A Bayesian-Based CFAR Detector for Pareto Type II Clutter

Graham V. Weinberg, Stephen D. Howard and Charlie Tran

National Security and Intelligence, Surveillance and Reconnaissance Division
Defence Science and Technology Group
P. O. Box 1500 Edinburgh, South Australia 5111, Australia
(Draft created at 10.22 am)
Graham.Weinberg@dst.defence.gov.au

Abstract—The Pareto class of distributions has been shown to be suitable intensity models for X-band maritime surveillance radar clutter. Hence there has been much interest in the construction of non-coherent detection processes with the constant false alarm rate (CFAR) property. As a result of the validity of the Pareto Type I model, as an approximation for the Pareto Type II, it has been possible to derive a large number of relevant detection processes. However, recent work has investigated whether it is possible to achieve CFAR with respect to the Pareto Type II class of models directly. It will be shown that the solution to this can be achieved via a Bayesian predictive inference approach. A detector, which is completely CFAR in Pareto Type II clutter, will be derived and its performance in radar clutter examined.

Index Terms—Non-coherent radar detection; Constant false alarm rate; Pareto Type II clutter; Bayesian detectors

I. INTRODUCTION

Since the introduction of the Pareto class of distributions, as a model for X-band maritime surveillance radar clutter, there has been a resurgence of interest in the development of non-coherent sliding window detection processes for target detection in such clutter [1]. Although thermal noise is present in a radar receiver, and can be modelled by an effective shape parameter as in [2], it is not considered in the analysis to follow. Validation of the Pareto model for X-band maritime surveillance radar data was first provided in [3], with subsequent studies of its validation examined in [4] and [5]. It has been found that the Pareto Type II model to Defence Science and Technology (DST) Group’s real data resulted in estimates of the Pareto scale parameter which were always bounded by unity. Since the Pareto Type II model was being fitted to the data, this discovery justified the application of the Pareto Type I model to the design of non-coherent radar detection schemes. As a result of this approximation, it was possible to produce a large number of sliding window detectors, with varying degrees of CFAR control [6] - [16]. Despite this, it has been of interest to investigate the design of sliding window CFAR detectors for operation in Pareto Type II clutter directly. Towards this aim, [17] investigated the development of detectors, for this clutter environment, via the transformation approach of [18]. It was shown that it was possible to achieve CFAR with respect to the Pareto shape parameter, but the resultant detectors required a priori knowledge of the Pareto scale parameter. Recently, [19] introduced a predictive inference approach for the design of such detectors. This allowed the construction of CFAR decision rules via Bayesian analysis. In the case of scale and power invariant distributions, as discussed in [19], it is possible to derive an optimal CFAR detector. Here optimality is in the sense of strong consistency [20]. Therefore it is of interest to derive the Bayesian based detector, in the spirit of [19], for the case of target detection in Pareto Type II clutter. Towards this objective, the paper is organised as follows. Section II introduces the Pareto Type II model and the formulation of the fundamental detection process. Section III discusses the choice of prior distribution for the predictive inference approach. Section IV then derives the Bayesian CFAR detector. Finally Section V provides examples of the performance of the Bayesian CFAR with reference to detectors derived in [17].

II. MODEL AND DETECTION PROCESS

Recall that the density of the Pareto Type II model is given by

\[ f(t; \alpha, \beta) = \frac{\alpha \beta^\alpha}{(t + \beta)^{\alpha+1}}, \]  \hspace{1cm} (1)

where \( \alpha > 0 \) is the shape parameter and \( \beta > 0 \) is the scale parameter. The support of this distribution is the interval \([0, \infty)\). Hence a fundamental difference between this distribution, and that of a Pareto Type I model, is that in the latter case the support begins at \( \beta \). However, the Pareto Type II model is the intensity distribution of a compound Gaussian process with inverse gamma texture [3], and therefore is consistent with the currently accepted radar clutter model phenomenology.

The detection processes of interest here, known as a sliding window process, assumes the existence of a series of independent and identically distributed clutter measurements, denoted by statistics \( Z_1, Z_2, \ldots, Z_N \), and referred to as the clutter range profile (CRP). A cell under test (CUT) is then considered, which is represented by \( Z_0 \). In practical implementation of this detection process, a series of guard cells is often applied to limit the effects of a range spread target. Based upon a normalised measurement of the clutter level, the existence of a target in the CUT is then investigated. It is
assumed that the CUT and CRP are also independent. Suppose that \( g(Z_1, Z_2, \ldots, Z_N) \) is the measurement of the clutter level, for some known function \( g \). Then if \( H_0 \) is the hypothesis that the CUT does not contain a target, and \( H_1 \) the hypothesis that the CUT contains a target embedded within clutter, then the test can be written

\[
Z_0 \xrightarrow{H_0} \tau g(Z_1, Z_2, \ldots, Z_N),
\]

(2)

where the notation employed in (2) means that \( H_0 \) is only rejected when \( Z_0 > \tau g(Z_1, Z_2, \ldots, Z_N) \). Such detection processes were first introduced in [21], in the context of exponentially distributed intensity clutter. The parameter \( \tau > 0 \) is known as a threshold multiplier, which is used to produce an adaptive threshold that maintains the desired probability of false alarm (Pfa) in homogeneous clutter. Detectors of the form (2) have been studied extensively in the case where the CRP is modelled by exponentially distributed clutter returns [22]. In fact, in such a situation it can be shown that the Pfa of (2), given by

\[
P_{fa} = P(Z_0 > \tau g(Z_1, Z_2, \ldots, Z_N)|H_0),
\]

(3)

does not depend on the clutter mean parameter, provided \( g \)

is a scale-invariant function. This implies that such detectors achieve CFAR in ideal situations [23]. The extension of detectors of the form (2) to non-exponentially distributed clutter scenarios often results in partial loss of the CFAR property. This is in particular true when (2) is applied in the case of Pareto Type II clutter [17].

### III. Selection of Prior Distribution

In the Bayesian approach of [19] one assumes that the clutter distributional parameters are realisations of stochastic variables, and the objective is to derive the Bayes predictive distribution of the CUT conditioned on the CRP [24]. This then permits a decision rule to be specified. However it is necessary to determine appropriate prior distributions for the variables modelling the clutter parameters. Hence the purpose of this section is to derive a suitable prior distribution in the context of target detection in a Pareto Type II clutter model environment. As is explained in [19], distributions which can be classified as scale and power invariant permit the selection of the right Haar measure as a non-informative prior distribution. Then one can appeal to the theory of strong consistency to argue that the resultant choice yields an optimal selection for the prior distribution. However in the case of Pareto Type II clutter this is not the case, since this distribution is only scale invariant. This implies, however, that if one assumes that the Pareto Type II shape parameter is known \( a \) \textit{a priori} then one can apply the right Haar measure as an improper prior for the Pareto scale parameter. For the case where both the clutter parameters are assumed unknown, it is necessary to examine the Fisher information matrix to determine a Jeffreys prior for the problem [25].

Towards this objective, one requires expectations of derivatives of the logarithm of the Pareto Type II density (1). Observe that

\[
\log f(t; \alpha, \beta) = \log \alpha + \alpha \log \beta - (\alpha + 1) \log (t + \beta).
\]

(4)

The Fisher information matrix, for a distribution with two unknown population parameters \( \alpha \) and \( \beta \) and density \( f(t; \alpha, \beta) \), is given by

\[
\mathcal{F} = \begin{bmatrix}
-\mathbb{E} \left( \frac{\partial^2 \log f}{\partial \alpha^2} \right) & -\mathbb{E} \left( \frac{\partial^2 \log f}{\partial \alpha \partial \beta} \right) \\
-\mathbb{E} \left( \frac{\partial^2 \log f}{\partial \beta \partial \alpha} \right) & -\mathbb{E} \left( \frac{\partial^2 \log f}{\partial \beta^2} \right)
\end{bmatrix},
\]

(5)

where the expectation is understood to be with respect to the underlying distributional model. By differentiating (1), and evaluating the expectation of the resultant expressions with respect to (1), one can show that (5) reduces to

\[
\det \mathcal{F} = \frac{1}{\alpha(\alpha+1)^2(\alpha+2)}.
\]

Hence evaluating the determinant of the matrix (6), one determines that the Jeffreys prior is proportional to

\[
\sqrt{\det \mathcal{F}} =: f(\alpha, \beta) = \frac{1}{\sqrt{\alpha(\alpha+1)^2(\alpha+2)}}.
\]

(7)

As will become evident, the prior (7) results in a detector with a high computational complexity, and so it is necessary to consider whether another prior is suitable. Note that in (7) the right Haar measure appears, in terms of the Pareto scale parameter, as the reciprocal of \( \beta \). Hence this can be used as a basis for the construction of a simpler prior.

As remarked previously, if one was to assume that the Pareto scale parameter \( \beta \) was known \( a \) \textit{priori}, then using the Fisher information matrix one can show that the Jeffreys prior is proportional to the reciprocal of the Pareto shape parameter. Consequently one can apply the prior proportional to

\[
f(\alpha, \beta) = \frac{1}{\alpha \beta},
\]

(8)

for the Pareto Type II case. This can be viewed as a simpler alternative to (7), used to facilitate the design of the Bayesian decision rule. It is interesting to observe that this prior is exactly that used in [19] for the Pareto Type I clutter environment. This prior also results in a substantial simplification of the resultant Bayesian decision rule, in contrast to that based upon (7). Hence in the following the prior (8) will be adopted for the construction of the Bayesian decision rule.

### IV. Bayesian Decision Rule

In order to produce a detection process which achieves the full CFAR property in Pareto Type II clutter it is necessary to derive the Bayes predictive inference density as in [19]. Suppose that \( A \) and \( B \) are random variables representing the Pareto shape and scale parameters respectively. Given the estimators of these parameters in the classical interpretation are independent [26] it will be assumed that these random variables are also independent. Here (8) will be applied as the non-informative prior. What is necessary is to determine the joint conditional distribution of the CUT under \( H_0 \) given
the CRP, which is the Bayes predictive distribution. This will then enable one to determine an expression for the Pfa for the sliding window detection process, based upon a Bayes decision rule.

For a series of independent and identically distributed Pareto Type II clutter returns with marginal probability distribution function (1), the joint density of the CRP is

\[
\begin{align*}
    & f_{Z_1, Z_2, \ldots, Z_N|A,B}(z_1, z_2, \ldots, z_N|\alpha, \beta) = \frac{\alpha^N B^\alpha N}{\prod_{j=1}^{N}(z_j + \beta)^{\alpha+1}}. \\
\end{align*}
\]

(9)

Hence, by an appeal to Bayes’ Theorem, the joint density of the Pareto parameters, conditioned on the CRP, is given by

\[
\begin{align*}
    & f_{A,B|Z_1, Z_2, \ldots, Z_N}(\alpha, \beta|z_1, z_2, \ldots, z_N) = \\
    & \frac{\alpha^N B^\alpha N}{(N-1)! I \prod_{j=1}^{N}(z_j + \beta)^{\alpha+1}}. \\
\end{align*}
\]

(10)

where (8) has been used as the non-informative prior. The integral with respect to \( \alpha \) can be evaluated by application of the gamma function, and so it can be shown that (10) reduces to

\[
\begin{align*}
    & f_{A,B|Z_1, Z_2, \ldots, Z_N}(\alpha, \beta|z_1, z_2, \ldots, z_N) = \\
    & \frac{\alpha^N B^\alpha N}{(N-1)! I \prod_{j=1}^{N}(z_j + \beta)^{\alpha+1}}. \\
\end{align*}
\]

(11)

where

\[
I = \int_0^\infty \frac{1}{\prod_{j=1}^{N}(z_j + \beta)} \left[ \log \left( \prod_{j=1}^{N} \frac{z_j + \beta}{\beta} \right) \right]^{-N} d\beta \beta. \\
\]

(12)

Hence the Bayes predictive distribution has density under \( H_0 \)

\[
\begin{align*}
    & f_{Z_0|Z_1, Z_2, \ldots, Z_N}(z_0|z_1, z_2, \ldots, z_N) = \\
    & \int_0^\infty \frac{1}{\prod_{j=1}^{N}(z_j + \beta)} \left[ \log \left( \prod_{j=1}^{N} \frac{z_j + \beta}{\beta} \right) \right]^{-N} d\beta \beta. \\
\end{align*}
\]

(13)

where the fact that the CUT, conditioned on \( \alpha \) and \( \beta \), has a Pareto Type II distribution has been applied. As before the integral with respect to \( \alpha \) can be evaluated to show that (13) reduces to

\[
\begin{align*}
    & f_{Z_0|Z_1, Z_2, \ldots, Z_N}(z_0|z_1, z_2, \ldots, z_N) = \\
    & \frac{N}{I} \times \\
    & \int_0^\infty \frac{1}{(z_0 + \beta) \prod_{j=1}^{N}(z_j + \beta)} \left[ \log(z_0 + \beta) - (N+1) \log(\beta) + \sum_{j=1}^{N} \log(z_j + \beta) \right]^{-N-1} d\beta \beta. \\
\end{align*}
\]

(14)

The probability of false alarm of the test \( Z_0 \overset{\text{Pfa}}{\geq} \tau \) is thus given by

\[
\begin{align*}
    & \text{Pfa} = \int_{\tau}^{\infty} f_{Z_0|Z_1, Z_2, \ldots, Z_N}(z_0|z_1, z_2, \ldots, z_N) \, dz_0 \\
    & = \frac{1}{I} \int_0^{\infty} \frac{1}{\prod_{j=1}^{N}(z_j + \beta)} \left[ \log(\tau + \beta) - (N+1) \log(\beta) + \sum_{j=1}^{N} \log(z_j + \beta) \right]^{-N} d\beta \beta. \\
\end{align*}
\]

(15)

where the double integral implicit in the above has been reduced to a single integral by switching the order of integration, performing integration with respect to \( z_0 \). By applying a change of variables with \( \phi = \beta^{-1} \) one can demonstrate that (15) reduces to

\[
\begin{align*}
    & \text{Pfa}(\tau) = \\
    & \int_0^{\infty} \frac{\phi^{N-1}}{\prod_{j=1}^{N}(\phi z_j + 1)} \left[ \log(\tau \phi + 1) + \sum_{j=1}^{N} \log(\phi z_j + 1) \right]^{-N} d\phi \\
    & \quad \times \int_0^{\infty} \frac{\phi^{N-1}}{\prod_{j=1}^{N}(\phi z_j + 1)} \left[ \sum_{j=1}^{N} \log(\phi z_j + 1) \right]^{-N} d\phi. \\
\end{align*}
\]

(16)

where the Pfa has been written as a function of \( \tau \). In principle, for a given Pfa, one must solve numerically for \( \tau \) in (16)

However, it is more convenient to observe that since the Pfa is a decreasing function of \( \tau \), the event \( z_0 > \tau \) is equivalent to \( \text{Pfa}(z_0) < \text{Pfa}(\tau) \). Since the latter is the desired Pfa one can instead compute \( \text{Pfa}(z_0) \) and reject \( H_0 \) if this is smaller than the desired Pfa.

Based upon this approach, one can conclude that the Bayesian test is to reject \( H_0 \) if the test statistic \( T(Z_0, Z_1, \ldots, Z_N) \) is negative, where

\[
\begin{align*}
    & T(Z_0, Z_1, \ldots, Z_N) = \\
    & \int_0^{\infty} \frac{\phi^{N-1}}{\prod_{j=1}^{N}(\phi z_j + 1)} \times \\
    & \left[ \log(Z_0 \phi + 1) + \sum_{j=1}^{N} \log(\phi z_j + 1) \right]^{-N} d\phi. \\
\end{align*}
\]

(17)

Although (17) provides a CFAR decision rule with respect to the Pareto Type II shape and scale parameter, it cannot be written in a closed form expression. Hence the test must be applied via numerical evaluation of the integral, noting its sign is only of interest. It is also worth noting that if one was to apply the non-informative prior (7), instead of (8), there would not be a reduction of the Bayesian decision rule to a single
integral. The resultant detector has been found to have a high computational complexity, making performance analysis very difficult.

The next section examines the performance of (17) relative to the decision rules introduced in [17].

V. PERFORMANCE ANALYSIS

In order to examine the performance of the Bayesian CFAR detector, clutter is simulated from a Pareto Type II distribution with parameters \( \alpha = 4.7241 \) and \( \beta = 0.0446 \). These parameters have been taken from fits of the Pareto Type II model to DST Group’s Ingara radar clutter. Ingara is an experimental imaging X-band radar [27], which has been deployed in a number of data gathering exercises [28], [29]. The particular data set from which the Pareto clutter parameters have been estimated is known as run 34683, obtained at an azimuth angle of \( 225^\circ \) and acquired with horizontal transmit and receive polarisation.

Several comparison detectors are also provided, which have been introduced in [17]. The first is a transformed geometric mean (GM) detector, with decision rule

\[
Z_0 > \frac{\beta}{\tau_0} \left( \prod_{j=1}^{N} \left( 1 + \frac{Z_j}{\beta} \right)^{\tau} - 1 \right), \tag{18}
\]

with \( \tau = P_{FA}^{-1/N} - 1 \).

A second decision rule, also examined in [17], is

\[
Z_0 > \frac{\beta}{\tau_0} \left( \left( 1 + \frac{Z_{(k)}}{\beta} \right)^{\tau} - 1 \right) \tag{19}
\]

where \( Z_{(k)} \) is the \( k \)th order statistic of the CRP, where \( \tau \) is determined by numerical inversion of

\[
P_{FA} = \frac{N! \Gamma(N-k+\tau+1)}{(N-k)! \Gamma(N+\tau+1)}, \tag{20}
\]

where \( \Gamma \) is the gamma function. This detector is known as an order statistic (OS) detector.

A third detector is the classical OS decision rule

\[
Z_0 > \frac{\beta}{\tau_0} \tau Z_{(k)}, \tag{21}
\]

whose threshold multiplier, for the clutter environment under consideration, is determined via inversion of

\[
P_{FA} = k \binom{N}{k} \int_0^1 (1-\Phi)^{k-1} \Phi^{N-k} \left( 1 + \tau \left( \Phi^{-\frac{1}{\beta}} - 1 \right) \right)^{-\alpha} d\Phi, \tag{22}
\]

It is clear that the first two detectors are CFAR with respect to the Pareto shape parameter but require \textit{a priori} knowledge of the Pareto scale parameter. The third detector is CFAR with respect to the Pareto scale parameter but requires \textit{a priori} knowledge of \( \alpha \).

An upper bound on performance is provided by a linear threshold detector, also examined in [17], with decision rule

\[
Z_0 > \frac{\beta(P_{FA})^{-\frac{1}{\beta}} - \beta}{\tau_0}, \tag{23}
\]

which requires \textit{a priori} knowledge of both Pareto clutter parameters.

For the simulations to follow, \( N = 32 \) and \( P_{FA} = 10^{-4} \). All OS-based decision rules use an index of \( k = 30 \).

![Performance in Homogeneous Clutter](image)

Fig. 1. Detector performance in homogeneous clutter, where \( N = 32 \) and \( P_{FA} = 10^{-4} \). All OS-based decision rules use an index of \( k = 30 \).
Fig. 2. Performance of the Bayesian CFAR when the CRP contains a single interfering target. The figure shows the effect on the detection process as the SCR of the interfering target is increased gradually.

so is expected to experience a severe detection loss in this context. Figure 2 shows the performance of (17) when it is subjected to 1, 5, 10, 15 and 20 dB interference. As can be observed, the Bayesian CFAR is not robust to strong interference, but only experiences a small loss, relative to the same detector in the absence of interference, when the interference SCR level is less than 5 dB. By contrast, the detector (18) is well-known to not experience a severe detection loss until the level of interference is around 10 dB, while the OS-based decision rules (19) and (21) have been designed to manage up to two interfering targets in the CRP, by virtue of the selection of the OS index to be \( k = N - 2 \) (curves omitted for brevity).

Figure 3 provides a receiver operating characteristic (ROC) perspective on the effects of interference on the Bayesian CFAR process. In this situation the target in the CUT has SCR of 30 dB and the resultant Pd is calculated, as a function of the Pfa, with all other parameters maintained as in the previous two figures.

To complete the numerical analysis of the proposed Bayesian CFAR, its behaviour during clutter power transitions is examined. This involves estimating the resultant Pfa, as the members of the CRP are increased gradually with higher power clutter [23]. Once the mid-point of the CRP is saturated with such clutter, the CUT is then considered to be also affected with the higher power clutter, resulting in a characteristic jump in the plots of the estimated Pfa as a function of the number of higher power clutter cells. As in [17] the increase in clutter power is generated by producing a new Pareto shape parameter \( \alpha_I \), while preserving the Pareto scale parameter \( \beta \), by solving the equation

\[
10^{0.1x} = \frac{(\alpha - 2)(\alpha - 1)}{(\alpha_I - 2)(\alpha_I - 1)},
\]

which is the ratio of clutter power in the two regions, and \( x \) is the clutter power level increase in dB. Figure 4 provides an example of false alarm regulation when the clutter power level has been increased by 1 dB. As can be observed, the Bayesian CFAR has the smallest variation from the design Pfa until the CUT is affected. Thereafter, it tends to increase the Pfa in much the same way as the decision rule (21).

VI. CONCLUSIONS

This paper derived a CFAR detector for operation in Pareto Type II clutter, using a Bayesian predictive inference approach. Unlike the decision rules introduced in [17] this detector does not require a priori knowledge of either Pareto clutter parameter, thus providing a practical CFAR detector for maritime surveillance radar applications. This detector is able to manage small interference and behaves reasonably during clutter power transitions.

REFERENCES


