EasyChair Preprint

№ 699

Analysis of Sea Clutter Distribution and Evaluate Ship Detection Performance for Sentinel-1 SAR Data

Yongxu Li, Xudong Lai, Jie Zhang, Junmin Meng, Genwang Liu and Xi Zhang

April 15, 2019

Analysis of Sea Clutter Distribution and Evaluation Ship Detection Performance for Sentinel-1 SAR Data

Yongxu Li [1,2], Xudong Lai [2], Jie Zhang [1], Junmin Meng [1], Genwang Liu [1], Xi Zhang [1]

[1]The First Institute of Oceanography, State Oceanic Administration, Qingdao, China

[2]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

Abstract—This paper statistically analyzed the sea clutter distribution and ship detection performance for Sentinel-1 synthetic aperture radar image. First, the goodness-of-fit of five commonly used distribution models were evaluated to find out the most suitable model and the Kullback-Leivler Distance was adopted to judge the fitting degree. Then construct a constant false alarm rate detector for ship detection. To measure the robustness of the detector, the figure of merit, the probability of detection and the false alarm rate were calculated to evaluate detector performance.

Index Terms—Ship detection, synthetic aperture radar, constant false alarm rate, statistical distribution.

I. INTRODUCTION

The monitoring of the ship target is one of the most important research fields. The information such as location and type of ship is widely used in maritime surveillance, maritime detection, traffic safety, fisheries control and so on. For synthetic aperture radar (SAR), the basic method of ship detection takes advantage of SAR images' feature that the backscattering signal from the ship is much stronger than the sea clutter background in most cases, and the detection was realized by searching pixels whose amplitudes are greater than a given threshold. The constant false alarm rate (CFAR) detector was widely used due to the variable threshold which determined by accurately describing the real-time dynamic sea clutter around the target. An appropriate model is essential to the detector [1-3].

However, properties like frequency, polarization, resolution, grazing angle and sea state have the influence on SAR imaging, therefore the sea clutter may fit different models and the suitable model of the given dataset needs to be judged [4-8]. In high-resolution SAR image, the K distribution becomes popular due to the compound formulation, which was introduced by Ward that enables both the small-scale and large-scale components of the sea clutter to be characterized [9,10]. The Weibull (WBL) distribution was used to model amplitude earlier, and the results show that it can fit most SAR images effectively. The Lognormal (LGN) distribution can achieve better goodness-of-fit even under the heterogeneous situation in high-resolution SAR images [11]. The computational complexity of G^0 distribution was reduced by eliminating an iterative computing step; meanwhile, it has been demonstrated with excellent performance in a heterogeneous sea surface environment [12]. The generalized gamma distribution (GΓD) was used for modeling many scenes of high-resolution SAR images and shows a better performance in most cases [13].

Therefore, the comparative analyses among those five commonly used distributions were carried out. The Kullback-Leivler (K-L) Distance was used to judge the goodness-of-fit. After the best fitting was found, the CFAR detector was

This work was supported by the Public Science and Technology Research Funds Projects of Ocean under grant 201505002.

constructed to carry out ship detection experiments. To evaluate CFAR detector performance, the ground truth (GT), which were determined via Automatic Identification System (AIS) data combined with manual interpretation, was used to validate the result, and the figure of merit (FOM), probability of detection (PoD) and false alarm rate (FAR) were calculated.

The remainder of this paper is organized as follows. Section II briefly describes the Sentinel-1 dataset. The experiment methodologies are presented in Section III, while Section IV shows the results. Finally, the conclusions are given in Section V.

II. SENTINEL-1 IW LEVEL-1 GRD DATA

Following the "open and free" data access policy, seven dual polarization images of Sentinel-1 IW Level-1 Ground Range Detected High resolution (GRDH) produce from October 2014 to January 2016 have been obtained in the three areas of the Strait of Malacca, shown in Fig.1. These images were acquired using vertical transmit, vertical receive (VV) polarization and vertical transmit, horizontal receive (VH) polarization in high resolution 20 m × 22 m.

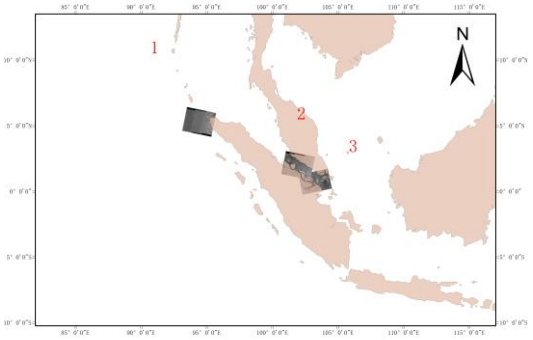

Fig.1. Geographical location

In order to avoid the influence of shoreline or island on the experimental results, 30 pairs of sub-images were extracted from the original data by artificial cutting. In addition to the goodness-of-fit analysis, whether the detector has the capability to provide reliable results is also the scope of the assessment. We verified 304 ship targets by AIS data and manual supervision to acquire GT as the reference. To acquire GT, 304 ship targets were verified by AIS data and manually supervision. Additional information about the dataset is shown in Table I.

TABLE I ADDITIONAL INFORMATION

No.	Acquire time	Location	GT	Sub-images
1	20150815_232741	1	9	2
2	20151026_232742	1	5	1
3	20141222_225535	2	74	6
4	20150316_225534	2	72	6
5	20150527_225538	2	69	6
6	20151018_225543	2	58	6
7	20160116_112449	3	17	3

III. EXPERIMENT METHODOLOGIES

A. *Distribution model*

In this section, these five distribution models: LGN, WBL, K-root, GΓD, G^0 were used to fit the sea clutter. After the distribution model determined, the Method of log-cumulants (MoLC) was used to solve the parameters of the distribution model by counting the pixels in the background box around the target to obtain the probability density function (PDF) of the distribution models. The PDFs and the corresponding equations for the models are shown in Table II [14].

TABLE II AMPLITUDE PDFS AND MOLC EQUATIONS OF THE MODELS

Model	PDF	MoLC equations
WBL	$f_A(x) = \frac{\gamma}{\sigma^\gamma}(x)^{\gamma-1}\exp[-(\frac{x}{\sigma})^\gamma], x,\gamma,\sigma > 0$	$k_1 = \log(\sigma) + \Phi_0(1)\gamma^{-1}$ $k_2 = \Phi_0(1,1)\gamma^{-2}$

LGN	$f_A(x) = \frac{1}{\sqrt{2\pi}\sigma x}\exp[-\frac{(\ln x - m)^2}{2\sigma^2}], \sigma > 0, m \in R$	$k_1 = m$ $k_2 = \sigma^2$
G⁰	$f_A(x) = \frac{2L^L(L-\alpha)}{\Gamma^\alpha\Gamma(L)\Gamma(-\alpha)}\frac{x^{2L-1}}{(\gamma + Lx^2)^{L-\alpha}}, x, L, \gamma > 0, \alpha < 0$	$2k_1 = \log\gamma/L + \Phi_0(L) - \Phi_0(-\alpha)$ $4k_2 = \Phi_0(1,L) + \Phi_0(1,-\alpha)$ $8k_3 = \Phi_0(2,L) - \Phi_0(2,-\alpha)$
K-root	$f_A(x) = \frac{4}{\Gamma(L)\Gamma(\gamma)}(\frac{L\gamma}{\mu})^{\frac{L+\gamma}{2}} x^{L+\gamma-1}K_{\lambda-L}(2x\sqrt{\frac{L\gamma}{\mu}}), x > 0$	$2k_1 = \log\mu/L\gamma + \Phi_0(L) + \Phi_0(\gamma)$ $4k_2 = \Phi_0(1,L) + \Phi_0(1,\gamma)$ $8k_3 = \Phi_0(2,L) + \Phi_0(2,\gamma)$
GΓD	$f_A(x) = \frac{\lvert\gamma\rvert\kappa^\kappa}{\sigma\Gamma(\kappa)}(\frac{x}{\sigma})^{\kappa\gamma-1}\exp[-\kappa(\frac{x}{\sigma})^\gamma], x \in R^+$	$k_1 = \log(\sigma) + (\Phi_0(\kappa) - \log(\kappa))\gamma^{-1}$ $k_i = \Phi_0(i-1,\kappa)\gamma^{-i}, i = 2,3\ldots$

B. *CFAR detector*

The CFAR detector is widely used in ship detection by comparing pixels with an adaptive threshold **T** to maintain a constant false alarm rate. **T** can be obtained by solving equation (1), according to PDF $f(x)$ and the desired probability of false alarm (p_{fa}). The former was obtained by modeling the distribution of sea clutter from the background cell, and the latter was based on experience [15].

$$p_{fa} = 1 - \int_0^T f(x)\,dx = \int_T^\infty f(x)\,dx \quad (1)$$

C. *Evaluation methodology*

To evaluate the fitting degree of different models, the K-L Distance was adopted to access the modeling accuracy, which is a measure of how one probability distribution is different from another reference, reference probability distribution PDF and histogram of sea clutter in this case. The lower the K-L Distance, the higher the fitting accuracy.

$$I(g, f) = \int g(x)\ln\frac{g(x)}{f(x)}dx \quad (2)$$

Where $f(x)$ is the PDF and g(x) is the histogram of sea clutter.

The detector's performance was quantitatively evaluated by comparing results with GT, and the following metrics were estimated:

FOM:

$$FOM = N_{tt} / (N_{gt}+N_{fa}) \quad (3)$$

PoD:

$$PoD = (N_{tt} / N_{gt}) \quad (4)$$

FAR:

$$FAR = (N_{fa} / N_{gt}) \quad (5)$$

Where Ntt is the total number of detection targets matching GT, Ngt is the total number of GT, Nfa is the total number of false alarms.

IV. RESULTS

A. *Analysis of goodness-of-fit*

The fitting results of five distributions were collected and illustrated as Fig.2. The vertical coordinate recorded the K-L Distance, the transverse coordinate indicates sample number that sorted by incident angle, results under VV are shown in Fig.2 (a) and under VH are shown in Fig.2 (b). As shown, there was no obvious correlation between incident angles with fitting accuracy. However, there are significant differences among polarizations. That is, for the same distribution model, the results under VH were significantly lower and were more uniform than those under VV. Furthermore, analyzing the differences between models, the other four distributions except WBL can hardly be discerned from the diagram.

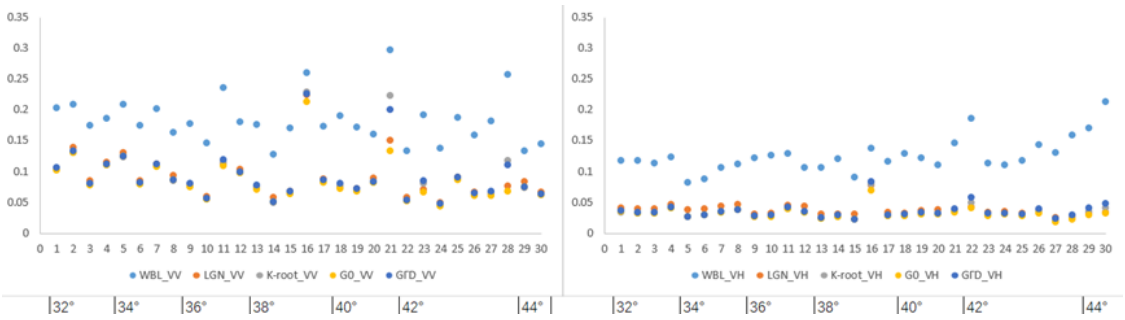

Fig.2. K-L distance of distributions under VV (left plot), VH (right plot)

To quantify the experimental results and find out the most accurate and robust distribution, the mean and the variance value of those results were counted and collected in Table III. Take LGN as an example, the mean and the variance value of VH were much smaller than those of VV. The former exhibited values of 0.037 and 0.009, while for the latter, they were 0.091 and 0.035, respectively. That means the fitting performance under VH is better than that of VV.

TABLE III THE MEAN AND THE VARIANCE OF K-L DISTANCE

	LGN	G^0	K-root	GΓD	WBL
VV_mean	0.091	**0.085**	0.092	0.093	0.184
VH_mean	0.037	**0.032**	0.033	0.035	0.125
ALL_mean	0.064	**0.058**	0.063	0.065	0.155
VV_variance	0.035	**0.033**	0.042	0.039	0.038
VH_variance	0.009	**0.009**	0.010	0.011	0.027
ALL_variance	0.037	**0.036**	0.042	0.041	0.044

Moreover, mark the best performance of each indicator in bold. As can be seen, the bold indicators are clustered in the third column of the table, which were the results for the G^0. That means G^0 has the best fitting with an overall average of 0.058 and an overall variance of 0.036. The results of LGN, K-root, and GΓD are basically identical, which are clearly better than the WBL. Therefore, G^0 was used to construct a CFAR detector for ship detection experiments and the detector performance was evaluated below.

B. *Detection performance*

The parameter that controls the performance of CFAR detectors is the p_{fa}. By recording the results of the detector under different p_{fa}, the curves of the FOM, the PoD and the FAR changing with p_{fa} were drawn, with transverse coordinate recording the values of p_{fa} (10^{-x}), and vertical coordinate recording the statistical results of detector's FOM, PoD, FAR respectively. The performance of the detector was directly shown in Fig.3.

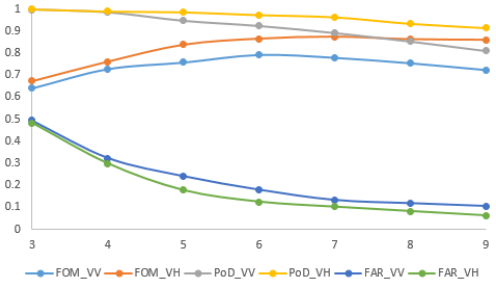

Fig.3. The FOM, POD, FAR curves of the detector

As can be seen, the results under VH were different from VV, having higher FOM, PoD and lower FAR, which means the detection performance of VH was better than VV.

In general, with the decrease of p_{fa}, the decreasing trend of PoD under VH was gentle, from 100% to 91.45%, down 8.55%. In contrast, FAR decreased significantly, from 48.36% to 6.25%, down 42.11%. FOM increased from 67.41% to 86.07%, rising 18.66%. However, there was no obvious change after p_{fa} was 10^{-7}. In the case of VV, the decreasing trend of PoD was greater than that of VH, from 99.79% to 81.06%, down 18.72%. Meanwhile, FAR decreased more obviously, from 55.96% to 11.91%, down 44.04%, FOM increased from 63.98% to the highest 79.34% when p_{fa} was 10^{-6}, rising 15.36%, then there was a slight decline from the peak value to 72.43%, down 6.91%.

In conclusion, the CFAR detector based on G^0 distribution performed notably well. The best performance of VH was when the p_{fa} value was 10^{-7}, and so of VV was when the p_{fa} value was 10^{-6}. Which has reference significance for the reasonable value of p_{fa}.

V. CONCLUSIONS

In this paper, five commonly used distribution models were used to fit the Sentinel-1 SAR image, and the K-L distance was calculated to quantitatively judge the matching degree. G^0 distribution was shown most fitting to the data, followed by LGN, K-root, GΓD, and WBL. Therefore, a CFAR detector was constructed with G^0 to carry out ship detection experiments, and the results showed that VH was more suitable for ship detection than VV. For VH, the recommended p_{fa} was 10^{-7}. For VV, the recommended p_{fa} was 10^{-6}.

ACKNOWLEDGEMENT

The authors are very grateful to the European Space Agency for providing the experimental dataset.

REFERENCES

[1] Pelich, R., Longepe, N., Mercier, G., Hajduch, G., & Garello, R. (2015). Performance evaluation of Sentinel-1 data in SAR ship detection. IGARSS 2015 - 2015 IEEE International Geoscience and Remote Sensing Symposium. IEEE.

[2] Daum, F. (2008). Radar handbook, 3rd edition (m.i. skolnik, ed; 2008) [book review]. IEEE Aerospace & Electronic Systems Magazine, 23(5), 41-41.

[3] Eldarymli, K., Mcguire, P., Power, D., & Moloney, C. R. (2013). Target detection in synthetic aperture radar imagery: a state-of-the-art survey. Journal of Applied Remote Sensing, 7(7), 071598-071598.

[4] Fiche, A., Sébastien Angelliaume, Rosenberg, L., & Khenchaf, A. (2018). Statistical analysis of low grazing angle high resolution X-band SAR sea clutter. Radar Conference. IEEE.

[5] Fiche, A., Angelliaume, S., Rosenberg, L., & Khenchaf, A. (2015). Analysis of x-band sar sea-clutter distributions at different grazing angles. IEEE Transactions on Geoscience and Remote Sensing, 53(8), 4650-4660.

[6] Xin, Z., Liao, G., Yang, Z., Zhang, Y., & Dang, H. (2017). Analysis of distribution using graphical goodness of fit for airborne sar sea-clutter data. IEEE Transactions on Geoscience and Remote Sensing, 1-10.

[7] Guillaume, H. (2016). AIS-based evaluation of target detectors and sar sensors characteristics for maritime surveillance. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 8(8), 3892-3901.

[8] Gao, G. (2010). Statistical modeling of sar images: a survey. Sensors (14248220), 10(1), 775-795.

[9] Jakeman, E., & Pusey, P. (2003). A model for non-rayleigh sea echo. IEEE Transactions on Antennas and Propagation, 24(6), 806-814.

[10] Ward, K., Tough, R., Watts, S., Ward, K., Tough, R., & Watts, S. (2013). Sea clutter:

scattering, the k-distribution and radar performance. Waves in Random & Complex Media, 17(2), 233-234.

[11] Goldstein, & G. (1973). False-alarm regulation in log-normal and weibull clutter. IEEE Transactions on Aerospace and Electronic Systems, AES-9(1), 84-92.

[12] Jung, C. H., Yang, H. J., & Kwag, Y. K. (2009). Local Cell-Averaging Fast CFAR for Multi-Target Detection in High-Resolution SAR Images. Asian-pacific Conference on Synthetic Aperture Radar.

[13] Qin, Xianxiang, Zhou, Shilin, Zou, & Huanxin, et al. (2013). A cfar detection algorithm for generalized gamma distributed background; in high-resolution sar images. IEEE Geoscience & Remote Sensing Letters, 10(4), 806-810.

[14] Cui, S., Schwarz, G., & Datcu, M. (2014). A comparative study of statistical models for multilook sar images. IEEE Geoscience and Remote Sensing Letters, 11(10), 1752-1756.

[15] Ji, Y., Zhang, J., Meng, J., & Zhang, X. (2010). A new cfar ship target detection method in sar imagery. Acta Oceanologica Sinica, 29(1), 12-16.