Statistical tests for a ship detector based on the Polarimetric Notch Filter

Armando Marino, *Member, IEEE*, Irena Hajnsek, *Fellow, IEEE*

Abstract

Ship detection is an important topic in remote sensing and Synthetic Aperture Radar has a valuable contribution, allowing detection at night time and with almost any weather conditions. Additionally, polarimetry can play a significant role considering its capability to discriminate between different targets. Recently, a new ship detector exploiting polarimetric information was developed, namely the Geometrical Perturbation Polarimetric Notch Filter (GP-PNF).

This work is focused on devising two statistical tests for the GP-PNF. The latter allow an automatic and adaptive selection of the detector threshold. Initially, the probability density function (pdf) of the detector is analytically derived. Finally, the Neyman-Pearson (NP) lemma is exploited to set the threshold calculating probabilities using the clutter pdf (i.e. a Constant False Alarm Rate, CFAR) and a likelihood ratio (LR).

The goodness of fit of the clutter pdf is tested with four real SAR datasets acquired by the RADARSAT-2 and the TanDEM-X satellites. The former images are quad-polarimetric, while the latter are dual-polarimetric HH/VV. The data are accompanied by the Automatic Identification System (AIS) location of vessels, which facilitates the validation of the detection masks. It can be observed that the pdf's fit the data histograms and they pass the two sample Kolmogorov-Smirnov and χ^2 tests.

Keywords

Synthetic Aperture Radar, Polarimetry, Ship Detection.

I. Introduction

Synthetic Aperture Radar (SAR) allows to measure the reflectivity maps at microwave frequencies of an observed scene. The strategic advantage of SAR in ship detection is the possibility to monitor at night time, under cloud cover and with meters resolution independently on the distance. For this reason, SAR was largely exploited in the past decades to monitor ships from satellites [1–14]. In SAR images, the main feature of ships is a relatively large backscattering signal, which is usually brighter in comparison with the sea background.

Armando Marino and Irena Hajnsek are with ETH Zurich, Institute of Environmental Engineering, Zurich, Switzerland (e-mail: marino@ifu.baug.ethz.ch). Irena Hajnsek is also with the German Aerospace Center (DLR), High Frequency Department, Oberpfaffenhofen, Germany.

This led to the idea of separating vessels from the background performing a statistical test on the intensity of the sea clutter. Once the pdf of the clutter intensity is known, the problem of selecting the detector threshold can be solved using the Neyman-Pearson lemma on the probability of detection (P_d) or false alarms (P_f) [15]. Unfortunately, the statistical distribution of ships is complicated to derive and therefore the tests are generally based only on P_f. In particular, the threshold is selected in order to keep P_f constant (constant false alarm rate, CFAR) [1–4, 7, 10, 12, 16, 17]. When the sea statistics are not perfectly fitting the data, the test can still be performed but it will not assure a constant P_f.

The detector exploited in this paper makes use of more polarimetric channels that can be acquired by all the recent SAR satellites. Regarding the benefits of polarimetry in the context of ship detection, it can readily be observed that the simple exploitation of the cross-polarized channel (HV) rather than the co-polarized ones (HH or VV) may increase substantially the detection performance [1, 18]. Several detectors were proposed in the recent years. Some of them exploit the different polarimetric channels as independent measurements of the same scene [4, 19, 20]. Another class of polarimetric detectors adds some physical rationale exploiting knowledge regarding the scattering. The idea behind these methodologies is that the differences between sea clutter and targets can be magnified if some specific aspects of the polarimetric return are observed. In this second category, there are algorithms with a detection role based on some rationale linked to the physical behavior of the sea clutter [21–25]. For instance, some algorithms rely on the assumption that the sea behaves as a Bragg surface [26–28]. The ship detector analyzed in this paper belongs to this physical category and is based on the assumption that the sea clutter (locally) presents an homogeneous polarimetric behavior. This is the *Geometrical Perturbation - Polarimetric Notch Filter* (GP-PNF) [29–32].

A very brief introduction to polarimetry is provided here with the mere purpose to show the tools that will be exploited in the following. A single target is defined as a deterministic target which does not change its polarimetric behavior in time/space. Therefore, it can be represented by a single scattering matrix or equivalently a single scattering vector [26, 33]:

$$\underline{k}_L = [HH, HV, VH, VV]^T , \tag{1}$$

where T stands for Transpose, H and V are for linear horizontal and vertical and the repeated letter is for transmitter-receiver. The previous is obtained using the Lexicographic basis set and HH, HV, VH

and VV are the scattering channels. In the case of a reciprocal medium and monostatic sensor, $HV = VH$ and $\underline{k}$ is three-dimensional complex (i.e. $\underline{k}_P \in \mathbb{C}^3$) [26]. Another largely exploited basis set to convert the scattering matrix into a scattering vector is the Pauli basis. In the reciprocal case, this is $\underline{k}_P = 1/\sqrt{2}\,[HH + VV, HH - VV, 2HV]^T$. The targets observed by a SAR system are often distributed over an area larger than the resolution cell and composed by different objects. For this reason, each pixel of such distributed targets may have a specific polarimetric behavior. Such targets take names of *partial targets* and they can be characterized exploiting the second order statistics [26]. In this context, a target covariance matrix can be estimated as $[C] = \langle \underline{k}\ \underline{k}^{*T} \rangle$, where $\langle . \rangle$ is used here as the finite averaging operator and $*$ stands for Conjugate. In case that the Pauli basis is exploited, the covariance matrix takes the name of *coherency* matrix.

II. Geometrical Perturbation - Polarimetric Notch Filter

The main idea of the GP-PNF is that the polarimetric responses of sea clutter and ships are different. The GP-PNF is based on the Geometrical Perturbation Filter [34–37]. The latter considers a perturbed version of the target to be detected and then it checks for the coherence between original and perturbed version in the data. The reader is redirected to [34, 35] for more information regarding the GPF.

A *feature partial scattering vector* is introduced [35]:

$$\begin{aligned} \underline{t} =& [t_1, t_2, t_3, t_4, t_5, t_6]^T = \\ =& [\langle |k_1|^2 \rangle, \langle |k_2|^2 \rangle, \langle |k_3|^2 \rangle, \langle k_1^{*T} k_2 \rangle, \langle k_1^{*T} k_3 \rangle, \langle k_2^{*T} k_3 \rangle]^T. \end{aligned} \tag{2}$$

The vector containing the second order statistics of the sea clutter is defined as $\underline{t}_{sea}$ [29, 30]. The sea clutter can be completely characterized by a vector in $\mathbb{C}^6$ (in case quad-pol data are available), while, vessels can have a large variety of polarimetric signatures depending on orientation, material and structure of the vessel. Therefore, it is not possible to characterize each possible polarimetric signature of ships with a single vector. The GP-PNF approach is to focus on targets that do not behave as the sea. For this reason, the GP-PNF is a heterogeneity detector, that is focused on targets which present polarimetric heterogeneity in the 6D complex subset. It is interesting to notice, that such strategy allows to detect targets with backscattering power (i.e. Trace of Covariance matrix) comparable to the one of the sea as long as they appear polarimetrically different. Geometrically, this means that it is focused on targets that live in the complement orthogonal subset to the sea vector (5 dimensional complex). Please note, such strategy has the advantage of not being related

to any assumption regarding the specific polarimetric signature of the sea, as long as this is stable in the training window exploited. The length of the vector (hereafter defined as power) representing the target in the complementary subset can be calculated as: $P_t = \underline{t}^{*T}\underline{t} - |\underline{t}^{*T}\hat{\underline{t}}_{sea}|^2 = P_{tot} - P_{sea}$. The final expression is:

$$\gamma_n = \frac{1}{\sqrt{1 + \frac{RedR}{\underline{t}^{*T}\underline{t} - |\underline{t}^{*T}\hat{\underline{t}}_{sea}|^2}}} > T. \tag{3}$$

where γ_n is the GP-PNF distance (i.e. the detector) and $RedR$ and T are two detector parameters. More details and justifications regarding the mathematical derivations can be found in [30]. From an implementation point of view, the vector $\underline{t}$ is estimated with a boxcar filter with a small window (e.g. 11x11 pixels), while $\underline{t}_{sea}$ is computed with a boxcar filter using a larger window (e.g. 51x51 pixels). It is important to keep in mind that this filtering methodology could be optimized in the future. For instance, for the smaller window an adaptive filtering could be employed that takes into account the heterogeneity (or non-stationarity) of the target observed [38]. In such a way, the polarimetric characteristic of each pixel of a vessel will be preserved providing a better discrimination with respect to the sea background. On the other hand, a non-local filter [39] could be used instead than the large window, producing better estimates of the clutter background. We leave all these analysis for future work.

Previously, the detector parameters were fixed following an asymptotic approach ($\langle [C]_{sea} \rangle = E\left[[C]_{sea}\right]$). However, Monte Carlo simulations were performed in [30], showing that the estimation of the local sea clutter power can improve the performances. Aim of next section is to derive the statistical distribution of γ_n for the sea clutter, in order to have a more rigorous test and to take into account the variability of the sea clutter.

III. Statistical characterization

The aim of this section is to derive an analytical expression for the probability density function (pdf) of γ_n (i.e. the output of the detector).

Before to start the derivation of the pdf, the test hypotheses are defined:

$$H_0 : sea\ clutter \tag{4}$$

$$H_1 : vessel$$

Initially, the conditional pdf of γ_n in the hypothesis H_0 (only presence of clutter) is derived. This can be written

as $f_{\Gamma_n}(\gamma_n|H_0)$ and represents the likelihood of an output γ_n given that the hypothesis H_0 is true. The problem is tackled decomposing the derivation in two parts. In the first, the attention is focused on finding the pdf of the random variable (r.v.) that is generating γ_n. Secondly, a transformation of random variables can be applied to derive the pdf of γ_n [40]. In this context, a possible candidate to be a generator of γ_n is the vector $\underline{t}$, since the pdf's of its components are well known in the hypothesis of complex Gaussian Single Look Complex (SLC) pixels [33]. However, magnitudes and inner products of $\underline{t}$ vectors are unknown. Besides, it would be beneficial to have a single r.v. generating γ_n, since this assures a much easier transformation. After an analysis of the final formula of γ_n, it appears that a good candidate to be the generator of γ_n is P_t (i.e. the target power).

A. *Distribution of P_t*

P_t is the squared norm of the vector in the subset complementary to the sea vector (i.e. the target subset). In the H_0 hypothesis, there is absence of any target and therefore the differences between $\underline{t}_{sea}$ (extracted in the training window) and $\underline{t}$ (extracted in the test window) are only due to estimation errors (e.g. due to finite number of samples). The training window contains a much larger amount of samples than the test window. Therefore, the underlying signature of the sea can be extracted with a much smaller estimation error in the training window. More details regarding training and test windows in practical scenarios are provided in a following section. The target vector in the complementary subset can be calculated as $\underline{t}_t = \underline{t} - \left(\hat{\underline{t}}_{sea}^{*T}\underline{t}\right)\hat{\underline{t}}_{sea}$.

The χ^2 theorem states that, given a Gaussian r.v. x, the test $\sum_{i=1}^{n}\frac{(x_i-E[x_i])^2}{VAR[x_i]}$ (with n number of realizations in the considered table, $E[.]$ is the expected value and $VAR[.]$ is the expected variance) has a χ^2 distribution. In order to be able to use such theorem two assumptions have to be made:

1. The expected value can be substituted by the components of $\underline{t}_{sea}$ estimated on large windows. Please note, this assumption requires that the number of pixels used to derive $\underline{t}_{sea}$ is big enough to have a very small variance.

2. $\left(\hat{\underline{t}}_{sea}^{*T}\underline{t}\right)\hat{\underline{t}}_{sea} \approx \underline{t}_{sea}$: This means that the sea vector in the training window is not largely different (in average) to the test vector $\underline{t}_{sea} \approx \underline{t}$. This assumption is true as long as the sea is homogeneous and $\underline{t}$ is obtained performing some averaging. Such assumption is expected from a detector based on second order statistics (i.e. the latter cannot be extracted with a single SLC pixel). Some analysis of the minimum number of pixels required is investigated in the section dedicated to the Monte Carlo simulation.

Following the two previous assumptions, the expression $\underline{t}_t = \underline{t} - \left(\underline{\hat{t}}_{sea}^{*T}\underline{t}\right)\underline{\hat{t}}_{sea} \approx \underline{t} - \underline{t}_{sea}$ can be written. The power of such vector can be considered: $\|\underline{t}_t\|^2 = \|\underline{t} - \underline{t}_{sea}\|^2$. The squared norm can be decomposed as the sum of the components of the vectors $\underline{t}$ and $\underline{t}_{sea}$ (6 dimensional in quad-polarimetry):$\|\underline{t}_t\|^2 = \sum_{j=1}^{6} |\underline{t}_j - \underline{t}_{jsea}|^2$.

In order to obtain the same formulation of the χ^2 theorem the denominator should contain the variances of each component. Dividing both expressions by the variance of the first component ($VAR[t_1]$) it can be obtained: $\frac{\|\underline{t}_t\|^2}{VAR[t_1]} = \sum_{j=1}^{6} \frac{\|\underline{t}_j - \underline{t}_{jSea}\|^2}{VAR[t_1]}$. Subsequently, a change of basis can be considered that makes equal the variances of each of the components. By definition the length of a vector is invariant to change of basis therefore such operation does not modify the value of P_t. Moreover, such operation can be always accomplished. A way to proceed may be to perform a whitening of the $\underline{t}_{sea}$ components. Interestingly, such operation is not necessary from a practical point of view, since the length of the vector is not influenced by such transformation and we are only interested in the length of the vector. Therefore, in the basis where $VAR[t_i] = VAR[t_j],\ \forall i, j = 1, ..., 6$, the previous expression will become:

$$\frac{\|\underline{t}_t\|^2}{VAR[t_1]} = \sum_{j=1}^{6} \frac{\|\underline{t}_j - \underline{t}_{jSea}\|^2}{VAR[t_j]}. \tag{5}$$

The final step in order to obtain a χ^2 distributed is that the components of the target vector are Gaussian distributed. The vector components are estimated performing some average, however, the dimension of the test windows may be not large enough for the theorem of the central limit to be valid. Fortunately, we are not interested in the single component, but in their sum after the change of bases (which is the operation that would allow the χ^2 theorem to be applicable). After the latter operation, the elements in Eq. 5 are linear combination of 6 r.v. with statistics similar (but not equal) to Gaussian. Such operation can increase up to 6 times the number of looks considered. This should make the central limit more valid also for a smaller number of initial average.

In the hypothesi of homogeneous clutter, it is possible to state that the estimation errors do not have a preferential polarimetric behavior, (i.e. they are polarimetrically white). This is because the sample mean estimator is the maximum likelihood estimator for the mean and it is unbiased. Therefore, under the previous hypothesis, the components of $\underline{t}_t$ are Gaussian zero mean. The actual number of samples necessary to have an adequate approximation of a Gaussian is investigated in the following with Monte Carlo simulations.

After all these considerations, Eq. 5 has a χ^2 distribution. The following step consists in multiplying both parts by $VAR[t_1]$. The result of scaling a χ^2 distribution is a Γ distribution, independently on the scaling factor

[40]. Therefore, the target power should be Γ distributed.

It is possible to conclude that $\|\underline{t}_t\|^2$ has a Γ distribution, under the assumption that the averaging windows are big enough. Interestingly, P_t has the same statistical behavior as the power of a SAR image, which is coherent with the idea that such parameter represents the power of a target in case of perfectly homogeneous distribution. It has also to be noticed, that for the sea clutter, the Γ distribution is not the most adequate statistical model [41]. Here, we only want to point out that the distribution of P_t resembles the one of an intensity image in case of homogeneous target.

B. Distribution of γ_n

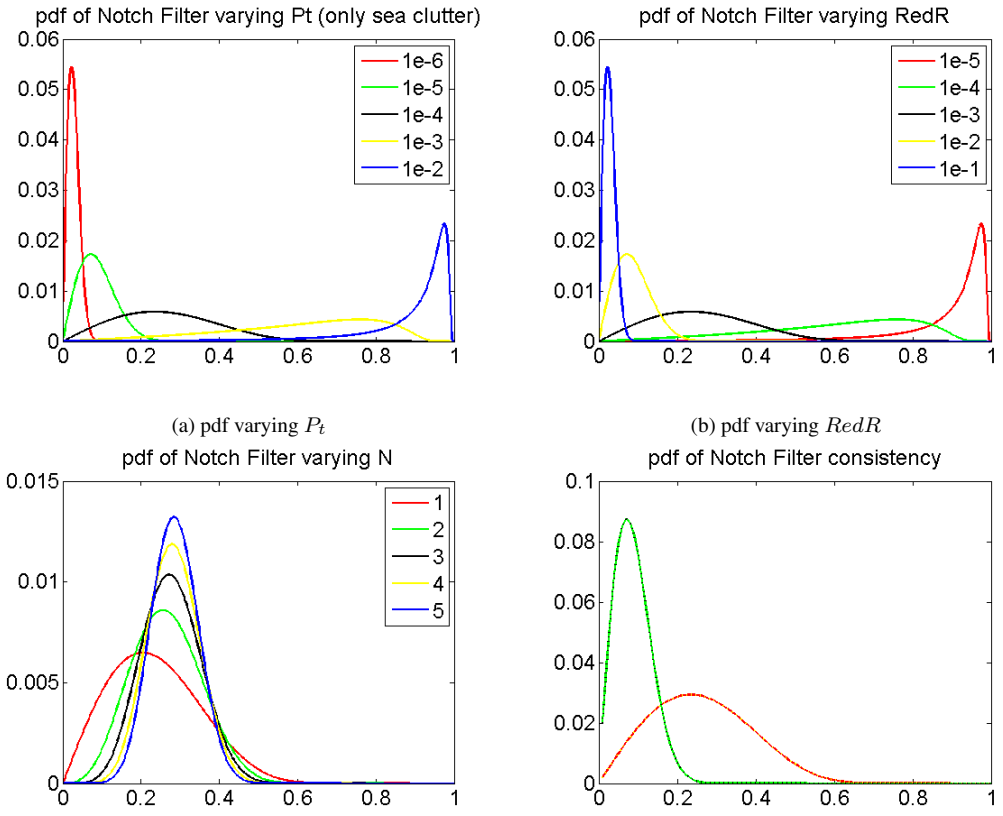
The theorem of transformation of random variables is employed to transform the pdf of P_t into the one of γ_n (the theorem can be interpreted as a change of variables for an integral) [40]. The transformation is the mathematical expression of the detector in Eq.3: $\gamma_n(p_t)$. The theorem states that $f_{\Gamma_n}(\gamma_n) = f_{P_t}(\hat{p}_t)\frac{\partial p_t(\gamma_n)}{\partial \gamma_n}$, where $\hat{p}_t$ is the solution of $\gamma_n(p_t)$. Additionally, the pdf is only valid where the solution of γ_n exists.

The mathematical derivation is presented in the Appendix, here only the final expression is provided:

$$f_{\Gamma_n}(\gamma_n) = \frac{2}{\Gamma(N)}\left(\frac{N}{\mu}\right)^N \left(RedR\frac{\gamma_n^2}{1-\gamma_n^2}\right)^{N+1} \frac{2}{RedR}\,\gamma_n^{-3}\times \\ \times exp\left[-RedR\frac{N}{\mu}\frac{\gamma_n^2}{1-\gamma_n^2}\right]\; rect\left[\gamma_n - \frac{1}{2}\right], \tag{6}$$

where N is the equivalent number of looks, $\mu = E[P_t]$ and $\Gamma[.]$ is the Gamma function. The previous parameters can be estimated as $N = \frac{\langle P_t\rangle^2}{(P_t - \langle P_t\rangle)^2}$ and $\mu = \langle P_t\rangle$ in the training window.

In order to have some insight on the dependency of the pdf with respect to its parameters, Figure 1 shows the analytic pdf varying $RedR$, N and P_t. The first plot is obtained fixing $N = 1$ and $RedR = 10^{-3}$. These are common values in real scenarios. The plots shows that if the power assigned to the target increases than γ_n has higher realizations. This means that the presence of targets or higher estimation errors produce higher values of the detector. The second test is obtained varying the value of $RedR$. This dependency leads to the idea that the $RedR$ could be adjusted depending on the value of P_t in the H_0 hypothesi in order to have a distribution with a mean value relatively small (e.g. 0.1), in order to facilitate the test when executed solving numerically the integrals. Theoretically, the selection of the threshold based on a statistical test is not influenced by the actual value or $RedR$, but since γ_n is contained between zero and one, it may be computationally disadvantageous



(a) pdf varying P_t

(b) pdf varying $RedR$

(c) pdf varying N

(d) test of consistency

Fig. 1. Plots of analytical pdf's. (a) pdf with $P_t = [10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$, $N = 1$ and $RedR = 10^{-3}$; (b) pdf with $P_t = 10^{-4}$, $N = 1$ and $RedR = [10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}]$; (c) pdf with $P_t = 10^{-4}$, $N = [1, 2, 3, 4, 5]$ and $RedR = 10^{-3}$; (d) test of consistency with green line: $RedR = 10$, $P_t = 1$; black cross: $RedR = 1$, $P_t = 0.1$; red line: $RedR = 100$, $P_t = 1$; yellow crosses: $RedR = 1$, $P_t = 0.01$; $N = 1$ for all the curves.

to have a threshold very close to 1, leading to higher quantization errors. The dependency on N shows that increasing N the variance of γ_n reduces as expected. Moreover, the mean appears to be unchanged.

In order to test the consistency of the pdf mathematical expression, P_t and $RedR$ are varied in order to obtain the same mean value of γ_n. Four curves are plotted using the couples of P_t and $RedR$ parameters $C_1 = [1, 10], C_2 = [1, 100], C_3 = [0.1, 1], C_4 = [0.01, 1]$. If the transformation is mathematically correct these set of values should couple in only two curves. The plots show that this is correct.

C. Likelihood ratio

The presence of some *a priori* information on the target of interest can improve the detection. Unfortunately, it is not easy to find an exact statistical distribution for vessels (hypothesis H_1). For this reason, only a very

general assumption is adopted here. This is that any vessel of interest has a power P_t that is higher than a minimum value, empirically found to be related to image artifacts: $P_t > P_t^{min}$. Clearly, this value is supposed to be dependent on the sensor and with a larger dataset it could be refined.

P_t^{min} corresponds to a minimum value for the GP-PNF: $\gamma_n^{min} = \left(1 + \frac{RedR}{P_t^{min}}\right)^{-0.5}$. The vessels pdf considers a uniform distribution between γ_n^{min} and 1: $rect\left[\frac{\gamma_n - (\gamma_n^{min}+1)/2}{1-\gamma_n^{min}}\right]$.

The likelihood ratio (LR) can be expressed as:

$$\Lambda = \frac{f_\Gamma(\gamma_n|H_1)}{f_\Gamma(\gamma_n|H_0)} = \tag{7}$$

$$= \frac{rect\left[\frac{\gamma_n - (\gamma_n^{min}+1)/2}{1-\gamma_n^{min}}\right]}{\frac{2}{\Gamma(N)}\left(\frac{N}{\mu}\right)^N \left(RedR\frac{\gamma_n^2}{1-\gamma_n^2}\right)^{N+1} \frac{2}{RedR}\,\gamma_n^{-3} exp\left[-RedR\frac{N}{\mu}\frac{\gamma_n^2}{1-\gamma_n^2}\right]\, rect\left[\gamma_n - \frac{1}{2}\right]}. \tag{8}$$

With such formulation, any value lower than γ_n^{min} will not provide any contribution to the Neyman-Pearson test (since it is multiplied by zero). On the other hand, the inverse of the clutter pdf will keep the Λ low when the probability of having clutter is high.

IV. Statistical tests

In this section two Neyman-Pearson tests are devised based on the expression of the clutter pdf and the LR.

A. Constant False Alarm Rate, CFAR

This test sets the threshold based on the clutter pdf in order to keep P_f constant. Probabilities can be calculated as integrals of pdf's therefore, in the hypothesis H_0, P_f can be calculated as:

$$P_f = \int_{T_n}^{1} f_\Gamma(\gamma_n|H_0)d\gamma_n, \tag{9}$$

where T_n is the threshold. Unfortunately, it was not possible to find an analytical solution for P_f, therefore the integrals are performed numerically exploiting cumulative sums.

B. Neyman-Pearson for likelihood ratio

This test sets the threshold fixing a size of the test α based on probabilities of LR. This is done inverting the integral:

$$\alpha = P\left(\Lambda_\Gamma \geq T_n|H_1\right) = \int_{T_n}^{1} \Lambda_\Gamma(\gamma_n|H_0)d\gamma_n. \tag{10}$$

The higher is the probability, the more the test increases the P_d.

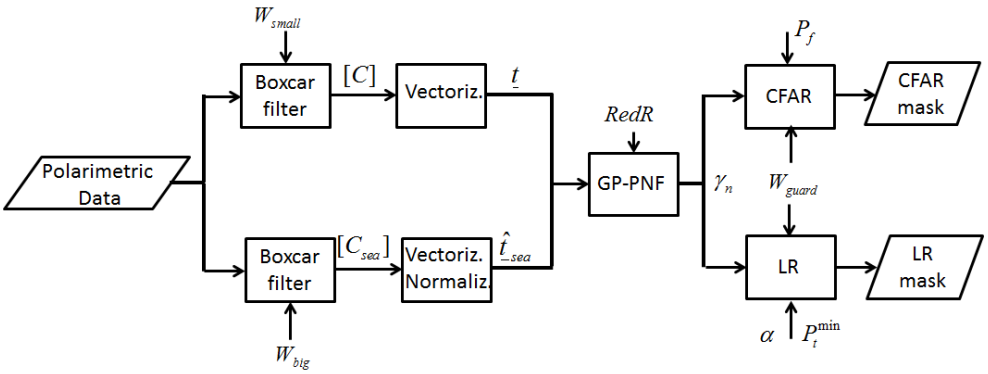


Fig. 2. Flow chart of the GP-PNF detector.

C. Local estimation with guard windows

In order to deal with a non-stationary sea clutter, the distribution parameters are estimated locally. In this context, bright vessels may bias the estimator providing thresholds much higher than necessary. Therefore, guard windows are exploited: the statistics are extracted in rings around a guard area (where pixels are rejected). The test area is inside the guard area. More details on the use of guard windows can be found in the literature [1, 15, 42].

To conclude this theoretical section, Figure 2 presents the flow chart of the GP-PNF including the statistical test.

V. Analysis of simulated data

Before to start testing the distributions on real data, it is interesting to understand if they are valid under ideal conditions. In order to achieve this, Monte Carlo simulations are performed.

A. Monte Carlo simulation

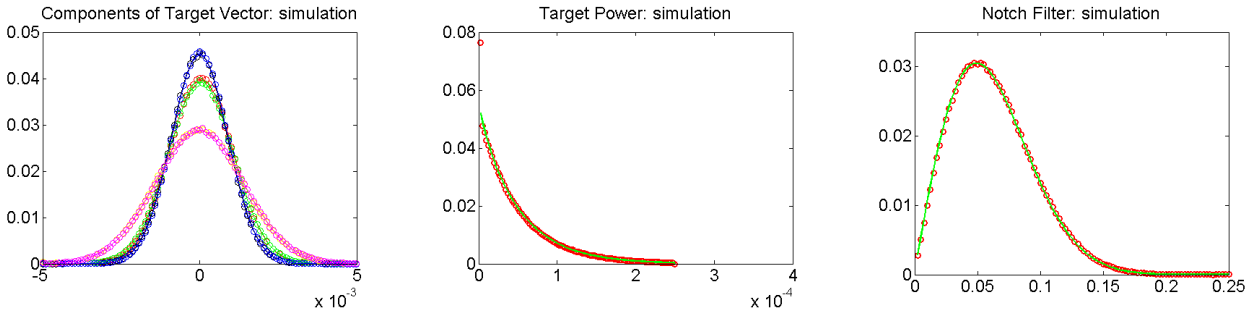
The simulations model the scattering vectors as 3 dimensional vectors with zero mean complex Gaussian components. Additionally, the scattering vector is colored with the polarimetric signature expected from the sea. This is obtained following the procedure described in [33, 38]: $\underline{k}_{sea} = [C]^{\frac{1}{2}}_{sea}\underline{k}$, where $[C]_{sea}$ is a covariance matrix of the sea and $\underline{k}$ is a standard complex Gaussian 3D vector. The polarimetric signature of the sea $[C]_{sea}$ is extracted from one the RADARSAT2 datasets introduced in the next section.

The homogeneity of the clutter is assured by the use of the same covariance matrix for each of the realizations. One dataset of 1000x1000 pixels is generated.

B. Test of fit for pdf

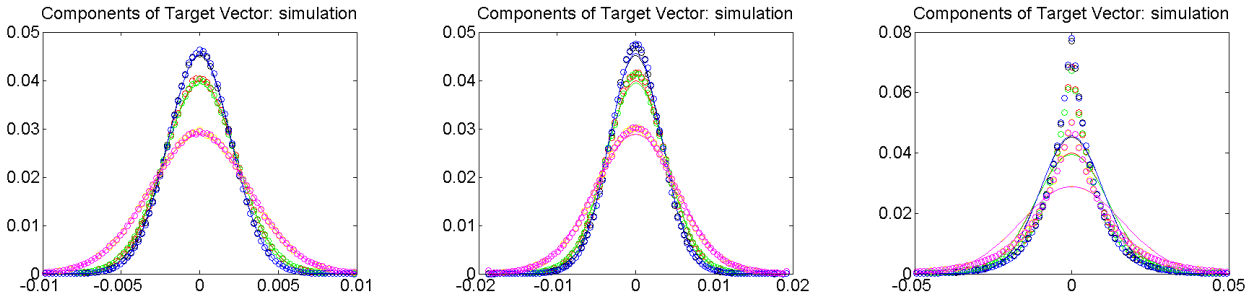
In this section, the derived distributions are fitted to the simulated data. The GP-PNF is executed employing a small window of 11x11 pixels (for estimating $\underline{t}$) and a large window of 51x51 pixels (for estimating $\underline{t}_{sea}$). Figure 3.a shows the data histograms of real and imaginary parts of the first three components of the target vector $\underline{t}_t$ (circles) and the fit with Gaussian distributions (solid lines). The color coding of the plot is the following: Red is $Re\{t_{t1}\}$, Green is $Im\{t_{t1}\}$, Black is $Re\{t_{t2}\}$, Blue is $Im\{t_{t2}\}$, Yellow is $Re\{t_{t3}\}$ and Magenta is $Im\{t_{t3}\}$, where the vector in the target space is defined as $\underline{t}_t = [t_{t1}, t_{t2}, t_{t3}, t_{t4}, t_{t5}, t_{t6}]^T$, Re stands for real part and Im for imaginary part. The basis used to represent $\underline{t}_t$ is selected pseudo-randomly, therefore the three variances of the components are not identical. On the other hand, the real and imaginary part for each component overlap almost completely (please note, when a color is not visible is because it overlaps with the corresponding real or imaginary part). It is possible to observe that the fitting with a Gaussian pdf is excellent. Figure 3.b presents the data histograms of the target power P_t (circles) and the fit with a Γ distribution (solid line). Again the fitting is good (except for the very first histogram bin). The final test is with γ_n. Again the fit seems excellent. It appears that all the exploited pdf are able to capture properly the data distribution for an ideal homogeneous sea clutter. In order to have a more quantitative analysis, two methodologies commonly used for testing the goodness-of-fit are exploited. These are the two sample Kolmogorov-Smirnov and the χ^2 tests. All the pdf's pass both the tests exploiting 100 test samples. It is interesting to evaluate which is the minimum number of samples that has to be used to obtain Gaussian distributions. Figure 4 shows the components of $\underline{t}_t$ when 5x5, 3x3 and 1x1 windows are used for the average. It is possible to observe that the 5x5 window provides an excellent estimation. The plots for P_t and γ_n are not provided for the sake of brevity, but they show good agreement. The estimation with 3x3 is not perfect and also P_t and γ_n show some divergence from the expected pdf (i.e. the estimated pdf has a larger variance than the date histogram). Nevertheless, such small windows may still be used in practical exercises provided that it is clear that the test is not a rigorous CFAR. For this reason, we would recommend to use the test with at least nine equivalent looks. Absence of average provides a rather poor result and the pdf has a much larger variance compared to the data histogram. The authors would therefore discourage from applying the test without any average for γ_n.

Finally, the CFAR test on $f_\Gamma(\gamma_n|H_0)$ and the NP test on the LR are applied to evaluate the detection capabil-



(a) $\underline{t}_t$ components (b) P_t (c) γ_n

Fig. 3. Fitting of derived pdf with Monte Carlo simulations. Circles: data histogram; Solid lines: fitted pdf's. (a) Real and Imaginary parts of $\underline{t}_t$ components and Gaussian; Red: $Re\{t_{t1}\}$; Green: $Im\{t_{t1}\}$; Black: $Re\{t_{t2}\}$; Blue: $Im\{t_{t2}\}$; Yellow: $Re\{t_{t3}\}$; Magenta: $Im\{t_{t3}\}$; The other components are omitted. (b) P_t and Γ distribution; (c) GP-PNF and derived pdf. 1000x1000 pixels.



(a) 5x5 (b) 3x3 (c) No average

Fig. 4. Fitting of derived pdf with Monte Carlo simulations: Real and Imaginary parts of $\underline{t}_t$ components and Gaussian. (a) 5x5 pixels average for test area; (b) 3x3 pixels average for test area; (c) No average for test area. 1000x1000 pixels.

ities. No detections could be identified in the entire simulated scene. The detection masks are not shown since they are black everywhere.

Considering that the fit of the distribution appears to be good in the ideal case, the more interesting and challenging scenario of real data is investigated in the following sections.

VI. Real data analysis

A. Presentation of the data

In order to test the fitting on real data several datasets, quad and dual polarimetric are employed. This allows the investigation of diversity in frequency and resolution. During the data acquisitions some validation data were acquired collecting the Automatic Identification System (AIS) positions of vessels, that were used to identify eventual false alarms and missing detections.

TABLE I

DETAILS ON FINE QUAD-POL RADARSAT-2 DATA. TIME IS IN UTC.

Date	Location	Beam	Incidence angle	Ground range res.	Wind speed	Ships with AIS
29/11/2013 (17:30)	North Sea	FQ12	$\sim 32^0$	$10.0\ m$ to $9.5\ m$	32 knots (NW)	11
09/02/2014 (17:30)	North Sea	FQ15	$\sim 35^0$	$9.2\ m$ to $8.8\ m$	35 knots (SW)	20

A.1 RADARSAT-2

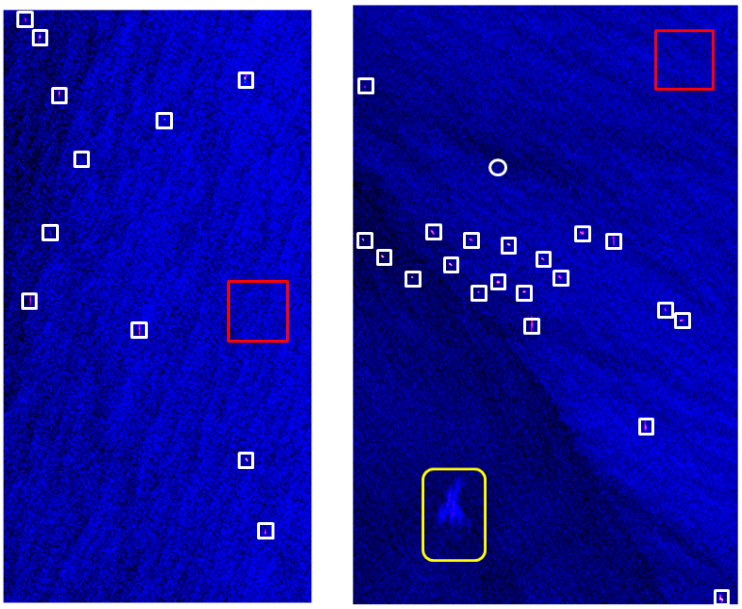
Two Fine Quad-polarimetric images were acquired during winter 2013/2014 in the North Sea. The data were collected under the SOAR project EI-5145. The central frequency is C-band ($5.4\ GHz$), while the chirp bandwidth is $30\ GHz$. The scenes are in Single Look Complex (SLC) format, covering approximately $25x25\ km$, with a slant range resolution of $5.2\ m$ and an azimuth resolution of $7.6\ m$. The image Noise Equivalent Sigma Zero (NESZ) is around $-36\ dB$. More details about the acquisitions are provided in Table I.

In total, 31 validated ships were observed with a variety of dimensions (ranging between 30 m to 200 m in length) and typology (e.g. fishing boats, cargos, etc).

In Figure 5, Pauli RGB color coding images of the two acquisitions are presented. The red is the intensity of $HH - VV$, the blue is $HH + VV$ and the green is HV. In the images, some ships can be identified as bright points, while others need some image zoom to be visible. The harsh weather conditions captured by the data show a strong sea clutter where several features can be observed. In the images the white rectangles represent validated vessels, while the white circle is an azimuth ambiguity. The large red rectangles are the areas used for testing the pdf's fit. In the lower left corner of the 09/02/2014 acquisition a large feature or image artifact of unknown origin can be observed (yellow box). The scene was specially selected to observe the capability of the local estimator to remove such sea clutter anomaly.

A.2 TanDEM-X

In order to test the fit of the pdf's in X-band and with dual polarimetric data, TanDEM-X images were acquired during winter 2012/2013. Two locations in the North Sea close to Aberdeen (Scotland) and Boknafjorden (Norway) were selected. For all the images, the azimuth resolution is $6.7\ m$, while the slat range resolution is $1.1\ m$ (i.e. the chirp bandwidth is $\sim 150\ MHz$). The swath width is $15\ km$ and the length of the



(a) 29/11/2013 (b) 09/02/2014

Fig. 5. RGB Pauli color composite images. RADARSAT-2, North Sea, Rotterdam. Red: $HH - VV$; Green: HV; Blue: $HH + VV$. The two acquisitions shown here correspond to 29/11/2013 and 09/02/2014. The images are visualized in dB, all the colors have the same normalization that is aimed at enhancing the contrast for sea features. The images are in radar coordinate and they represent areas on the ground that are approximately $25 \times 25\ km$.

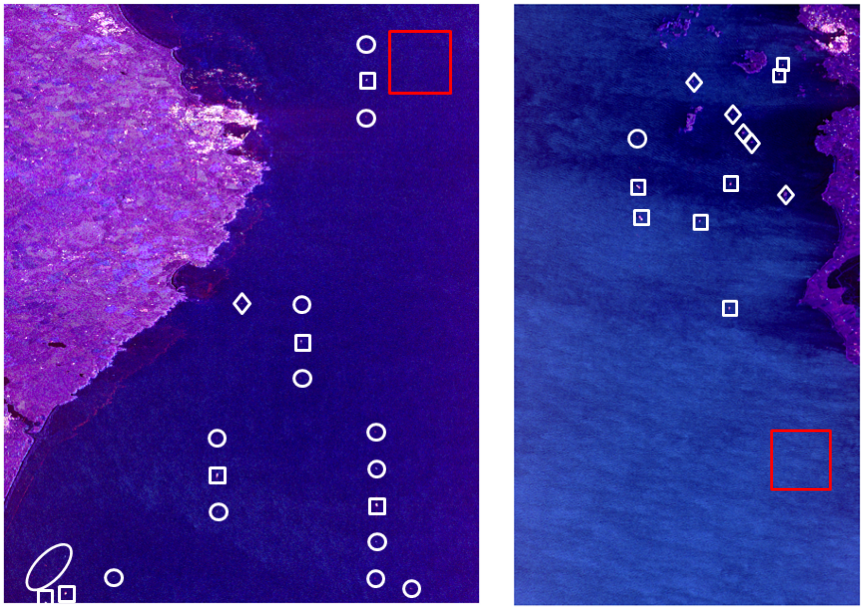
TABLE II

DETAILS ON HH/VV TANDEM-X IMAGES EXPLOITED IN THE COMPARISON. TIME IS IN UTC.

Date	Location	Beam	Incidence angle	Ground range res.	Wind speed	Ships with AIS
03/12/2012 (06:33)	Aberdeen	stripFar008	$\sim 33.5^0$	$2.1\ m$	13 knots (SE)	6+1 buoy
21/12/2012 (06:33)	Boknafjorden	stripNear008	$\sim 31.9^0$	$2.1\ m$	15 to 23 knots (SE)	7

strip is $50\ km$. The image Noise Equivalent Sigma Zero (NESZ) should be around $-21\ dB$. Two polarimetric channels HH/VV were considered. Table II summarizes further details regarding the acquisitions.

The analyzed images contain 13 validated vessels and one buoy. The RGB composite image is shown in Figure 6. Since quad-polarimetric data are not available a different color coding is exploited: red is the intensity of $HH - VV$, green is the magnitude of the correlation between HH and VV ($|\langle HH \cdot VV^* \rangle|$) and blue is the intensity of $HH + VV$. White rectangles and circles represent again vessels and azimuth ambiguities respectively. The white diamond in the Aberdeen scene is a buoy 2×2 meter in dimensions, while the diamonds in the Boknafjorden scene are small islands/rocks.



(a) Aberdeen (03/12/2012) (b) Boknafjorden (21/12/2012)

Fig. 6. RGB Pauli color composite images. TanDEM-X. Red: $HH - VV$; Green: $|HH \cdot VV^*|$; Blue: $HH + VV$. The two acquisitions shown here correspond to 03/12/2012 (Aberdeen), 21/12/2012 (Boknafjorden). The images are visualized in dB, all the colors have the same normalization that is aimed at enhancing the contrast for sea features. The images are in radar coordinate and they represent areas on the ground that are approximately $15x50\ km$.

A.3 AIS data

In the four scene exploited, 44 vessels had an operating AIS. Matching the AIS positions with points on radar images is not a trivial task. There are several factors that impede an easy matching. Initially, the ship GPS position is not updated continuously, but with intervals that can be up to 15min in the areas under analysis. In such time gap, the ship can travel significantly. Additionally, a ship moving along the range direction will be mislocated in the SAR image. In this analysis, several AIS positioning were recorded starting from 20min before and after the actual acquisitions and this time series of positions were used to take into account possible delays of the AIS in transmitting the vessel location.

In the Aberdeen scenes, the most of the vessels are multipurpose vessels, providing services to the oil rigs and their length ranges from $75\ m$ and $122\ m$. Interestingly, in such scene there is also a small buoy ($2 \times 2\ m$) close to the shoreline. In the Boknafjorden dataset, there is a variety of multipurpose and tankers. One of the vessels is a standby safety vessel of $47\ m$. Three vessels are between $50\ m$ and $100\ m$ long and two are around $265\ m$ long. In the RADARSAT-2 2013 dataset, the vessels are either oil tankers or general cargo. Two

TABLE III

DETECTION PARAMETERS. WINDOWS DIMENSIONS ARE GIVEN IN PIXELS.

Data	α	P_t^{min}	P_f	W_{small}	W_{big}	W_{guard}	W_{ring}
RADARSAT-2	0.9	$3e-4$	$1e-6$	11x11	51x51	51x51	71x71
TerraSAR-X	0.9	$3e-4$	$1e-6$	33x33	151x151	151x151	201x201

vessels have length less than $100\ m$, five are between $100\ m$ and $160\ m$ and the last is $230\ m$. Finally, in the RADARSAT-2 2014 scene, the most of the vessels are oil tankers or carriers. Two vessels are less than $100\ m$, six are between $100\ m$ and $150\ m$, ten are between $150\ m$ and $200\ m$ and the other three are more than $200\ m$.

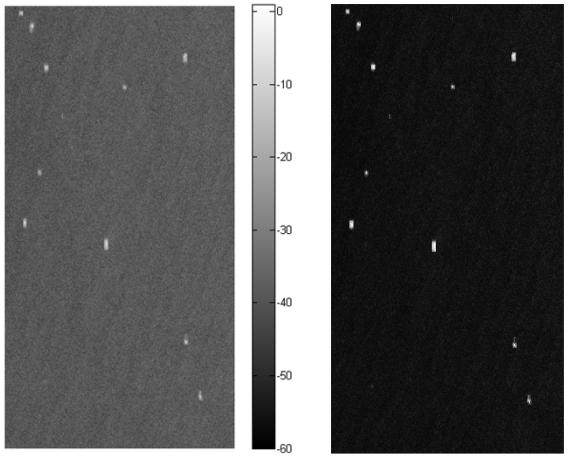
B. GP-PNF results

In order to gain a better understanding of the detector, P_t and γ_n are analyzed. Additionally, it is interesting to know if γ_n is able to provide a good visual contrast that may be beneficial in case the detection masks are visually inspected by an analyst. The parameters used in all the following analysis are listed in Table III. In the table the window sizes are given in pixels. W_{small} is the window used for generating $\underline{t}$, while W_{big} is used to calculate $\hat{\underline{t}}_{sea}$. W_{guard} represents the guard window and W_{ring} expresses how far the training ring area goes after the guard window.

B.1 RADARSAT-2

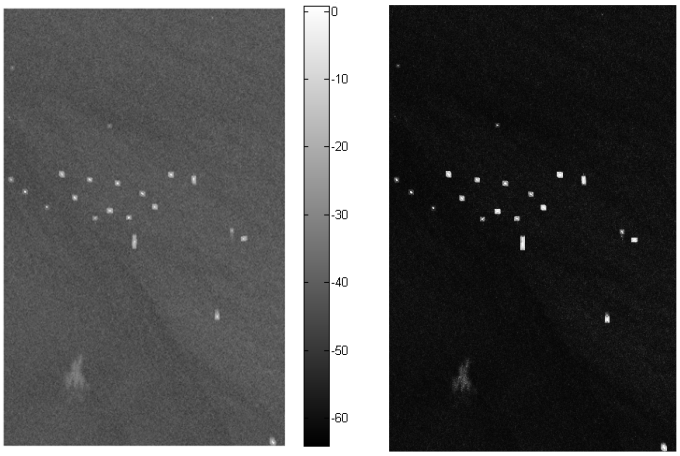
The images of P_t and γ_n are presented in Figure 7 and 8. The GP-PNF employs a test window of 11x11 pixels and a training window of 51x51 pixels [30]. The value of $RedR$ is equal to 0.1. This value is different from the one previously exploited and it was chosen merely because it provides a distribution of γ_n (for the sea clutter) around 0.1.

As expected, the output of P_t and γ_n are fairly similar, this is because the two images are completely correlated (i.e. they are linked by a deterministic transformation). All the validated vessels can be visually identified in the γ_n images and an eventual manual setting of the threshold seems to be relatively trivial. It can be observed that some areas on the sea can have values of P_t as small as $-60dB$. The fact that P_t is smaller than the noise floor is justified by the notch filtering procedure that only considers the component of the target



(a) P_t (b) γ_n

Fig. 7. P_t and γ_n images. RADARSAT-2, Rotterdam (21/12/2012). (a) Target Power, P_t; (b) GP-PNF γ_n scaled between 0 and 1.



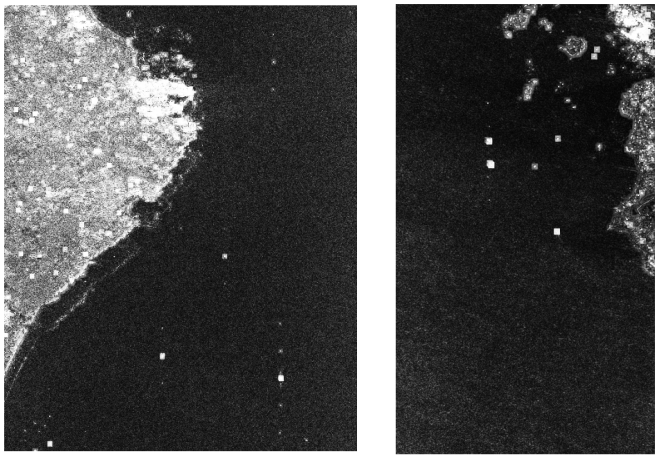
(a) P_t (b) γ_n

Fig. 8. P_t and γ_n images. RADARSAT-2, Rotterdam (09/02/2014). (a) Target Power, P_t; (b) GP-PNF γ_n scaled between 0 and 1.

orthogonal to the clutter background. It is also clear that P_t could be used alone for producing images with enhanced contrast between sea clutter and vessels.

C. TanDEM-X

The output of the GP-PNF on dual-polarimetric TanDEM-X data is presented in Figure 9. The images of P_t are omitted for the sake of brevity. Several vessels are visible in the Pauli RGB images. The resolution of TanDEM-X is higher allowing larger averaging. The GP-PNF exploits 33x33 pixels for the test and 151x151 pixels for the training. Using large windows without losing any vessels helps removing more speckle and it eliminates small heterogeneity and artifacts that may affect the results as false alarms. The $RedR$ used for



(a) Aberdeen (b) Boknafjorden

Fig. 9. γ_n images scaled between 0 and 1. TanDEM-X, (a) Aberdeen (03/12/2012); (b) Boknafjorden (21/12/2012).

such images is 0.001. This is because the clutter background for TanDEM-X is much lower than in the case of RADARSAT-2, due to the much higher average exploited. Again the selection of $RedR$ is merely related with setting the γ_n distribution around 0.1.

D. Test of fit for distributions

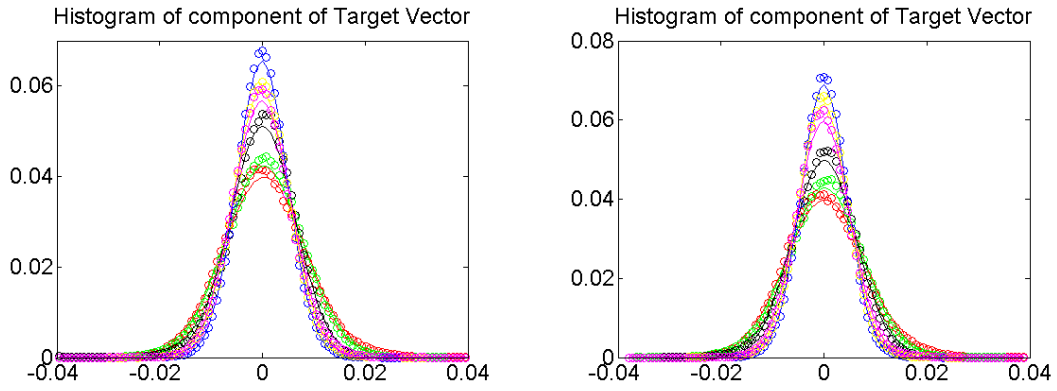
D.1 RADARSAT-2

The area used to derive the histograms and the distribution parameters are indicated by red rectangles in the Pauli RGB images. Such areas were selected since they appear relatively homogeneous, however some heterogeneity can still be observed.

The first test considers the components of the target vector $\underline{t}_t$. The plots of the real and imaginary parts of the first three components of $\underline{t}_t$ in a randomly generated basis are presented in Figure 10.

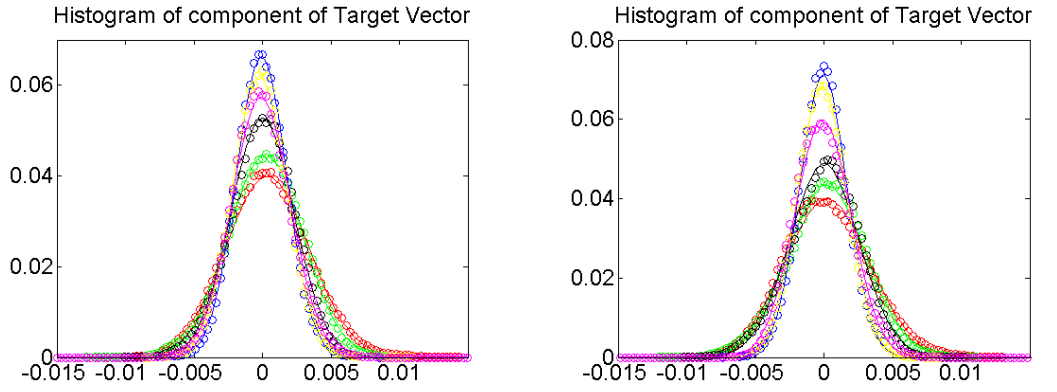
Again, circles represent data histograms and solid lines are the fitted Gaussian distributions. It can be observed that the zero mean Gaussian distribution fits adequately the histograms even though it is possible to observe that the theoretical pdf are slightly more disperse than the data. This may be due to the fact that the 11x11 pixels considered for average are not independent and therefore they correspond to a much smaller average. Also, data heterogeneity may impact the estimation of the distribution parameters.

In order to check that the number of pixels is the main cause of the difference between histograms and pdf a test is performed using 31x31 pixels for $\underline{t}$ and 151x151pixels for $\underline{t}_{sea}$. The results (Figure 11) show that the fit



(a) 29/11/2013 (b) 14/02/2014

Fig. 10. Fit of Gaussian distribution with real and imaginary parts of the first three components of $\underline{t}_t$. RADARSAT-2. Red: $Re\{t_{t1}\}$; Green: $Im\{t_{t1}\}$; Black: $Re\{t_{t2}\}$; Blue: $Im\{t_{t2}\}$; Yellow: $Re\{t_{t3}\}$; Magenta: $Im\{t_{t3}\}$. Boxcar used for test area: 11x11 pixels. (a) 29/11/2013; (b) 09/02/2014.



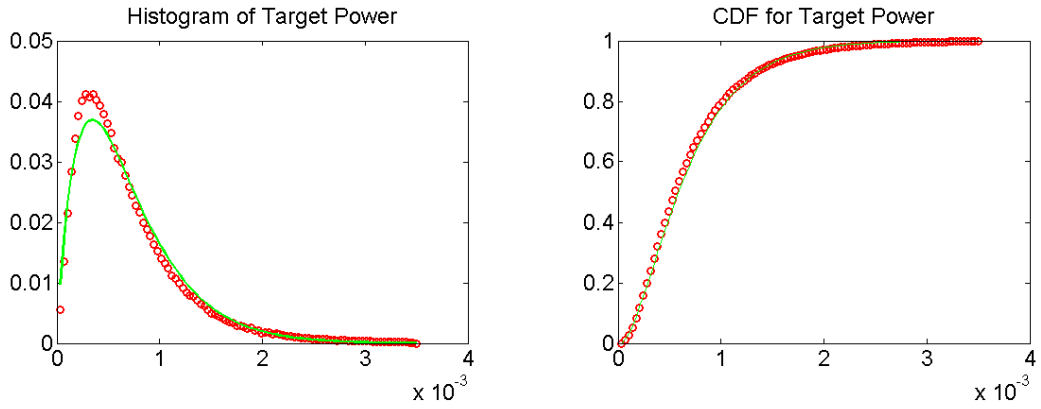
(a) 29/11/2013 (b) 14/02/2014

Fig. 11. Fit of Gaussian distribution with real and imaginary parts of the first three components of $\underline{t}_t$. RADARSAT-2. Red: $Re\{t_{t1}\}$; Green: $Im\{t_{t1}\}$; Black: $Re\{t_{t2}\}$; Blue: $Im\{t_{t2}\}$; Yellow: $Re\{t_{t3}\}$; Magenta: $Im\{t_{t3}\}$. Boxcar used for test area: 31x31 pixels. (a) 29/11/2013; (b) 09/02/2014.

improves, as for the simulated data.

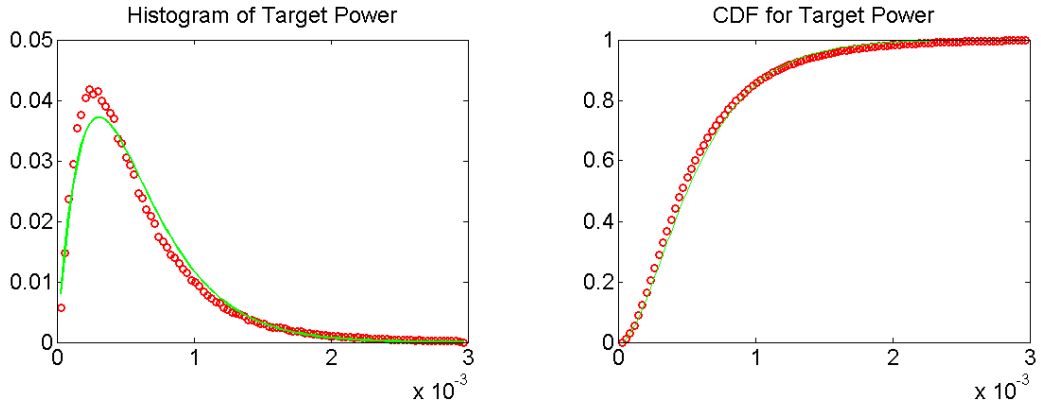
Figure 12 and 13 present the normalized histograms of P_t and the fitting with Γ distributions (exploiting 11x11 and 31x31 boxcar windows respectively).

The fit of the pdf is visually adequate, but not perfect. In particular, it appears that the peak of the distribution is slightly higher, which is indicative that the theoretical pdf's have a larger variance. This is in line with the previous analysis of the target components. Another possible reason for such mismatch is a wrong estimation of N. The latter is obtained assuming a homogeneous Gaussian scattering, therefore small heterogeneity in



(a) pdf (b) CDF

Fig. 12. Fit of Gamma distribution with P_t. RADARSAT-2, (29/11/2013). (a) Γ pdf; (b) Γ CDF.



(a) pdf (b) CDF

Fig. 13. Fit of Gamma distribution with P_t. RADARSAT-2, (09/02/2014). (a) Γ pdf; (b) Γ CDF.

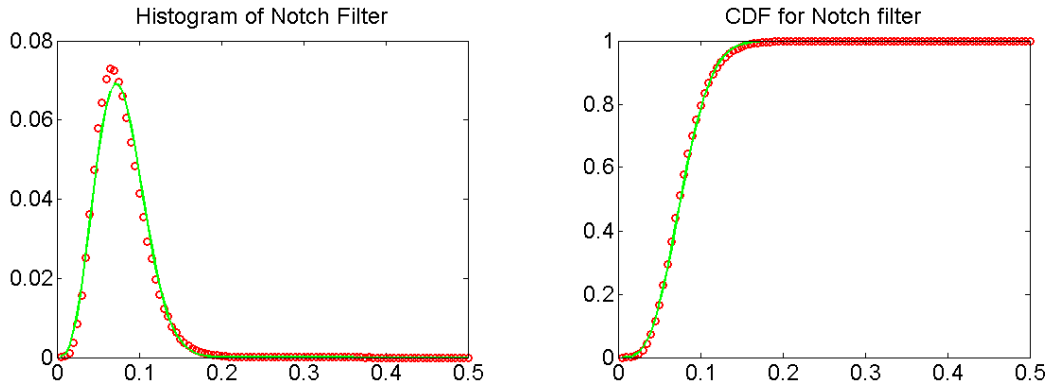
the data can introduce errors in estimating N which have impact on the pdf variance. To sum the integrals numerically, cumulative sums are exploited. For this reason it is interesting to understand the impact of such estimation errors on the Cumulative Density Function (CDF). These are plotted in Figure 12 and 13.

To extract some quantitative results, two goodness-of-fit tests, the Kolmogorov-Smirnov and the χ^2 test (with 50 samples and a size of 0.05) are exploited. The tests were passed for both the histograms.

Finally, the test of fit for the output of the Notch Filter is investigated. This is the most important analysis, since the final statistical test is set on γ_n. Figure 14 and 15 present the pdf and CDF of γ_n.

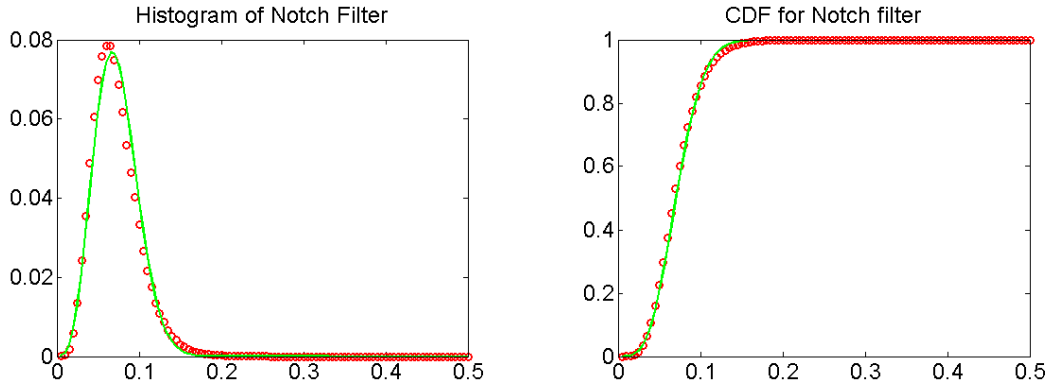
Interestingly, the distribution appears to have a good fit and the previous problem of larger variance of P_t seems not to affect significantly the distribution of γ_n. As for the previous case, the two goodness-of-fit tests were passed.

As a final remark, a larger variance of the estimated pdf will have the effect to have a CFAR test that is more



(a) pdf (b) CDF

Fig. 14. Fit of GP-PNF distribution with γ_n. RADRASAT-2, (29/11/2013). (a) derived pdf; (b) derived CDF.



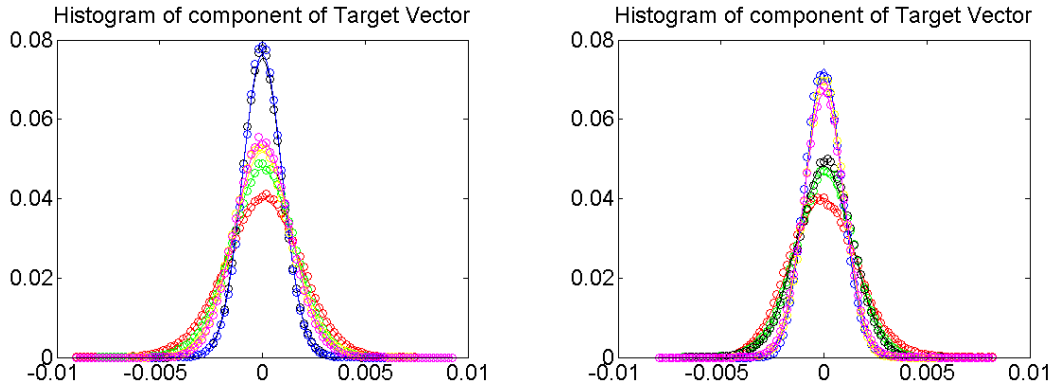
(a) pdf (b) CDF

Fig. 15. Fit of GP-PNF distribution with γ_n. RADARSAT-2, (09/02/2014). (a) derived pdf; (b) derived CDF.

conservative. This means that the threshold will be set higher rather than lower, which provides an actual P_f smaller than the selected one. P_f clearly decreases at the expenses of a decrease of P_d as well.

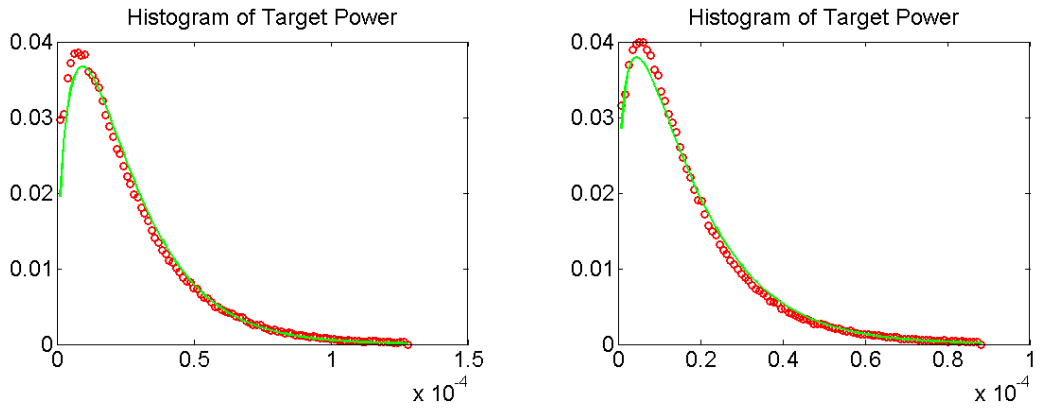
E. TanDEM-X

Figure 16, 17 and 18 present the fit of the data histograms with the derived pdf's. The CDF are omitted for the sake of brevity. The fits appear better than with RADARSAT-2 data. This is due to a higher number of samples used for the average (allowed by the higher resolution of TanDEM-X).



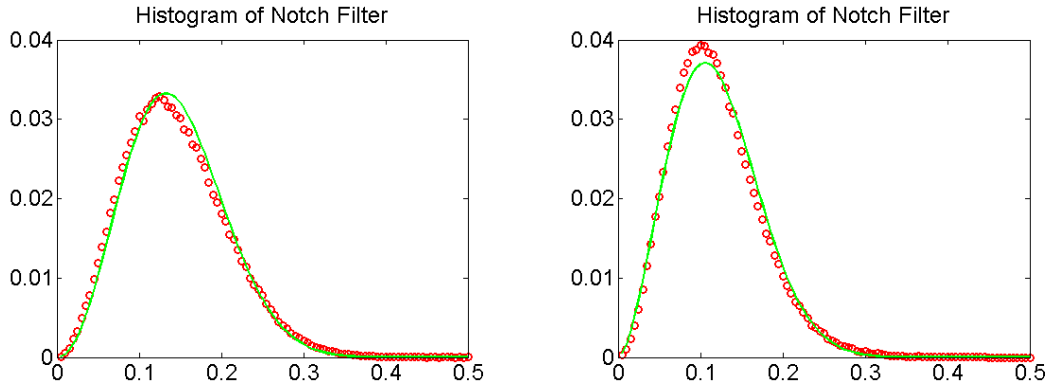
(a) Aberdeen (b) Boknafjorden

Fig. 16. Fit of Gauss distribution with real and imaginary parts of $\underline{t}_t$ components: TanDEM-X. (a) Aberdeen; (b) Boknafjorden.



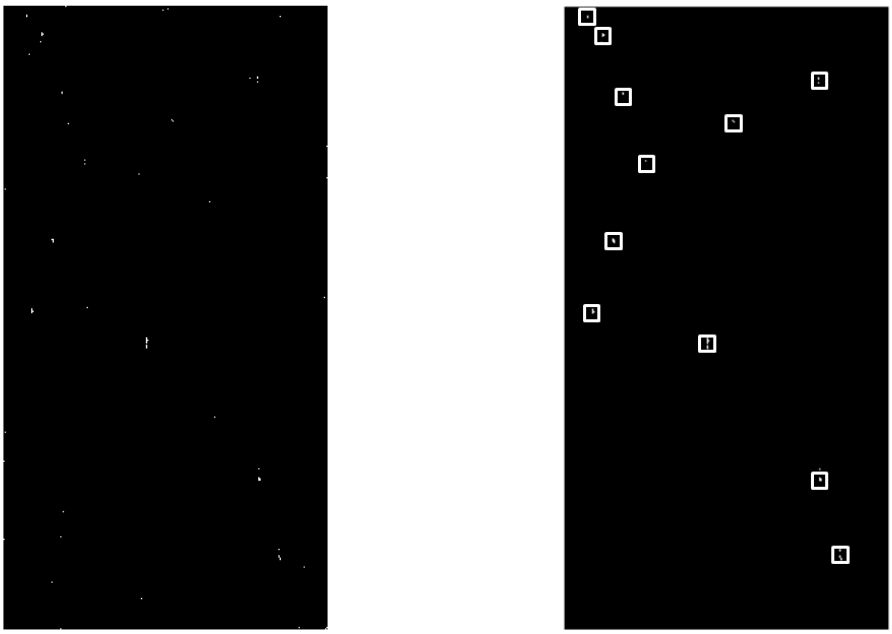
(a) Aberdeen (b) Boknafjorden

Fig. 17. Fit of Gamma distribution with P_t. TanDEM-X. (a) Aberdeen; (b) Boknafjorden.



(a) Aberdeen (b) Boknafjorden

Fig. 18. Fit of GP-PNF distribution with γ_n. TanDEM-X. (a) Aberdeen; (b) Boknafjorden.



(a) CFAR with $P_f = 10^{-6}$ (b) NP with $\alpha = 0.9$

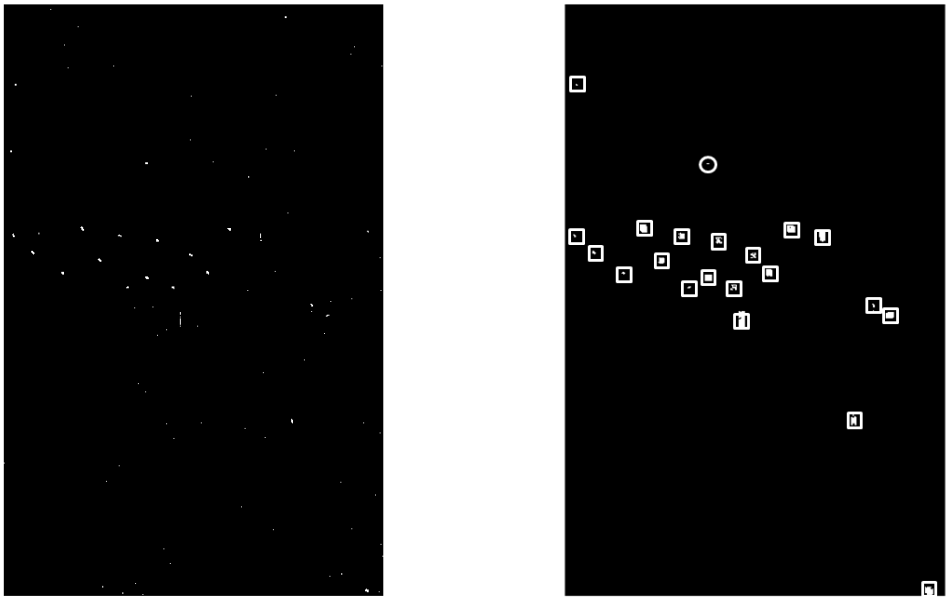
Fig. 19. Detection masks using local statistical tests on the GP-PNF. RADARSAT-2, (2013/11/29) (a) CFAR on pdf; (b) N-P on LR.

F. Detection results

F.1 RADARSAT-2

The CFAR is performed with $P_f = 10^{-6}$, while the LR exploits a size $\alpha = 0.9$ and $P_t^{min} = 3 \cdot 10^{-4}$. The result of the detection exploiting guard windows is presented in Figure 19 and 20. The guard area is 40 pixels wide (around the test area) and the training area is a ring 20 pixels large (around the guard area). This provides 2000 samples to estimate $\underline{t}_{sea}$.

It can be observed that the CFAR test presents some false alarms. The images are roughly composed of 3000x5000 pixels, which provide around 15 million pixels. Considering the setting $P_f = 10^{-6}$, around 15 false alarms are expected. They mostly come as single points and therefore they could be eliminated with a morphological filter, nevertheless it is valuable to also have a solution that does not rely on morphological filters. Moreover, testing the detector with lower values of P_f showed that few false alarms are not eliminated unless the value of P_f becomes unreasonably small (e.g. 10^{-20}). From this, it could be concluded that such points do not belong to the sea clutter distribution, but they are outliers, i.e. consequence of some small heterogeneity



(a) CFAR with $P_f = 10^{-6}$ (b) NP with $\alpha = 0.9$

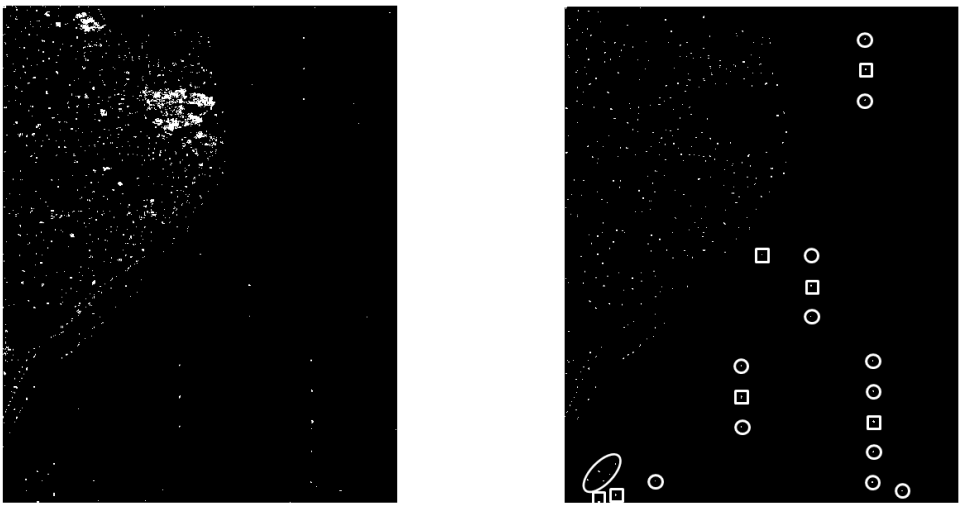
Fig. 20. Detection masks using local statistical tests on the GP-PNF. RADARSAT-2, (2014/02/09) (a) CFAR on pdf; (b) N-P on LR.

or image artifacts (that are always present in SAR images). Since they have a different distribution they are conceptually targets for the statistical tests. To reject them from the masks some a-priori information regarding "proper targets" is needed. This is the reason why a statistical test based on a likelihood ratio is exploited. The empirical value used here for P_t^{min} is $3 \cdot 10^{-4}$, since this was showing good rejection of artifacts. The result of the LR mask illustrates how all the false alarms are eliminated and none of the targets are lost (i.e. perfect detection performance with $P_f = 0$ and $P_d = 1$). It has also to be said that one azimuth ambiguity is detected. This means that another detection stage aimed at cleaning azimuth ambiguities has to be carried out [43].

G. TanDEM-X

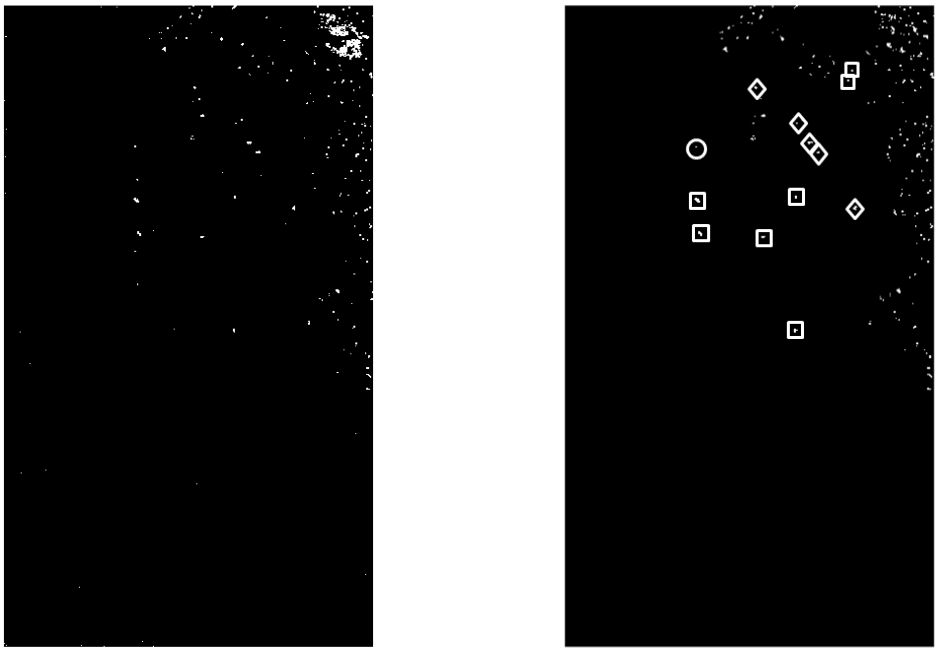
This section presents the results of the detection on TanDEM-X data (Figure 21 and 22). Considering the resolution is different, the guard window now is 120 pixels around the test area and the training area is 60 pixels around the guard area.

Again, it appears that the CFAR presents some false alarms, due to the large amount of pixels in the image and the presence of image artifacts. In order to remove them the LR test can be applied. The value of P_t^{min} is again chosen equal to $3 \cdot 10^{-4}$. Clearly, the value of P_t^{min} may depend on the sensor and the dimension of the



(a) CFAR with $P_f = 10^{-6}$ (b) NP with $\alpha = 0.9$

Fig. 21. Detection masks using local statistical tests on the GP-PNF. TanDEM-X Aberdeen. (a) CFAR on pdf; (b) N-P on LR.



(a) CFAR with $P_f = 10^{-6}$ (b) NP with $\alpha = 0.9$

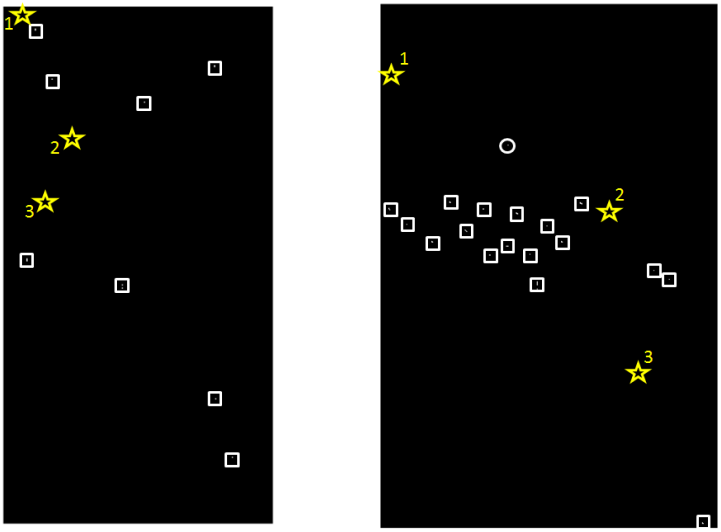
Fig. 22. Detection masks using local statistical tests on the GP-PNF. TanDEM-X Boknafjorden. (a) CFAR on pdf; (b) N-P on LR.

test window (since larger windows are able to average out more point-like anomalies), but here the same value is used for simplicity and the issue of optimizing it depending on the specific detection task is left as a future work. The detection result is again excellent, with $P_f = 0$ and $P_d = 1$. However, as mentioned previously, azimuth ambiguities are detected as for the previous case.

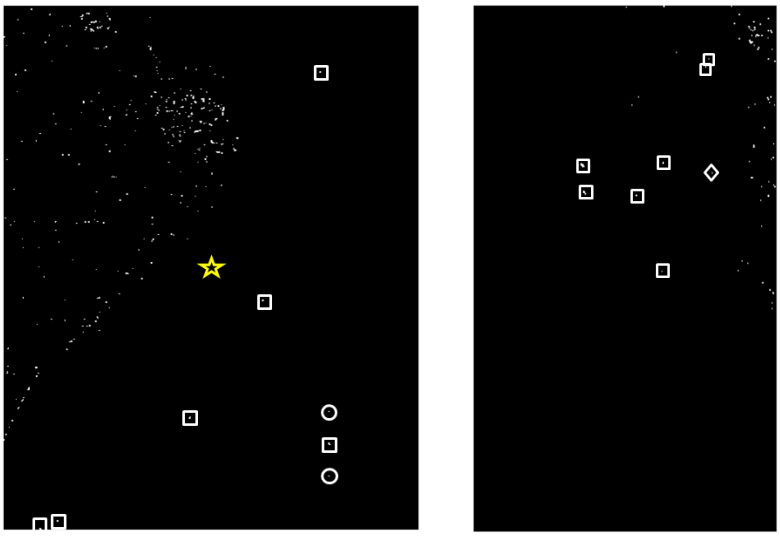
As a final remark, it is interesting to notice that γ_n^{min} resulting from the use of P_t is actually lower than the mean of the γ_n histogram in the red rectangle. This means that using γ_n^{min} as a brute threshold (without exploiting the LR test) would result in an enormous number of false alarms.

H. Comparison with single channel detectors

In this section, a comparison of the new statistical tests with a CFAR applied on single channel intensity images is presented. The distribution exploited for the analysis is the K-distribution, the probability of false alarms is set to $P_f = 10^{-4}$ and the integral are solved analytically. Such detectors were selected because the K-distribution was observed to model the sea clutter accurately [1, 44] and the numerical solution does not imply any assumption that may not be fulfilled in these specific datasets. The CFAR test was carried out on SLC intensity of each of the polarimetric channels separately. Image filtering was not changing dramatically the detection masks unless the average was more than 15x15 pixels (when the performance were getting lower). Figure 23 shows the detection masks for the four scenes choosing the polarization channel that gives the best detection (the other masks are omitted for the sake of brevity). Finally, all the detection results are summarized in Table IV. In Figure 23 the stars indicate vessels that were not detected (i.e. miss-detections). With TerraSAR-X data, the detection performance appear relatively similar with only one target missing in the HH CFAR mask. This is a small metallic buoy 2×2 m large. A reason for such similar performance is that the areas were the vessels are located in the TerraSAR-X scenes present a sea state that is not very high. On the other hand, on the RADARSAT-2 datasets, where the sea state was rougher, the performances are different. Here several miss-detection can be counted. In the scene from 2013, one missing vessel labelled as '2' is a 105 m long oil tanker travelling at about 20 km/h, while the vessels '1' and '3' were not providing any information regarding type and dimensions, but only location (we could guess that they are smaller boats). In the 2014 scene, the missing vessel labelled as '1' is a 44 m long trawler, '2' is a cargo 80 m long travelling at 10 km/h and '3' is an oil tanker 140 m long travelling at 21 km/h.



(a) HV RADARSAT-2 2013 (b) HV RADARSAT-2 2014



(c) HH TerraSAR-X Aberdeen (d) HH TerraSAR-X Boknafjorden

Fig. 23. Detection masks using a local CFAR tests on the intensity of single images. (a) RADARSAT-2, 2013, HV channel; (b) RADARSAT-2 2014, HV channel; (c) TerraSAR-X Aberdeen, HH channel; (d) TerraSAR-X Boknafjorden, HH channel.

VII. Discussions

Few conclusions can be drawn regarding the fit of the derived pdf's.

- Simulated data: The fit with Monte Carlo simulation is excellent, suggesting that the analytical derivation using the proposed assumptions is valid.
- Number of looks exploited: From simulations and tests on real data, the derived pdf's seem to not fit properly the data histograms when not enough samples are used. In particular, it appeared that the fitting was good with 25 independent samples and started to show some problem with less than nine independent samples. This

TABLE IV

DETECTION RESULTS (NUMBER OF DETECTIONS / TOTAL NUMBER).

Detector	Aberdeen	Boknafjorden	RS2 2013	RS2 2014
GP-PNF	$7/7$	$7/7$	$11/11$	$20/20$
CFAR HV	-	-	$8/11$	$17/20$
CFAR HH	$6/7$	$7/7$	$7/11$	$17/20$
CFAR VV	$6/7$	$7/7$	$5/11$	$16/20$

expected behavior has the consequence of broadening the estimated pdf and it is an indicator that the statistical tests cannot be used rigorously if the average is higher than nine equivalent looks. In a CFAR test, the effect of the broadened variance leads to a more conservative setting of the threshold that will be higher than necessary. This will return a P_f lower than the selected one (which is not problematic), but it may also lower P_d. On the other hand, it has to be said that smaller averages would facilitate the detection of weak targets, which eventually may compensate the lost due to an higher threshold. In order to be able to state precisely which the smallest window that should be used is, much more data has to be analyzed. Therefore, this is left as future work.

Some conclusions could also be drawn regarding the results provided by the detection masks.

- The CFAR test seems to show several false alarms in all the detection exercises (except the TanDEM-X Aberdeen test site where the wind conditions were lower). The false alarm rate is fixed to $P_f = 10^{-6}$. One explanation to such false alarms is the large amount of pixels in the image (tens of millions). However, some of the false alarms cannot be removed unless the value of P_f becomes extremely small. Therefore, we could conclude that such pixels do not belong to the same distribution of the clutter, but are outliers. A reason for such outliers can be image artifacts. To reject such pixels, it is possible to exploit some erosion morphological filter. However, in order to do not rely on morphological filters another solution is proposed, considering a Neyman-Pearson test on the likelihood ratio (LR).
- The LR test is based on the idea that a point to be called target should present a minimum power P_t^{min}. The selection of P_t^{min} depends on the likeliness of having artifacts in the image (closeness to a city), the window size (which helps averaging them out) and the sensor characteristics. The value used in this paper is the same

for both satellites and was empirically derived. It could be improved once much larger datasets are analyzed.

- The comparison with single channels detectors showed that the use of polarimetric data is especially beneficial when the detection is aimed at small or fast moving targets in high sea clutter. In situations where the sea clutter is low and the vessels do not move fast, probably a single channel detector may be sufficient.

VIII. Conclusions

In this work, two statistical tests for the ship detector based on the Geometrical Perturbation - Polarimetric Notch Filter (GP-PNF) were devised. The probability density function pdf of γ_n (i.e. the output of the GP-PNF detector) was derived. The pdf of P_t (i.e. the power of the target in the polarimetric subset perpendicular to the sea clutter) was analytically derived as a Γ distribution (provided that the averaging exploited contains more than nine independent pixels). Secondly, the pdf of γ_n can be derived with a transformation of random variables. Two Neyman-Pearson tests were proposed to set the threshold on γ_n following a constant false alarm rate (CFAR) or using a likelihood ratio (LR).

The pdf's and statistical tests were tested with Monte Carlo simulations and real data. Two RADARSAT-2 fine quad-polarimetric and two TanDEM-X dual-polarimetric HH/VV acquisitions were considered. The data presented fairly high values of sea clutter, which were beneficial to test the performances of the tests in more challenging scenarios.

The fit of the pdf's showed good visual results, with theoretical pdf's that were following narrowly the data histograms. In order to obtain some quantitative results, the two sample Kolmogorov-Smirnov and the χ^2 goodness-of-fit tests were executed and passed. The statistical tests were finally used to obtain detection masks. The CFAR test presents some false alarms, probably linked to the presence of small scale heterogeneity or artifacts. On the other hand, the LR test presents a perfect detection performance (if the azimuth ambiguities are not considered) with $P_f = 0$ and $P_d = 1$. Some azimuth ambiguities were detected since they are replicas of ship signatures and therefore they triggered detection. This means that some post- or pre-processing algorithms should be carried out to remove azimuth ambiguities.

IX. Appendix

The derivation of $f_{\Gamma_n}(\gamma_n)$ as a transformation of $f_{P_t}(p_t)$ is provided in this section. The theorem of transformation of random variables states that $f_{\Gamma_n}(\gamma_n) = f_{P_t}(\hat{p}_t)\frac{\partial p_t(\gamma_n)}{\partial \gamma_n}$, where $p_t(\gamma_n)$ is the solution of $\gamma_n(p_t)$

457 and ∂ denotes derivative.

458 The solution of $\gamma_n(p_t)$ is

$$\hat{p_t} = RedR\frac{\gamma^2}{1-\gamma^2}. \tag{11}$$

459 The pdf of P_t is a Γ and it can be written as

$$f_{P_t}(p_t) = \frac{1}{\Gamma(N)}\left(\frac{N}{\mu}\right)^N p_t^{N-1} exp\left[-N\frac{p_t}{\mu}\right]. \tag{12}$$

460 After substituting $\hat{p}_t$ the expression becomes

$$f_{P_t}(p_t) = \frac{1}{\Gamma(N)}\left(\frac{N}{\mu}\right)^N \left(RedR\frac{\gamma_n^2}{1-\gamma_n^2}\right)^{N-1} exp\left[-RedR\frac{N}{\mu}\frac{\gamma_n^2}{1-\gamma_n^2}\right]. \tag{13}$$

461 The derivative of the transformation $\gamma_n(p_t)$ is

$$\frac{\partial\gamma_n(p_t)}{\partial\gamma_n} = \frac{2RedR}{\gamma_n^3}\left(\frac{\gamma_n^2}{\gamma_n^2-1}\right)^2 \tag{14}$$

462 Multiplying together the last two expressions it is possible to obtain:

$$f_{\Gamma_n}(\gamma_n) = \frac{1}{\Gamma(N)}\left(\frac{N}{\mu}\right)^N \left(RedR\frac{\gamma_n^2}{1-\gamma_n^2}\right)^{N-1} exp\left[-RedR\frac{N}{\mu}\frac{\gamma_n^2}{1-\gamma_n^2}\right]\frac{2RedR}{\gamma_n^3}\left(\frac{\gamma_n^2}{\gamma_n^2-1}\right)^2 \tag{15}$$

463 Such expression can be simplified obtaining:

$$f_{\Gamma_n}(\gamma_n) = \frac{2}{\Gamma(N)}\left(\frac{N}{\mu}\right)^N \left(RedR\frac{\gamma_n^2}{1-\gamma_n^2}\right)^{N+1} \frac{2}{RedR}\,\gamma_n^{-3} exp\left[-RedR\frac{N}{\mu}\frac{\gamma_n^2}{1-\gamma_n^2}\right]\,rect\left[\gamma_n-\frac{1}{2}\right], \tag{16}$$

464 ACKNOWLEDGMENTS

465 RADARSAT-2 Data and Products © MacDonald, Dettwiler and Associates Ltd. (2013-2014) - All Rights
466 Reserved. RADARSAT is an official trademark of the Canadian Space Agency. the data were acquired under
467 the SOAR Education International Initiative (SOAR-EI-5145). TanDEM-X data were provided courtesy of
468 DLR in the framework of the proposal POLIN_1734.

469 REFERENCES

470 [1] D. J. Crisp, "The State-of-the-Art in ship detection in Synthetic Aperture Radar imagery," *Australian Government Department of*
471 *Defence*, 2004.

[2] K. Eldhuset, "An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 4, pp. 1010 – 1019, July 1996.

[3] C.C. Wackerman, K.S. Friedman, W.G. Pichel, P. Clemente-Colon, and X. Li, "Automatic detection of ships in RADARSAT-1 SAR imagery," *Canadian Journal of Remote Sensing*, vol. 27, 2001.

[4] M. Yeremy, G. Geling, M. Rey, B. Plache, and M. Henschel, "Results from the CRUSADE ship detection trial: polarimetric SAR," *Proceeding on IGARSS 2002, Toronto, Canada, 24-28 June*, 2002.

[5] G Ferrara, M. Migliaccio, F. Nunziata, and A. Sorrentino, "Generalized-K (GK)-Based observation of metallic objects at sea in full-resolution Synthetic Aperture Radar (SAR) data: A multipolarization study," *IEEE Journal of Oceanic Engineering*, vol. 36(2), pp. 195–204, 2011.

[6] S. N. Anfinsen and C Brekke, "Statistical models for constant false alarm rate ship detection with the sublook correlation magnitude," *IEEE Geoscience and Remote Sensing Letters*, pp. 5626–5629, 2012.

[7] M. Brizi, P. Lombardo, and D. Pastina, "Exploiting the shadow information to increase the target detection performance in SAR images," *International Conference on Radar Systems, RADAR 1999. Brest, Germany, April*, 1999.

[8] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, and B. van Schie, "Ship surveillance with TerraSAR-X," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 9, no. 3, pp. 1092–1103, March Mar. 2011.

[9] M.F. Fingas and C.E. Brown, "Review of ship detection from airborne platforms," *Canadian Journal of Remote Sensing*, vol. 27(4), pp. 379–385, 2001.

[10] K.S. Friedman, C. Wackerman, Funk, W.G. Pichel, P. Clemente-Colon, and X. Li, "Validation of a CFAR vessel detection algorithm using known vessel locations," *IEEE 2001 International Geoscience and Remote Sensing Symposium (IGARSS'01), 9-13 July*, vol. 4, pp. 1804–1806, 2001.

[11] J. Gower and S. Skey, "Evaluation of RADARSAT ScanSAR for observing wind, slicks and fish-boats," *Canadian Journal of Remote Sensing*, vol. 26, pp. 484–493, 2000.

[12] G. Margarit, J. A. Barba Milanés, and A. Tabasco, "Operational ship monitoring system based on synthetic aperture radar processing," *Remote Sensing*, vol. 1(3), pp. 375–392, 2009.

[13] N. Robertson, P. Bird, and C. Brownsword, "Ship surveillance using RADARSAT ScanSAR images," *In Alliance for Marine Remote Sensing (AMRS) Workshop on Ship Detection in Coastal Waters*, 2000.

[14] J.A. Johannessen, "Coastal observing systems: The role of synthetic aperture radar," *Johns Hopkins APL Technical Digest*, vol. 21(1), pp. 41–48, 2000.

[15] S. M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice Hall, Upper Saddle River, US, 1993.

[16] P. W Vachon, "Ship detection in synthetic aperture radar imagery," *Proceedings OceanSAR, St. John s, NL, Canada*, 2006.

[17] P. Iervolino, R. Guida, and P. Whittaker, "NovaSAR-S and maritime surveillance," *IEEE Geoscience and Remote Sensing Symposium, 21-26 July, Melbourne, Australia*, 2013.

[18] D.J. Crisp and T. Keevers, "Comparison of ship detectors for polarimetric sar imagery," *OCEANS 2010 IEEE - Sydney*, pp. 1–8, 2010.

[19] C. Liu, P. W. Vachon, and G. W. Geling, "Improved ship detection using polarimetric SAR data," *IGARSS Geoscience and Remote Sensing Symposium, Anchorage, USA*, vol. 3, pp. 1800–1803, September 2004.

[20] R. Ringrose and N. Harris, "Ship Detection Using Polarimetric SAR Data," *SAR Workshop: CEOS Committee on Earth Observation Satellites*, 2000.

[21] F. Nunziata, M. Migliaccio, and C.E. Brown, "Reflection symmetry for polarimetric observation of man-made metallic targets at sea," *IEEE Journal of Oceanic Engineering*, vol. 37, no. 3, pp. 384–394, July 2012.

[22] R. Shirvany, M. Chabert, and J.-Y. Tourneret, "Ship and oil-spill detection using the degree of polarization in linear and hybrid/compact dual-pol SAR," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2012.

[23] R. Touzi, "On the use of polarimetric SAR data for ship detection," *IGARSS Geoscience and Remote Sensing Symposium*, vol. 2, pp. 812–814, 1999.

[24] W.L. Cameron, N.N. Youssef, and L.K. Leung, "Simulated polarimetric signatures of primitive geometrical shapes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 3, pp. 793–803, May 1996.

[25] D. Velotto, F. Nunziata, M. Migliaccio, and S. Lehner, "Dual-polarimetric TerraSAR-X SAR data for target at sea observation," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 5, pp. 1114–1118, Sep. 2013.

[26] S. R. Cloude, *Polarisation: Applications in Remote Sensing*, Oxford University Press, Oxford, UK, 2009.

[27] C. R. Jackson and J. R. Apel, *Synthetic Aperture Radar Marine Users Manual*, U.S. Department of Commerce, National Oceanic and Atmospheric Administration (NOAA), 2004.

[28] W. Alpers, *Imaging ocean surface waves by synthetic aperture radar: a review*, Ellis Horwood Ltd., Cambridge, UK, 1983.

[29] A. Marino, M. Sugimoto, K. Ouchi, and I. Hajnsek, "Validating a notch filter for detection of targets at sea with ALOS-PALSAR data: Tokyo Bay," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 12, pp. 4907–4918, Dec. 2014.

[30] A. Marino, "A notch filter for ship detection with polarimetric SAR data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 3, pp. 1219 – 1232, June 2013.

[31] A. Marino and N. Walker, "Ship detection with quad polarimetric TerraSAR-X data: an adaptive notch filter," *Proc. on IGARSS11*, 2011.

[32] A. Marino, N. Walker, and I. H. Woodhouse, "Ship detection using SAR polarimetry. The development of a new algorithm designed to exploit new satellite SAR capabilities for maritime surveillance," *Proceedings on SEASAR, Frascati, Italy, January*, 2010.

[33] J. S. Lee and E. Pottier, *Polarimetric radar imaging: from basics to applications*, CRC Press, Taylor & Francis Group, 2009.

[34] A. Marino, *A New Target Detector Based on Geometrical Perturbation Filters for Polarimetric Synthetic Aperture Radar (POL-SAR)*, Springer-Verlag, 2012.

[35] A. Marino, S. R. Cloude, and I. H. Woodhouse, "Detecting depolarized targets using a new geometrical perturbation filter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3787–3799, October 2012.

[36] A. Marino, S. R. Cloude, and I. H. Woodhouse, "A polarimetric target detector using the Huynen Fork," *IEEE Transaction on Geocience and Remote Sensing*, vol. 48, no. 5, pp. 2357–2366, May 2010.

[37] A. Marino, S. R. Cloude, and J. M. Lopez-Sanchez, "A new polarimetric change detector in radar imagery," *IEEE Transactions on Gescience and Remote Sensing*, vol. 51, no. 5, pp. 2986 – 3000, 2013.

[38] J.S. Lee, M.R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution," *International Journal of Remote Sensing*, vol. 15(11), pp. 2299–2311, 1994.

[39] J Chen, Y. Chen, W. An, Yi. Cui, and J. Yang, "Nonlocal filtering for polarimetric SAR data: A pretest approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49(5), 2011.

[40] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, Mc Grow Hill, 2002.

[41] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26(6), pp. 764–773, 1988.

[42] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 19(4), pp. 608–621, 1983.

[43] A. M. Guarnieri, "Adaptive removal of azimuth ambiguities in SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 625–633, March 2005.

[44] R. L. Paes, J. A. Lorenzzetti, and D. F. M. Gherardi, "Ship detection using Cosmo-Skymed SAR images on Rio de Janeiro area," *SERFA08*, 2008.