

# COMPARISON OF MODEL SELECTION TECHNIQUES FOR SEAFLOOR SCATTERING STATISTICS

Derek R. Olson  
Marc Geilhufe

Naval Postgraduate School (NPS), Monterey, California, USA  
Norwegian Defence Research Establishment (FFI), Kjeller, Norway

## 1 INTRODUCTION

Acoustic measurements of seafloor backscattering are a source of unwanted sound in seafloor object detection<sup>1-3</sup>, but also provide a rich set of information regarding the seafloor properties and structure<sup>4-7</sup>. The intensity in a sonar image (i.e. a spatial map of measured backscattering) is typically characterized by a random process<sup>7</sup>. There are a variety of metrics, or features, that can be used to describe this random process, including the autocorrelation function, power spectrum<sup>8</sup>, wavelet decomposition<sup>9</sup>, gray-level co-occurrence matrix<sup>6,10</sup>, the mean intensity (scattering cross section)<sup>11,12</sup> and in general, the intensity probability density function<sup>13-15</sup>.

It was found that for complex scattering environments (such as rocky seafloors, and man-made structures), a mixture pdf was most appropriate<sup>14,16</sup>, which was justified by the non-stationary character of the acoustic data. Each sample of the data was modeled as being drawn from a finite number of distributions, e.g. either from the seafloor or man-made structure, or from horizontal or vertical facets.

In general, the number of components that make up a non-stationary sonar image is unknown, and must be selected prior to choosing a model and estimating the parameters. The more model parameters are used (i.e. more components, or a more complex statistical model for each component), the better the data will be fit, but the parameters may lose meaning. In this work, we explore the use of several model selection techniques based on Bayesian statistics, primarily the Bayesian information criterion (BIC) and Akaike information criterion (AIC). These techniques penalize more complex models in different ways. We also use the log-likelihood (LL) to characterize the model-data fit.

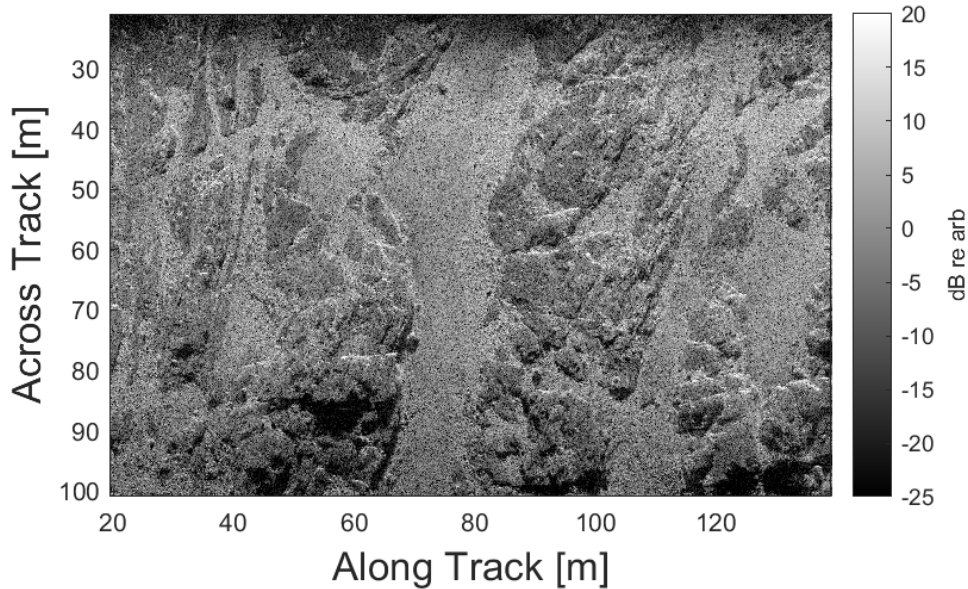
This paper is organized as follows. A description of the sonar data used in this work and example images are given in Section 2. The background statistical modeling and model selection techniques are given in Section 3. Results are presented and discussed in Section 4. Conclusions are given in Section 5.

## 2 DATA

The sonar measurements used in this work are synthetic aperture sonar (SAS) images collected off the coast of Bergen, Norway by the Norwegian Defence Research establishment (FFI). The platform used for these measurements is the HUGIN-HUS autonomous underwater vehicle (AUV), using a HISAS-1032 interferometric SAS. This sonar system has a center frequency of 100 kHz and a bandwidth of 30 kHz. The beamformed data is oversampled on a grid with  $2 \times 2$  cm resolution. Data that is used to fit mixture models is decimated by a factor of 6 in each dimension to reduce the computational complexity of parameter estimation, while reducing the correlation between samples due to the point spread function of the sonar system.

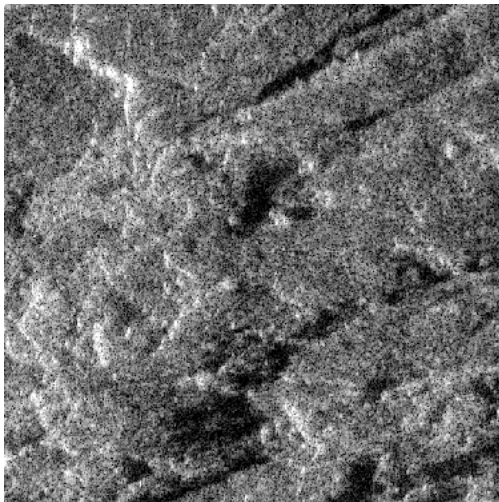
An example image is shown in Fig. 1. The image consists of an exposed rock outcrop, with sedimented areas in between. To show the detailed environmental structure, two tiles are plotted in Fig. 2, both of which are  $600 \times 600$  pixels, or 12 m per side.

These tiles show that the rock structure consists of a low intensity uniform scatterer that varies continuously due to undulations in the rock structure. These continuous variations in intensity are

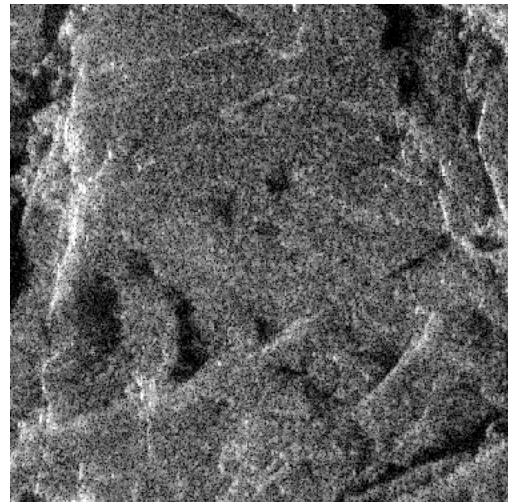


**Figure 1:** An example SAS image, plotted as a function of along track distance on the horizontal axis and across track distance on the vertical axis. The color scale is in decibels referenced to an arbitrary pixel pressure, since the system is uncalibrated.

**(a) Tile 1**



**(b) Tile 2**



**Figure 2:** Two tiles from Fig. 1, plotted on a decibel scale with 40 dB of dynamic range. Each image is 600×600 pixels.

punctuated by bright and dark lines due to fractures, and steps created by glacial erosion. These features are distinguishable due to their different intensities, and the SAS system likely has a high enough resolution that discrete scatterers in the environment are distinguishable. Therefore a mixture model is appropriate for modeling the pdf of the ensemble consisting of the pixels from each tile.

### 3 BACKGROUND

In this section, the basic definitions, statistical models, and model selection definitions are given. We assume data is given in the form of an  $N_1 \times N_2$  array of intensity values, where the intensity samples are statistically uncorrelated with each other. The total number of samples is  $N = N_1 N_2$ .

The random variable for intensity is denoted  $I$ , and the amplitude random variable is  $A = \sqrt{I}$ . We denote the probability of this variable by  $p(a)$ , with  $a$  being a member of the population.

We use two different distributions to build up a mixture model: a Rayleigh distribution for amplitude, and a K-distribution. A Rayleigh model for amplitude pdf has the form

$$p_R(a|\lambda_0) = 2a/(\lambda_0)e^{-a^2/\lambda_0}, a \geq 0, \quad (1)$$

where  $\lambda_0$  is the mean square value of the pdf. The expected value of the intensity is  $E[a^2] = \lambda_0$ .

A K-distribution for the scattered field amplitude has the form

$$p_K(a|\lambda, \alpha) = 4/(\sqrt{\lambda}\Gamma(\alpha))(a/\sqrt{\lambda})^\alpha K_{\alpha-1}(2a/\sqrt{\lambda}), a > 0, \quad (2)$$

where  $\Gamma(\cdot)$  is the gamma function,  $\lambda$  is the scale parameter of the K distribution, and  $\alpha$  is the shape parameter. The expected value of the intensity for this model is  $E[I] = \alpha\lambda = \sigma_K$ , where  $\sigma_K$  is the average intensity. In the results below, the K distribution is parameterized using the pair  $(\sigma_K, \alpha)$  rather than the shape parameter. When used in a mixture distribution, the parameters have subscripts to denote which K component the parameters correspond to.

Mixture models are formed by a weighted sum of individual pdf components. The physical meaning of this type of model is that every measurement in a population, or sample, can be identified with one of the  $M$  components. The weights of the distributions,  $w_m$  are normalized such that  $\sum w_m = 1$ , and therefore the weights can be interpreted as the fraction of pixels corresponding to each component.

The form of the mixture models used here is

$$p(a|\theta) = w_0 p_R(a|\lambda_0) + \sum_{m=1}^{M-1} p_K(a|\sigma_{K_m}, \alpha_m) \quad (3)$$

where  $M$  is the number of mixture components,  $\theta$  is a vector of length  $k$ , consisting of the parameters of the model consisting of  $w_m$ ,  $\lambda$  and the average powers and shape parameters of the K distributions. The parameters of this mixture model are found using the expectation-maximization (EM) algorithm<sup>17</sup>. This method maximizes a slightly altered version of the log-likelihood for each component, but asymptotically maximizes the likelihood function for mixture models<sup>17</sup>.

Since the number of components that constitutes the environment is in general unknown, model selection techniques are used to pick the number of K distribution components. As the number of components,  $M$ , increases, it is better able to match the pdf of the measurement, but at the cost of more uncertainty per parameter<sup>14</sup>. The likelihood function,  $\ell(\theta|a)$ , is a common metric for model-data fit. It is defined for  $N$  independent samples by,

$$\ell(\theta|a) = \prod_{n=1}^N p(a_n|\theta) \quad (4)$$

where  $a_n$  is the  $n$ -th member of the ensemble, and again  $\theta$  is the parameter vector. The parameter vector  $\hat{\theta}$  that maximizes  $\ell$  is called the maximum likelihood estimate. It is common to work with the log-likelihood,  $\mathcal{L} = \log(\ell)$ , which is given as

$$\mathcal{L}(\theta|a) = \sum_{n=1}^N \log(p(a_n|\theta)) \quad (5)$$

As stated before, more complex models typically result in a more uncertain maximum likelihood function, so  $\mathcal{L}$  cannot be used as a basis on which a decision about  $M$  can be made. There exist various model selection techniques, some of which are based on Bayesian concepts<sup>18</sup>. Two simple metrics are based on the log-likelihood function, but with an additional penalty that depends on the number of parameters,  $k$ . The Bayesian information criterion (BIC) is defined as

$$BIC = -2\mathcal{L}(\hat{\theta}) + \log(N)k \quad (6)$$

with smaller  $BIC$  preferred. This criterion results from an asymptotic (i.e. large  $N$ ) Gaussian approximation of the posterior probability density (ppd) function of  $\theta$  given the data. The penalty to the log-likelihood is therefore a function of the number of parameters, which is due to the ppd becoming narrower as  $N$  becomes asymptotically large. Another model selection criterion is the Akaike information criterion (AIC), due to Akaike<sup>19</sup>, and also described by Gelman<sup>18</sup>. It is given by the formula

$$AIC = -2\mathcal{L}(\hat{\theta}) + 2k \quad (7)$$

This criterion is simpler than the BIC and penalizes the log-likelihood independently of the number of samples.

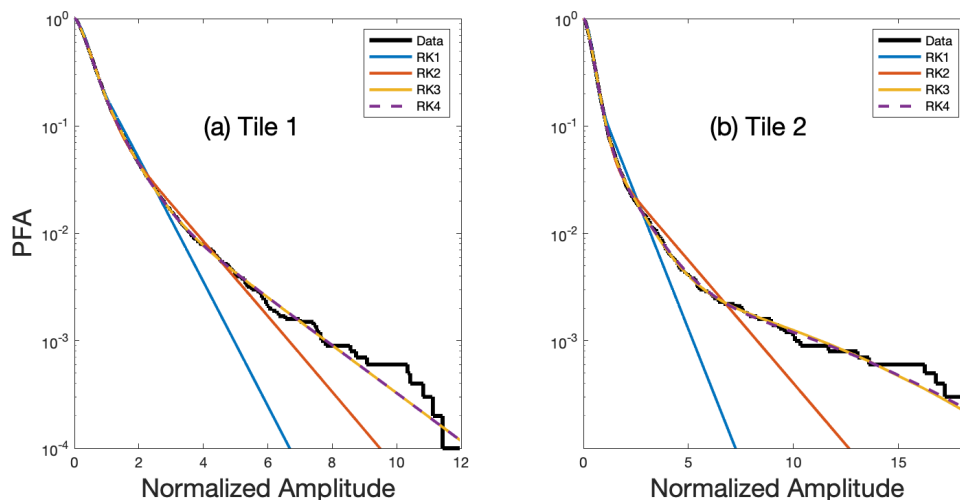
Both the AIC and BIC are based on the log-likelihood function evaluated at  $\hat{\theta}$ , and are therefore quite sensitive to the numerical estimate. Additionally, point estimates do not contain any information about the ppd as a whole. Other information criteria such as the deviance information criterion and the Wantanabe-Akaike information criterion<sup>18</sup> are not studied here, but we consider them as fruitful areas for future work. Another possibility is to use trans-dimensional Monte-Carlo methods to estimate the model with the highest posterior probability<sup>20</sup>.

## 4 RESULTS AND DISCUSSION

R-K mixture models were fit to the amplitude data in Figs. 2 (a) and (b), using between  $M = 2$  and  $M = 5$  components. This means that the maximum number of K-distribution components was 4. The model-data fit is shown graphically in Fig. 3 for both tiles in terms of the log of the probability of false alarm (PFA), also called the exceedence distribution function (EDF). The PFA is related to the cumulative distribution function (CDF) through  $PFA = 1 - CDF$ , and is a common method of presenting sonar reverberation statistics<sup>21</sup>. The data in Tile 1 shows a slight “knee” near a normalized amplitude of 2, 6, and 8. These changes in slope of the log-PFA indicate different components that make up the model. The log-PFA of Tile 2 has more pronounced knees in the curve, near the normalized amplitudes of 2, 4, and 12, although the last one is more uncertain due to the finite sample size being more evident at high amplitudes (i.e. the PFA curve becomes more stair-case like, rather than a smooth curve).

Model-data fits for both tiles are poor for both R-K1 and R-K2. This behavior is likely due to the fact that there are not enough components to fit the data. R-K3 and R-K4 fit the data much better, but are almost the same. This behavior indicates that a 5-component model does not provide significantly better fit than the 4-component model. From a visual standpoint, it makes sense to use a 4-component model for both of these data sets. This hypothesis will be compared to the results of the more formal model selection techniques.

Model selection criteria, the log-likelihood (LL), BIC, and AIC are shown for both tiles in Table 1. The maximum LL value for Tile 1 is a tie between the 4- and 5-component models, and the maximum LL for Tile 2 is for the 4-component model. Intuitively, the more complex model should have a higher likelihood function, but here, this may be an issue with the numerical parameter estimates. Additionally, as discussed in Section 3, one cannot base a model selection decision purely on the maximum of the likelihood function. Further work on refining these estimate should be made. In terms of the AIC, the smallest value occurs for the 4-component model for both datasets. However, for the BIC, the smallest value occurs for the 3-component model for Tile 2, and the 4-component model for Tile 1. The AIC penalizes model complexity only slightly, whereas the BIC has a severe penalty for complexity for large number of data samples. Since the number of data samples was



**Figure 3:** The probability of false alarm (PFA) for the data from snippets 1 and 2, compared to the various mixture models explored here.

Model	Tile 1			Tile 2		
	LL	AIC	BIC	LL	AIC	BIC
R-K1	-4321	8652	8688	-885	1780	1810
R-K2	-4157	8331	8389	-847	1711	1759
R-K3	-4137	8297	8377	-839	1700	1766
R-K4	-4137	8301	8404	-840	1709	1793

**Table 1:** Model selection results for both image tiles. The log-likelihood (LL), Akaike information criterion (AIC), and the Bayesian information criterion (BIC) are shown.

about  $10^4$ , the penalty for model complexity is about  $0.5 \log 10^4 \approx 4.6$  times larger for the BIC as it is for the AIC. We conclude that the 3-component model is preferred by BIC for Tile 2 because of the much larger penalty for model complexity.

## 5 CONCLUSION

We presented a statistical model for SAS images of complex, non-stationary, rocky seafloors. This model consisted of a Rayleigh distribution, plus an unknown number of K distributions. The number of K-distributions was selected using model selection techniques, including the AIC and BIC. For the AIC, the 4-component model was selected as the most appropriate for two image tiles used here. For the BIC, different number of components were chosen for each image tile. It is likely that a different number of components was chosen by the BIC due to its more severe penalty for model complexity. Future work should include more robust model selections techniques, such as the deviance information criterion, which employs a Monte-Carlo Markov chain sampling of the distribution parameters. This work can also be used to partition an image into different scattering components, which can aid in estimates of image complexity, and may also be the basis for quantitative seafloor remote sensing of geological and/or biological parameters.

## 6 ACKNOWLEDGMENTS

This research was supported by grant number N00014-23-WX01149 from the Office of Naval Research under the Young Investigator Program. The authors thank FFI's HUGIN-HUS operator group for collecting the data used in this study.

## 7 REFERENCES

1. D. P. Williams, "Fast unsupervised seafloor characterization in sonar imagery using lacunarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6022–6034, Nov 2015.
2. R. Quinn, *Acoustic Remote Sensing in Maritime Archaeology*, B. Ford, D. L. Hamilton, and A. Catsambis, Eds. Oxford University Press, sep 2012.
3. A. P. Galusha, J. M. Keller, A. Zare, and G. Galusha, "A fast target detection algorithm for underwater synthetic aperture sonar imagery," in *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIII*, J. C. Isaacs and S. S. Bishop, Eds. SPIE, apr 2018.
4. A. Zare, N. Young, D. Suen, T. Nabelek, A. Galusha, and J. Keller, "Possibilistic fuzzy local information c-means for sonar image segmentation," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, nov 2017.
5. J. Peeples, W. Xu, and A. Zare, "Histogram layers for texture analysis," *IEEE Transactions on Artificial Intelligence*, vol. 3, no. 4, pp. 541–552, aug 2022.
6. P. Blondel and O. G. Sichi, "Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf," *Applied Acoustics*, vol. 70, no. 10, pp. 1288–1297, oct 2009.
7. W. K. Stewart, D. Chu, A. Malik, S. Lerner, and H. Singh, "Quantitative seafloor characterization using a bathymetric sidescan sonar," *IEEE J. Ocean Eng.*, vol. 19, pp. 599–610, 1994.
8. J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-6, no. 4, pp. 269–285, 1976.
9. D. P. Williams, "Unsupervised seabed segmentation of synthetic aperture sonar imagery via wavelet features and spectral clustering," in *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE, nov 2009.
10. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, nov 1973.
11. D. R. Jackson and M. D. Richardson, *High-Frequency Seafloor Acoustics*. New York, NY: Springer, 2007.
12. D. R. Olson, A. P. Lyons, and T. O. Sæbø, "Measurements of high-frequency acoustic scattering from glacially eroded rock outcrops," *J. Acoust. Soc. Am.*, vol. 139, no. 4, pp. 1833–1847, 2016.
13. A. P. Lyons and D. A. Abraham, "Statistical characterization of high-frequency shallow-water seafloor backscatter," *The Journal of the Acoustical Society of America*, vol. 106, no. 3, pp. 1307–1315, sep 1999.
14. D. R. Olson, A. P. Lyons, D. A. Abraham, and T. O. Sæbø, "Scattering statistics of rock outcrops: Model-data comparisons and bayesian inference using mixture distributions," *J. Acoust. Soc. Am.*, vol. 145, no. 2, pp. 761–774, 2019.
15. R. C. Gauss, J. M. Fialkowski, D. C. Calvo, R. Menis, D. R. Olson, and A. P. Lyons, "Moment-based method to statistically categorize rock outcrops based on their topographical features," in *OCEANS 2015 - MTS/IEEE Washington*, Oct 2015, pp. 1–5.
16. D. Abraham, J. Gelb, and A. Oldag, "Background and clutter mixture distributions for active sonar statistics," *IEEE J. Ocean Eng.*, vol. 36, no. 2, pp. 231–247, April 2011.
17. 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 (Methodological)*, vol. 39, no. 1, pp. 1–22, sep 1977.

18. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*. Chapman and Hall/CRC, nov 2013.
19. H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, dec 1974.
20. J. Dettmer, S. E. Dosso, and C. W. Holland, "Trans-dimensional geoaoustic inversion," *The Journal of the Acoustical Society of America*, vol. 128, no. 6, pp. 3393–3405, dec 2010.
21. A. Lyons, S. Johnson, D. Abraham, and E. Pouliquen, "High-frequency scattered envelope statistics of patchy seafloors," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 4, pp. 451–458, oct 2009.