Sea clutter contamination test with log-cumulants

Ding Tao[a], Anthony P. Doulgeris[b] and Camilla Brekke[c]

University of Tromsø, Department of Physics and Technology, NO-9037 Tromsø, Norway

ABSTRACT

In maritime applications involving estimation of radar sea clutter properties, non-sea-clutter targets and transitions between statistically different oceanographic conditions in the estimation window may lead to inaccurate modeling. Referring to mixtures in the estimation window as *contamination*, this work introduces a novel sea clutter contamination test based on log-cumulants from Mellin kind statistics [1]. It measures the significant deviation in log-cumulant space due to the contamination, and appears to be an effective tool for improving the sea clutter estimation or to be a direct first-stage target detector. The proposed contamination test is examined with real single look complex (SLC) fine resolution quad-polarimetric Radarsat-2 synthetic aperture radar (SAR) measurements, from the Norwegian Sea, under various oceanographic conditions.

Keywords: Synthetic aperture radar, Target detection, Sea clutter, Contamination, Log-cumulant.

1. INTRODUCTION

Synthetic aperture radar (SAR) is a well-known remote sensing technology, capable of producing imagery in all weather conditions independent of daylight. Polarimetric SAR (PolSAR) can provide information about the target in the intensity measurements and the polarimetric properties of the electromagnetic wave [2]. Because of the complementary information contained in the different channels, PolSAR is of interest in the field of maritime applications, e.g. in ship detection and investigation of oceanographic conditions. Thus the sea clutter statistical properties are usually considered essential information. For example, in conventional constant false alarm rate (CFAR) ship detection, a joint task is required where background sea clutter statistics are estimated before target detection. However, the large scale background sea clutter, measured in a fixed relative window, is often a mixture including non-background targets having different statistical properties. This kind of contamination will lead to inaccurate sea clutter estimation and reduction in sensitivity of the background statistical models. In this study, the log-cumulant based method from Mellin kind statistics [1] is investigated due to its ability to distinguish the contaminated area in log-cumulant space. Thereupon, a novel contamination test based on log-cumulant is proposed. Its test mechanism is demonstrated using the log-cumulant diagram [3]. Experiments are performed on several Radarsat-2 quad-polarimetric SAR datasets under various oceanographic conditions.

This paper is organized as follows. Section 2 provides a brief descriptions of PolSAR sea clutter statistics, basics about log-moment/log-cumulant from Mellin kind statistics, and interpretation of the log-cumulant diagram. Section 3 is dedicated to introduce the contamination test mechanism and strategy. In Section 4, the contamination test is examined with real sea clutter datasets under various oceanographic conditions. Finally, Section 5 presents the main conclusions derived from the study and thoughts about the prospective work.

---

a. Email: tao.ding@uit.no, Telephone: +47 - 473 20898;
b. Email: anthony.p.doulgeris@uit.no, Telephone: +47 - 776 45177;
c. Email: camilla.brekke@uit.no, Telephone: +47 - 776 46297.

SAR Image Analysis, Modeling, and Techniques XII, edited by
Claudia Notarnicola, Simonetta Paloscia, Nazzareno Pierdicca, Proc. of SPIE
Vol. 8536, 85360I · © 2012 SPIE · CCC code: 0277-786/12/$18 · doi: 10.1117/12.976879

2. SEA CLUTTER STATISTICS

2.1 Polarimetric SAR sea clutter modeling

To obtain the sea clutter statistical properties, accurate modeling is required for different types of datasets and sea conditions. For low resolution PolSAR imagery, sea clutter is commonly assumed to be homogenous (fully developed speckle), which indicates that there are large number of discrete scatterers per resolution cell. This assumption usually leads to the complex multivariate Gaussian distribution for single look complex (SLC) data, and the complex Wishart distribution is used for multi look complex (MLC) data. Recent PolSAR sensors with very high resolution, noticeably decrease the number of scatterers in each resolution cell. The significance of using non-Gaussian models increases as the imagery resolution gets higher. Thus, the multivariate product model has been introduced by Yueh et al. [4,5], which decomposes the scattering vector $\boldsymbol{s}$ from a quad-polarimetric SLC measurement (satisfying the independent and identically distributed assumption) as

$$\boldsymbol{s} = [s_{hh}\ s_{hv}\ s_{vh}\ s_{vv}]^T = \sqrt{\tau}\boldsymbol{x}\ , \tag{1}$$

where T denotes the transpose operator, the s_{bq} are complex scattering coefficients subscripted with transmit polarization b and receive polarization q, that can be horizontal (h) or vertical (v), the texture τ, representing spatial variability of the reflectivity, is a real positive random variable with unit mean and probability density function (pdf) $p_\tau(\tau)$, while speckle is represented as a d-dimensional circular complex Gaussian vector, $\boldsymbol{x} \sim \mathcal{N}_d^{\mathbb{C}}(0, \Sigma)$, with zero mean and covariance matrix Σ, which is independent of τ.

The $\mathcal{K}$ distribution [6] is one of the accurate distribution models for high-resolution sea clutter SAR data, which can be described as the multivariate product model. Thus, when the texture variable and speckle vector are assumed to be gamma distributed and multivariate Gaussian distributed respectively, the scattering vector becomes multivariate $\mathcal{K}$-distributed. The recent $\mathcal{G}^0$ distribution [7] and $\mathcal{U}$ distribution [8] can also be derived from the product model.

2.2 Log-moments and log-cumulants from Mellin kind statistics

As we know, radar image data on a logarithmic scale provides an efficient dynamic range of pixel intensities and good contrast. In a previous study by Nicolas [1], a novel statistical theory named second kind statistics or Mellin kind statistics was introduced along with the log-moment and log-cumulant. For the product model, the Mellin transform has the similar convolution properties as the traditional Fourier transform for the additive model [9].

In this study, univariate log-moments, μ, and log-cumulants, κ, of the SAR intensity measurements are required for the contamination test described in Section 3, which can be obtained from the Mellin kind characteristic functions. They are essentially moments and cumulants computed on a logarithmic scale. The rth-order log-moment of the intensity measurements I is defined as

$$\mu_r\{I\} \triangleq E\{(\log I)^r\}\ , \tag{2}$$

where $E\{\dots\}$ is the expectation operator. The corresponding rth-order log-cumulant $\kappa_r\{I\}$ is usually calculated from the linear combination of the log-moments up to order r. The first four orders of log-moment to log-cumulant transformations are:

$$\kappa_1 = \mu_1\ , \tag{3}$$
$$\kappa_2 = \mu_2 - \mu_1^2\ , \tag{4}$$
$$\kappa_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3\ , \tag{5}$$
$$\kappa_4 = \mu_4 - 4\mu_1\mu_3 - 3\mu_2^2 + 12\mu_1^2\mu_2 - 6\mu_1^4\ . \tag{6}$$

In practice, sample intensity log-cumulants are estimated from the data via the sample intensity log-moment equation

$$\langle \mu_r\{I\} \rangle = \frac{1}{n}\sum_{i=1}^{n}(\log I_i)^r\ , \tag{7}$$

where n is the number of samples for estimation.

For a single continuous variable, mean (brightness), variance, skewness and kurtosis (peakedness) are commonly used to describe the shape of its pdf, which can be related to the first four orders of log-cumulant, respectively. Note that, higher order log-cumulants are known to have larger estimation variance. Also, in real PolSAR measurements, the brightness (represented by the first order log-cumulant) is usually considerably different between the co- and cross-pol channels. It must be taken into account when comparing results from multiple polarimetric channels. In this study, only second and third order log-cumulants (which are independent of absolute brightness) are applied at the moment, which simplifies the contamination test mechanism, have reasonable variance and avoids the brightness discrepancy between channels.

2.3 Log-cumulant diagram

From the previous study [3], the log-cumulant diagram displays a multi-dimension space, where each dimension corresponds the log-cumulant of one particular order. This is the preferred tool to visualize the second and third order log-cumulant space used in this study.

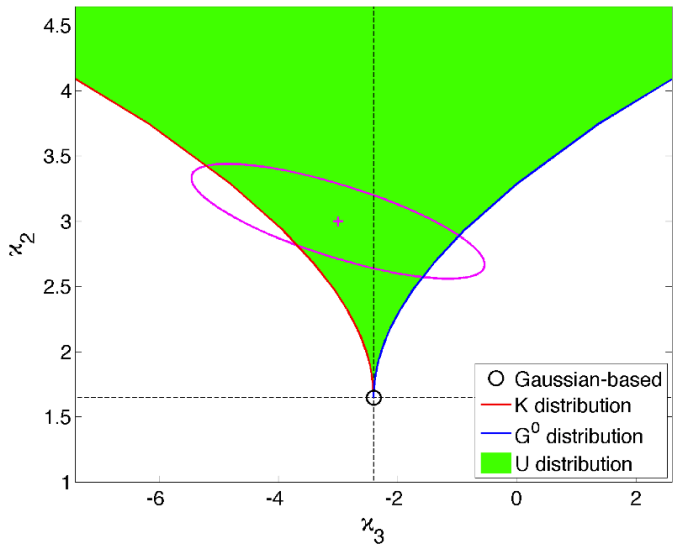

Figure 1. A pure sea clutter example in the single look univariate log-cumulant diagram. The third-order log-cumulant is plotted against the second-order log-cumulant. The + sign and the ellipse indicate the expected mean and variance of the sample estimator.

Figure 1 shows the single look univariate log-cumulant diagram plotting the third-order against the second-order log-cumulant. A pure sea clutter example is plotted in magenta. The magenta cross marks a location in the log-cumulant space, which defines the shape of the clutter probability distribution's variance and skewness in terms of log-intensity. The magenta ellipse represents the variability due to the finite number of random measurements. The ellipse size depends on both the number of samples used to estimate the log-cumulants and the selected significant level. When another clutter area's log-cumulant measurements appear inside the magenta ellipse in the log-cumulant diagram, its statistical properties (shape of probability distribution) are considered to be similar to the plotted clutter example. However, in the opposite case, if the log-cumulant measurements show obvious deviations from a selected reference (they appear outside the variability ellipse), it most likely means that the two clutters have significantly different statistics.

Also in the log-cumulant diagram, the colored background framework represents different model families, i.e. the Gaussian-based, $\mathcal{K}$, $\mathcal{G}^0$ and $\mathcal{U}$ distributions. The Gaussian-based distribution is centered with the crossing of the dashed lines at a point (black circle). Considering the product model, textured distributions are multiplied to the Gaussian distribution representing speckle, which results in additive log-cumulants. Therefore, textured data deviate from the Gaussian-based distribution point in the log-cumulant diagram. Thus, the background framework becomes a useful guide to understand the degree and type of textured distribution.

3. CONTAMINATION TEST

In this section, a contamination test is introduced accompanied by a real world example. Figure 2 shows the SPAN image of the selected Radarsat-2 scene on a logarithmic scale. Two 50×20 pixel test areas (red and blue rectangles) are selected to demonstrate the sea clutter contamination mechanism. Area 1 represents contaminated sea clutter containing a non-background target (a ship). A randomly selected area 2 represents some uncontaminated background clutter for comparison.

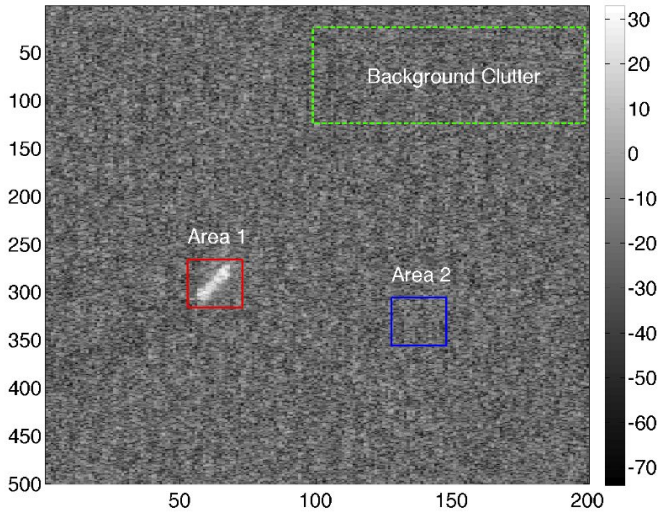

Figure 2. SPAN image of the selected Radarsat-2 scene on a logarithmic scale.

As demonstrated in Figure 2, the background sea clutter is often mixed with non-background targets. The mean-normalized single look intensity histograms of the HH and HV polarizations for the two selected test areas are shown in Figure 3 and 4. Significant differences between the two test areas can be observed by regarding the shape of each histogram, e.g. variance and skewness. The histograms from the contaminated area 1 (red) have heavier tails and larger skewness compared to area 2 (blue). It is evident that they do have different statistical properties.

From section 2, log-cumulants appear to be an effective tool to measure the shape differences of probability distributions. In the following experiments, the second and third order log-cumulants are measured for each polarimetric channel of the two test areas and a reference background sea clutter area. The reference background sea clutter is selected as a 100×100 pixels green dashed-line rectangle area in the upper-right corner of Figure 2. An 8×8 pixel sliding window is applied in the computation of the log-cumulants. Figure 5 shows the

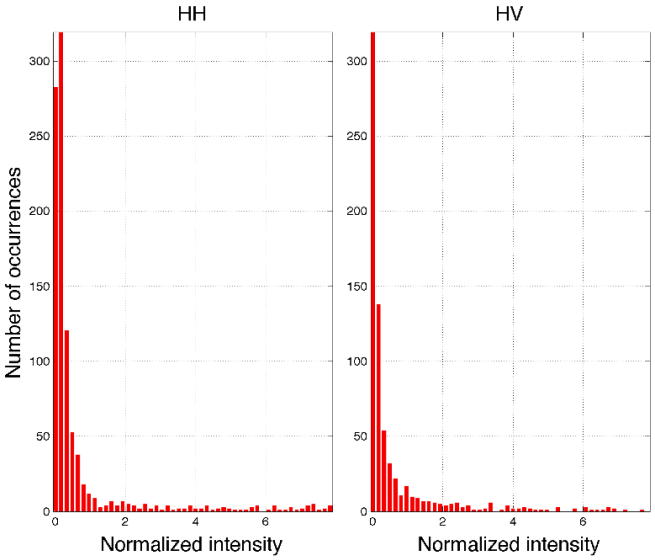

Figure 3. Normalized single look intensity histograms of HH and HV polarizations for test area 1.

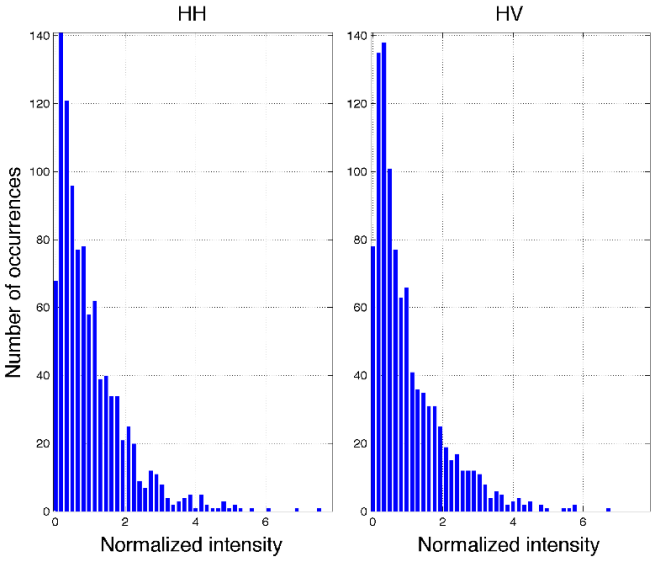

Figure 4. Normalized single look intensity histograms of HH and HV polarizations for test area 2.

log-cumulant measurements in the log-cumulant diagrams plotting the third-order against the second-order log-cumulants. Red and blue dots represent the log-cumulants calculated from the sliding window neighborhood around each pixel of the test area 1 and 2, respectively. The magenta cross and ellipse indicate the selected reference background clutters location and variability in the log-cumulant space. Note that the variability are calculated with a 99.999% statistical significance level. Only the HH and HV polarizations are shown here to represent the co- and cross-pol channels, assuming the shape similarity between the VV and HH channels and the VH and HV channels, observed in Figures 3 and 4.

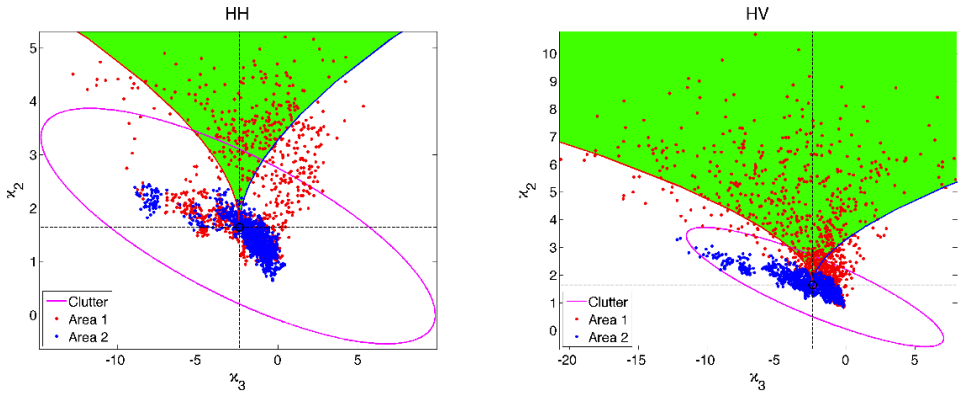

Figure 5. Log-cumulant diagrams. An 8×8 pixel sliding window is applied. The ellipse shows the 99.999% significant level. Left: HH polarization. Right: HV polarization.

Compared to the selected reference background sea clutter, the results from the uncontaminated area 2 (blue dots) are mostly centralized inside the magenta ellipse. However, the results from the contaminated area 1 (red dots) are significantly deviated from the reference background. Cross-pol channels show more obvious deviations than co-pol channels, since the co-pol channels have less sensitivity to the non-background targets. Notice that the blue dot cloud is not shown homogeneously as we expected, but shows some local proupings. This is mainly caused by the correlation between each sliding window. However, the results still clearly show that the log-cumulant has the ability to distinguish the contaminated area from the uncontaminated background.

Therefore, the statistical contamination test is designed to exploit this observed effect. Note that this test is capable of handling any order of log-cumulant, but only second and third order log-cumulant will be discussed in this study. The test strategy is listed as follows: (Functions and assumptions are based on the previous study [3].)

(1) Calculate the sample log-cumulants up to third order and form the sample log-cumulant vector $\langle \mathbf{k} \rangle_i$ with a sliding window approach for each pixel, i.

(2) Calculate the mean log-cumulant vector $\bar{\mathbf{k}}_{ref}$ from these estimates within the manually selected reference area.

(3) Calculate the sample log-cumulant covariance matrix from the sample estimates in the reference area as

$$\mathbf{K} = \frac{N}{N-1} \sum_{i=1}^{N} [(\langle \mathbf{k} \rangle_i - \bar{\mathbf{k}}_{ref})(\langle \mathbf{k} \rangle_i - \bar{\mathbf{k}}_{ref})^T] \,, \tag{8}$$

where N is the number of samples of the selected reference background area.

(4) Calculate the testing values for each pixel as

$$Q(i) = n(\langle \mathbf{k} \rangle_i - \bar{\mathbf{k}}_{ref})^T \mathbf{K}^{-1} (\langle \mathbf{k} \rangle_i - \bar{\mathbf{k}}_{ref}) \,, \tag{9}$$

where n is the number of sample applied in the sliding window.

(5) Repeat steps 1 to 4 for all polarimetric channels.

(6) Assuming the Chi-square approximation [3], $Q \xrightarrow{\mathcal{D}} \mathcal{X}^2(p)$, where $\mathcal{X}^2(p)$ denotes a central chi-square distribution with $p = 2$ degrees of freedom (the dimension of the log-cumulant vector). Calculate the testing threshold with a given statistical significant level (99.999% in this study) for the chi-square distribution.

(7) Compare the testing values for each pixel to the testing threshold per channel. Thus four polarimetric channels will provide an output with four layers, which can be combined to one four-level output. The output pixel with higher level means that it has larger possibility belonging to a non-background target.

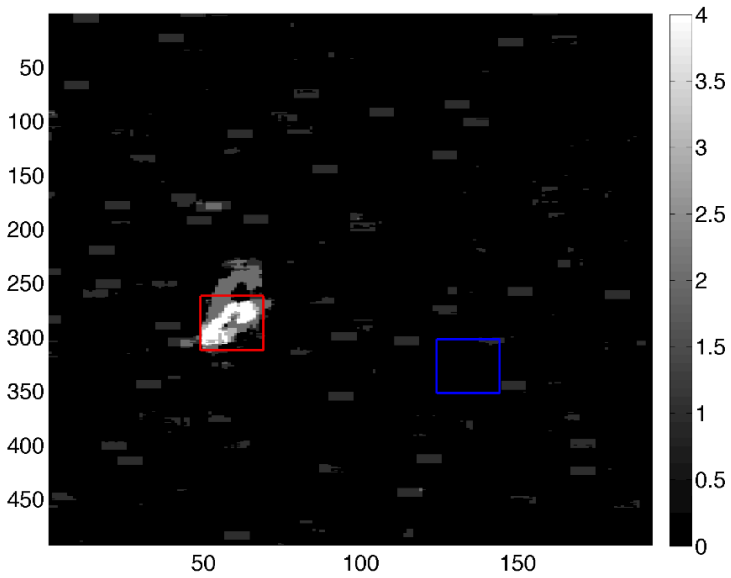

Figure 6. Four-level contamination test result. An 8×8 pixel sliding window is applied.

The contamination test result of the selected Radarsat-2 scene is shown in Figure 6, where an 8×8 pixel sliding window is applied, and the test area 1 and 2 are marked as red and blue rectangles, respectively. The non-background target is very well detected in all polarimetric channels, which in shown as the level 4 pixels in the figure. There are some false alarms observed from multiple channels, but none of them are found with level 4. Other sliding window sizes, e.g. 16×16 pixels and 32×32 pixels, have also been tested in this study. Similar log-cumulant diagrams and contamination test results are shown in Figures 7 to 10. The contamination ratio, which is the measure of the fraction of contaminated pixels versus total pixels of the sliding window, is an important factor to consider when choosing the window size. From the figures, larger sliding windows can provide larger deviations in the log-cumulant space, which is good for the statistical test. However, the larger sliding window loses the details of the non-background target and will increase the risk of missing any small targets. Therefore, an 8×8 pixel sliding window will be applied for the rest experiments.

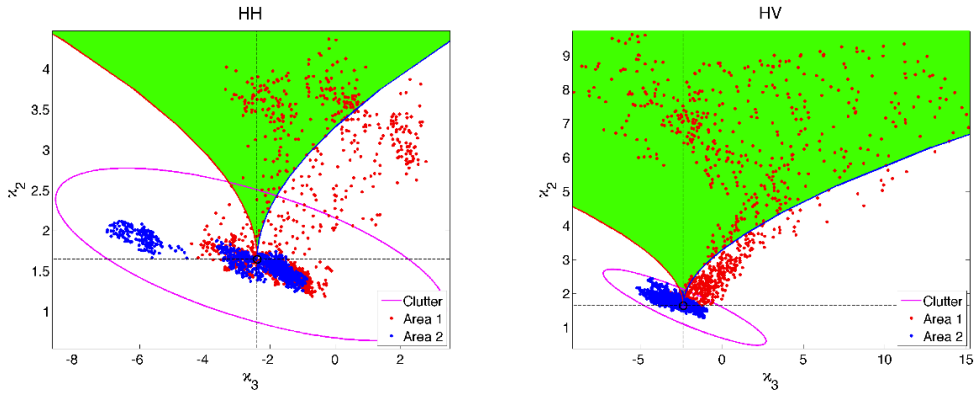

Figure 7. Log-cumulant diagrams. A 16×16 pixel sliding window is applied. The ellipse shows the 99.999% significant level. Left: HH polarization. Right: HV polarization.

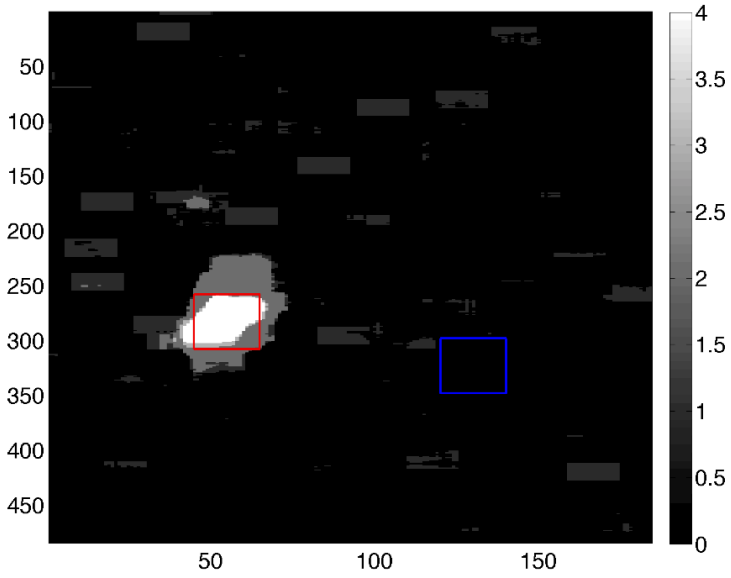

Figure 8. Four-level contamination test result. A 16×16 pixel sliding window is applied.

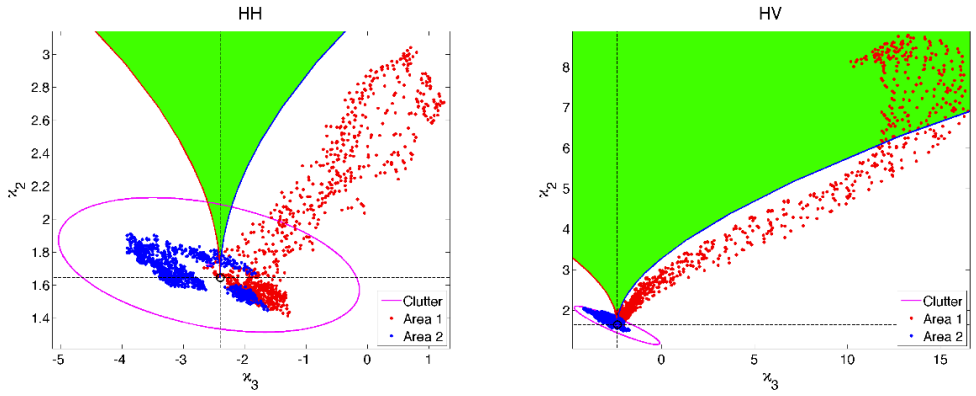

Figure 9. Log-cumulant diagrams. A 32×32 pixel sliding window is applied. The ellipse shows the 99.999% significant level. Left: HH polarization. Right: HV polarization.

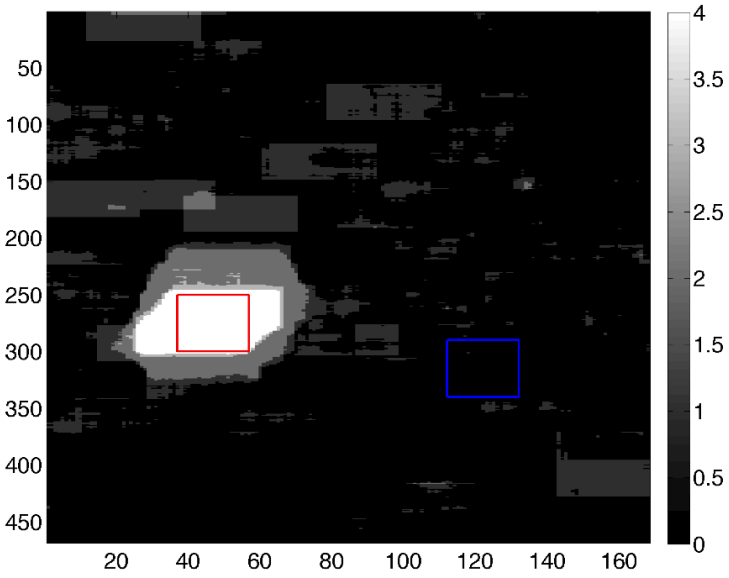

Figure 10. Four-level contamination test result. A 32×32 pixel sliding window is applied.

4. EXPERIMENTS

In this section, the proposed contamination test is examined based on three quad-polarimetric SAR scenes from August 5, March 26 and March 16, 2010. Detailed information on the datasets are given in Table 1. Corresponding Automatic Identification System (AIS) [10] data from Statoil Marin are included in this study, which provides the coordinates of the registered ships in the system. Three 2500×500 pixels sub-scenes are selected, one from each dataset, covering various oceanographic conditions, i.e. smooth sea surface on August 5 and more complicated oceanographic conditions on March 26 and March 16.

Table 1. Radarsat-2 dataset information.

Date	Resolution (azimuth $\times$ range)	Incidence angle range	Number of targets in selected region
2010-08-05	4.8 m $\times$ 08.0 m	$25.8^\circ \sim 27.6^\circ$	4
2010-03-26	5.2 m $\times$ 11.8 m	$35.5^\circ \sim 37.0^\circ$	5
2010-03-16	5.1 m $\times$ 11.8 m	$40.2^\circ \sim 41.6^\circ$	5

Figures 11, 12 and 13 show the SPAN images of the selected Radarsat-2 scenes, on a logarithmic scale, together with the results from the four-level contamination tests. A sliding window of 8×8 pixels is applied in the experiments. All known targets ($T_1 \ldots T_i$), whose locations are marked with red rectangles, are detected in every polarimetric channel. There are some false alarms observed in some of the individual polarimetric channels, but very few are detected across all four polarimetric channels. By combining information from all the channels, the background sea clutter can be more accurately estimated by eliminating the potential targets detected. For the more complicated oceanographic conditions on March 26 and March 16, 2010, both bright and dark areas are visible. This is problematic, e.g. for target detection, due to the transitions of different statistical properties. However, by applying a small sliding window, the effect around the gradual and weak transitions can be limited. Potentially, the contamination test can be applied to detect the boundary of the mixed oceanographic conditions, which have been observed when using the larger sliding windows, or the brightness change by testing more than the shape of probability distribution.

5. CONCLUSIONS

This work introduces an effective log-cumulant based test for sea clutter contamination. Background clutter mixed with non-background targets gives contaminated areas, leading to inaccurate sea clutter estimation due to a changing shape of the mixed probability distribution. In log-cumulant space, the mixture induces deviation from the background model. The test results show that the log-cumulant is a sensitive tool that may help to detect mixtures and thus allow improved background sea clutter estimation for subsequent applications or direct first-stage target detection. Further studies regarding the contamination ratio and information from the first order log-cumulant, are necessary to design a more efficient test strategy.

ACKNOWLEDGMENTS

All Radarsat-2 Data and Products ©MacDONALD, DETTWILER AND ASSOCIATES LTD. 2010 - All Rights Reserved. The Automatic Identification System (AIS) data are provided by Statoil Marin.

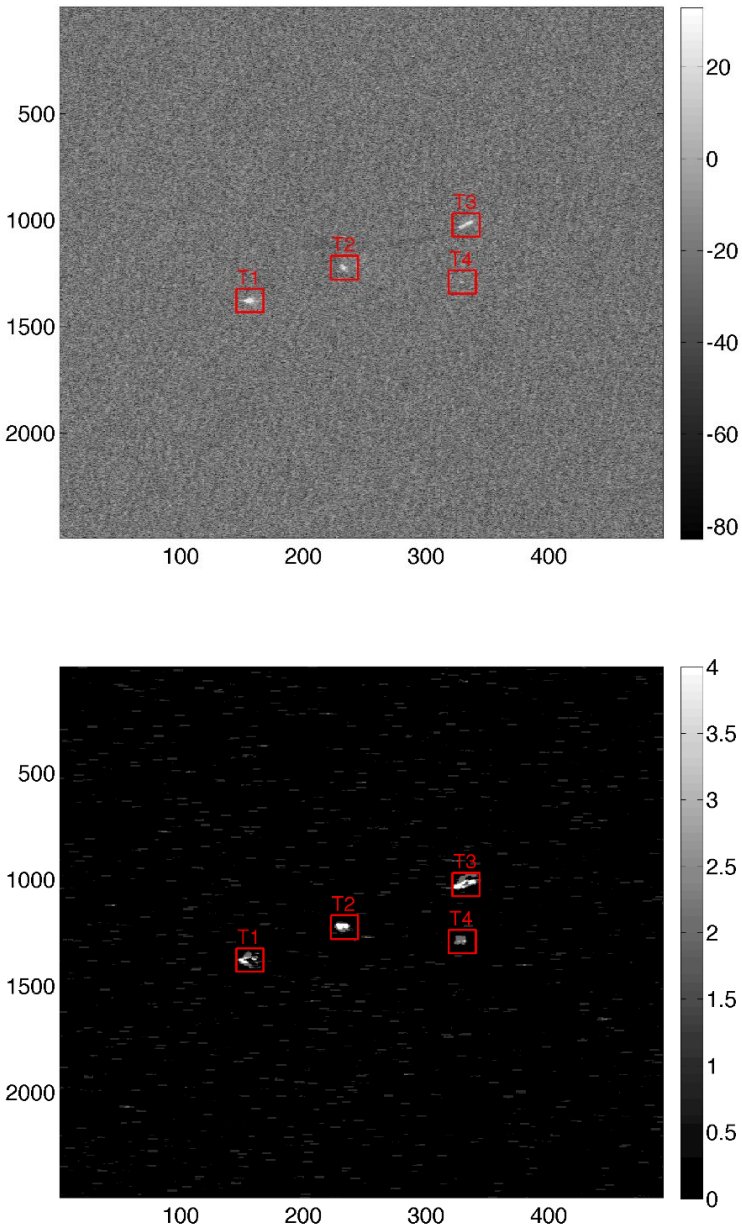

Figure 11. Upper: SPAN image of the selected Radarsat-2 scene from 2010-08-05 on a logarithmic scale. Bottom: Four-level contamination test result with an 8×8 pixel sliding window.

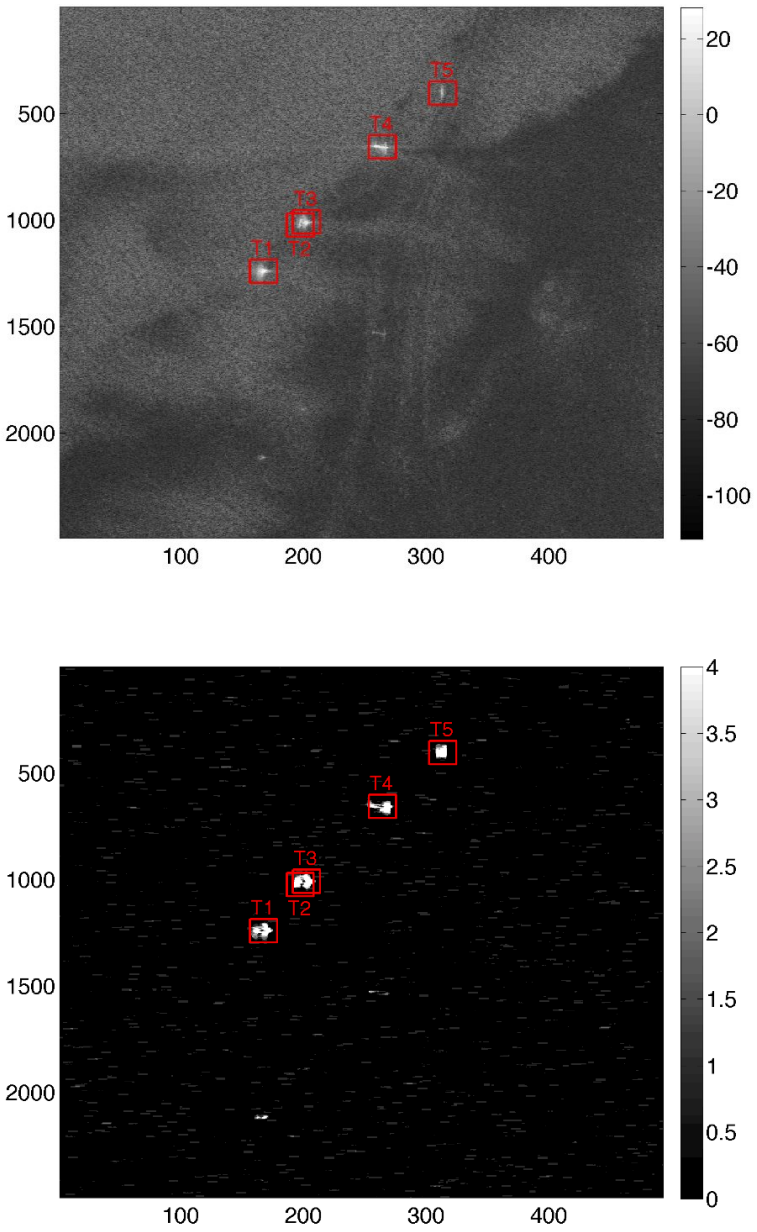

Figure 12. Upper: SPAN image of the selected Radarsat-2 scene from 2010-03-26 on a logarithmic scale. Bottom: Four-level contamination test result with an 8×8 pixel sliding window.

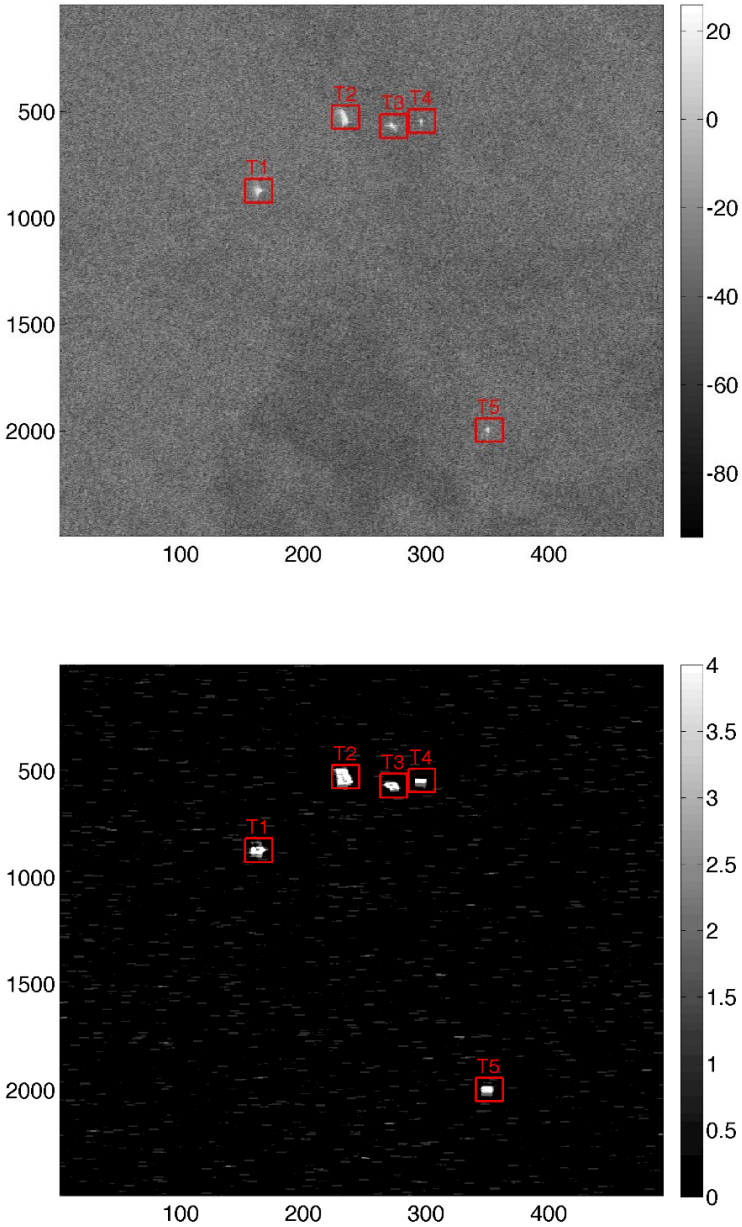

Figure 13. Upper: SPAN image of the selected Radarsat-2 scene from 2010-03-16 on a logarithmic scale. Bottom: Four-level contamination test result with an 8×8 pixel sliding window.

REFERENCES

[1] Nicolas, J.-M., "Introduction aux statistiques de deuxième espèce: applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar," *Traitement du Signal* **19**(3), 139–167 (2002).

[2] van Zyl, J. and Yunjin, K., [*Synthetic Aperture Radar Polarimetry*], JPL Space Science and Technology Series, Wiley (2011).

[3] Anfinsen, S. and Eltoft, T., "Application of the matrix-variate Mellin transform to analysis of polarimetric radar images," *Geoscience and Remote Sensing, IEEE Transactions on* **49**, 2281–2295 (Jun 2011).

[4] Yueh, S. H., Kong, J. A., Jao, J. K., Shin, R. T., and Novak, L. M., "K-distribution and polarimetric terrain radar clutter," *J. Electrom. Waves Applic.* **3**(8), 747–768 (1989).

[5] Yueh, S., Kong, J., Shin, R., and Zebker, H., "Statistical modeling for polarimetric remote sensing of earth terrain," in [*Geoscience and Remote Sensing Symposium, 1990. IGARSS '90. 'Remote Sensing Science for the Nineties'., 10th Annual International*], 157–160 (May 1990).

[6] Jakeman, E. and Pusey, P., "A model for non-Rayleigh sea echo," *Antennas and Propagation, IEEE Transactions on* **24**, 806–814 (Nov 1976).

[7] Freitas, C. C., Frery, A. C., and Correia, A. H., "The polarimetric G distribution for SAR data analysis," *Environmetrics* **16**(1), 13–31 (2005).

[8] Bombrun, L. and Beaulieu, J.-M., "Fisher distribution for texture modeling of polarimetric SAR data," *Geoscience and Remote Sensing Letters, IEEE* **5**, 512–516 (Jul 2008).

[9] Anfinsen, S. N., "On the supremacy of logging," in [*Proc. 5th Int. Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLinSAR2011)*], (Jan 2011).

[10] International Maritime Organization, "http://www.imo.org/ourwork/safety/navigation/pages/ais.aspx," (Last checked on 4th Sep 2012).