

Research Article

Analysis of a Pareto Mixture Distribution for Maritime Surveillance Radar

Graham V. Weinberg and Daniel Finch

Electronic Warfare and Radar Division, Defence Science and Technology Organisation, P.O. Box 1500, Edinburgh, SA 5111, Australia

Correspondence should be addressed to Graham V. Weinberg, graham.weinberg@dsto.defence.gov.au

Received 24 April 2012; Accepted 6 August 2012

Academic Editor: Igor Djurović

Copyright © 2012 G. V. Weinberg and D. Finch. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Pareto distribution has been shown to be an excellent model for X-band high-resolution maritime surveillance radar clutter returns. Given the success of mixture distributions in radar, it is thus of interest to consider the effect of Pareto mixture models. This paper introduces a formulation of a Pareto intensity mixture distribution and investigates coherent multilook radar detector performance using this new clutter model. Clutter parameter estimates are derived from data sets produced by the Defence Science and Technology Organisation's Ingara maritime surveillance radar.

1. Introduction

In recent years at the Australian Defence Science and Technology Organisation (DSTO), clutter mixture models have been analysed, since they can provide tighter fits to spiky clutter in the distribution's upper tail region [1, 2]. The main application has been to improve the fit of the K-Distribution to real data, resulting in the KK-Distribution introduced in [1]. This mixture model involves two separate K-Distributions combined in the amplitude domain. The justification for this is that one component of the mixture models the Bragg/whitecap scatterers, while the other accounts for sea spikes. It was also shown in [1] that the KK-Distribution modelled real polarimetric high-resolution sea clutter returns better than a single K-Distribution, for the case of horizontally polarised clutter, which tends to be spikier than vertically polarised returns.

Recently, the Pareto distribution has been proposed as an alternative clutter model and has also been validated for high-resolution radar clutter returns [3, 4]. This intensity model has been found to fit both low and high grazing angle clutter returns, for X-band radar operating in a maritime surveillance environment. It has been found to match the performance of the K-Distribution and to be almost as accurate as the KK-model. There are examples where it

matches the performance of the KK-Distribution exactly [4]. The advantage of using a Pareto clutter model is that it is simpler mathematically, resulting in more basic and efficient detection schemes [5]. Thus it is of interest to examine whether, in comparison to the success of the KK-Distribution, a Pareto mixture model yields any improvement in clutter modelling and detection performance.

The paper is structured by introducing the Pareto distribution in Section 2 and illustrating why it is becoming of interest to the radar signal processing community. Section 3 proposes a Pareto mixture model and describes the result of fitting it to DSTO's high-resolution radar clutter sets. Section 4 derives the Neyman-Pearson optimal detector, based upon this clutter model. In addition, a generalised likelihood ratio test (GLRT) detector is produced. Finally, Section 5 gives some examples of detector performance curves, with parameters estimated from real clutter returns.

2. The Pareto Distribution

As maritime surveillance radar resolution has improved over the years, the backscattering from the sea surface, known as sea clutter, has been found to deviate significantly from the traditional Gaussian model [6]. As a consequence

of this, many models of clutter have been proposed and analysed, including the Lognormal [7], Weibull [8] and the K-Distribution [9, 10]. The KK-Distribution [11, 12] was proposed to see whether more accurate clutter fits could be achieved, especially for the very spiky horizontally polarised clutter returns. This model proved to be very successful but has a number of serious shortcomings. The first is that the model is a 5-parameter family, although [1] reduces this to 4, based upon DSTO X-band sea clutter data. Previous models studied in radar signal processing have been two parameter distributions. The complexity of the K- and KK-Distribution clutter models also results in difficult mathematical expressions for the coherent multilook detectors. A good example of this is the detector introduced in [13], which shows the KK-Distribution coherent multilook detector is analytically difficult to work with.

The Pareto distribution [14, 15] is named after the Italian economist Vilfredo Pareto (15 July 1848–19, August 1923) [16, 17] and is a power law probability distribution that has been found to be an excellent model of long-tailed phenomena [18]. Its original application was in the modelling of income over a population [19]. It has been used in the modelling of actuarial data; an example is in excess of loss quotations in insurance [20]. Its usefulness as a model for long-tailed distributions has resulted in applications to a number of diverse areas including physics, hydrology, and seismology [18]. It has also found suitable applications in engineering, such as internet teletraffic modelling [21]. From a radar perspective, its simplicity has resulted in the construction of much simpler detection schemes [5, 22, 23].

In the intensity domain, the Pareto distribution has a power law density given by

$$f_{\alpha_p, \beta_p}(t) = \frac{\alpha_p \beta_p^{\alpha_p}}{t^{\alpha_p+1}}, \quad (1)$$

for $t \geq \beta_p$. The parameter $\alpha_p > 0$ is called the distribution's shape parameter, which governs the tail behaviour. Parameter $\beta_p > 0$ is the scale parameter, which determines where the distribution's support begins. Its mean is given by $\mathbb{E}(X) = \alpha_p \beta_p / (\alpha_p - 1)$, while its variance is $\text{var}(X) = \beta_p^2 \alpha_p / (\alpha_p - 1)^2 (\alpha_p - 2)$. Hence for finite first and second moments, it is required that $\alpha_p > 2$.

Figure 1 shows an example of correlated Pareto intensity clutter, generated using the compound Gaussian model defined in [5]. The figure shows the clutter in both the time and frequency domain. In this case, the Pareto shape parameter is $\alpha_p = 25$, while its scale parameter is $\beta_p = 1$. The compound Gaussian model has a clutter covariance matrix Σ that is Toeplitz, with factor $\kappa = 0.01$, resulting in weak clutter returns. This means the (i, j) th entry of Σ is given by $\kappa^{|i-j|}$, for i and j in the set $\{1, 2, \dots, N\}$ with N the number of rows/columns of Σ .

Figure 2 shows the same Pareto distribution simulated with stronger clutter returns (Toeplitz factor $\kappa = 0.99$). In this case we see the correlation having a significant effect on the clutter returns.

These figures show the difficulty of performing target detection in a spiky clutter environment. Both Figures 1

and 2 illustrate that many small targets would easily be missed because of the spikiness of the clutter. In addition, false detections are likely for the same reason. Hence it is important to try to find detectors which allow increased probability of detection, even if there are only minor gains in performance. Small gains are significant for detection of small targets in spiky clutter environments.

3. A Pareto Mixture Model

The extension of the Pareto model (1) to a mixture model can be based upon the definition of the KK-Distribution in [1]. The Pareto mixture distribution, in the intensity domain, is defined by

$$f_{(\alpha_{pp_1}, \alpha_{pp_2}, \beta_{pp}, k)}(t) = (1 - k) \frac{\alpha_{pp_1} \beta_{pp}^{\alpha_{pp_1}}}{t^{\alpha_{pp_1}+1}} + k \frac{\alpha_{pp_2} \beta_{pp}^{\alpha_{pp_2}}}{t^{\alpha_{pp_2}+1}}, \quad (2)$$

where β_{pp} is the distribution's scale parameter (so that $t \geq \beta_{pp}$) and α_{pp_1} and α_{pp_2} are two shape parameters. These three parameters are all assumed to be positive real numbers. The mixing coefficient is $k \in [0, 1]$. The standard Pareto distribution density (1) is recovered by setting k to zero or one in (2). In contrast to the KK-Distribution, the Pareto mixture was chosen to have a single-scale parameter, and two shape parameters. The idea behind this was that since the scale parameter determines the support, it would be useful to have only one such parameter and allow for greater distributional fitting through two shape parameters. In MATLAB, an algorithm was created to optimise the fit of the Pareto mixture to real data, in a similar vein to MATLAB's automatic Pareto fit function *gpfit*, using maximum likelihood estimation. This algorithm fitted the density (2) to individual data sets based upon the four distributional parameters (α_{pp_1} , α_{pp_2} , β_{pp} , and k).

Figure 3 shows an example of the fit of a Pareto mixture distribution to real data, using empirical distribution functions. The data used for this purpose is the DSTO Ingara maritime sea clutter trials data, which is fully polarised high-resolution X-band clutter returns. Details of the data and radar can be found in [1, 24]. This data was collected during a trial in 2004, in the Southern Ocean, and has been well documented. A description of the radar and clutter analysis can be found in [1], while the fitting of the Pareto distribution to this data is described in [4]. Key points about the trial are that the radar operated in a circular spotlight mode, scanning the same patch of sea surface at different azimuth angles. The Ingara radar is a fully polarimetric, operating in X-band [24]. No targets were present in the trial, so that the returns are pure clutter.

This figure shows the Ingara data corresponding to run 34690, at azimuth angle of 255° , with horizontal polarisation, as considered in [4]. The fitted Pareto parameters are $\alpha_p = 3.191$ and $\beta_p = 0.0047$, while the Pareto mixture parameters are $\alpha_{pp_1} = 3.6637$, $\alpha_{pp_2} = 7.0837$, $\beta_{pp} = 0.0083$, and $k = 0.5402$. For comparison, a K- and KK-Distribution were also fitted, using the parameter estimation guidelines in [1]. The fitted K-Distribution has scale parameter $c = 74.45$

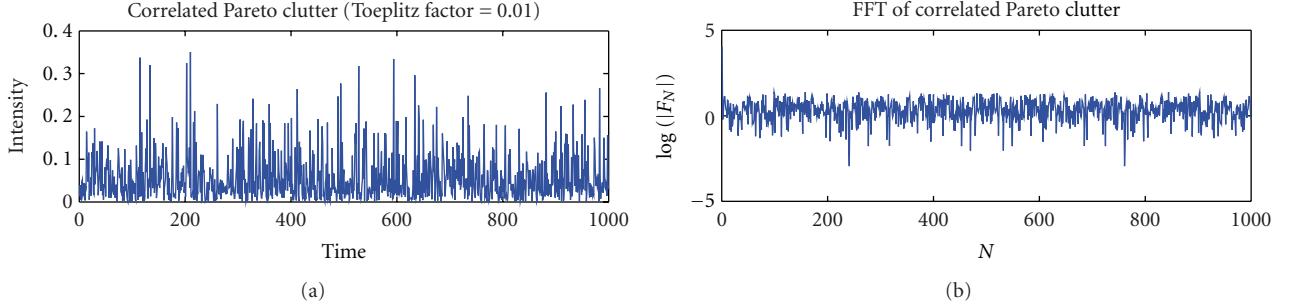


FIGURE 1: An example of correlated Pareto clutter returns, where the correlation is small. (a) shows the clutter in the time domain, while the (b) is the result of a fast Fourier transform (FFT) applied to the same clutter.

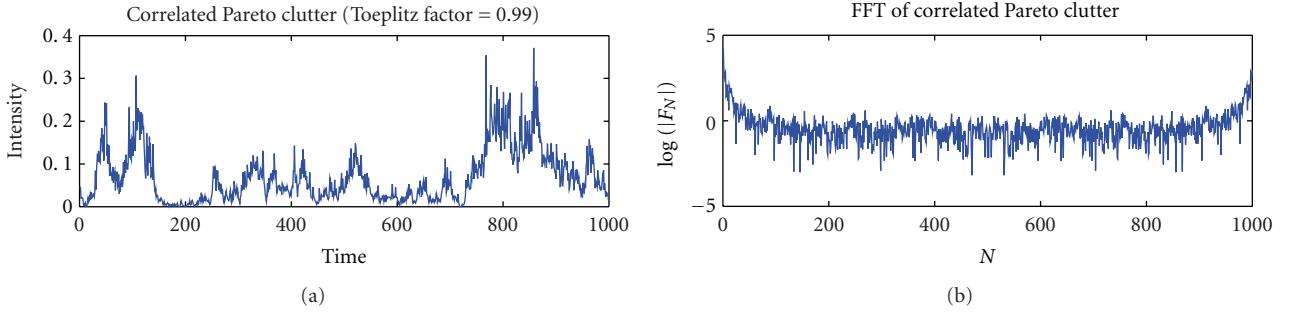


FIGURE 2: Correlated Pareto clutter, where the correlation is strong.

and shape parameter $\nu = 2.863$, while the KK-Distribution has scale parameters $c_1 = 71.45$ and $c_2 = 344.39$, the same shape parameter and mixing coefficient $k = 0.01$. In this case the K- and KK-Distributions are extremely close together, while the Pareto mixture (denoted PP in Figure 3) fits the data more closely. Under magnification, the KK-Distribution is closer to the data than the K-Distribution, especially in the upper tail region. The Pareto mixture clearly fits the data better than a single Pareto distribution. This was observed in all the Ingara data examined, implying that the Pareto mixture model gave a better fit to the data.

4. Detector Performance

Neyman-Pearson detectors [15] can be constructed using the Pareto mixture distribution as an intensity clutter model. The formulation of the basic detection problem is outlined briefly, following the set up in [5]. The radar return is denoted \mathbf{z} , which is a vector of length N , corresponding to the number of looks. The coherent multilook detection problem is stated as the statistical hypothesis test $H_0 : \mathbf{z} = \mathbf{c}$ against $H_1 : \mathbf{z} = R\mathbf{p} + \mathbf{c}$, where all complex vectors are $N \times 1$. The vector \mathbf{p} is the Doppler steering vector, whose components are given by $\mathbf{p}(j) = e^{j2\pi f_D j}$, for $j \in \{1, 2, \dots, N\}$, where f_D is the normalised target Doppler frequency ($-0.5 \leq f_D \leq 0.5$) and is assumed to be completely known. The complex random variable R represents the target, and $|R|$ is its amplitude. The clutter vector \mathbf{c} will be modelled as a spherically invariant random process (SIRP). This means we assume that the clutter takes the form $\mathbf{c} = S\mathbf{g}$, where \mathbf{g}

is a zero mean complex Gaussian process with covariance matrix Σ and S is a univariate nonnegative random variable with density f_S . Useful references on the SIRP approach include [25, 26]. As in [5], it will be assumed that Σ is semi-positive definite, implying that a Cholesky factorisation can be applied to its inverse. Consequently, there exists a matrix A such that $\Sigma^{-1} = A^H A$, where A^H is the Hermitian transpose. This enables a whitening approach to be applied to the original hypothesis test, since SIRPs are unaffected by a linear transformation [25]. As a result of this, the original hypothesis test can be re-expressed as $H_0 : \mathbf{r} = \mathbf{n}$ versus $H_1 : \mathbf{r} = R\mathbf{u} + \mathbf{n}$, where $\mathbf{r} = Az$, $\mathbf{n} = Ac$ and $\mathbf{u} = Ap$. The advantage of this is that the clutter process $\mathbf{n} = SA\mathbf{g}$ is, when conditioned on S , a complex Gaussian process because $A\mathbf{g}$ is multidimensional complex Gaussian process, with zero mean but covariance the $N \times N$ identity matrix. This enables the corresponding densities to be written down more simply. In particular, since $f_{n|S}(\mathbf{x}) = (1/\pi^N s^{2N}) e^{-s^{-2}\|\mathbf{x}\|^2}$, it can be shown that

$$f_n(\mathbf{x}) = \frac{1}{\pi^N} h_N(\|\mathbf{x}\|^2), \quad (3)$$

where

$$h_N(x) = \int_0^\infty s^{-2N} \exp\left(-\frac{x}{s^2}\right) f_S(s) ds \quad (4)$$

is called the characteristic function [25].

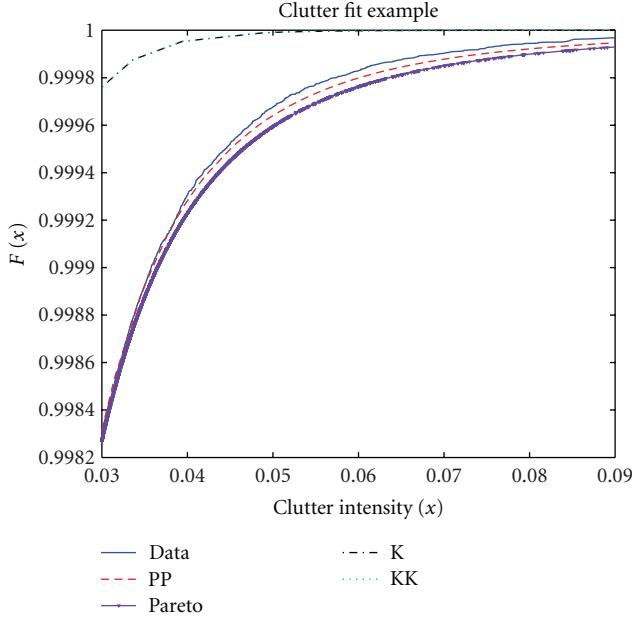


FIGURE 3: Clutter fit example, showing the empirical distribution functions $F(x)$ for real Ingara data compared to the fitted Pareto, Pareto mixture (PP), K- and KK-Distributions, as a function of intensity x . Here the latter two coincide almost exactly, while the Pareto mixture provides the best fit.

For the case of a fixed target model (so that R is constant), the density under H_1 can be shown to be

$$f_{H_1}(\mathbf{r}) = \frac{1}{\pi^N} h_N(\|\mathbf{r} - R\mathbf{u}\|^2). \quad (5)$$

To extend this to the case where R is unknown the GLRT uses the approximation:

$$\|\mathbf{r} - R\mathbf{u}\|^2 \approx \|\mathbf{r}\|^2 - \frac{|\mathbf{u}^H \mathbf{r}|^2}{\|\mathbf{u}\|^2} \quad (6)$$

in the density under H_1 .

In order to examine the performance of detectors based upon a Pareto mixture distribution assumption, it is necessary to specify the SIRP that generates the desired marginal intensity distributions. To generate a Pareto mixture SIRP, one uses the characteristic function:

$$h_N(p) = (1 - k)\zeta(p, \alpha_{pp_1}, \beta_{pp}, N) + k\zeta(p, \alpha_{pp_2}, \beta_{pp}, N), \quad (7)$$

where ζ is defined by

$$\zeta(p, \alpha, \beta, N) = \frac{\beta^\alpha \Gamma(N + \alpha)}{\Gamma(\alpha)[p + \beta]^{N+\alpha}}. \quad (8)$$

It can be shown that the SIRP with characteristic function (7) has Pareto mixture marginal intensity distributions. The Neyman-Pearson optimal detector [15] is given by the ratio of densities under the respective hypotheses:

$$L(\mathbf{r}) = \frac{f_{H_1}(\mathbf{r})}{f_{H_0}(\mathbf{r})} \stackrel{H_1}{\gtrless} \tau, \quad (9)$$

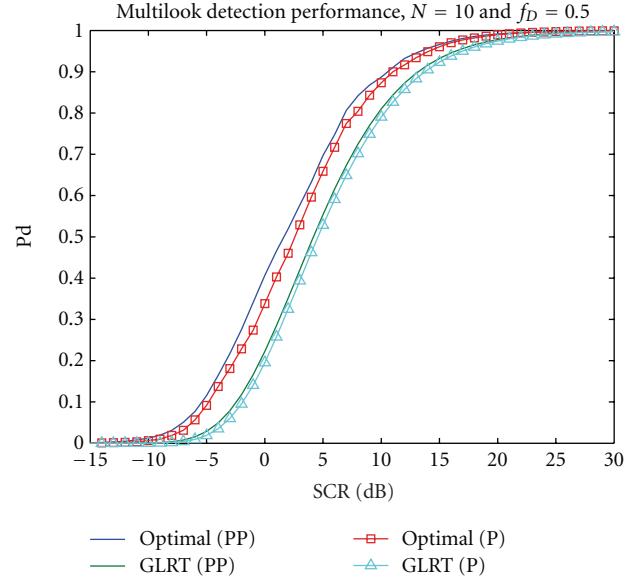


FIGURE 4: Comparison of detector performance curves for Pareto and Pareto mixture models. The probability of detection (P_d) is plotted as a function of the signal to clutter ratio (SCR). The optimal detectors are denoted Optimal (PP) for the mixture model and Optimal (P) for the standard Pareto counterpart. The GLRT for the Pareto mixture is denoted GLRT (PP), while the same for a single Pareto model is GLRT (P). The figure shows that the mixture model provides only a very small detection gain.

where $L(\mathbf{r})$ is the likelihood function and τ is the detection threshold. The notation (9) means we reject the null hypothesis H_0 (return is pure clutter) if the likelihood exceeds the threshold (set by the probability of false alarm). When we reject H_0 we accept the alternative hypothesis H_1 , implying that there is strong evidence of a target present in the radar return \mathbf{r} . The densities (3) and (5), together with (7), can be applied to (9) to construct the optimal detector explicitly. The GLRT is obtained explicitly by applying (6) to (5) prior to it being applied to (9).

5. Detector Performance Considerations

Using the DSTO Ingara data sets, the performance of the optimal detector, and GLRT, for Pareto mixture clutter models was assessed. Due to the fact that the Ingara data is pure clutter, a synthetic Gaussian target model has been used throughout, as in [5]. Pareto and Pareto mixture clutter models have had their parameters estimated from Ingara data sets. The detector performance curves have then been generated using Monte Carlo simulation, using 10^6 runs in each case.

The comparison of detectors based upon the Pareto and Pareto mixture clutter models produced some very interesting results. In most cases it was observed that the mixture distribution assumption did not improve detector performance significantly. This is illustrated in Figure 4, which is a detection performance, plotting the probability of detection (P_d) as a function of the target signal to clutter

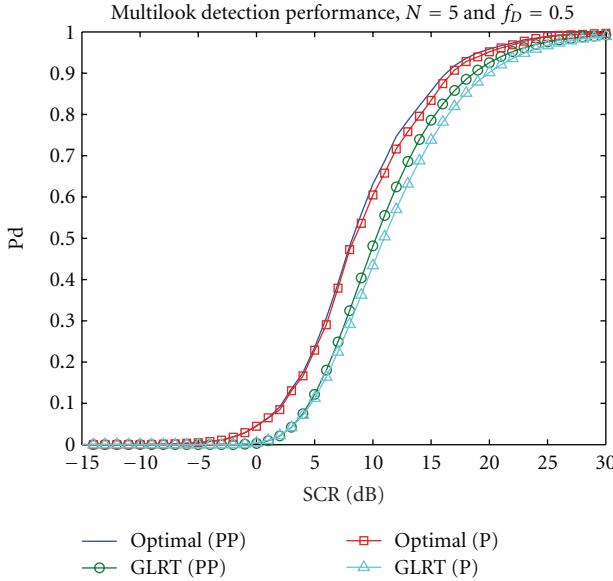


FIGURE 5: A second example of detector performance, with same legend as for Figure 1. In this case, the clutter strength has been increased (Toeplitz factor of 0.7) and the number of looks decreased. As in the previous figure, the detection gain provided by a mixture model is very small.

ratio (SCR). In this example, the mixture distribution has $\alpha_{pp_1} = 20$, $\alpha_{pp_2} = 100$, $\beta_{pp} = 0.05$, and $k = 0.8$. The single Pareto distribution has $\alpha_p = 2.347$ and $\beta_p = 0.001277$. Additionally, $N = 10$ looks and the normalised Doppler frequency is $f_D = 0.5$. The false alarm probability has been set to 10^{-6} . The simulated clutter has been generated with a Toeplitz covariance matrix with $\kappa = 0.4$. This example corresponds to a typical horizontally polarised Ingara data set. Figure 4 shows the optimum detector result as well as the GLRT detector. As can be observed, although the mixture model provided a better clutter fit, there is only a minor gain in detection performance.

Figure 5 is a second example of detector performance. Here, the same clutter parameters and normalised Doppler frequency have been used as for Figure 4. However, the number of looks is $N = 5$ and the Toeplitz factor has been increased to 0.7. Additionally, the false alarm probability has been reduced to 10^{-7} . Comparing the pairs of decision rules, we again observe that the mixture model does not provide a significant improvement in detection performance.

Varying the number of looks, normalised Doppler frequency and Toeplitz covariance factor did not significantly alter the results. Hence it appears that although the KK-mixture distribution provides significant detection performance relative to that provided by a detector based upon a single K-Distributed clutter assumption [11], the same is not repeated when the Pareto assumption is applied. This implies a single Pareto distribution is sufficient from a detection performance perspective.

6. Conclusions

Although a Pareto mixture model can provide a better fit to the upper tail region in real radar clutter returns, this has been observed to not significantly improve detection performance, which is the primary application of radar. Increasing the number of degrees of freedom in a model adds computational cost, slowing the radar's ability to perform its operation in real time. Hence it is computationally better to use a nonmixture model when designing radar detection schemes, under the basic Pareto clutter model assumption.

References

- [1] Y. Dong, "Distribution of X-band high resolution and high grazing angle sea clutter," DSTO Report RR-0316, 2006.
- [2] G. V. Weinberg, "Neyman-pearson optimal and suboptimal detection for signals in general clutter mixture distributions," DSTO Report RR-0363, 2011.
- [3] M. Farshchian and F. L. Posner, "The pareto distribution for low grazing angle and high resolution X-band sea clutter," in *Proceedings of the IEEE International Radar Conference 2010, RADAR 2010*, pp. 789–793, May 2010.
- [4] G. V. Weinberg, "Assessing Pareto fit to high-resolution high-grazing-angle sea clutter," *Electronics Letters*, vol. 47, no. 8, pp. 516–517, 2011.
- [5] G. V. Weinberg, "Coherent multilook radar detection for targets in pareto distributed clutter," *Electronics Letters*, vol. 47, no. 14, pp. 822–824, 2011.
- [6] N. Levanon, *Radar Principles*, Wiley, New York, NY, USA, 1988.
- [7] G. V. Trunk and S. F. George, "Detection of Targets in Non-Gaussian Sea Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 6, no. 5, pp. 620–628, 1970.
- [8] D. C. Schleher, "Radar detection in weibull clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-12, no. 6, pp. 736–743, 1976.
- [9] E. Jakeman and P. N. Pusey, "A model for non-rayleigh sea echo," *IEEE Transactions on Antennas and Propagation*, vol. AP-24, no. 6, pp. 806–818, 1976.
- [10] Y. Dong, "Optimal coherent radar detection in a K-distributed clutter environment," *IET Radar, Sonar and Navigation*, vol. 6, no. 5, pp. 283–292, 2012.
- [11] L. Rosenberg, D. J. Crisp, and N. J. Stacy, "Analysis of the KK-distribution with medium grazing angle sea-clutter," *IET Radar, Sonar and Navigation*, vol. 4, no. 2, pp. 209–222, 2010.
- [12] Y. Dong and B. Haywood, "High grazing angle X-band sea clutter distributions," in *Proceedings of the RADAR 2007—The Institution of Engineering and Technology International Conference on Radar Systems*, October 2007.
- [13] G. V. Weinberg, "Coherent multilook radar detection for targets in KK-distributed clutter," in *Digital Communication*, C. Palanisamy, Ed., chapter 9, Intech, Rijeka, Croatia, 2012.
- [14] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*, Wiley, New York, NY, USA, 3rd edition, 2000.
- [15] G. P. Beaumont, *Intermediate Mathematical Statistics*, Chapman and Hall, London, UK, 1980.
- [16] G. Bruno, "Pareto, Vilfredo," in *The New Palgrave: A Dictionary of Economics*, vol. 5, pp. 799–804, 1987.
- [17] R. Cirillo, *The Economics of Vilfredo Pareto*, Frank Cass Publishers, London, UK, 1978.
- [18] I. B. Aban, M. M. Meerschaert, and A. K. Panorska, "Parameter estimation for the truncated pareto distribution," *Journal*

- of the American Statistical Association, vol. 101, no. 473, pp. 270–277, 2006.
- [19] A. V. Asimit, E. Furman, and R. Vernic, “On a multivariate pareto distribution,” *Insurance*, vol. 46, no. 2, pp. 308–316, 2010.
 - [20] M. Rytgaard, “Estimation in the pareto distribution,” *ASTIN Bulletin*, vol. 20, no. 2, pp. 201–216, 1990.
 - [21] E. Chlebus and R. Ohri, “Estimating parameters of the pareto distribution by means of zipf’s law: application to internet research,” in *Proceedings of the IEEE Globecom*, pp. 1039–1043, December 2005.
 - [22] X. Shang and H. Song, “Radar detection based on compound-gaussian model with inverse gamma texture,” *IET Radar, Sonar and Navigation*, vol. 5, no. 3, pp. 315–321, 2011.
 - [23] K. J. Sangston, F. Gini, and M. S. Greco, “New results on coherent radar target detection in heavy-tailed compound-gaussian clutter,” in *Proceedings of the IEEE International Radar Conference*, pp. 779–784, May 2010.
 - [24] N. J. S. Stacy, D. Crisp, A. Goh, D. Badger, and M. Preiss, “Polarimetric analysis of fine resolution X-band SAR sea clutter data,” in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS ’05)*, pp. 2787–2790, July 2005.
 - [25] M. Rangaswamy, D. Weiner, and A. Oeztuerk, “Non-Gaussian random vector identification using spherically invariant random processes,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 1, pp. 111–124, 1993.
 - [26] K. Yao, “Radar detection based on compound-Gaussian model with inverse gamma texture,” *IEEE Transactions on Information Theory*, vol. IT-19, no. 5, pp. 600–608, 1973.

