Ship Detection in SAR Imagery

James K.E. Tunaley

Abstract---**As a part of Maritime Domain Awareness, there is a requirement to detect ships in satellite-borne Synthetic Aperture Radar (SAR) images, which provide wide area ocean surveillance. When ship detection is implemented using a Constant False Alarm Rate (CFAR), statistical theory can be employed to ensure that proper parameters are used to find the thresholds for detection; inaccuracy in parameter estimation tends to lead to threshold bias and should be compensated. Otherwise, in spatially varying clutter, the practical performance of automatic ship detectors is likely to be compromised by increases in the false alarm rate and/or by reduction of the probability of detection.**

Index Terms---**K-distribution, sea clutter, target detection, CFAR, detection threshold, SAR.**

I. Introduction

SHIP detection from images derived from satellite borne Synthetic Aperture Radar (SAR) is important because the area coverage rate can be very large; the satellites provide a wide area surveillance capability. The topic has been reviewed in some depth by Crisp [1], who discusses and compares a number of ship detection software systems. A more recent review with a Chinese perspective is provided in [2]; this attempts to summarize the status and the directions of future work. Other workers have contributed in [3]–[8]. Some of this cited work includes discussions on the benefits of polarimetric data and sub-aperture processing.

Ships must be detected against a background of radar sea clutter. The K-distribution can be used to describe the intensity statistics of radar sea clutter [9]–[11] and is often employed in the automatic detection of ships from satellite radar imagery. The technique involves the specification of a Constant False Alarm Rate (CFAR), which implies that a threshold of detection is set according to the local statistics of the clutter at each point in the image plane. Detection losses due to three simple detection schemes are discussed by Armstrong and Griffiths [12] and their work also addresses the effects of errors in K-distribution parameter estimation. The K-distribution has been discussed in detail by Redding [13]. The focus of [13] is on distribution parameter estimation. Other authors [14] have applied neural networks to optimize parameter extraction.

Published December 2nd 2010. J.K.E. Tunaley is President, London Research and Development Corporation, Ottawa, ON, Prof. Emeritus, Physics Dept., University of Western Ontario, London, ON, Adjunct Prof., Physics Dept., Royal Military College, Kingston, ON. E-mail: jtunaley@london-research-and-development.com

Though the K-distribution is usually an excellent empirical match to clutter statistics, the required FAR per pixel or resolution cell is extremely small, say 10^{-9}, and the question of whether the distribution is an adequate representation of reality must be asked. In effect, a detection threshold far out in the tail of the distribution is required but must be determined mainly by the shape of the probability density function (pdf) near its peak; the accuracy with which the distribution represents the clutter is quite critical.

Fortunately it has been shown that the K-distribution arises from a modulation of Gaussian complex clutter return [9]. The Gaussian part can be attributed to a large number of contributions from randomly phased Bragg waves by virtue of the Central Limit Theorem. A number of discrete reflection components, again randomly phased, would also be very likely to contribute a Gaussian complex signal.

The modulating signal appropriate to the K-distribution is gamma distributed. The gamma probability density is one sided; over a wide range of parameters it exhibits a sharp peak. Tunaley [15] has recently shown that, for the purposes of determining a detection threshold, the details of the randomizing distribution may not be critical; this adds considerable support to the use of the distribution to represent sea clutter. For example, if the modulations arise from a great number of large scale ambient ocean surface gravity waves of random phase, the Central Limit Theorem likely applies and the gamma assumption can be discarded. However, even the assumption of a normal or Gaussian modulation distribution function seems not to be important. It may also be noted that gamma-like modulation can also arise from simple stochastic models [16].

It is not clear that all ocean features give K-distributed returns; for example, possible exceptions are some wind generated features and shoals [1]. Though a physical and statistical justification now exists, some post-processing is likely to be required. In the final analysis the performance of a ship detection software application based on the K-distribution will determine its validity.

The K-distribution and approximations to it depend on three parameters, namely the mean, μ, the order or shape parameter, ν, and a parameter, L, which is the number of independent SAR looks and is determined essentially by the satellite radar hardware and the image processor; it should be known ahead of time. When applied to sea clutter, high values of ν represent smooth Rayleigh distributed amplitude clutter (or exponentially distributed intensities) whereas low values of ν (perhaps as low as 0.1) represent spiky clutter.

Significant research effort has been spent in parameter estimation for the K-distribution, which is relevant not only to radar sea clutter but also to sonar signals and radio communication.

Using the Cramér-Rao lower bound, Blacknell [17] contends that the ordinary mean estimator and the average of the logarithms of the samples are close to optimal for finding the parameters. A similar general approach is used here but numerical methods are employed and no attempt is made to employ Blacknell's approximations to the K-distribution that are valid at large L.

Abraham and Lyons [18], [19] have analyzed a bootstrap method [20] for generating the K-distribution parameters. Some of their treatments parallel those used here but their emphasis is on the parameters themselves rather than the detection thresholds; it is not clear whether some of their approximations are appropriate to the tails of the distribution that are important for target detection. They also discuss the trade off between performance and computational complexity and point out that many calculations can be done ahead of time. Lookup tables can be employed for operational code.

As shown in texts on statistics, for example [20] and [21], the extraction of the parameters in a pdf can be based on Maximum Likelihood (ML). However, mainly because of the Bessel function in the K-distribution, this is difficult to implement except by numerical means. Therefore efforts have been made to introduce sub-optimal statistical methods to minimize the computational load.

Dong and Crisp [23] have compared several schemes of simple parameter estimation with the ML approach. This is based largely on simulation and involves various averages based on the logarithm of the data. They find that a combination of expectations, $E(X)$ and $E(\log(X))$, where X is the signal intensity, provides the best estimates of the mean and K-distribution order parameter. Their study also demonstrated the effect of bias.

Perhaps the simplest statistic that might be used to extract the order parameter, ν, is the ratio of the clutter variance to the square of the mean. This type of approach falls under a category known as the Method Of Moments (MOM).

As can be deduced from the Appendix, the variance of the intensity is given by:

$$\sigma^2 = \mu^2 \frac{1+L+\nu}{L\nu} \tag{1}$$

This can be rearranged to yield the order parameter in terms of the mean, μ, and variance, σ^2:

$$\nu = \frac{L+1}{L\sigma^2/\mu^2 - 1} \tag{2}$$

Unfortunately, since ν is not linearly related to σ^2, significant bias will be introduced. (Later we use $1/\nu$ as the parameter of interest.) Furthermore, though MOM simplifies computation, when the measured variance of the clutter happens by chance to fall to low values, negative order parameters are predicted. If the clutter is truly Rayleigh, the probability of this is about 0.5. However, this is not the critical problem implied by some authors [18] as ν can just be set to infinity.

Since the MOM approach can be far from optimal, other statistics, T, [22]–[25], have been proposed, which involve the logarithm of the data or a product:

$$T = [X^r \log(X)], \tag{3}$$

where X is the intensity of a pixel, r is some constant and the square brackets indicate an average over a block of pixels. A study by Hu [25] suggests that improved parameter estimation is possible when $r = 0.5$.

Because the distribution parameters must be extracted from a finite number of clutter samples, there is uncertainty in the parameter values. As will be shown, this translates into increased detection thresholds; significant bias is introduced as a result of the non-linear functionality between the parameters and the threshold. The probabilities of false alarm and detection may be compromised and the ship detection performance may be degraded.

II. Theory

The K-distribution for the signal intensity, X, of a point in a single look image with mean, μ, and order (shape) parameter ν is given by [17]:

$$P\{X > x\} = \frac{2}{\Gamma(\nu)}\left(\frac{\nu x}{\mu}\right)^{\nu/2} K_\nu\left(2\sqrt{\frac{\nu x}{\mu}}\right) \tag{4}$$

where K is a modified Bessel function of the second kind; its properties can be found in [26]. For multi-look images, the distribution itself cannot be expressed easily in closed form but the pdf is [13]:

$$p(x) = \frac{2}{x\Gamma(L)\Gamma(\nu)}\left(\frac{L\nu x}{\mu}\right)^{(L+\nu)/2} K_{|\nu-L|}\left(2\sqrt{\frac{L\nu x}{\mu}}\right) \tag{5}$$

where L is the number of independent looks, which is usually not an integer because sub-aperture images are typically somewhat correlated. It should be noted that the Bessel function is an even function of its order.

The parameters that must be estimated in practice are μ and ν and some care in the choice of a statistic is required. Not all statistics are capable of providing accurate estimates but we can at least determine the best result that might be obtainable. This not only provides a standard for evaluating the methods that might be employed but also indicates fundamental limits to practical detection performance.

The minimum covariance matrix for a parameter vector $\boldsymbol{\theta}$ (in this case with components μ and ν) can be found using the Cramér-Rao inequality [20], which provides the variance of an unbiased estimator. This involves the efficient score vector, $\boldsymbol{U}$, and for a single random pixel intensity, X, its components are given by:

$$U_i(X,\theta_i) = \frac{\partial \log(p(x,\theta_i))}{\partial \theta_i}. \tag{6}$$

It is easily shown from the normalization condition on x that the expectation of $\boldsymbol{U}$ with respect to x is zero. In our case, p is the K probability density.

The Fisher information matrix, $\boldsymbol{i}(\nu)$ is given by:

$$i_{nm}(\nu) = E(U_n U_m)\,. \qquad (7)$$

where E indicates an expectation value over x. The minimum covariance matrix for $\boldsymbol{\theta}$ is just the reciprocal of the Fisher information matrix [20].

The Fisher information can be calculated numerically from (5) and the result for μ (i_{11}, say) is shown in Fig. 1. Fig. 2 shows the Fisher information for ν (i_{22}, say). The calculation is straightforward when $\nu > 1$ but, if $\nu < 1$, the pdf tends to infinity at the origin. Therefore asymptotic forms must be found for this region and the integral implied by (7) must be evaluated analytically for small x. To verify that the calculations were satisfactory, the normalization, mean, variance and fourth moments of the K-distribution were calculated and compared with the theoretical values. It was also verified that the expectation of the efficient score was close to zero. Some details are provided in the Appendix.

The cross terms in the Fisher information are usually quite small and, when the inverse is taken, barely affect the minimum variances. For example, the difference between $(i_{11})^{-1}$ and $(\boldsymbol{i}^{-1})_{11}$ is typically less than a few percent but can be up to 15% in spiky clutter.

When the clutter statistics are derived from a block of N independent pixels, The Fisher information is the sum of that from each pixel so that, for constant distribution parameters, it increases by a factor N. Therefore the minimum variances and covariances decrease by N in the estimates of the parameters. Moreover, as N increases, the distribution of a parameter estimate approaches the normal distribution for the same reason as for the sum of any independent identically distributed random variables. For large N, the distribution of the estimate is almost entirely described by a mean and variance. Fig. 3 shows the ratio of the minimum Standard Deviation (SD) of the estimate of the mean as a function of ν for various L and $N = 256$. This was obtained by inverting the information matrix and is appropriate to a true mean of one.

Clutter spikiness increases the variance of the clutter intensity, which leads to a greater standard deviation in the estimate of the mean. Thus, as might be expected, the minimum standard deviation of μ increases as ν decreases. It too decreases with the number of looks because multi-looking is essentially pre-averaging.

As far as we can tell, the results for the mean with $L = 1$ are identical to those presented by Blacknell [17]. For $L > 1$, Blacknell has treated the sub-aperture multi-looks as independent samples; for example when $L = 4$, we interpret N samples as N independent samples each of 4 sub-apertures, which would correspond to $4N$ samples in Blacknell's paper.

A mean is usually found by averaging pixel intensities. The SD of the estimate of the true mean is related to the variance according to (1). In fact the SD of the mean derived from the Fisher information is only very slightly less than that derived by the traditional method; the difference is typically about 1%. Therefore there is no requirement for a special treatment of estimates of the mean pixel intensity; the usual method provides essentially optimal accuracy. This is consistent with the simulations of other authors, e.g. [13].

It can be seen from Fig. 4 that the minimum SD of ν tends to increase rapidly as ν increases. This is because the pdf is not sensitive to ν when ν is large, so that the range over which values of ν cannot be distinguished increases. Even with 256 independent samples, the standard deviation is a significant fraction of the expected value for most order parameters encountered in practice ($\nu > 1$) and may even exceed it.

We find that the minimum SDs of ν appear to be similar to those of Blacknell when $\nu > 1$ and $L = 1$ but is much smaller than his for $\nu < 1$. This may be due to the limitations of his calculation where approximations were required in this latter regime.

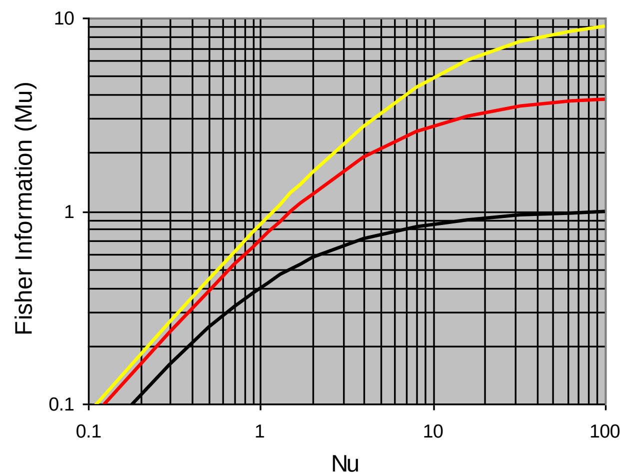

Fig. 1. Fisher Information for the mean as a function of order parameter for the K-distribution for different numbers of looks; $L = 1$ (—), $L = 4$ (—), $L = 10$ (—).

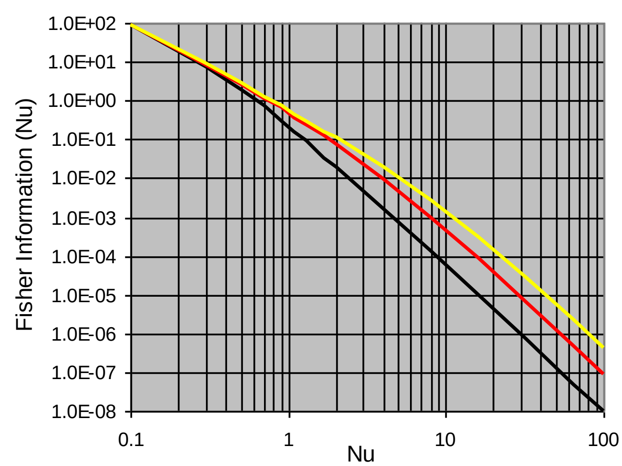

Fig. 2. Fisher Information for the order parameter as a function of order parameter for the K-distribution for different numbers of looks; $L = 1$ (—), $L = 4$ (—), $L = 10$ (—).

It is useful to compare the SDs of estimates based on practical statistics of the MOM type, such as the ratio of the mean to the variance, with the ideal SD provided by the Cramér-Rao bound. For comparison purposes, we assume that the mean is known exactly. The order parameter, ν, can be found using (2). By differentiation, it is easily shown that:

$$Var(\nu) = \left(\frac{\nu^2 L}{L+1}\right)^2 \frac{Var(\sigma^2)}{\mu^4} \quad (8)$$

Combining this with the result from equation (22) in the Appendix, we find for N independent pixel intensities:

$$Var(\nu) = \frac{\nu(\nu+1)}{N\mu^4 (L+1)L}((L+3)(L+2)(\nu+3)(\nu+2) \\ -(L+1)(\nu+1)L\nu) \quad (9)$$

Taking the square root of the variance on the left hand side, the SD is plotted in Fig. 5 for N = 256. It is clear that this is up to an order of magnitude greater than the theoretical minimum in Fig. 4 and that an estimate derived from σ^2/μ leaves much room for improvement. We note again that sub-optimality might be at least partly compensated by increasing N.

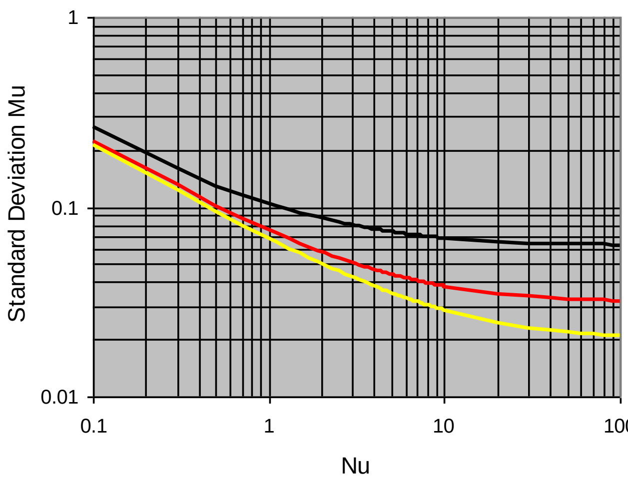

Fig. 3. Minimum standard deviation for the estimated mean as a function of order parameter for the K-distribution for N = 256 and looks L = 1 (—), L = 4 (—), L = 10 (—).

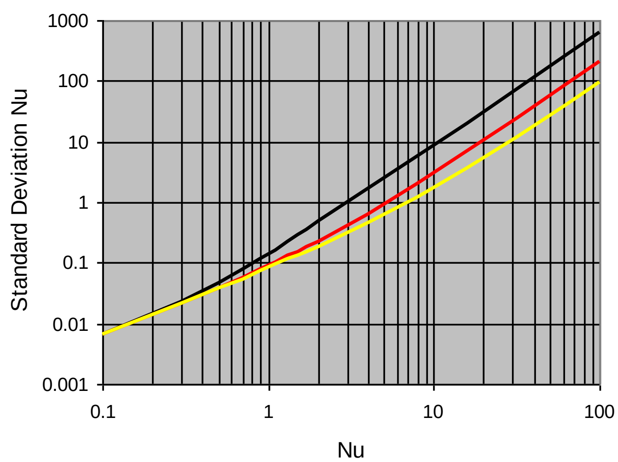

Fig. 4. Minimum standard deviation for the estimated ν as a function of ν for the K-distribution for N = 1000 and looks L = 1 (—), L = 4 (—), L = 10 (—).

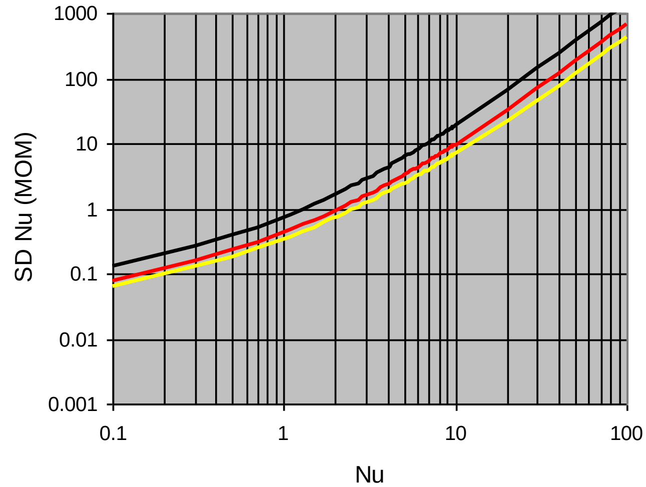

Fig. 5. MOM standard deviation for the estimated ν as a function of ν for N = 256 and looks L = 1 (—), L = 4 (—), L = 10 (—).

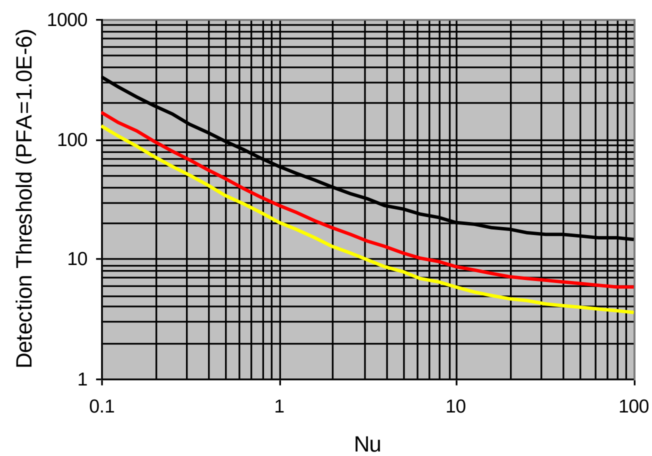

Fig. 6. Ideal detection threshold based on the K-distribution for looks L = 1 (—), L = 4 (—), L = 10 (—). The PFA is 10^{-6}.

III. Probabilty Of False Alarm

In practice a Probability of False Alarm (PFA) is specified and the threshold of detection is calculated as a multiple of the mean. Figs. 6 and 7 show the ideal detection threshold as a multiple of the mean as a function of ν and L; μ and ν are supposed known exactly as if the number of independent pixels were infinite. The PFAs are 10^{-6} and 10^{-9} respectively. Though estimates of the mean will generally be unbiased, estimates of the order parameter may not be. Bias and uncertainty in the order parameter will generally cause bias in the detection threshold. When the variance of the sample is used to estimate the order parameter, bias in the variance can be removed by Bessel's correction, in which the sum of the squares of the deviations is divided by (N-1) rather than N. It is also possible to correct for bias in an estimate of the variance of the variance (by dividing by (N+1) rather than N) but if $N \geq$

100, these corrections make a negligible difference. The uncertainty in the parameter estimate is much more serious.

To illustrate the process, consider a single look image of Rayleigh clutter so that the pixel intensities are exponentially distributed with a normally distributed randomizing of the mean:

$$PFA = P\{X > t\} = \int_{-\infty}^{\infty} e^{-t/\mu} \frac{e^{-(\mu-\mu_0)^2/(2\sigma_\mu{}^2)}}{\sqrt{2\pi\sigma_\mu{}^2}} d\mu \tag{10}$$

where μ_0 is the overall mean and $\sigma_\mu{}^2$ is its variance, which is actually equal to $\mu_0{}^2/N$. (When $N >> 1$, the lower limit of the integral can be extended to -∞.) This must be solved for the threshold, t. When $N >> 1$, the normal distribution is very concentrated about μ_0 and a steepest descents approach can be used.

The integrand is written as exp(f) where:

$$f = -\frac{1}{2}\log(2\pi\sigma_\mu^2) - \frac{t}{\mu} - \frac{(\mu-\mu_0)^2}{2\sigma_\mu^2} \tag{11}$$

Differentiating with respect to μ yields:

$$\begin{aligned} f' &= \frac{t}{\mu^2} - \frac{\mu-\mu_0}{\sigma_\mu^2} \\ f'' &= -\frac{2t}{\mu^3} - \frac{1}{\sigma_\mu^2} \end{aligned} \tag{12}$$

The integrand peaks when $f' = 0$, which occurs when:

$$\sigma_\mu^2 t = \mu^2(\mu - \mu_0) \tag{13}$$

Clearly, when N is large, the left hand side is very small and, to a first approximation, we have:

$$\mu = \mu_0 + \sigma_\mu^2 t / \mu_0^2 \tag{14}$$

where t can be replaced by its value appropriate to zero variance, namely -μ_0 log(PFA).

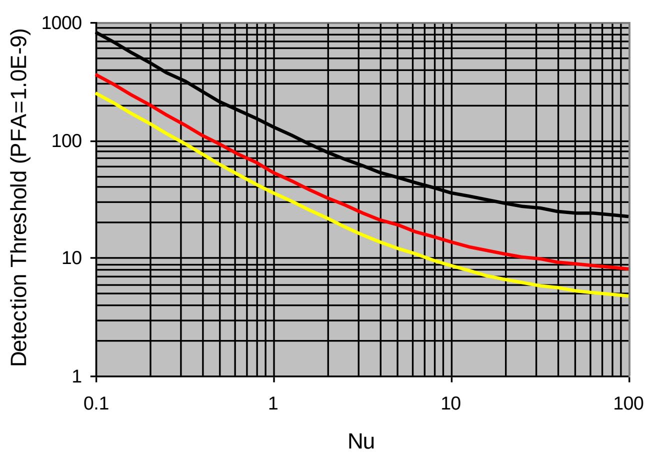

Fig. 7. Ideal detection threshold based on the K-distribution for looks L = 1 (—), L = 4 (—), L = 10 (—). The PFA is 10^{-9}, which is more realistic for SAR.

After integrating, setting $\sigma_\mu{}^2 = \mu_0{}^2/N$ and assuming $N >> 1$, the result is:

$$\begin{aligned} t &\approx -(\mu_0 - \sigma_\mu^2 \log(PFA)/\mu_0)\log(PFA) \\ &= -\mu_0 \log(PFA)(1 - \log(PFA)/N) \end{aligned} \tag{15}$$

For example, if PFA = 10^{-9}, μ_0 = 1 and N = 100, we find that t ≈ 25.0. This compares with 20.7 if no correction is made for the variance; the difference is very significant in terms of false alarms.

Though uncertainty in the mean will play an important role when N is small, it will often be overshadowed by uncertainty in the order parameter, especially in spiky clutter. In general, the exponential distribution in (10) can be replaced by an integral over the density. For large N, the PFA is then given approximately by:

$$PFA = \int_t^{\infty} dx \int_{-\infty}^{\infty} p(x,\nu) \frac{e^{-(\nu-\nu_0)^2/(2\sigma_\nu^2)}}{\sqrt{2\pi\sigma_\nu^2}} d\nu \tag{16}$$

where $\sigma_\nu{}^2$ is the variance of ν, which will be inversely proportional to N. As before, the principal assumption would be that N is sufficiently large, that the Central Limit Theorem applies and that the statistics are asymptotically normal for the order parameter so that it is characterized solely by its mean and variance. However, rather than randomize over ν, which is biased and causes problems with negative values, it is preferable to randomize over $1/\nu$, which ranges from zero (Rayleigh clutter) to infinity (spiky clutter). This is implied by (2) in which $1/\nu_0$ is proportional to σ^2. When $1/\nu_0$ is small, negative values are suppressed by setting $1/\nu = 0$.

The integrals corresponding to (16) were calculated numerically to yield a PFA for a given threshold. The randomization was actually over $1/\nu$ rather than ν, which avoided bias due to the non-linear relationship in (2). The variance of $1/\nu$ was determined from (1) and (22), i.e.

$$\begin{aligned} Var(1/\nu) &= \left(\frac{L}{L+1}\right)^2 Var(\sigma^2) \\ &= \frac{(\nu+1)}{L(L+1)\nu^3}((L+3)(L+2)(\nu+3)(\nu+2) \\ &\qquad - (L+1)(\nu+1)L\nu) \end{aligned} \tag{17}$$

This corresponds to a MOM approach. The calculation was assisted by the approximation in [15]. It was observed that in spiky clutter the PFA can increase by over two orders of magnitude even when N = 1000.

To assess the effect on ship detection performance we need to hold the PFA constant and find the detection threshold. This was accomplished by adding a binary search to the previous PFA calculations. For a PFA of 10^{-9} and for values of N ranging from 100 to ∞ the results are shown in Fig. 8. When N = 100, the threshold must be increased substantially; when N = 10,000, there is up to a 7% increase in the detection threshold relative to the case of no uncertainty.

IV. Discussion

It has been assumed that radar sea clutter statistics are closely approximated by the K-distribution. Then it has been shown that the usual estimate of the mean as the sum of independent values divided by N provides an estimate that is essentially optimal and, of course, it is unbiased. However, with $N = 256$ independent samples, the mean can be estimated only to within a few percent and this will affect the actual detection threshold, which is inversely proportional to the mean.

In very spiky clutter, $\nu < 1$, the pdf goes to infinity at the origin. This leads to the unexpected result that the order parameter component of the Fisher information tends to be concentrated not in the distribution tails but near the origin. This explains why moment techniques designed to extract the order parameter are ineffective in spiky clutter; they ignore this regime. It also explains why methods based on the logarithm of the data are an improvement.

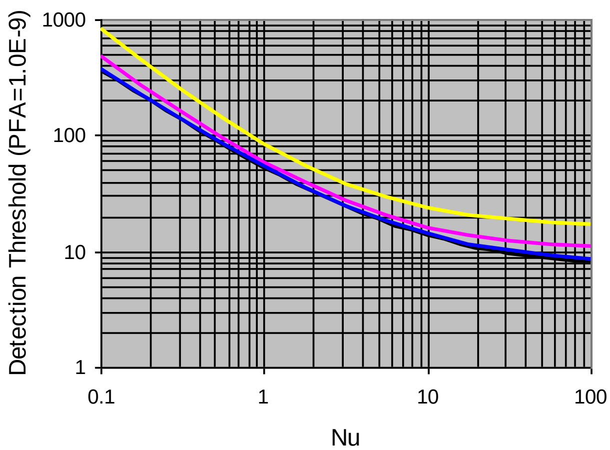

Fig. 8. Detection thresholds for PFA = 10^{-9}, L = 4 and N = 100 (—), N = 1000 (—), N = 10,000 (—) and $N = \infty$ (—).

The advantage of the MOM approach to the estimation of ν is that it is easy to calculate and can be unbiased. The theory is also simple compared with other approaches and functions well in smooth clutter. However, the variations in the estimates of ν create a significant upward bias in the estimate of the detection threshold for a given PFA; the threshold may be increased dramatically for values of $N < 1,000$. This will typically result in an increase in the false alarm rate or a reduction in the probability of detecting a ship.

In practice both μ and ν will be uncertain and both these uncertainties will result in an increase in the detection threshold. However, the uncertainty in ν is likely to dominate upward shifts in detection threshold.

The ship detection process must take into account variations in the clutter field. These additional variations will have a similar effect to the basic uncertainties in parameter estimation. The radar backscatter from the ocean surface depends mainly on the wind speed. Though it is often almost constant over quite large areas, it can vary spatially over scales of less than 1 km. For a SAR with a resolution of 75 m, only $N \approx 100$ independent samples may be available to estimate the K-distribution shape parameters. This represents another fundamental limitation to ship detection in SAR images.

Finally, the estimation of the clutter parameters must take into consideration the presence of a ship. This is discussed in [1]. (One technique is to employ a ship sized exclusion zone at the pixel in question and sample a ring of pixels around it.) The requirement for large N may produce problems in high density shipping areas. This is a topic for another study.

V. Conclusions

The statistics of K-distributed clutter have been studied in the context of parameter extraction. In spiky clutter the shape of the pdf is important not only in the tails but also near the origin of the distribution.

Uncertainty in the non-linear relationship between the parameters of the K-distribution and the CFAR detection threshold biases the detection threshold upwards. This implies that the false alarm rate can significantly increase relative to the situation where parameters can be determined accurately. On the other hand, if the threshold increases are compensated by an increase in the number of independent pixel intensities, N, used for the estimation, the ability to handle variations in clutter intensity will be affected.

Using MOM, the effects on detection threshold may be sufficiently minor to be acceptable when $N > 1000$ and especially if $N \geq 10,000$. However, it may be prudent to adapt N to the clutter variations and to accept the degradation in ship detection performance in regions where the variations are large or over small spatial scales.

Any prediction of ship detection performance should reflect the size of the block of pixels used.

It will be useful to examine approaches other than those mentioned here to optimize performance without a large penalty in computational complexity. It is emphasized that the time needed to run a software application to detect ships is yet another parameter in the design trade-off of an operational ship detection system.

References

[1] D. J. Crisp, "The State-of-the-Art in Ship Detection in Synthetic Aperture Radar Imagery", Intelligence, Surveillance and Reconnaissance Division Information Sciences Laboratory, DSTO–RR–0272, May, 2004.

[2] Wang Juan, Sun Lijie, and Zhang Xuelan, "Study evolution of ship target detection and recognition in SAR imagery", Proc. International Symposium on Information Processing (ISIP'09), Huangshan, P. R. China, pp. 147–150, August 21–23, 2009.

[3] P.W. Vachon, "Ship Detection in Synthetic Aperture Radar Imagery", Proc. OceanSAR 2006 – Third Workshop on Coastal and Marine Applications of SAR, St. John's, NL, Canada, pp 1–10, October, 2006.

[4] J.K.E. Tunaley, "Algorithms for ship detection and tracking using satellite imagery". Proc. International Geoscience and Remote Sensing Symposium (IGARSS 2004), on CD-ROM, 20–24 Sept. 2004, Anchorage, Alaska.

[5] H. Greidanus, P. Clayton, M. Indregard, G. Staples, N. Suzuki, P. Vachon, C. Wackerman, T. Tennvassas, J. Mallorqui, N. Kourti, R. Ringrose, and H. Melief, "Benchmarking operational

SAR ship detection", Proc. International Geosci. Remote Sens. Symposium, (IGARSS 2004*)*, on CD-ROM, 20-24 September 2004, Anchorage, Alaska.

[6] T.N. Arnesen, R.B. Olsen, P.W. Vachon, Polarization-Dependent Signatures of Ships in ENVISAT AP Mode Data, Proc. Envisat & ERS Symposium, Salzburg, Austria 6–10 September, 2004 (ESA SP-572, April 2005).

[7] H. Greidanus, N. Kourti, Findings of the DECLIMS Project – Detection and Classification of Marine Traffic from Space, Proceedings of SEASAR 2006, pp 23–26 Frascati, Italy, January, 2006.

[8] Haiyan Li, Yijun He and Wenguang Wang, "Improving Ship Detection with Polarimetric SAR based on Convolution between Co-polarization Channels", *Sensors*, vol. 9, pp 1221–1236, 2009.

[9] E. Jakeman and P.N. Pusey, "A Model for non-Rayleigh Sea Echo", *IEEE Trans. Antennas Propag*., vol. AP-24, pp 806–814, November, 1976.

[10] K.D. Ward, Compound Representation of High Resolution Sea Clutter, Electron. Lett., vol. 17, no. 16, pp 561–563, August, 1981.

[11] K.D. Ward, S. Watts, R.J.A. Tough, *Sea clutter: Scattering, the K-Distribution and Radar Performance*, Institution of Engineering and Technology, London, UK, 2006.

[12] B.C. Armstrong and H.D. Griffiths, "CFAR detection of fluctuating targets in spatially correlated K-distributed clutter", *IEE Proc.-F*, vol. 138, no. 2, pp 139–152, April, 1991.

[13] N.J. Redding, "Estimating the Parameters of the K Distribution", Surveillance Systems Division Electronics and Surveillance Research Laboratory, DSTO, Technical Report 0839, July, 1999.

[14] M. Jahangir, D. Blacknell and R.G.White, Accurate approximation to the optimum parameter estimate for K-distributed clutter, *IEE Proc-Radar, Sonar Navig*., vol. 143, no. 6, pp 383–390, December, 1996.

[15] J.K.E. Tunaley, "K-Distribution Approximation", London Research and Development Corp. Report, June, 2010.

[16] P. Fayard and T.R. Field, "Inference of a generalised texture for a compound – Gaussian clutter", *IET Radar Sonar Navig*., vol. 4, no. 2, pp. 187–194, 2010.

[17] D. Blacknell, "Comparison of parameter estimators for K-distribution," *IEE Proc. - Radar, Sonar and Navig*., vol. 141, no. 1, pp. 45-52, February, 1994.

[18] D.A. Abraham and A.P. Lyons, "Bootstrapped K-Distribution Parameter Estimation", Proc. Oceans Conference, Boston, Massachusetts, September, 2006.

[19] D.A. Abraham and A.P. Lyons, "Reliable methods for estimating the K-distribution shape parameter", *IEEE J. Oceanic Engineering*, vol. 35, no. 2, pp. 288-302, April, 2010.

[20] D.R. Cox and D.V. Hinkley, *Theoretical Statistics*, Chapman and Hall, New York, 1986.

[21] A. Stuart and J.K. Ord, *Kendall's Advanced Theory of Statistics*, vol. 2, 5^{th} Ed., Arnold, London, 1991.

[22] C.J. Oliver, "Optimum texture estimators for SAR clutter", *J. Phys. D: Appl. Phys*., vol. 26, pp. 1824-1835, 1993.

[23] Y. Dong and D. Crisp, "Comparison of Estimation Schemes for the K Distribution Shape Parameter", International RADAR Conference, pp. 6–11, 2009.

[24] D. Blacknell and R.J.A. Tough, "Parameter estimation for the K-distribution based on [z log(z)]", *IEE Proc. Radar Sonar Navig***.,** vol. 148, no**.** 6, pp 309-312, December, 2001.

[25] Wenlin Hu, "Estimation of K-Distribution Parameters Using [z^rlog(z)]", International RADAR Conference, 2009.

[26] N.W. McLachlan, *Bessel functions for Engineers*, 2^{nd} Edition, Oxford, 1961.

Appendix

The moments of the K-distribution can be found using a relation from [26], namely:

$$\int_0^\infty z^\mu K_\nu(z)dz = 2^{\mu-1}\Gamma\left(\frac{\mu+\nu+1}{2}\right)\Gamma\left(\frac{\mu-\nu+1}{2}\right) \quad (18)$$

where μ is now just a constant. Thus the moments, m_n, of the K-pdf in (5) but with unit mean are given by:

$$m_n = \frac{\Gamma(n+L)\Gamma(n+\nu)}{(L\nu)^n\Gamma(L)\Gamma(\nu)} \quad (19)$$

For example we have:

$$\begin{aligned} m_0 &= 1 \\ m_1 &= 1 \\ m_2 &= \frac{(1+L)(1+\nu)}{L\nu} \end{aligned} \quad (20)$$

The variance of the variance, with known mean, can be found in terms of the moments. By integrating over the pdf, it is easy to show that the variance of the variance is given by:

$$\begin{aligned} Var(\sigma^2) &= E\left((x^2-m_1^2)^2-(m_2-m_1^2)^2\right) \\ &= m_4 - m_2^2 \end{aligned} \quad (21)$$

Therefore, from (19) we find that:

$$\begin{aligned} Var(\sigma^2) = \frac{(L+1)(\nu+1)}{L^3\nu^3}&((L+3)(L+2)(\nu+3)(\nu+2) \\ &-(L+1)(\nu+1)L\nu) \end{aligned} \quad (22)$$

When the order parameter is small, the pdf goes to infinity at $x = 0$. However, there exists an asymptotic expansion of the pdf, which is valid for $x^2 << 4|\nu - L|$. This is given by:

$$p(x) \to \frac{\Gamma(L-\nu)}{x\,\Gamma(L)\Gamma(\nu)}\left(\frac{L\nu x}{\mu}\right)^\nu, L>\nu \quad (23)$$

It is based on [26] and for small z and non-integer ν:

$$K_\nu(z) \to \frac{\pi}{2\sin(\pi\nu)}\left(\frac{1}{\Gamma(1-\nu)}\left(\frac{z}{2}\right)^{-\nu} - \frac{1}{\Gamma(1+\nu)}\left(\frac{z}{2}\right)^{\nu}\right). \quad (24)$$

After integration we have for small x:

$$\int_0^x p(x)dx \to \frac{\Gamma(L-\nu)}{\nu\,\Gamma(L)\Gamma(\nu)}\left(\frac{L\nu x}{\mu}\right)^\nu, L>\nu \quad (25)$$

Clearly this integral is finite for small x (and by continuity also applies to integer ν).

From (23) we can express the derivative of the logarithm of the pdf with respect to ν; close to the origin we have:

$$\frac{\partial \log p(x)}{\partial \nu} = C(L,\nu) + \log x \quad (26)$$

where

$$C(L,\nu) = 1 - \psi(\nu) - \psi(L-\nu) + \log(L\nu) \quad (27)$$

Here the digamma function, ψ, is defined by:

$$\psi(z) = \frac{1}{\Gamma(z)}\frac{d\Gamma(z)}{dz} \quad (28)$$

For example the contribution to the expectation of the efficient score by the distribution close to the origin is given by:

$$\int_0^x p(x)\frac{\partial \log p(x)}{\partial \nu}dx \rightarrow \frac{\Gamma(L-\nu)(L\nu)^{\nu}x^{\nu}}{\nu\Gamma(L)\Gamma(\nu)}\left(C(L,\nu)+\log x-\frac{1}{\nu}\right) \tag{29}$$

This is necessary to verify the integrity of the numerical calculations.

We also need the Fisher information i_{22} for small x, which is given by:

$$\int_0^x p(x)\left(\frac{\partial \log p(x)}{\partial \nu}\right)^2 dx \rightarrow \frac{\Gamma(L-\nu)(L\nu)^{\nu}}{\Gamma(L)\Gamma(\nu)}\int_0^x x^{\nu-1}(C+\log x)^2\,dx, \quad L>\nu \tag{30}$$

This can be integrated to yield:

$$i_{22} \rightarrow \frac{\Gamma(L-\nu)(L\nu x)^{\nu}}{\nu\,\Gamma(L)\Gamma(\nu)} \times\left(C^2-\frac{2C}{\nu}+\frac{2}{\nu^2}+2\left(C-\frac{1}{\nu}\right)\log x+\log^2 x\right) \tag{31}$$

We include here some detailed results that will be useful for similar calculations involving the estimator log(X). A calculation of the mean and variance of log(X) is not straightforward. The key is to note that the K-distribution arises from a multiplicative process in which speckle is modulated by a random texture function that is independent from it. In the present context, both random variables in the multiplication are gamma distributed. Therefore, when logarithms are taken the result is a sum and the Laplace transforms (or characteristic functions) are multiplied together. The moments of a K-distributed variable can then be found by the usual process of differentiating the Laplace transform.

The Laplace transform of a pdf, $p(x)$, is given by:

$$\overline{p}(s)=\int_0^\infty p(x)e^{-sx}\,dx\,, \tag{32}$$

where s is the transform variable. Now suppose a well behaved non-linear transformation of x to z is required, where $z = z(x)$. The relationship of $p(x)$ to $p(z)$ is given by:

$$p(z)=p(x)\frac{dx}{dz} \tag{33}$$

Therefore the Laplace transform of $p(z)$ becomes:

$$\overline{p}(s)=\int p(z)e^{-sz}\,dz =\int_0^\infty p(x)e^{-sz}\,dx \tag{34}$$

For the case of the transformation z = -log(x), we have:

$$\overline{p}(s)=\int_0^\infty p(x)e^{s\log(x)}\,dx \tag{35}$$

If $p(x)$ is a gamma pdf, we have:

$$\overline{p}(s)=\int_0^\infty \frac{\nu}{\Gamma(\nu)}\left(\frac{\nu x}{\mu}\right)^{\nu-1}e^{-\nu x}x^s\,dx =\left(\frac{\mu}{\nu}\right)^s\frac{\Gamma(s+\nu)}{\Gamma(\nu)} \tag{36}$$

It follows that when the two logarithmic variates are added in the compound K-distribution, the Laplace transform is given by:

$$\overline{p}(s)=\left(\frac{\mu}{\nu}\right)^s\frac{\Gamma(s+\nu)}{\Gamma(\nu)}L^{-s}\frac{\Gamma(s+L)}{\Gamma(L)} \tag{37}$$

By differentiating with respect to s and then setting s to zero, it is easy to show that [17]:

$$\begin{aligned}E(\ln X)&=\ln(\mu)+\psi(\nu)-\ln(\nu)+\psi(L)-\ln(L)\\ Var(\ln X)&=\psi'(\nu)+\psi'(L)\end{aligned}, \tag{38}$$

where ψ' is the derivative of ψ, namely a polygamma function.

When the logarithm of the intensity is averaged over a clutter cell, the distribution of the average tends to the normal distribution and is characterized by a mean and variance. Therefore, if there are N independent estimates of the logarithm, the variance of the average becomes (second line of (38)):

$$Var(E(\ln X))=\frac{\psi'(\nu)+\psi'(L)}{N} \tag{39}$$

However when $N >> 1$, by differentiation of the first line of (38) with respect to ν, we have

$$Var(E(\ln(X))\approx\left(\psi'(\nu)-\frac{1}{\nu}\right)^2 Var(\nu)\,, \tag{40}$$

so that [17]:

$$Var(\nu)\approx\frac{\psi'(\nu)+\psi'(L)}{N}\left(\psi'(\nu)-\frac{1}{\nu}\right)^{-2}. \tag{41}$$