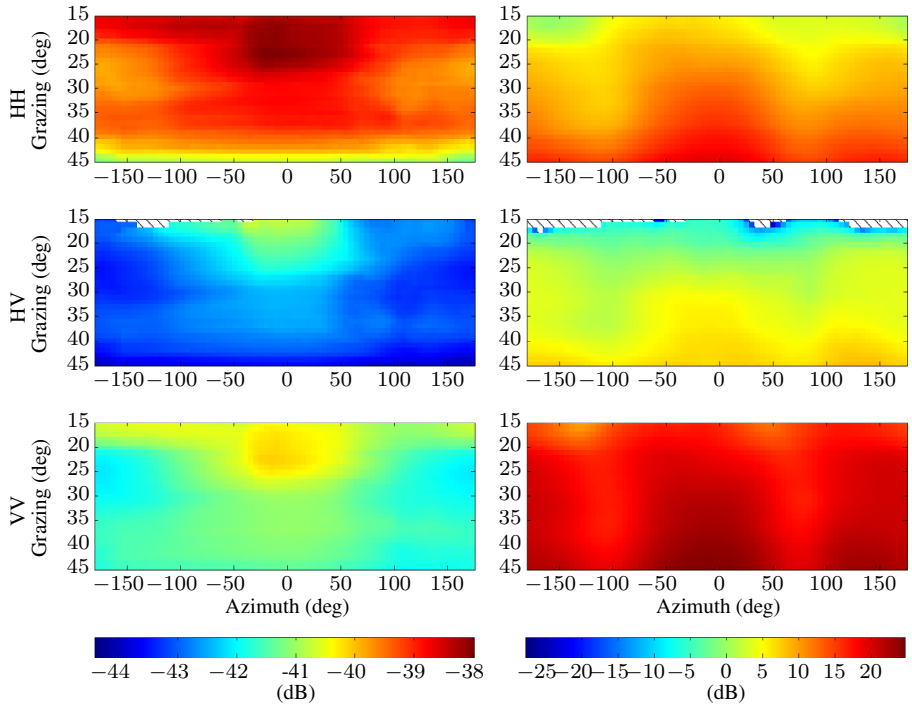

Figure 17: *F35 noise comparison, left: noise power, right: CNR. Hashed areas indicate regions with poor or missing data.*

For the ratio of means estimate, $\hat{\rho}$ shown in Figure 20, the largest ratios are observed in the HH channel at low grazing angles. A trend is observed primarily for the HH channel, where the magnitude decreases almost linearly as the grazing angle gets larger. This indicates that the sea-spike component is more prevalent at lower grazing angles, as is consistent with the literature, [Ward, Tough & Watts 2006]. For the VV channel, the ratio of means is very low with slight peaks in the cross wind directions. This effect is minor however, indicating that the K-distribution is a reasonably good fit to the sea-clutter backscatter received in the VV channel.

The second set of results is for the F9 data set and the same series of plots are now presented. Figure 21 shows the calculated noise power in dB on the left and the CNR in dB on the right. For the noise power, the HH channel shows the largest power level with peak around -34 dB. The VV and HV channels however show a much lower level down around -42.5 dB. This is a significant difference and is due to the calibration solutions which were generated for the 2006 trial. The CNR on the right also shows a number of regions where the noise is dominant, i.e. CNR < 0 dB. These are primarily in the HH and HV channels, while the VV channel has a larger CNR with a maximum of approximately 22 dB. Figure 22 shows a comparison of the mean with and without accounting for thermal noise. These results mirror what is seen in the first data set, with the dynamic range greatly improved for the lower power levels in both the HH and HV channels.

Results for the shape and ratio of means are shown in Figures 23 and 24. The first result shows a significant decrease in the shape for the HV channel, where the thermal noise was greatest. For the underlying shape from the KK estimate, the detail is again improved with significantly larger values in the VV channel. There is an artifact in the upwind direction of the VV channel, due to the averaging process. The ratio of means also shows large values at low grazing angles in the HH channel with a decreasing linear scale as the grazing angle gets larger. Again, for the VV channel,

there are slight peaks in the cross wind directions with the overall level being quite low indicating a near K return from the sea surface.

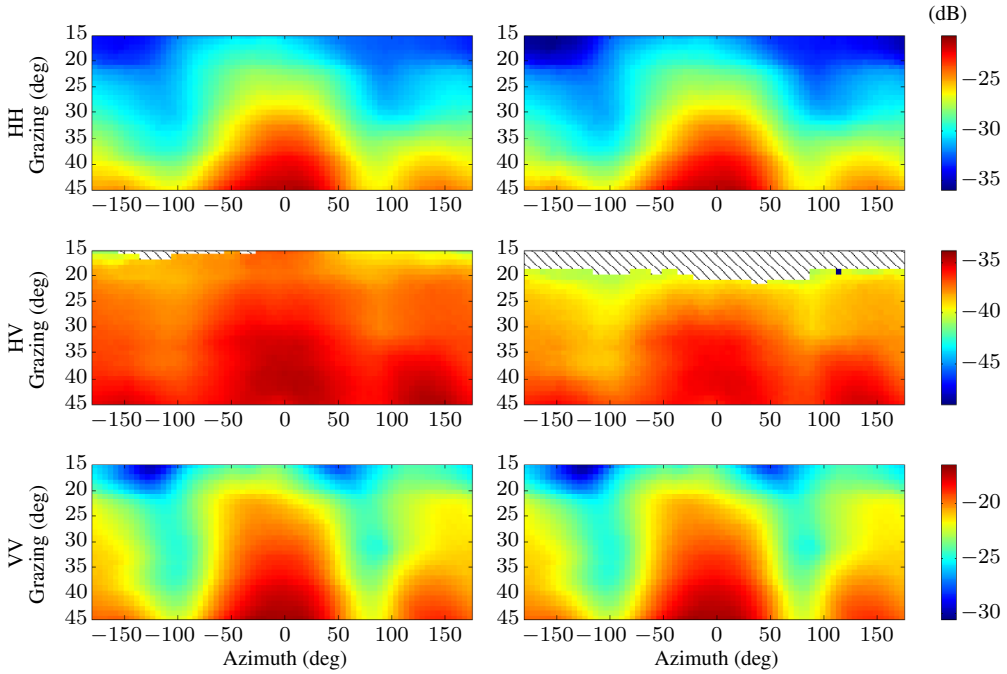

Figure 18: *F35 mean comparison, left: without noise, right: with noise. Hashed areas indicate regions with poor or missing data.*

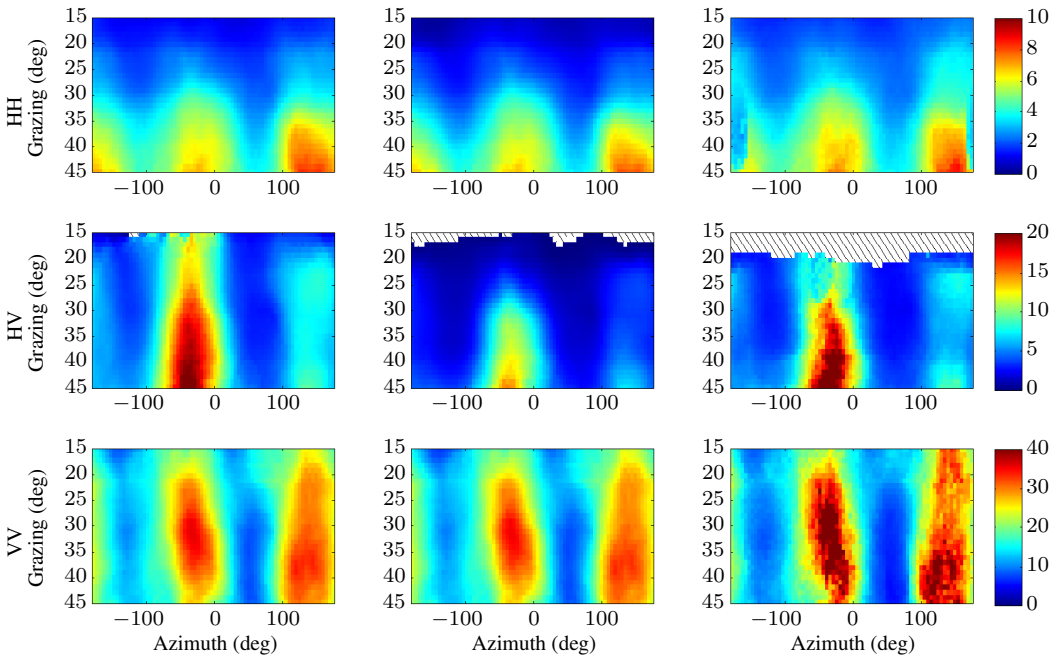

Figure 19: *F35 shape comparison, left: K-distribution, middle: K-distribution with noise, right: KK-distribution with noise. Hashed areas indicate regions with poor or missing data.*

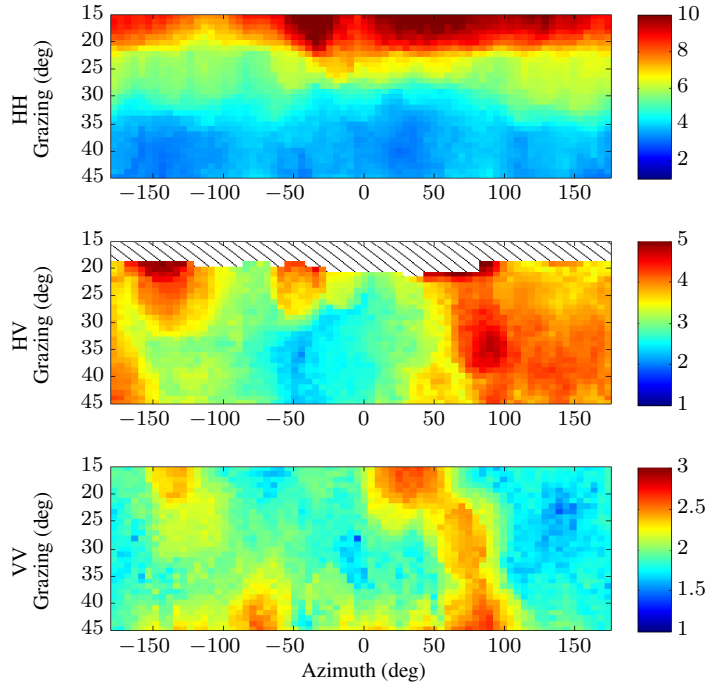

Figure 20: *F35 KK ratio of means. Hashed areas indicate regions with poor or missing data.*

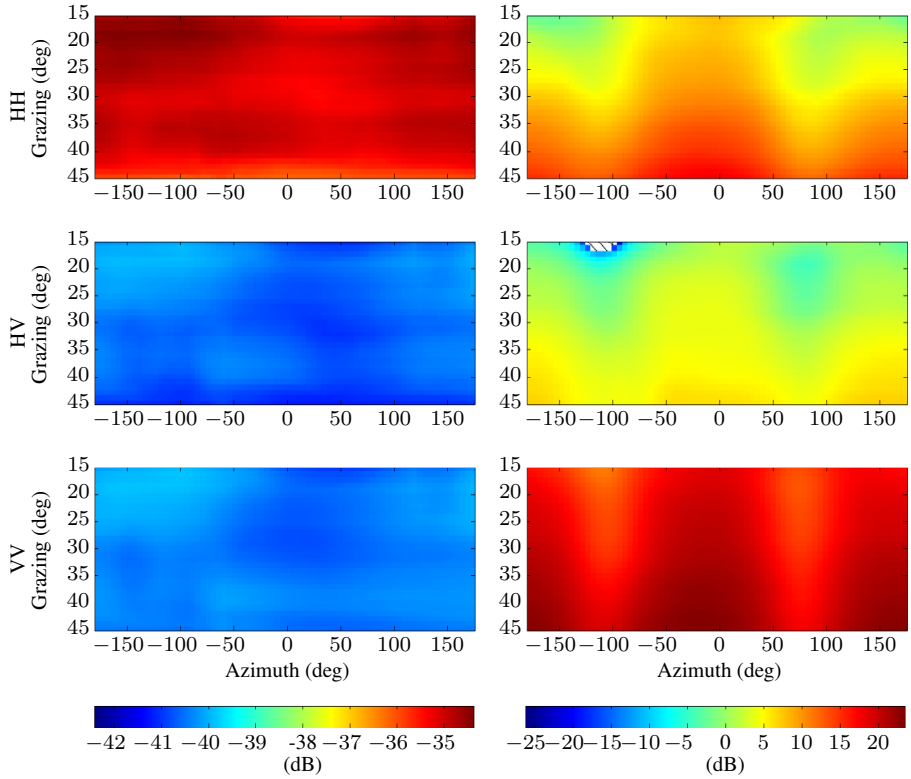

Figure 21: *F9 noise comparison, left: noise power, right: CNR. Hashed areas indicate regions with poor or missing data.*

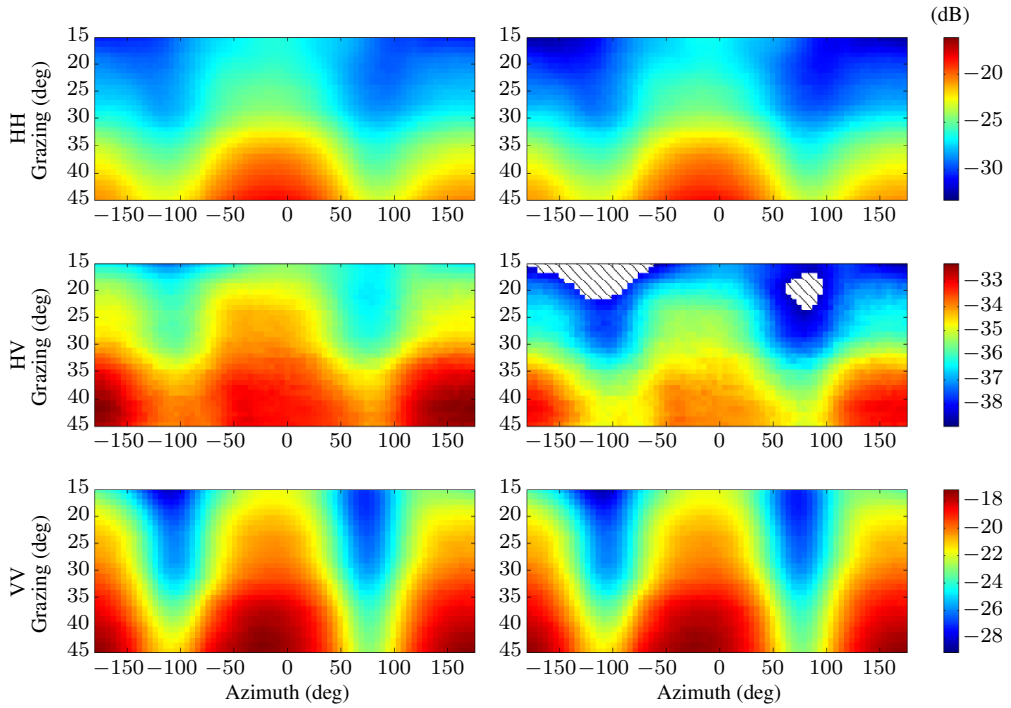

Figure 22: *F9 mean comparison, left: without noise, right: with noise. Hashed areas indicate regions with poor or missing data.*

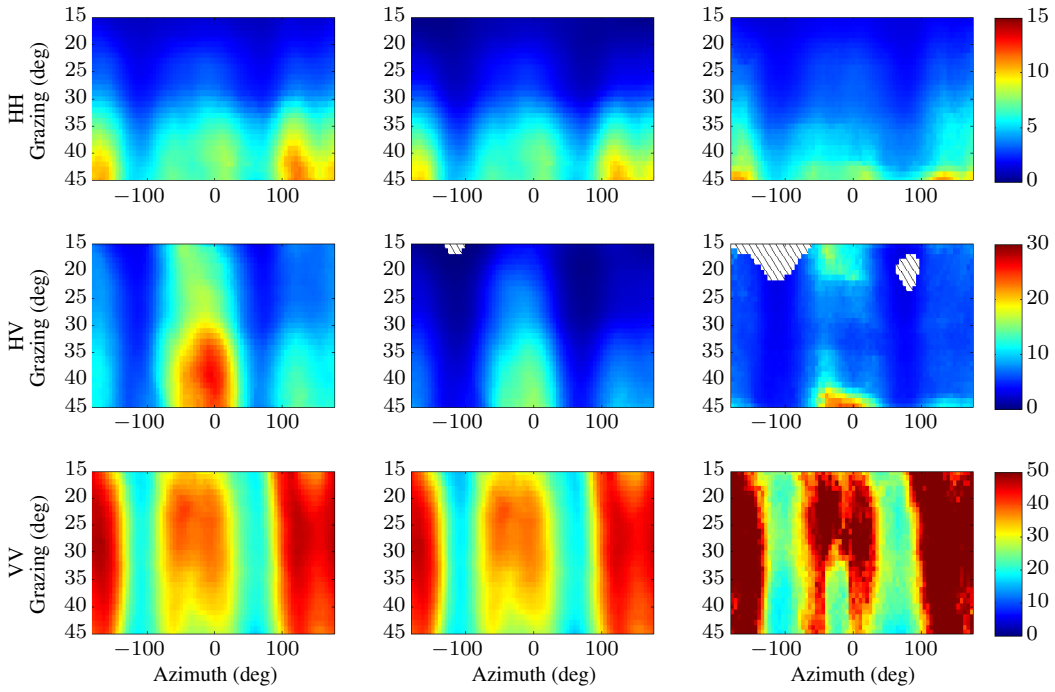

Figure 23: *F9 shape comparison, left: K-distribution, middle: K-distribution with noise, right: KK-distribution with noise. Hashed areas indicate regions with poor or missing data.*

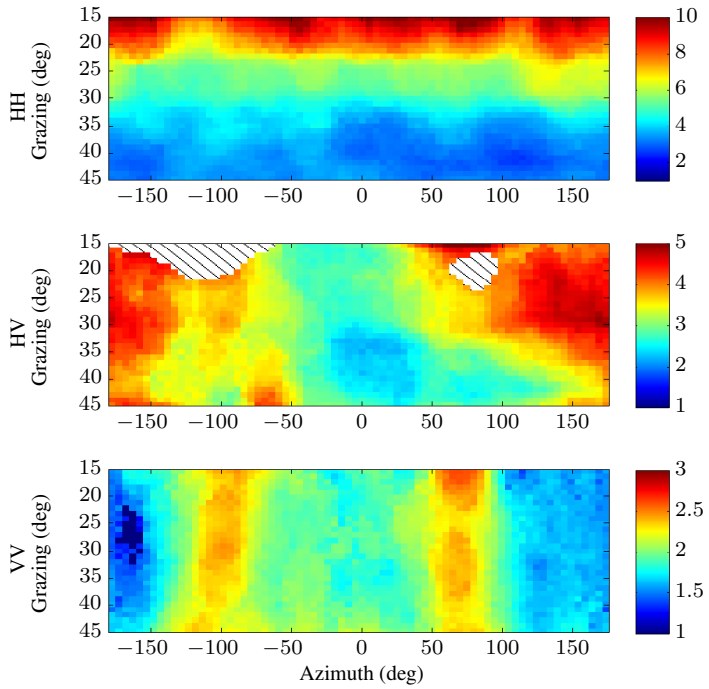

Figure 24: *F9 KK ratio of means. Hashed areas indicate regions with poor or missing data.*

5.2.2 Multi-look

The second group of results focusses on 2 and 4-look data which correspond to range resolutions of 1.5 and 3 m respectively. Examples of the 4-look PDFs for each polarisation are shown in Figure 25. The data is from the upwind direction with a 30° grazing angle and fits are shown for both the K-distribution and the KK model with thermal noise. Similarly to the single-look results, its clear that the K-distribution does not accurately match the HH and to a lesser extent, the HV and VV channels with the mismatched tail on the right hand side. Compared to Figure 8, multi-looking has also changed the distribution by reducing the dynamic range. The K-distribution fit for HH channel also shows a slight mismatch on the left hand side. This minor mismatch is not seen in the KK-distribution fit however.

The next group of results focusses on the shape and ratio of means for the KK-distribution fit plus noise. Figure 26 shows the shape with the 2-look results on the left and 4-look on the right. When compared with the single-look results in Figure 19, the HH and HV shapes have increased slightly for 2-looks and more significantly for 4-looks, while the VV channel has slightly increased in the cross wind directions. The increase in shape is due to the increase in the number of scattering contributions over the larger resolution cell leading to a stronger Rayleigh component in the PDF. This effect can also be seen in the ratio of means in Figure 27, which is reduced as the number of looks has increased. To visualise this, Figure 28 shows a direct comparison of multi-look PDFs for the F35 HH channel. When compared with the simulated results in Figure 6, the deviation between the two K-components is less pronounced due to this reduction in the ratio of means.

Similar results can also be seen for the F9 data set in Figures 29 and 30. The first figure shows the shape for the 2 and 4-look results with increases for the HH and HV channels. The ratio of means also shows consistent results with decreasing values as the number of looks increases.

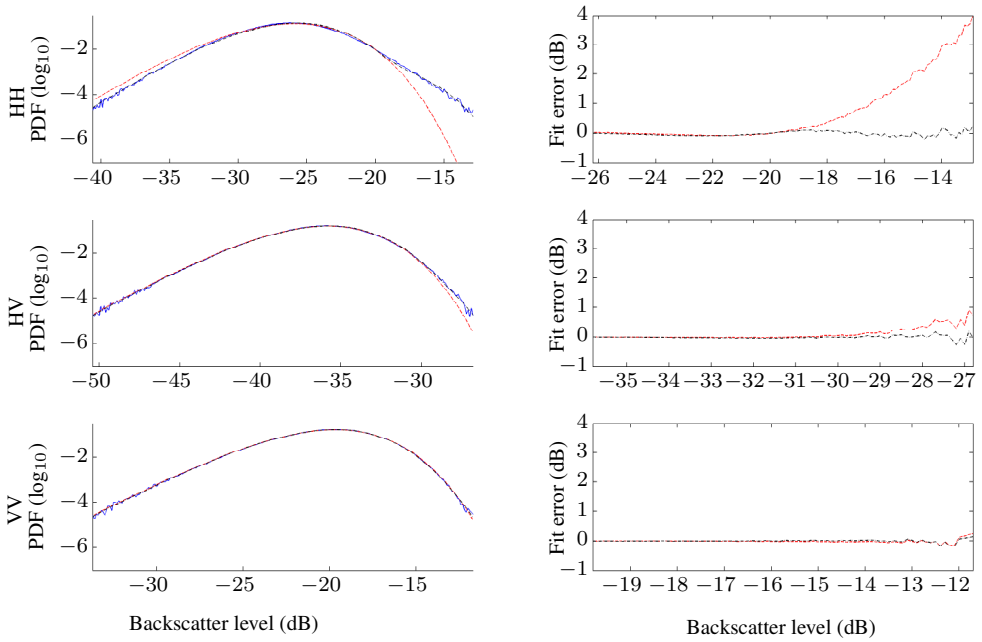

Figure 25: *F35 4-look probability distributions (log_{10}), 30° grazing, upwind, (—) data, (- -) K, (-.-) KK. Left: model fit, right: fit error between data and model fit.*

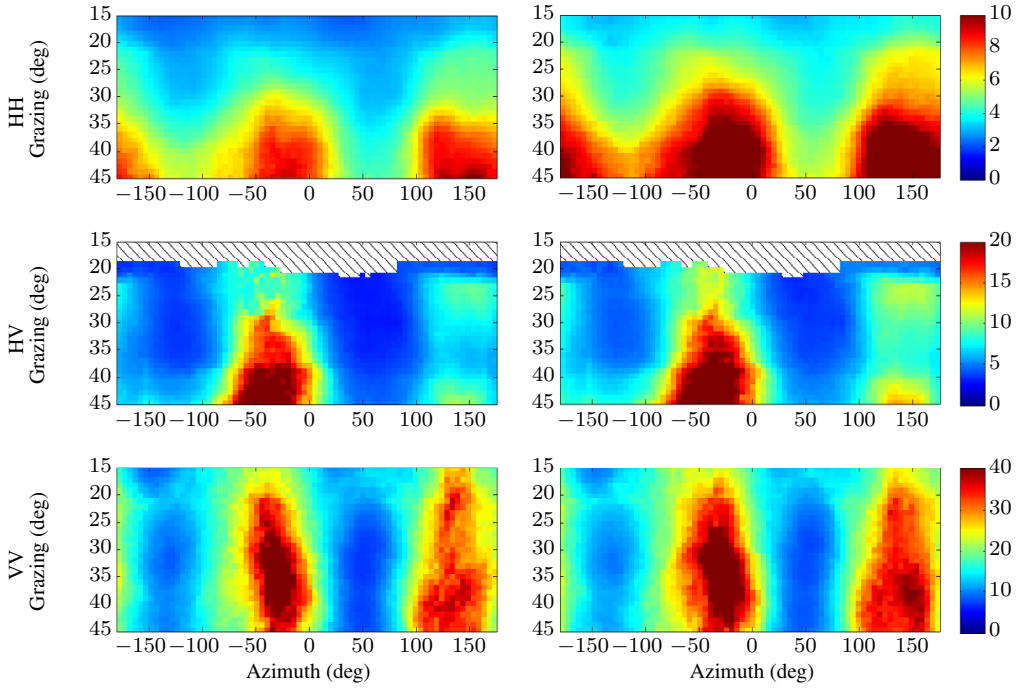

Figure 26: *F35 multi-look shape, comparison for the KK-distribution, left: 2-look, right: 4-look. Hashed areas indicate regions with poor or missing data.*

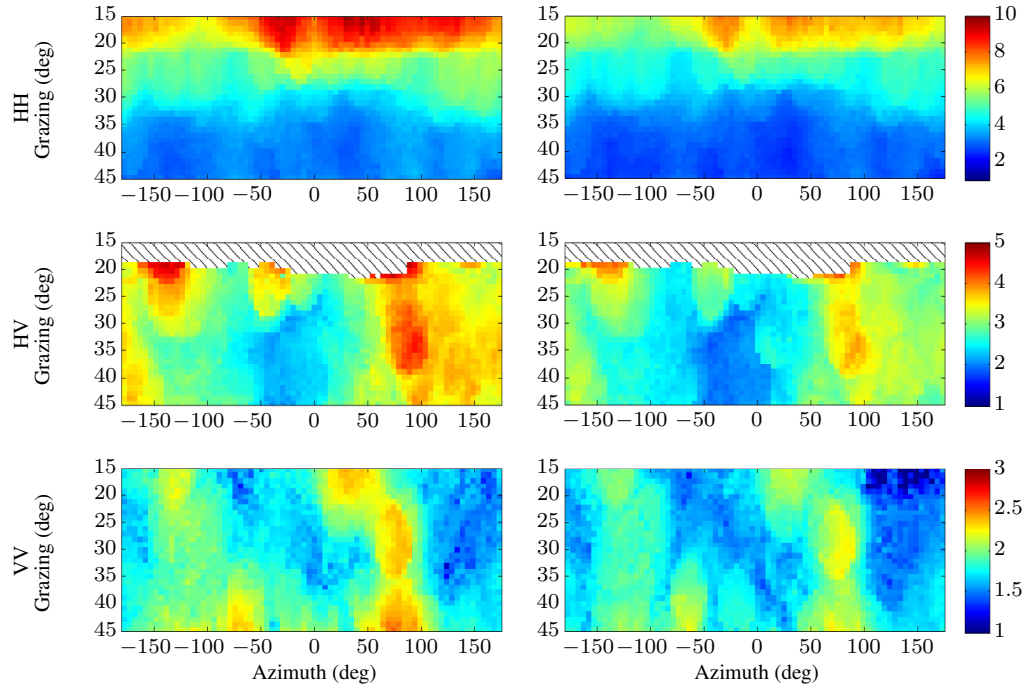

Figure 27: *F35 multi-look KK ratio of means, left: 2-look, right: 4-look. Hashed areas indicate regions with poor or missing data.*

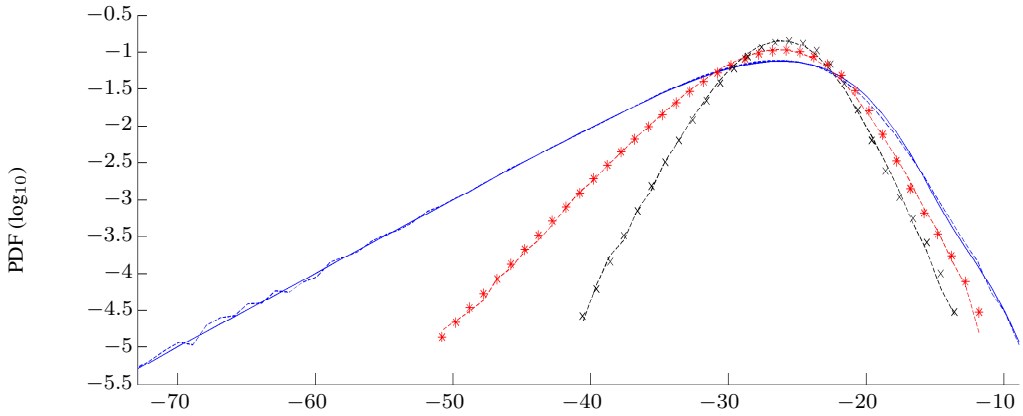

Figure 28: *F35 multi-look probability distributions, data and model fit (log_{10}) for HH polarisation, 30° grazing, upwind. KK model fit: (—) 1 look, (-*-) 2 looks, (-×-) 4 looks. Dashed lines show the data.*

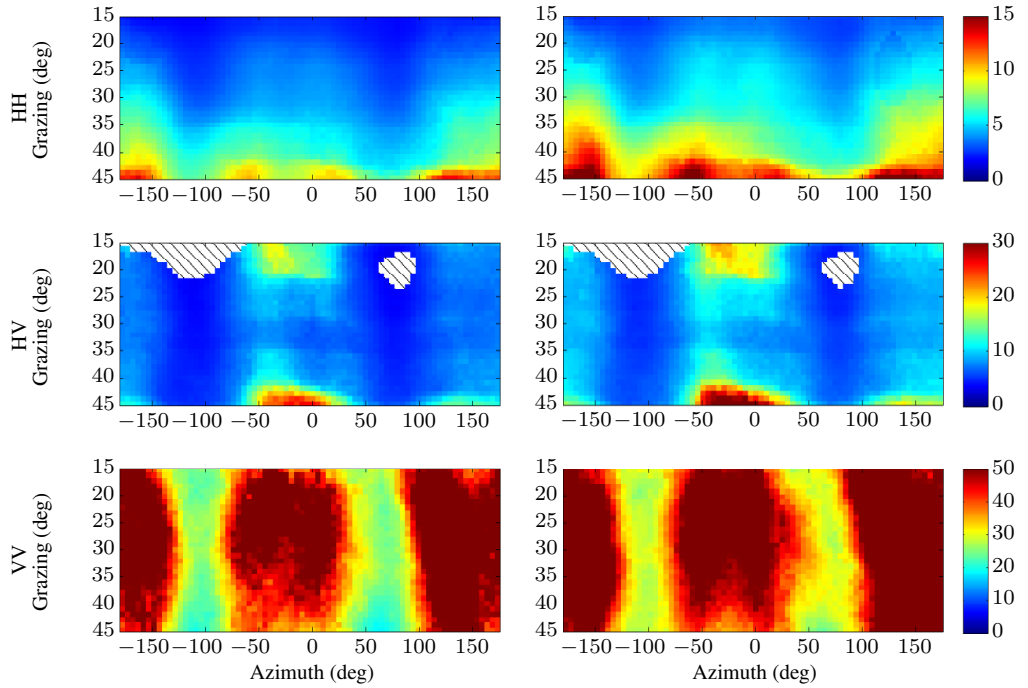

Figure 29: *F9 multi-look shape, comparison for the KK-distribution, left: 2-look, right: 4-look. Hashed areas indicate regions with poor or missing data.*

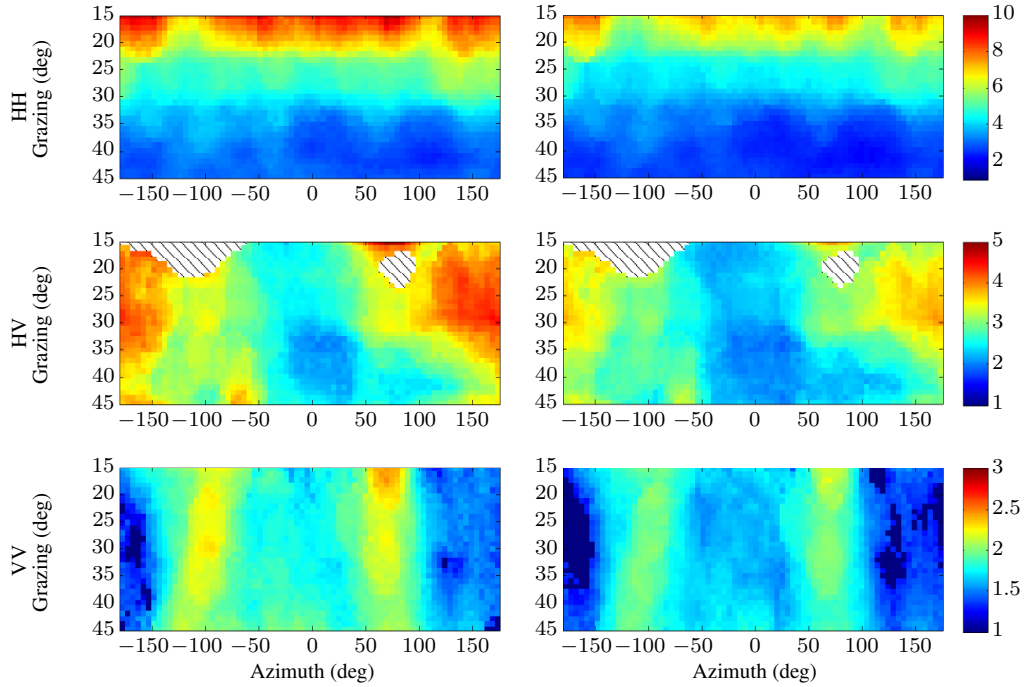

Figure 30: *F9 multi-look KK ratio of means, left: 2-look, right: 4-look. Hashed areas indicate regions with poor or missing data.*

5.3 CFAR threshold errors

Developing a good model for the clutter plus noise probability distribution is important to accurately determine the detection threshold. There is a large amount of literature addressing this problem, a good summary of which can be found in [Ward, Tough & Watts 2006]. The detection threshold is used to separate clutter plus noise from potential targets present in a received radar echo. Determining the optimal threshold can be a complicated procedure and may depend on requirements for the probability of false alarm P_{fa} and detection, P_{d}. For the analysis in this section, only the P_{fa} is relevant as it is related to the clutter plus noise distribution by P_{fa} = 1-CDF.

5.3.1 Single-look

The probability of false alarm and the relative threshold errors are now calculated for the single-look data sets F35 and F9. The left side of Figure 31 shows the F35 data set with P_{fa} on the left hand side in the upwind direction and 30° grazing. The data is shown in blue, the K-distribution fit in red and the KK-distribution fit including thermal noise is shown in black. Clearly, the K-distribution underestimates the true CFAR threshold with an approximate error γ at $P_{\text{fa}} = 10^{-6}$ of 2.8 dB for HH, 1.3 dB for HV and 0.3 dB for VV. The KK-distribution fit is again shown to be more accurate for each polarisation. To further quantify the error from the K-distribution, the right side of Figure 31 shows the relative threshold error as a function of grazing angle for 3 constant false alarm rates, $10^{-4}, 10^{-5}$ and 10^{-6}. For both the HH and HV channels, there is a decreasing error as the grazing angle increases with a maximum error of nearly 4 dB in the HH channel, while the VV channel is relatively constant with only a small error. For all the polarisations, the results for $P_{\text{fa}} = 10^{-4}$ show a smaller error than both $P_{\text{fa}} = 10^{-5}$ and $P_{\text{fa}} = 10^{-6}$ which are very similar. Using a constant false alarm rate of 10^{-6}, Figure 32 shows the relative threshold error calculated over the entire range of azimuth and grazing angles. The mismatch for the K-distribution HH channel is clearly worse for lower grazing angles with a maximum of 6 dB. This is consistent with the estimate for the ratio of means in the previous section demonstrating the region where the effect of the sea-spike component is greatest. For the HV and VV channels however, the relative threshold errors are quite variable. The equivalent mismatch in the KK estimate is very small over the entire region with a mean error on the order of 10^{-2} dB.

For the second data set F9, similar results are shown for the HH and VV channels in Figures 33 and 34. In the first figure, the main differences are the maximum error which is now at 4.3 dB in the HH channel. Figure 34 reveals the greatest threshold error in the HH channel is at low grazing, while the HV channel has largest error in the downwind direction. Interestingly for the VV channel, the result matches the previous data set with the largest error in the cross wind directions.

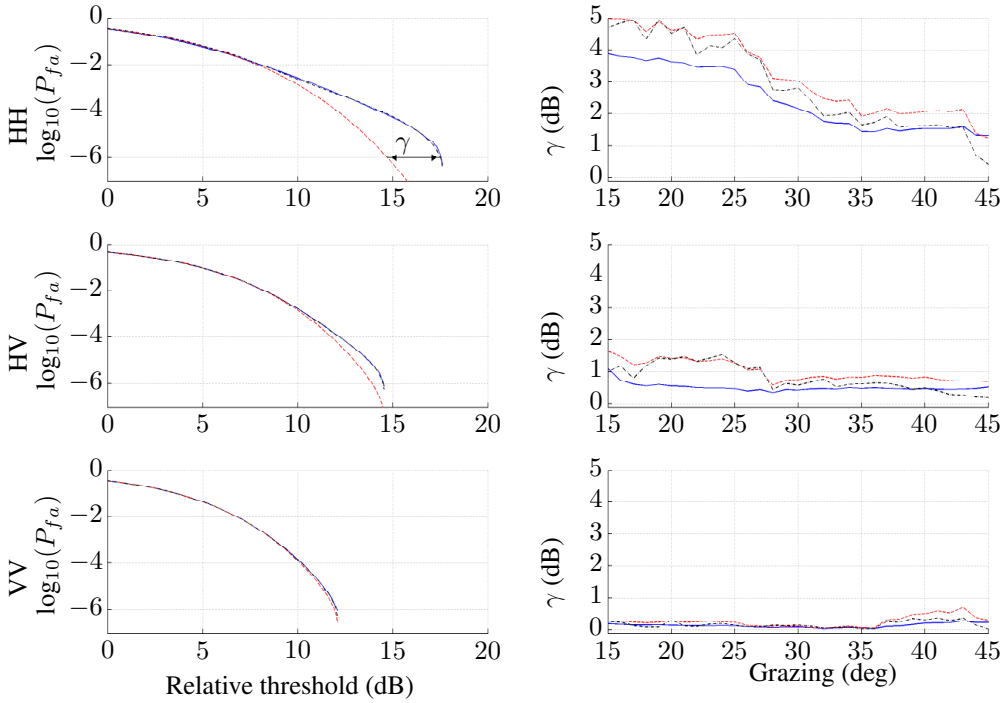

Figure 31: *F35 upwind comparison plots. Left: false alarm rate (log_{10}) at 30° grazing, (—) data, (- -) K, (-.-) KK. Right: K threshold error between the fit and data, P_{fa} = (—) 10^{-4}, (- -) 10^{-5}, (-.-) 10^{-6}.*

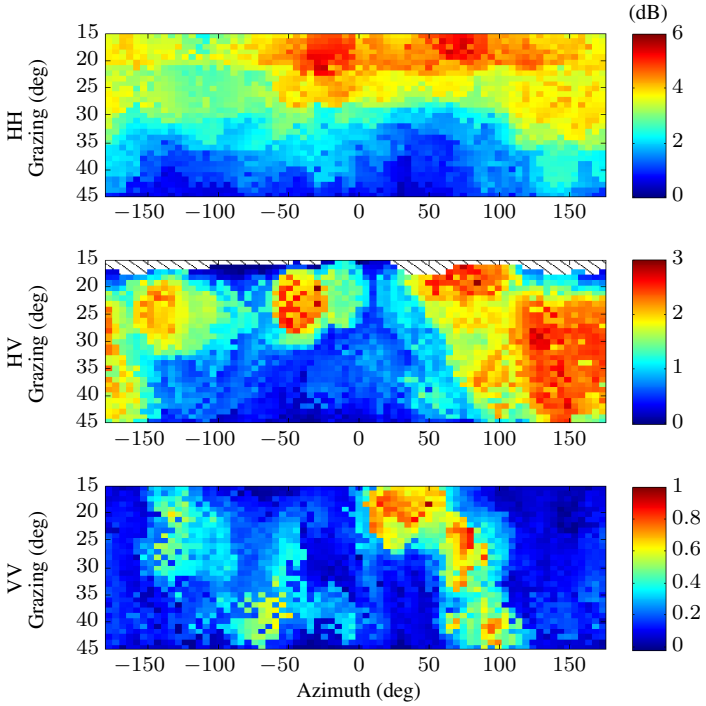

Figure 32: *F35 threshold error between the K-distribution fit and data, $P_{fa} = 10^{-6}$. Hashed areas indicate regions with poor or missing data.*

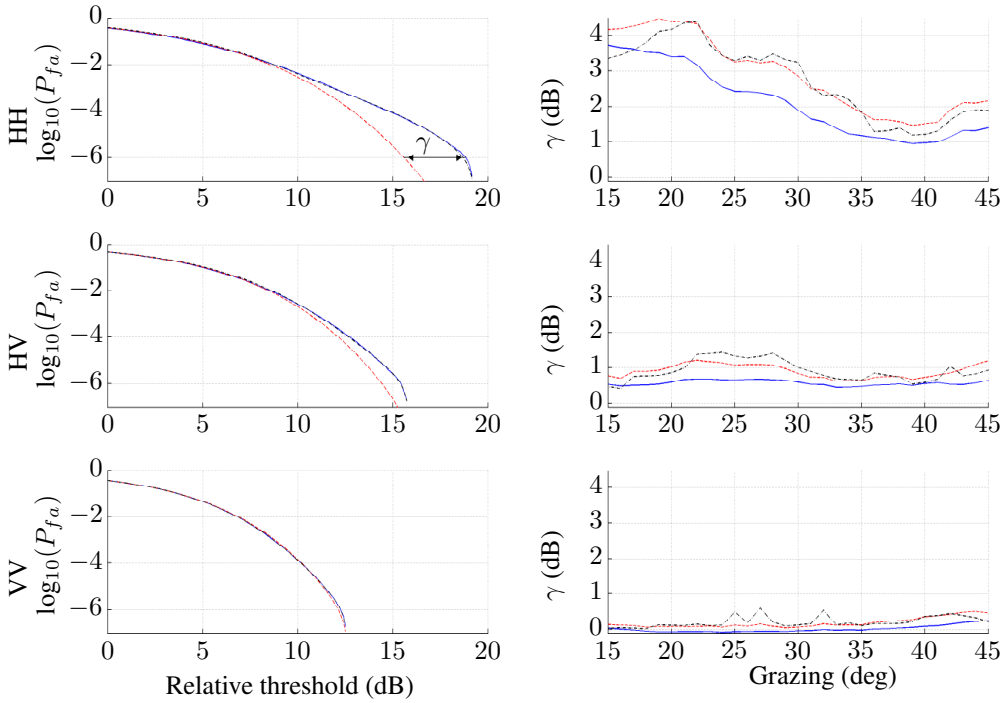

Figure 33: *F9 upwind comparison plots. Left: false alarm rate (log_{10}), 30° grazing, (—) data, (- -) K, (-.-) KK. Right: K threshold error between the fit and data, P_{fa} = (—) 10^{-4}, (- -) 10^{-5}, (-.-) 10^{-6}.*

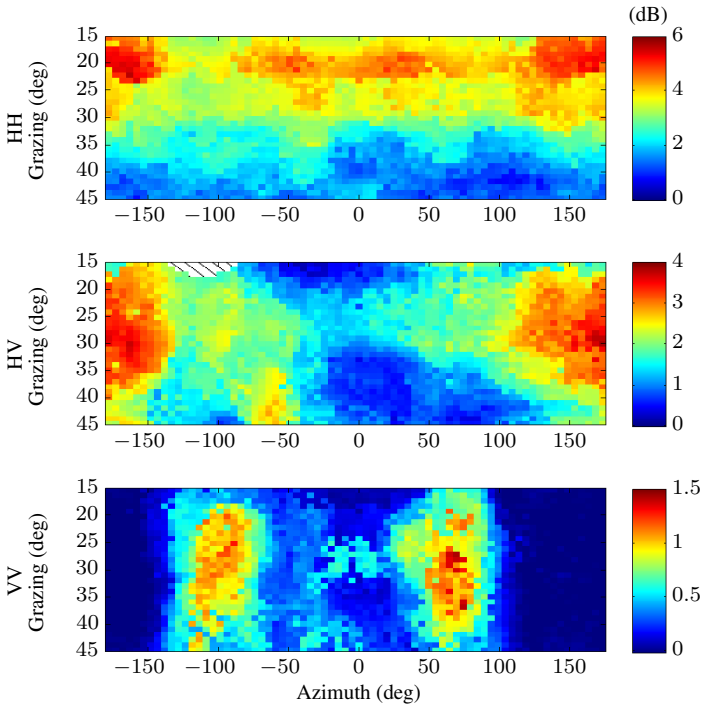

Figure 34: *F9 threshold error between the K-distribution fit and data, $P_{fa} = 10^{-6}$. Hashed areas indicate regions with poor or missing data.*

5.3.2 Multi-look

This section now looks at the CFAR thresholds for the multi-look data. For 2-look data, Figure 35 shows the probability of false alarm on the left hand side for each polarisation. The K-distribution again underestimates the data value with an approximate error at $P_{\text{fa}} = 10^{-6}$ of 2.5 dB for HH, 0.8 dB for HV and 0.2 dB for VV. Interestingly, this is very similar to the single-look relative threshold error, with the biggest difference now being that the relative threshold for both the data and fits has decreased by approximately 2 dB for each of the channels due to the reduced dynamic range that arises from the multi-look averaging. The right hand side of this figure then shows the relative threshold error for the K-distribution as the grazing angle increases. These results are again very similar to the single-look case to within 0.5 dB. This result is then repeated for the 4-look case in Figure 36. At $P_{\text{fa}} = 10^{-6}$, the relative threshold errors over grazing remain very similar with a further 2 dB difference in the relative threshold. The relative threshold error is again very similar to both the single and 2-look results to within 0.5 dB.

The results shown for the second data set F9 in Figures 37 and 38 are consistent with the ones discussed for F35. The threshold errors remain the same as the single-look case with only the relative threshold decreasing as the number of looks increases.

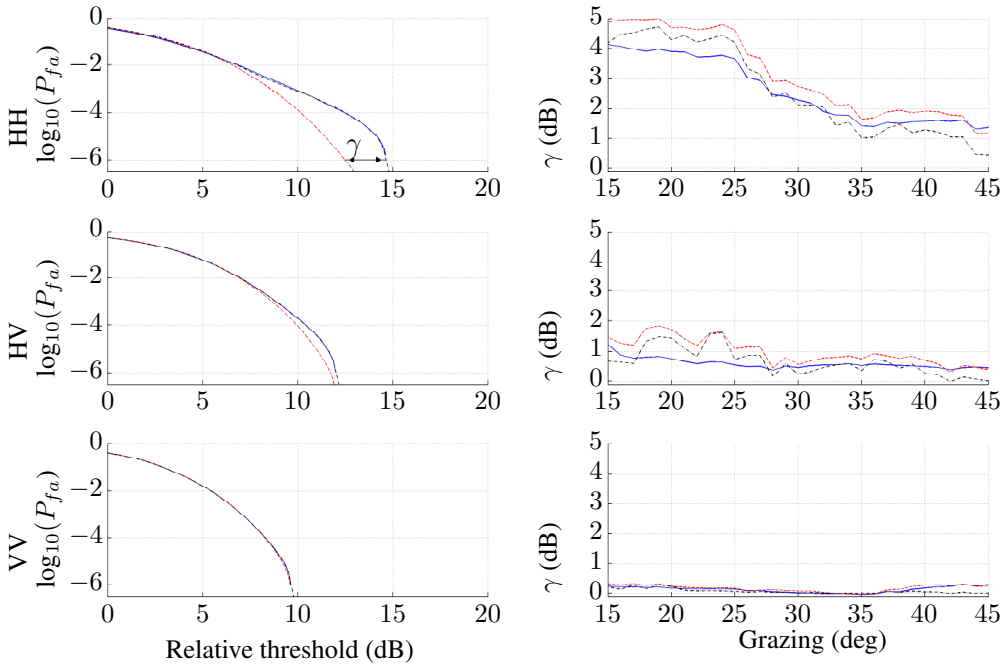

Figure 35: *F35 2-look upwind comparison plots. Left: false alarm rate (log_{10}) at 30° grazing, (—) data, (- -) K, (-.-) KK. Right: K threshold error between the fit and data, P_{fa} = (—) 10^{-4}, (- -) 10^{-5}, (-.-) 10^{-6}.*

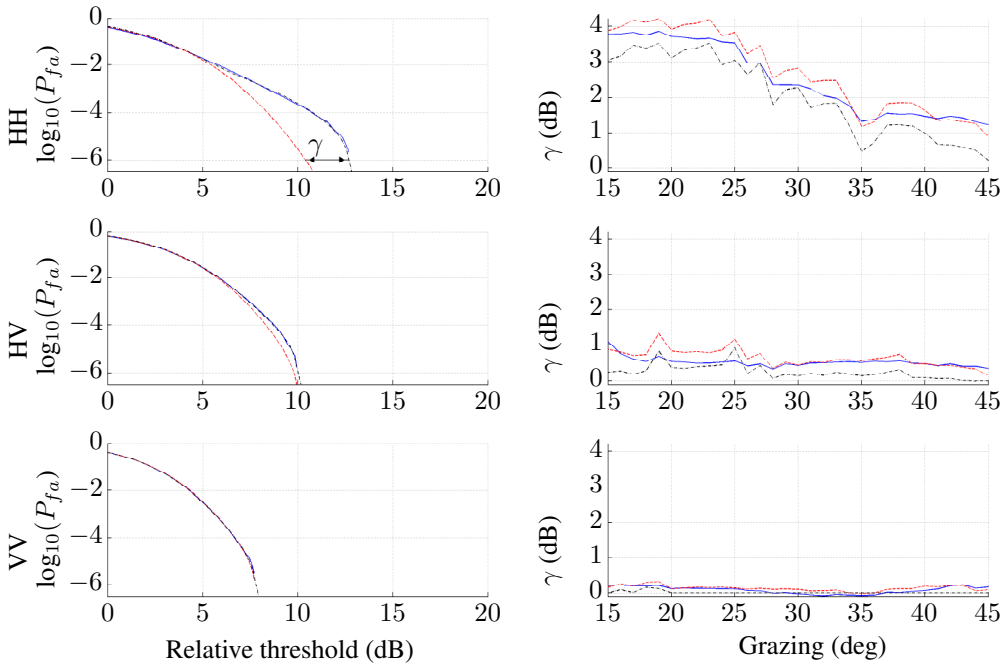

Figure 36: *F35 4-look upwind comparison plots. Left: false alarm rate (log_{10}) at 30° grazing, (—) data, (- -) K, (-.-) KK. Right: K threshold error between the fit and data, P_{fa} = (—) 10^{-4}, (- -) 10^{-5}, (-.-) 10^{-6}.*

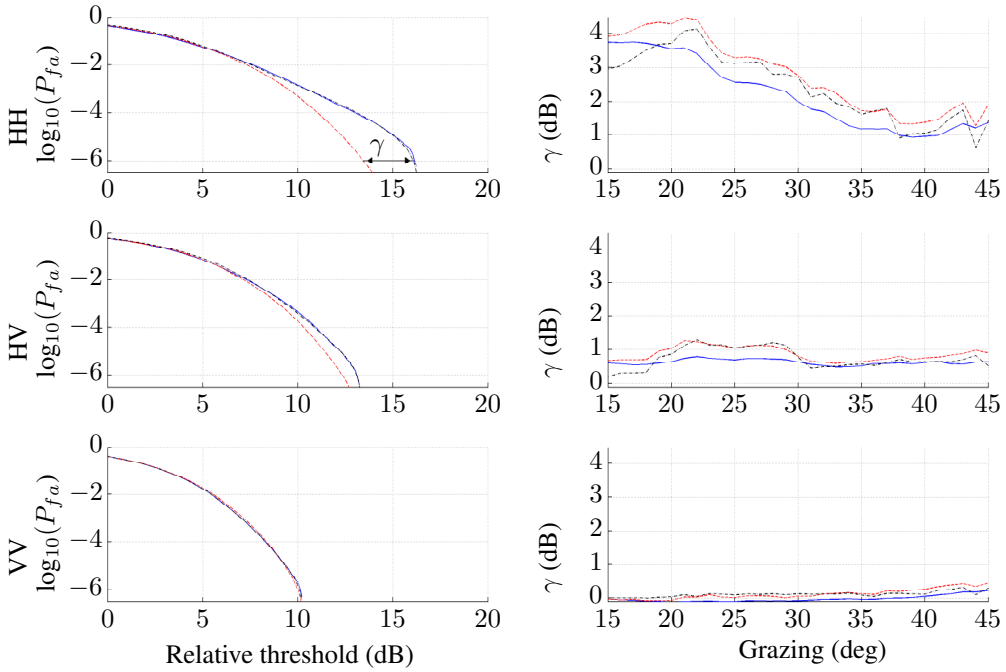

Figure 37: *F9 2-look upwind comparison plots. Left: false alarm rate (log_{10}) at 30° grazing, (—) data, (- -) K, (-.-) KK. Right: K threshold error between the fit and data, P_{fa} = (—) 10^{-4}, (- -) 10^{-5}, (-.-) 10^{-6}.*

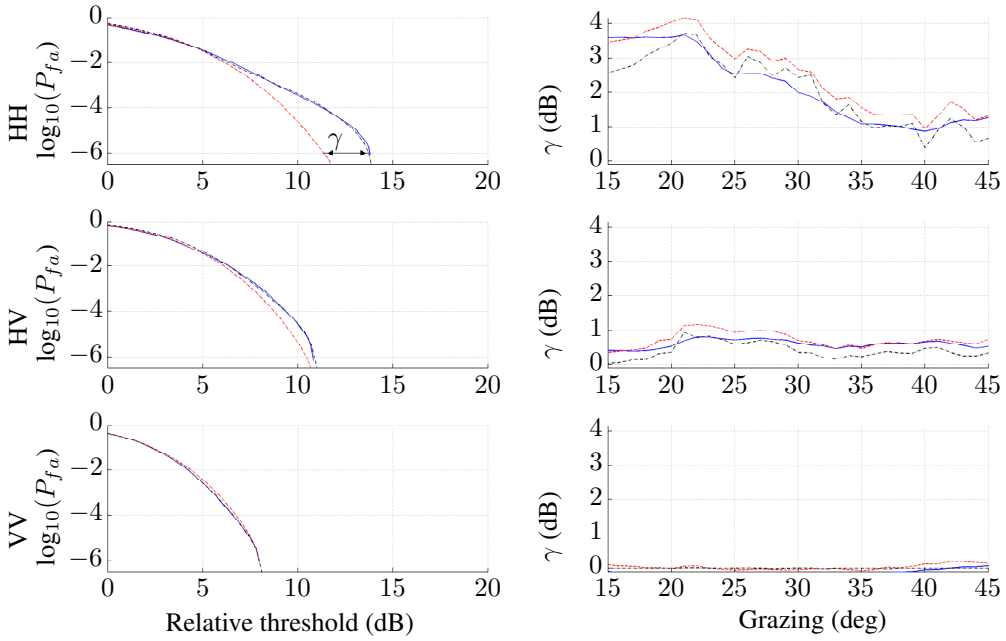

Figure 38: *F9 4-look upwind comparison plots, Left: false alarm rate (log_{10}) at 30° grazing, (—) data, (- -) K, (-.-) KK. Right: K threshold error between the fit and data, P_{fa} = (—) 10^{-4}, (- -) 10^{-5}, (-.-) 10^{-6}.*

5.3.3 Threshold error comparison

To further quantify the false alarm errors, Table 6 shows a comparison of the mean and variance of the relative threshold errors calculated over all the azimuth and incidence directions. To compare against the original KK formulation with no thermal noise, [Dong 2006], a fit is found using the same KK-distribution fitting method described in this report combined with Dong's criteria for choosing the underlying shape and mean. For the K-distribution, errors are clearly worse for the HH and HV channels, while the mean VV error is small. In contrast, both KK-distribution fits perform well with the newer version performing slightly better in most cases. There is also very little difference between the different number of looks. This demonstrates that accounting for thermal noise in the model estimate does not necessarily improve the radar performance from a CFAR perspective at the CNRs consistent with false alarm rates lower than 10^{-4}.

Table 6: *F35 relative threshold errors in dB at $P_{fa} = 10^{-6}$ - mean / variance.*

Polarisation	K	Dong's KK	KK w. noise
HH (1-look)	3.21 / 3.06	0.63 / 0.22	0.63 / 0.21
HV (1-look)	2.8 / 1.48	1.66 / 0.78	1.39 / 0.41
VV (1-look)	0.28 / 0.045	0.10 / 0.0057	0.096 / 0.0056
HH (2-look)	2.87 / 2.93	0.63 / 0.22	0.63 / 0.22
HV (2-look)	2.57 / 1.42	1.65 / 0.77	1.42 / 0.44
VV (2-look)	0.19 / 0.027	0.088 / 0.0052	0.086 / 0.0052
HH (4-look)	2.45 / 2.49	0.61 / 0.21	0.61 / 0.21
HV (4-look)	2.29 / 1.16	1.64 / 0.79	1.44 / 0.44
VV (4-look)	0.12 / 0.013	0.083 / 0.0052	0.084 / 0.0052

6 Conclusion

This report has focussed on modelling the sea-clutter distribution using data acquired by the DSTO Ingara radar, with the ultimate goal of accurately defining a CFAR threshold. The commonly used K-distribution was shown to not always be accurate enough for modelling the probability distribution, especially in the horizontally polarised channels at lower grazing angles. An alternative, known as the KK-distribution was used instead to improve the fit in these regions. The model uses two K-distributions to model both the Bragg and non-Bragg components, with the overall clutter distribution being the sum/mixture of the two components. Although, the choice of the second component is arbitrary and has no direct physical basis, the KK-distribution considerably improves the fit for the extended tails of the PDF that arise from backscatter off sea-spikes on the ocean surface.

The main contribution from this report is an extension of the KK model to include multiple looks and thermal noise, thereby allowing estimates of the underlying mean and shape. A technique for fitting this distribution to the data and estimating the parameters was presented in Section 3. Rather than fitting all five parameters, constraints were introduced to overcome difficulties in fitting the full PDF, and the data was fitted in the $\log_{10}$ domain in order to focus on the tail of the distribution. Following this, a sea-clutter simulation was presented in Section 4 to verify the KK-distribution fitting procedure and perform two tests measuring the effect of thermal noise and sea-spikes on the estimated parameters. The first test used a KK model for the data and compared the threshold error of a typical K-distribution with varying mean separation and CNR. A good match was observed when the mean separation was 1, but when the mean separation increased to 15, the threshold error increased up to 8.5 dB. The variation in threshold error however was not as significant in the region where the clutter mean was larger than the thermal noise, CNR $>$ 0 dB. The second test then used the simulation to verify the relationship between the CNR, the effective shape when thermal noise is present and the underlying shape. As this was derived using the moments of the K-distribution, it was shown not to be accurate when the distribution has an extended tail due to sea-spikes. The new KK estimation method was however able to accurately determine the underlying shape.

Section 5 then applied the new KK model with data collected by the DSTO Ingara radar. Both the single and multi-look data showed a poor match for the K-distribution in the HH and HV channels. When the modified KK-distribution is used however, the underlying mean was revealed for the HH and HV channels, particularly when the CNR was low. When comparing the shape estimates, the KK-distribution was also able to provide a better estimate, especially when the thermal noise becomes significant. The mean ratio had its highest values in the low grazing angle regions for the HH channel which is consistent with the large tail observed in the distribution. As the number of looks increased, the shape increased slightly for the HH and HV channels, and was relatively constant for the VV channel. This is due to the increase in the number of scattering contributions over the larger resolution cell leading to a stronger Rayleigh component in the PDF. This effect was also seen in the ratio of means, which decreased as the number of looks increased.

To assess the threshold required for CFAR target detection, the false alarm rate was calculated from the distributions. The false alarm results showed a good match between the data and the KK-distribution fit, while the K-distribution mismatched the data from approximately 10^{-2} and below. The mismatch was highest in the lower grazing region of the HH channel and decreased almost linearly as the grazing angle increased. As the number of looks increased, the detection

threshold was reduced, while the threshold error remained similar. Finally, a comparison between the original formulation of the KK-distribution and the modified version showed only a small improvement in the threshold error. This demonstrates that the parameters estimated with and without thermal noise are both able to construct the PDF accurately enough for target detection purposes. However, estimates of the underlying mean and shape are important as it will enable the KK-distribution to be used as a proxy for data in radar detection performance studies.

Acknowledgements

The authors would like to thank the Ingara team for their work developing, testing and operating the radar. This work would not be possible without your effort.

References

Antipov, I. (1998) *Simulation of Sea Clutter Returns*, Technical Report DSTO-TR-0679, DSTO.

Blacknell, D. & Tough, R. J. A. (2001) Parameter estimation for the K-distribution based on [z log(z)], *IEE Proceedings of Radar, Sonar and Navigation* **148**(6), 309–312.

Crisp, D. J., Kyprianou, R., Rosenberg, L. & Stacy, N. J. (2008) Modelling X-band sea clutter at moderate grazing angles, *in IEEE International Radar Conference*, pp. 596–601.

Crisp, D. J., Preiss, M. & Goh, A. S. (2008) *Ingara Medium-High Incidence Angle Polarimetric Sea Clutter Measurements and Analysis for the MAST'06 Trial*, Technical Report In preparation, DSTO.

Crisp, D. J., Stacy, N. J. & Goh, A. S. (2006) *Ingara Medium-High Incidence Angle Polarimetric Sea Clutter Measurements and Analysis*, Technical Report DSTO-TR-1818, DSTO.

Dong, Y. (2006) *Distribution of X-Band High Resolution and High Grazing Angle Sea Clutter*, Research Report DSTO-RR-0316, DSTO.

Farshchian, M. & Posner, F. L. (2010) The Pareto distribution for low grazing angle and high resolution X-band sea clutter, *in IEEE Radar Conference*, pp. 789–793.

Gini, F., Greco, M. V., Farina, A. & Lombardo, P. (1998) Optimum and mismatched detection against K-distributed plus gaussian clutter, *IEEE Transactions on Aerospace and Electronic Systems* **34**(3), 860–876.

Jakeman, E. & Pusey, P. N. (1976) A model for non-Rayleigh sea echo, *IEEE Transactions on Antennas and Propagation* **AP-24**(6), 806–814.

Lombardo, P., Oliver, C. J. & Tough, R. J. A. (1995) Effect of noise on order parameter estimation for K-distributed clutter, *IEE Proceedings of Radar, Sonar and Navigation* **142**(1), 33–40.

Long, M. W. (2001) *Radar Reflectivity of Land and Sea - Third Edition*, Artech House.

Middleton, D. (1999) New physical-statistical methods and model for clutter and reverberation: The KA-distribution and related probability structures, *IEEE Journal of Oceanic Engineering* **24**(3), 261–284.

Parfitt, A. & Nikolic, N. (2001) A dual-polarised wideband planar array for X-band synthetic aperture radar, *in IEEE Antennas and Propagation Society International Symposium*, Vol. 2, pp. 464–467.

Quegan, S. (1994) Unified algorithm for phase and cross-talk calibration of polarimetric data: Theory and observations, *IEEE Transactions on Geoscience and Remote Sensing* **32**(1), 89–99.

Rosenberg, L. (2012) Persistent sea-spike detection in medium grazing angle X-band sea-clutter, *in European SAR conference*, pp. 203–206.

Rosenberg, L. & Bocquet, S. (2013) The Pareto distribution for high grazing angle X-band sea-clutter, *in IEEE Proceedings on Geoscience and Remote Sensing*, pp. 4209–4212.

Rosenberg, L., Crisp, D. J. & Stacy, N. J. (2008) Characterisation of low-PRF X-band sea-clutter Doppler spectra, *in IEEE International Radar Conference*, pp. 100–105.

Rosenberg, L., Crisp, D. J. & Stacy, N. J. (2010) Analysis of the KK-distribution with medium grazing angle sea-clutter, *IET Proceedings of Radar Sonar and Navigation* **4**(2), 209–222.

Rosenberg, L. & Stacy, N. J. (2008) Analysis of medium angle X-band sea-clutter Doppler spectra, *in IEEE Radarcon Conference*.

Stacy, N. J. S., Badger, D. P., Goh, A. S., Preiss, M. & Williams, M. L. (2003) The DSTO Ingara airbone X-band SAR polarimetric upgrade: first results, *in IEEE International Geoscience and Remote Sensing Symposium*, Vol. 7, pp. 4474 – 4476.

Ward, K. D. (1981) Compound representation of high resolution sea clutter, *Electronic Letters* **17**(16), 561–563.

Ward, K. D. & Tough, R. J. A. (2002) Radar detection performance in sea clutter with discrete spikes, *in International Radar Converence*, pp. 15–17.

Ward, K. D., Tough, R. J. A. & Watts, S. (2006) *Sea Clutter: Scattering, the K-Distribution and Radar Performance*, The Institute of Engineering Technology.

Watts, S. (1987) Radar detection prediction in K-distributed sea clutter and thermal noise, *IEEE Transactions on Aerospace and Electronic Systems* **AES-23**(1), 40–45.

Weinberg, G. V. (2011) *An Investigation of the Pareto Distribution as a Model for High Grazing Angle Clutter*, Technical Report DSTO-TR-2525, DSTO.

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA				1. CAVEAT/PRIVACY MARKING
2. TITLE Analysis of High Grazing Angle Sea-clutter with the KK-Distribution			3. SECURITY CLASSIFICATION Document (U) Title (U) Abstract (U)	
4. AUTHORS Luke Rosenberg, David J. Crisp and Nick J. Stacy			5. CORPORATE AUTHOR Defence Science and Technology Organisation PO Box 1500 Edinburgh, South Australia 5111, Australia	
6a. DSTO NUMBER DSTO–TR–2915	6b. AR NUMBER 015-796		6c. TYPE OF REPORT Technical Report	7. DOCUMENT DATE November 2013
8. FILE NUMBER 2011/1158667/1	9. TASK NUMBER AIR7000	10. TASK SPONSOR DGAD	11. No. OF PAGES 42	12. No. OF REFS 25
13. URL OF ELECTRONIC VERSION `http://www.dsto.defence.gov.au/publications/scientific.php`			14. RELEASE AUTHORITY Chief, National Security and ISR Division	
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT *Approved for Public Release* OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111				
16. DELIBERATE ANNOUNCEMENT No Limitations				
17. CITATION IN OTHER DOCUMENTS No Limitations				
18. DSTO RESEARCH LIBRARY THESAURUS Sea-clutter, Probability distribution, radar				

19. ABSTRACT

An estimated probability distribution of the backscatter is commonly used to determine the threshold for distinguishing targets from clutter at a given false alarm rate. Data collected at high grazing angles ($15^\circ - 45^\circ$) by the Defence Science Technology Organisation's Ingara fully polarimetric X-band radar demonstrates that the commonly used K-distribution is not always adequate for modelling the probability distribution. This is especially the case for the horizontal polarisation when sea-spikes can cause high false alarm rates. An alternative proposed as a more accurate model is known as the KK-distribution. The analysis presented in this report describes this model with the addition of multiple looks and a thermal noise component to produce an estimate of the underlying mean and shape. This then enables the KK-distribution to be used as a proxy for data in radar detection performance studies. The threshold required to achieve a constant false alarm rate is then studied and compared with that obtained from the K-distribution.