K-Rayleigh Mixture Model for Sparse Active Sonar Clutter

D. A. Abraham
CausaSci LLC
P.O. Box 5892
Arlington, VA 22205
Email: abrahad@ieee.org

J.M. Gelb and A.W. Oldag
Applied Research Laboratories
The University of Texas
10000 Burnet Road
Austin, TX 78758
Email: gelb and oldag @arlut.utexas.edu

***Abstract*— The mixture of a Rayleigh probability density function (PDF) and a K PDF is proposed for representing active sonar data comprising clutter sparsely observed in a Rayleigh-distributed background. While both the Rayleigh and K distributions have been shown to accurately represent the statistics of certain types of active sonar data, it is common to observe data containing both echoes from distinct clutter objects and more diffuse, Rayleigh-distributed reverberation. While the K distribution can often still capture the behavior of such data, the K-Rayleigh mixture is seen to provide improved PDF fits and inference on the clutter statistics. A parameter estimation algorithm for the K-Rayleigh mixture PDF based on the expectation-maximization (EM) technique is proposed and shown to provide adequate performance in representing the PDF of very heavy tailed real sonar data.**

I. Introduction and Background

The performance of active sonar systems operating in shallow water is often limited by false alarms termed clutter that arise from a variety of geologic, biologic, and anthropogenic sources. Most research in the area represents the phenomena statistically as a departure from the traditionally assumed Rayleigh probability density function (PDF) for the matched-filter envelope. Both empirical analysis and theoretical models show that the K distribution can represent a wide variety of active sonar clutter over diverse frequencies and bandwidths. As might be expected, the best fits are observed when the data used in the estimation are statistically stationary, with the converse true as illustrated in [1], [2] where the K-distribution fit degraded as the data became subjectively less stationary or exceptionally heavy tailed. Analyzing data after normalization of the average power level can mitigate the non-stationarity induced by the decay in reverberation power level with range. However, the effects of varying environmental conditions may not be reliably removed. For example, recent research illustrates the impact on clutter statistics of non-stationarity induced by patchy seafloors [3], data containing fish echoes [4], and rippled seafloors [5]. Chotiros [3] proposes, but does not implement or evaluate, the use of a mixture of K distributions to represent data containing scattering from multiple seafloor sediment types. Stanton and Chu [4] found that data containing background and echoes from fish were best fit by a Rayleigh mixture when the fish were schooling and therefore unresolvable or the mixture of a Rayleigh component and a power-law PDF when they were isolated, with the power-law PDF representing the randomness induced by the fish being randomly displaced from the main response axis of the sonar beamformer. In [5] it was shown that non-stationarity induced by a deterministic scaling related to the shape of seafloor ripples resulted in a smaller equivalent K-distribution shape parameter (α_1) than that of the underlying backscattering (α_0); that is, the non-stationary data appear heavier tailed. The reduction depends on the variation in the amplitude scaling in a predictable manner, but more importantly it obfuscates estimation of α_0, which can be related to the number of independent, exponentially distributed, scatterers in a sonar resolution cell [6].

The focus of this paper is on a non-stationarity condition frequently encountered in low- or mid-frequency active sonar: sparse clutter events in an analysis window having a Rayleigh-distributed background. Isolated rock outcrops or fish, shipwrecks, and mud-volcano fields are examples of clutter sources leading to such a characterization. Treating the analysis window as stationary typically leads to a probability-of-false-alarm (P_{fa}) curve having a knee, with the background distribution dominating below the knee and the clutter-response distribution dominating above the knee. As a first step toward an appropriate statistical representation of data combining a Rayleigh background and sparse, heavier-tailed clutter, a mixture of a K distribution and a Rayleigh background is proposed and analyzed. The K-Rayleigh mixture distribution differs from that of a K-distributed complex random variable plus a Gaussian (i.e., Rayleigh-distributed envelope) which results in a convolution of the individual PDFs.

The PDF of the mixture for matched-filter intensity data is simply

$$f_X(x) = \rho f_E\left(x;\sigma_R^2\right) + (1-\rho) f_K\left(x;\alpha,\sigma_K^2\right) \tag{1}$$

where $\rho \in [0,1]$ is the proportion of time the Rayleigh distributed envelope is observed, f_E is the exponential PDF (i.e., the PDF of the square of a Rayleigh-distributed random variable) parameterized by its power (σ_R^2),

$$f_E(x;\sigma_R^2) = \frac{1}{\sigma_R^2} \mathrm{e}^{-x/\sigma_R^2}, \tag{2}$$

and f_K is the K PDF parameterized by its shape (α) and power

978-1-4244-5222-4/10/$26.00 ©2010 IEEE

(σ_K^2) parameters,

$$f_K\left(x;\alpha,\sigma_K^2\right) = \frac{2\alpha}{\sigma_K^2\Gamma(\alpha)}\left(\frac{\alpha x}{\sigma_K^2}\right)^{\frac{\alpha-1}{2}} K_{\alpha-1}\left(2\sqrt{\frac{\alpha x}{\sigma_K^2}}\right) \quad (3)$$

The scale parameter of the K distribution is related to the power and shape by $\lambda = \sigma_K^2/\alpha$. An example of simulated K-Rayleigh mixture data is shown in Fig. 1 where $\rho = 0.6$, $\alpha = 5$, and $\sigma_K^2/\sigma_R^2 = 10$. To make the data more realistic, they were generated through an underlying two-state Markov chain switching between the two PDF models such that the probability of staying in the K-distribution state was 0.9.

Compared with using a single K distribution to represent sparse clutter, the K-Rayleigh mixture is expected to provide a better fit to the matched-filter envelope PDF, as well as a means for obtaining better estimates of the K shape parameter characterizing the clutter portion of the data. Unfortunately, the well-known problems in parameter estimation for the K distribution are compounded by the additional parameters introduced by the mixture model. The iterative expectation-maximization (EM) algorithm [7] is often used to obtain maximum-likelihood estimates (MLEs) of the parameters in mixture models. In the K-Rayleigh mixture, however, the maximization step of the EM algorithm would be particularly cumbersome, requiring a two-dimensional search at each iteration for the K distribution parameters. As an alternative, an efficient and robust method-of-moments estimator [8] is substituted for the MLE in the maximization step in the EM algorithm. The initialization, EM iteration, and termination are described in Sect. II. Although the iteration may not converge to the true MLE, it is seen to provide adequate estimation performance. In Sect. III, the improved representation of the mixture model over a single K distribution is illustrated through P_{fa} analysis of real active sonar data and the improved estimation of the clutter shape parameter for sparsely cluttered analysis windows is illustrated through simulation analysis.

II. Parameter estimation

Let the matched-filter intensity data $X_1,\ldots,X_n$ be independent and identically distributed according to the K-Rayleigh mixture PDF defined in (1). No closed form solution exists for maximum-likelihood estimation (MLE) of the parameters for a K distribution, so none will exist for the mixture. The MLE for the K distribution is typically discarded in favor of the method of moments owing to the computational requirements of a two-dimensional search where each combination of $\left(\alpha,\sigma_K^2,x_i\right)$ considered requires evaluation of the K Bessel function. However, method-of-moments estimators are plagued by a not insignificant probability that the moment equations are not invertible, resulting in a failure to provide an estimate. Using the K distribution as part of a mixture compounds these issues. The approach taken in this paper uses the EM algorithm structure commonly applied to mixture distributions [7] for all but the maximization step over the K-distribution likelihood. Rather than attempting a search over $\left(\alpha,\sigma_K^2\right)$ at each EM iteration to maximize the K likelihood, a reliable method-of-moments estimator [8] is employed to provide parameter estimates used to approximate the maximum likelihood. Such a sub-optimal implementation of the EM algorithm is not guaranteed to converge nor to necessarily provide the MLE. However, as will be seen in the following section, the approach provides adequate estimation performance for fitting real-data distributions and improved estimation of the clutter parameters compared with using the K distribution alone.

A. *Initial parameter estimates*

Since the EM algorithm does not necessarily converge to a global maximum, obtaining a good starting point for the iteration is worth the effort. In many cases, the clutter component of a data sample arises from discrete objects in the ocean environment (e.g., fish, mud volcanoes, ship wrecks). Coupled with the spreading effects of the typical underwater environment, it is reasonable to expect that clutter data appears as clustered events. As such, the following initialization algorithm is proposed.

1) Set the threshold
$$h_0 = \frac{\log(n)}{\log(2)} X_{\mathrm{med}} \quad (4)$$
where X_{med} is the median of the data $X_1,\ldots,X_n$. This provides a P_{fa} of approximately $1/n$ in Rayleigh-distributed data. As such, if h_0 is greater than $X_{\max} = \max\{X_1,\ldots,X_n\}$, set $h_0 = X_{\max}$.
2) For each index i satisfying $X_i \geq h_0$, accrue without repetition the data from indices $i - i_c$ to $i + i_c$ into the set $\mathcal{I}_K$ where i_c is the number of independent time samples to cluster. The data defined by the indices in $\mathcal{I}_K$ are initially assumed to be K distributed. In this paper $i_c = 2$ for all cases considered.
3) Estimate the K distribution parameters from the data $\{X_i : i \in \mathcal{I}_K\}$ using the method proposed in [8] yielding $\left(\hat{\alpha},\hat{\sigma}_K^2\right)$.
4) Estimate the proportion $\hat{\rho} = 1 - n_K/n$ where $n_K = \#\{\mathcal{I}_K\}$ is the number of elements in the set $\mathcal{I}_K$.
5) Estimate the power in the Rayleigh component as
$$\hat{\sigma}_R^2 = \frac{1}{(n-n_K)}\left[\left(\sum_{i=1}^{n} X_i\right) - n_K\hat{\sigma}_K^2\right]. \quad (5)$$

B. *EM iteration*

Given the parameter estimates $\left(\hat{\rho},\hat{\sigma}_R^2,\hat{\alpha},\hat{\sigma}_K^2\right)$, the EM iteration [7] requires the following steps

1) Evaluate the likelihoods for the intensity data samples X_i, for $i = 1,\ldots,n$,
$$L_{R,i} = f_E\left(X_i;\hat{\sigma}_R^2\right) \quad (6)$$
and
$$L_{K,i} = f_K\left(X_i;\hat{\alpha},\hat{\sigma}_K^2\right). \quad (7)$$
Owing to the inaccuracy in evaluating the K Bessel function for large order and large argument, the exponential PDF is used when $\hat{\alpha} \geq 100$,
$$L_{K,i} = f_E\left(X_i;\hat{\sigma}_K^2\right) \quad (8)$$

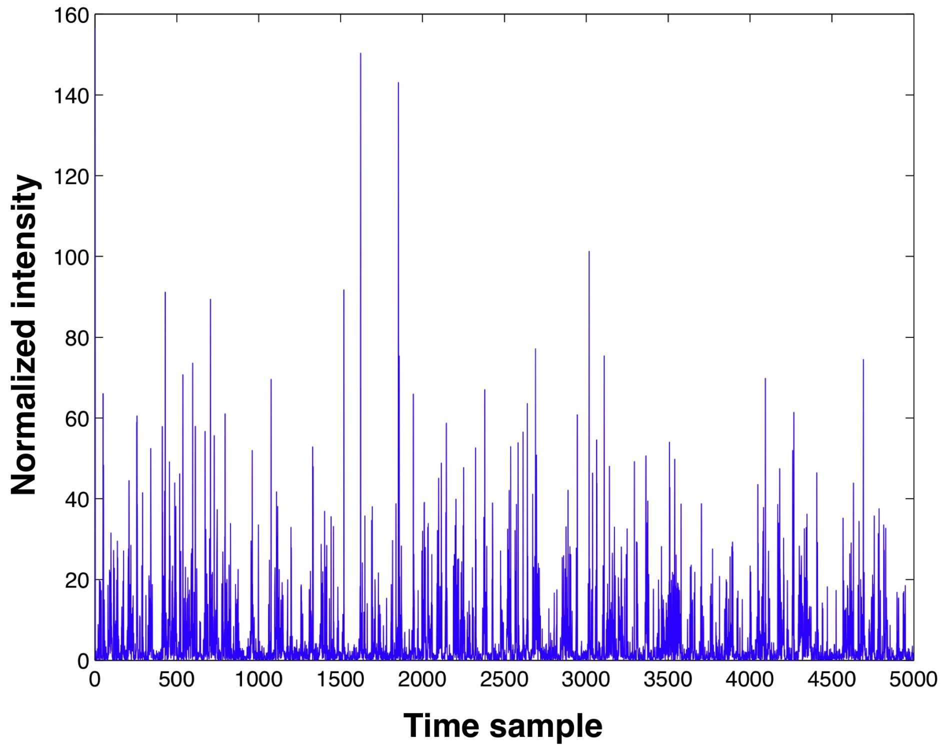

Fig. 1. Example of simulated sparsely cluttered active sonar data.

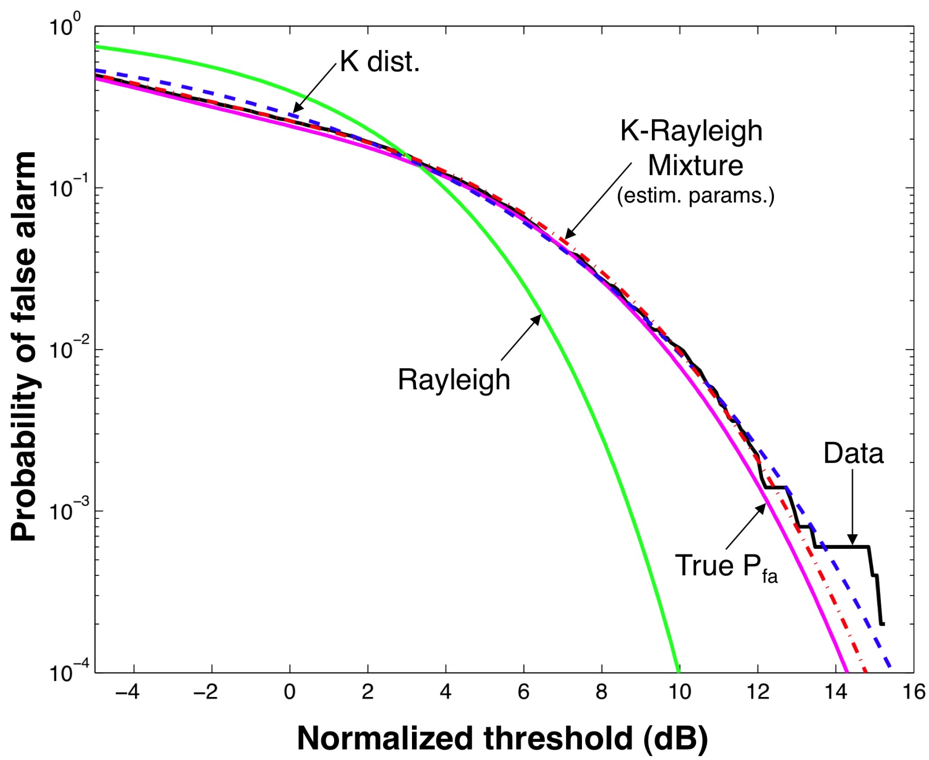

Fig. 2. Model and sample P_{fa} estimated from simulated data.

2) Compute the weights

$$W_{R,i} = \frac{\hat{\rho} L_{R,i}}{\hat{\rho} L_{R,i} + (1-\hat{\rho}) L_{K,i}} \tag{9}$$

and

$$W_{K,i} = 1 - W_{R,i}. \tag{10}$$

3) Estimate the parameters

$$\hat{\rho} = \frac{1}{n} \sum_{i=1}^{n} W_{R,i} \tag{11}$$

and

$$\hat{\sigma}_R^2 = \frac{1}{n\hat{\rho}} \sum_{i=1}^{n} W_{R,i} X_i \tag{12}$$

and obtain the K-distribution parameter estimates $\left(\hat{\alpha}, \hat{\sigma}_K^2\right)$ as described next in Sect. II-C.

C. Adaptation of K-distribution parameter estimation to EM

The analytical approximation technique described in [8] must be adapted for use in the EM iteration where the weights $W_{K,i}$ define belief in the K model for sample i. Since the estimation only requires the use of the sample mean, variance and coefficient of kurtosis formed from the envelope data ($Y = \sqrt{X}$), the adaptation only requires

$$\hat{\mu}_Y = \frac{1}{n(1-\hat{\rho})} \sum_{i=1}^{n} W_{K,i} \sqrt{X_i} \tag{13}$$

$$\hat{\sigma}_Y^2 = \frac{1}{n(1-\hat{\rho})} \sum_{i=1}^{n} W_{K,i} \left(\sqrt{X_i} - \hat{\mu}_Y\right)^2 \tag{14}$$

and

$$\hat{\eta}_4 = \frac{1}{n(1-\hat{\rho})\hat{\sigma}_Y^4} \sum_{i=1}^{n} W_{K,i} \left(\sqrt{X_i} - \hat{\mu}_Y\right)^4 \tag{15}$$

The estimates $\left(\hat{\alpha}, \hat{\sigma}_K^2\right)$ may then be formed from (13)–(15) as shown in [8] where the standard method-of-moments estimator is used when it exists.

D. Terminating the iteration

Since the parameter estimate used in Step 3 of the EM iteration in Sect. II-B does not implement the MLE for the K-distribution parameters, the total log-likelihood function (LLF)

$$l = \sum_{i=1}^{n} \log\left[\hat{\rho} L_{R,i} + (1-\hat{\rho}) L_{K,i}\right] \tag{16}$$

may not increase at each step. In evaluating the performance of the algorithm, the iteration at times reached a maximum and then decreased and at times continued to increase with diminishing increments. For the limited analysis of the algorithm presented in this paper, the iteration was terminated when the LLF peaked, when the maximum relative change in the parameters from one iteration to the next was less than 10^{-4}, or when 200 iterations was reached.

III. Data analysis examples

A. P_{fa} estimation

The sample P_{fa} of the simulated data found in Fig. 1 is shown in Fig. 2 along with the true P_{fa} and that obtained through fitting the Rayleigh, K, and K-Rayleigh mixtures. The true and estimated parameters for the various models are found in Table I where σ_X^2 is the total power. Note that the power in the K component of the mixture is 10 times that of the Rayleigh component and the total power in the model is 1 as if the data were normalized. The data are clearly heavy tailed compared with the Rayleigh distribution; however, the tail behavior is still captured adequately when the K distribution is used by itself. As seen in Table I, this requires a very low shape parameter ($\hat{\alpha} = 0.782$) compared with the true value of $\alpha = 5$. Using the K-Rayleigh mixture provides a better overall fit to the data (both the tail and the lower data values are fit better than the K) and results in a shape parameter estimate closer to the true value ($\hat{\alpha} = 3.5$).

PDF	ρ	σ_R^2	α	σ_K^2	σ_X^2
K-R Mixture (truth)	0.6	0.217	5	2.17	1
K-R Mixture	0.55	0.219	3.50	2.13	1.08
K	-	-	0.782	1.08	1.08
Rayleigh	-	1.08	-	-	1.08

TABLE I

Parameters obtained from simulation example ($n = 5000$).

B. Clutter parameter inference

As seen in the previous example, the shape parameter necessary to fit the heavier tails arising from a mixture of K-distributed clutter sparsely appearing in a Rayleigh-distributed background can be significantly lower than the true value. To examine the extent to which this occurs, the PDF of the shape parameter estimates obtained over 5000 trials using only $n = 1000$ samples per trial is shown in Fig. 3.

The PDF of the shape parameter estimate obtained by using the K distribution alone and applied to all the data in each trial is shown in the top graph of the figure. A very peaky distribution around a median value much lower than the true value is observed, indicating that the error in the estimate is severe and any inference based on this estimate would likely produce poor results. The second graph shows the PDF of the shape parameter estimate formed through the initialization routine described in Sect. II-A, which is seen to be closer to the true value, yet still biased low. The third graph shows the estimate obtained through the EM algorithm iteration described in Sect. II-B. Although the PDF is now more widely spread (a typical occurrence as α increases), the median is closer to the true value.

The estimate obtained by applying the method-of-moments technique described in [8] only to the data clairvoyantly known to follow the K distribution is shown in the bottom graph. The spread is not quite as large as for the K-Rayleigh mixture model, which must estimate which data follow the K distribution, and the median is very close to the true value. This clairvoyant estimation performance is essentially the best that might be obtained and illustrates that using the K-Rayleigh mixture is a good step toward this optimum performance.

C. Real data

The estimation algorithm described in Sect. II is applied to a sample of the "compact non-stationary" class clutter data from [2] where it was described as being "distinct dense discrete clumps." The data are received on a mid-frequency sonar using a 200-Hz bandwidth hyperbolic-frequency-modulated transmit waveform and are beamformed, matched filtered, and normalized prior to the following P_{fa} analysis. A sample of size $n = 11987$ independent samples was used to estimate the P_{fa} through a histogram and three PDF models: Rayleigh, K, and a K-Rayleigh mixture. As seen in Fig. 2 and described in [2], the data are extremely heavy tailed with both the Rayleigh and K distributions failing to represent the high P_{fa} at the higher threshold values. However, the K-Rayleigh mixture provides a very accurate fit to the data, not only in the tails, but also at the lower threshold values.

The parameter values obtained for each model are shown in Table II. The power parameters are exactly one as the data sample was normalized using the power estimated from all samples. Note that the shape parameter obtained for the K distribution used singly is higher ($\hat{\alpha} = 1.33$) than that obtained by the mixture model ($\hat{\alpha} = 0.653$), contrary to the examples shown earlier in this section. This may be explained by noting the poor fit of the K distribution to the data in the tails. Fitting the tails of the sample exceedance distribution around the $P_{fa} = 10^{-4}$ region with a single K distribution would require a shape parameter of approximately 0.15 and result in a very poor fit to the P_{fa} at lower threshold values.

PDF	ρ	σ_R^2	α	σ_K^2	σ_X^2
K-R Mixture	0.938	0.717	0.653	5.28	1
K	-	-	1.33	1	1
Rayleigh	-	1	-	-	1

TABLE II

PARAMETERS OBTAINED FROM REAL DATA EXAMPLE ($n = 11987$).

IV. CONCLUSIONS

The K-Rayleigh mixture model has been proposed for representing the statistics of active sonar clutter data, particularly when the character of the clutter is such that it occurs sparsely in the analysis window. A parameter estimation technique based on the EM algorithm was proposed and shown to perform adequately. Simulation and real data analysis illustrate the efficacy of the model in fitting data and the improved ability of the mixture to estimate parameters related to the clutter statistics when the analysis window is corrupted by Rayleigh-distributed background samples.

Although a Rayleigh mixture (i.e., a weighted combination of Rayleigh PDFs with different powers) has been shown to represent clutter statistics, its extreme flexibility [9] and potential for requiring many components make the K-Rayleigh mixture more appealing. Similarly, a mixture combining a Rayleigh PDF and the generalized Pareto distribution [10], [2], which has been shown to fit very heavy tailed data when they are thresholded, may provide more flexibility in the tails than the K-Rayleigh mixture while still maintaining a four-parameter model. The use of a four-parameter model, however, requires a larger amount of data for accurate parameter estimation than simpler models. The GPD-Rayleigh mixture and degradation in estimation performance as the analysis window decreases will be the subject of future research.

ACKNOWLEDGMENT

This work was sponsored by the Office of Naval Research under contract numbers N00014-09-C-0318 and N00014-06-G-0218-34.

REFERENCES

[1] J. M. Gelb, R. E. Heath, and G. L. Tipple, "Characterization of clutter by class in mid-frequency active sonar," in *Proceedings of International Symposium on Underwater Reverberation and Clutter*, September 2008, pp. 303–310.

[2] J. M. Gelb, R. E. Heath, and G. L. Tipple, "Statistics of distinct clutter classes in mid-frequency active sonar," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, April 2010.

[3] N. P. Chotiros, "Non-Rayleigh distributions in underwater acoustic reverberation in a patchy environment," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, April 2010, in press.

[4] T. K. Stanton and D. Chu, "Non-Rayleigh echoes from resolved individuals and patches of resonant fish at 2-4 kHz," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, April 2010.

[5] A. P. Lyons, D. A. Abraham, and S. F. Johnson, "Modeling the effect of seafloor ripples on synthetic aperture sonar speckle statistics," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, April 2010.

[6] D. A. Abraham and A. P. Lyons, "Novel physical interpretations of K-distributed reverberation," *IEEE Journal of Oceanic Engineering*, vol. 27, no. 4, pp. 800–813, October 2002.

[7] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 2, pp. 1–38, 1977.

[8] D. A. Abraham and A. P. Lyons, "Reliable methods for estimating the K-distribution shape parameter," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, April 2010.

[9] M. Gu and D. A. Abraham, "Using McDaniel's model to represent non-Rayleigh reverberation," *IEEE Journal of Oceanic Engineering*, vol. 26, no. 3, pp. 348–357, July 2001.

[10] B. R. La Cour, "Statistical characterization of active sonar reverberation using extreme value theory," *IEEE Journal of Oceanic Engineering*, vol. 29, no. 2, pp. 310–316, 2004.

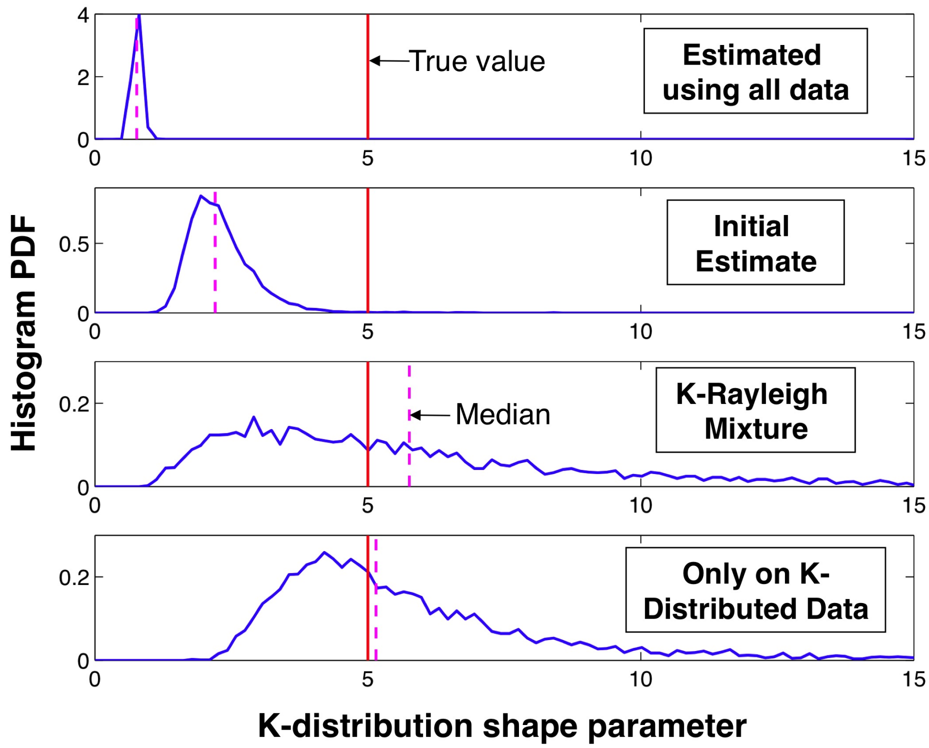

Fig. 3. Histogram PDFs of shape parameters (5000 trials) estimated from data samples ($n = 1000$) using various models.

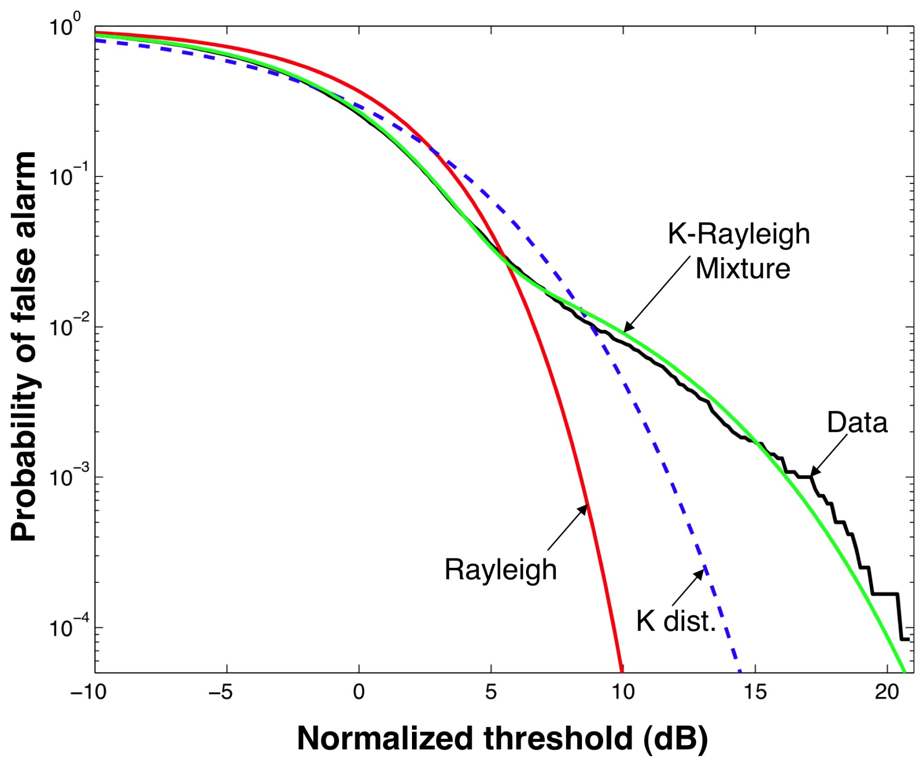

Fig. 4. Model and sample P_{fa} estimated from real sonar data.