Multimodel CFAR Detection in Foliage Penetrating SAR Images

ALESSIO IZZO
University of Strathclyde, Glasgow, U.K.

MARCO LIGUORI
Università degli Studi del Sannio, Benevento, Italy

CARMINE CLEMENTE, Member, IEEE
University of Strathclyde, Glasgow, U.K.

CARMELA GALDI, Member, IEEE
MAURIZIO DI BISCEGLIE
Università degli Studi del Sannio, Benevento, Italy

JOHN J. SORAGHAN, Senior Member, IEEE
University of Strathclyde, Glasgow, U.K.

A multimodel approach for constant false alarm ratio (CFAR) detection of vehicles through foliage in foliage penetrating synthetic aperture radar images is presented. Extreme value distributions and location scale properties are exploited to derive an adaptive CFAR approach that is able to cope with different forest densities. Performance analysis on real data is carried out to estimate the detection and false alarm probabilities in the presence of a ground truth.

I. INTRODUCTION

The fundamental feature of foliage penetrating (FOPEN) radars is the capability to collect returns from scatterers under foliage [1]. This is achieved by using fairly low carrier frequencies (in the ultrahigh-frequency (UHF) and very-high-frequency (VHF) bands) that are able to penetrate the vegetating layer. The ability to "see" through foliage canopies makes FOPEN radar a powerful tool for military purposes and, more interestingly, if synthetic aperture radar (SAR) techniques are used, a FOPEN SAR sensor becomes able to detect, track, and recognize vehicles hiding in forests [1].

Due to the nature of the imaged scene, several issues are still under investigation for the complete and reliable exploitation of such sensors. In particular, canopies and hidden vehicles are not the only possible reflecting targets in a forest scene; trunks are present and contribute significantly to the intensity of the signal returned to the radar. Detections will be affected from reflections from trunks if an accurate strategy of control of false alarms is not provided. Solutions or partial solutions to this problem have been provided in the literature. Clutter modeling has been identified as a viable solution to mitigate tree and trunks detections, physical, statistical, and the combination of the two approaches were used to model forest clutter in FOPEN SAR [2]–[6]. The models proposed in [2]–[4] consider electromagnetic modeling of forests to extract deterministic clutter models. These models are useful in terms of understanding of the scattering physics but are not applicable in statistical detection frameworks. The model presented in [5], introduces statistical properties in the model in [3], however, this model is not robust with respect to the presence of tree trunks dominating a scene. In [6], a model for VHF clutter generation was proposed, integrating both background scatterers and large-amplitude discrete clutter (trees). Despite its flexibility, the model proposed in [6] does not have a closed-form solution, and therefore, it is not suitable to derive detectors with false alarm rate control. Different detection approaches have been evaluated in the literature; in [7], a two parameters constant false alarm ratio (CFAR), a change detection and multiaperture detection approaches have been evaluated, with the conclusion that a standard CFAR approach does not perform as well as the other two approaches. In [8], a rank-order filters-based approach was used to address the poor performance of a traditional CFAR detector, finally, the case of slow moving target detection was investigated in [9], where the authors proposed an alternative approach based on along track interferometry.

The main problem in VHF SAR clutter is that the data samples are highly inhomogeneous and variable. In contrast to corner reflector effects at microwaves frequencies, targets do not exhibit the high peak response in FOPEN SAR images.

Due to the lack of accurate analytic models for FOPEN SAR clutter, a robust algorithm is required against large variability of amplitude and distribution model of the received data.

Starting from a statistical modeling of the clutter, in this paper, we introduce a novel framework based on multi-model CFAR detection in adverse clutter conditions exhibited from FOPEN SAR images.

The multimodel approach derives by a generalization of the single model. In selecting a specific model for target detection, the designer is assuming a risk; the risk of model mismatch is mitigated thanks to the flexibility of our approach that can be used for any type of location-scale (LS) distributions. The derisk is obtained through the exploitation of a goodness of fit-based clutter model selection. [10].

In our approach, the clutter is statistically modeled by light and heavy-tailed distributions that belongs (directly or after transformations) to the LS family. The light and heavy-tailed distributions are consistent with the nature of the scattering from different forest scenarios [1], [6], whereas the LS family is a requirement of the CFAR detection algorithm [11] that is exploited in this paper.

The CFAR detection algorithm introduced in [12] and [13] and applied to high-resolution SAR images in [11] can be embedded in a framework that is able to select the best LS distribution to be used to compute the adaptive threshold and that will ensure the constant false alarm rate in the highly inhomogeneous FOPEN SAR image environment. Generally, it is not easy to prove that a goodness-of-fit-test-based CFAR detector has the CFAR property, but when the data distribution has some special properties, we can prove that the resultant detector is CFAR. One of these cases is when the assumed distribution is a member of the LS family.

This paper is organized as follows. Section I introduces the multimodel CFAR detection framework, addressing the specific cases of Gumbel for maximum and Weibull distributed background. The performance in terms of distribution fitting of the two aforementioned models are assessed and discussed in Section II using real VHF FOPEN SAR data. Section III discusses the algorithm performance on real data, demonstrating the capability of the proposed approach to control the false alarm probability and to detect extended targets hiding in foliage.

Notation: Throughout this paper, we use (lower case) boldface for vectors **x**, and (upper case) for matrices **X**. The transpose and inverse matrix operators are denoted by the symbol $(\cdot)^T$ and $(\cdot)^{-1}$, respectively. After sorting and censoring operations, the vector $\mathbf{x}_i$ is denoted as $\tilde{\mathbf{x}}_i$; instead $\tilde{\mathbf{x}}_i'$ represent the vector $\tilde{\mathbf{x}}_i$ after the transformation function $f_i(\cdot)$ related by ith distribution.

II. MULTIMODEL CFAR DETECTOR IN LS ENVIRONMENT

The architecture of the proposed algorithm is depicted in Fig. 1.

The algorithm has been designed in order to ensure higher robustness and reliability of the results, with respect to the single-model approach introduced in [11], by considering K possible statistical distributions for the background. The multimodel approach stems as a generalization of the

single-model approach where, by selecting a specific distribution model, it is assumed a risk of model mismatch. This risk is mitigated thanks to the flexibility of the approach that is applicable to any family of LS distributions.[1]

In FOPEN SAR, the multimodel detector automatically searches for the distribution that best fits the real data in a specific reference window, introducing robustness with respect to nonaccurate *a priori* knowledge of trees density in a given cell.

The problem can be posed in terms of the binary hypothesis test as

$$
\begin{cases}
H_{B|D_i} : & X(m, n) \sim F_i(x), \quad \text{Target absent} \\
H_{T|D_i} : & X(m, n) \sim F_{T_i}(x), \quad \text{Target present}
\end{cases}
\quad (1)
$$

where (m, n) represents the cell under test. In the $H_{B|D_i}$ hypothesis the probability distribution of the background sample is one of a selected set $\{F_i(x)\}_{i=1...K}$ of distributions with cardinality K. The probability distribution is selected among the K distributions of LS type using a minimum distance nonparametric test. In the hypothesis $H_{T|D_i}$, the presence of a target in the background is modeled by the distribution $F_{T_i}(x)$, that is unknown. Afterwards, a CFAR procedure is applied for the selected background hypothesis.

With reference to Fig. 1, it is assumed that the following statements are applicable.

1) The data window for the estimation of the background statistics is composed of $N \times N$ intensity samples obtained from the complex SAR image.
2) The statistical characterization of the background is chosen according to a best-fit strategy between the empirical cumulative distribution function (CDF) of the data and the CDFs from a set of model distributions. The fitting strategy is based on the Lilliefors test, a modified Kolmogorov–Smirnov test accounting for the case of unknown distributional parameters [14].

The usual outcome of the Lilliefors test, with a given significance level α, is as follows.
a) H_0: The selected distribution is compatible with data, i.e., the p-value is greater than α.
b) H_1: The null hypothesis is rejected, i.e., the p-value is less than α.

However, we are here interested in fitting the best distribution model rather than determining if the H_0 hypothesis holds given a significance level α. Therefore, we use the p-value to determine the *a posteriori*

---

[1] A random variable X is of the LS type, with *location* parameter $\theta_L \in \mathbb{R}$ and *scale* parameter $\theta_S > 0$, if any variate of the family can be obtained by an affine transformation of the standardized variate X_0, namely

$$
X = \theta_S X_0 + \theta_L
$$

where it is clearly understood that the standardized variate corresponds to $\theta_S = 1$ and $\theta_L = 0$. As a consequence, the CDF has the property

$$
F_X(x; \theta_L, \theta_S) = F_{X_0} \left(\frac{x - \theta_L}{\theta_S} \right)
$$

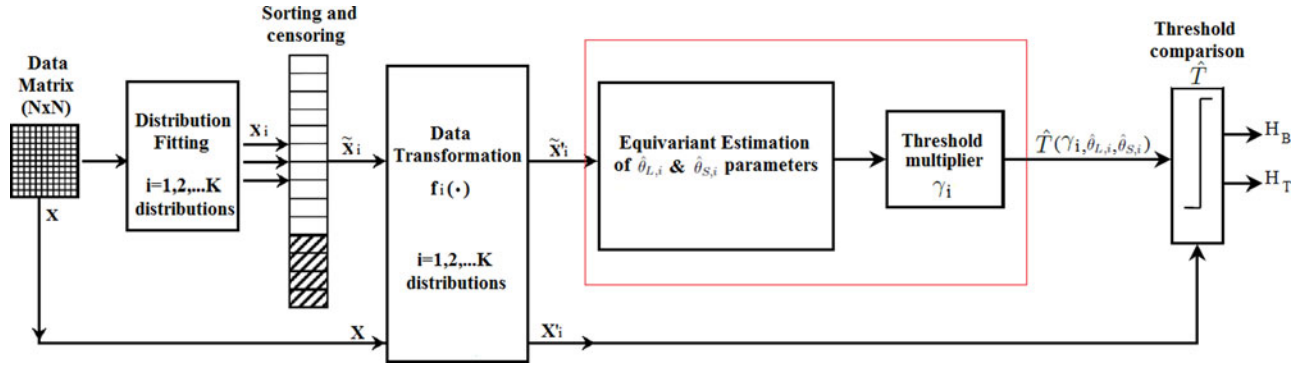where $F_{X_0}(\cdot)$ is the distribution of the standardized variate X_0 [11].

Fig. 1. Architecture of the multimodel CFAR system in LS environment.

significance level for the distribution model that best fits the data.

3) The background is highly nonhomogeneous due to the presence of tree trunks and foliage [6] and statistical distributions with light or heavy tail characteristics are preferred in such case. To our knowledge, confirmed by simulation assessments, best performance in terms of goodness of fit are provided by
   a) the Gumbel for maximum CDF

$$F\left(\mathbf{x}; \theta_L, \theta_S\right) = \exp\left[-\exp\left(-\frac{x - \theta_L}{\theta_S}\right)\right]$$

$$\text{with} \quad \theta_L \in \mathbb{R}, \quad \theta_S > 0 \qquad (2)$$

   which belongs to an LS family;
   b) the Weibull CDF

$$F\left(\mathbf{x}; \kappa, \lambda\right) = \begin{cases} 1 - e^{-\left(\frac{x}{\lambda}\right)^k}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \qquad (3)$$

   that is an LS (Gumbel) distribution after a logarithmic transformation. Location and scale parameters are derived from scale and shape parameters of the Weibull CDF as in [11] and [12].

   Let ρ_i be p-value associated to the i-th distribution $F_i(x)$, for each reference window, the selection rule for the statistical distribution is

$$F_i(x) : i = \arg\max_{i=1,2,\ldots,K} \rho_i. \qquad (4)$$

4) Data are organized into a vector $\mathbf{x}_i$ that is sorted and censored. After censoring, that consists in removing the largest r samples in the vector $\mathbf{x}_i$, the new vector $\tilde{\mathbf{x}}_i$ is produced. This transformation is required in order to control the self-masking from targets inside the reference window [15].

5) The multimodel CFAR algorithm is applied by exploiting, for each reference window, the parameters $(\gamma_i, \hat{\theta}_{L,i}, \hat{\theta}_{S,i})$ of the specific distribution, after a proper transformation to get an LS distribution. For the case at hand, the data transformation follows the rule

$$f_i\left(\cdot\right) = \begin{cases} 1, & \text{for } i = 1 \\ ln\left(\cdot\right), & \text{for } i = 2. \end{cases} \qquad (5)$$

The best linear unbiased estimates of the location and scale parameters are obtained by minimizing the variance of the estimators subject to the constraints of unbiasedness [13]. It is found that

$$\begin{pmatrix} \hat{\theta}_{L,i} \\ \hat{\theta}_{S,i} \end{pmatrix} = \left(\mathbf{H}_i^T \mathbf{C}_{0,i}^{-1} \mathbf{H}_i\right)^{-1} \mathbf{H}_i^T \mathbf{C}_{0,i}^{-1} \tilde{\mathbf{x}}_{\mathbf{i}}' \qquad (6)$$

where $\tilde{\mathbf{x}}_{\mathbf{i}}'$ is a vector of ordered and censored samples from an LS distribution, $\mathbf{H}_i = (1\, \boldsymbol{\mu}_{0,i})$ and $\boldsymbol{\mu}_{0,i}$ and $\mathbf{C}_{0,i}$ are the mean vector and the covariance matrix of the standardized vector $\tilde{\mathbf{x}}_{0,i}'$, respectively, and whose expressions are reported in Appendix A. Location and scale parameters related to the specific distribution under test can be thus estimated through (6), and the threshold for CFAR detection can be derived as

$$\hat{T}\left(\gamma_i, \hat{\theta}_{L,i}, \hat{\theta}_{S,i}\right) = \hat{\theta}_{S,i}(\tilde{\mathbf{x}}_{\mathbf{i}}')\, \gamma_i + \hat{\theta}_{L,i}(\tilde{\mathbf{x}}_{\mathbf{i}}'). \qquad (7)$$

The quantities in (11), in Appendix A, can be used to evaluate, for any value of the censoring depth r, the mean vector and the covariance matrix of the censored data. In other words, the "censored" mean vector as well as the "censored" covariance matrix can be obtained from the "uncensored" expressions by discarding the last r elements and the last r rows and columns, respectively. However, the evaluation of the adaptive threshold in (7) requires also the knowledge of the *threshold multiplier* γ_i, that is evaluated according to the desired false alarm probability.

The threshold multiplier is the solution of the equation

$$P_{\text{FA}}|D_i = Pr\left\{\frac{X_i'(m, n) - \hat{\theta}_{L,i}}{\hat{\theta}_{S,i}} > \gamma_i \,|\, H_{B|D_i}\right\} \qquad (8)$$

which is the $(1 - P_{\text{FA}})$-quantile of the normalized test statistic $[(X_i'(m, n) - \hat{\theta}_{L,i})/\hat{\theta}_{S,i}]$, where $X_i'(m, n)$ is the sample under test. Obviously, if the statistical distribution of the test statistic is known, then γ_i can be determined. Unfortunately, this distribution cannot be evaluated in a closed
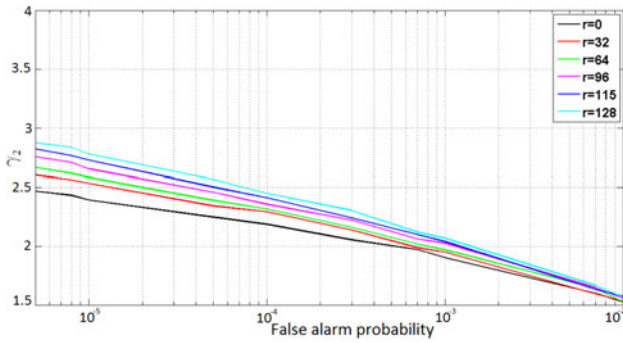
Fig. 2. Threshold multiplier values versus P_{FA}, for a Log-Weibull distribution, with $M = N \times N = 256$ and various censoring depth r.

Fig. 3. Empirical CDF compared to the theoretical Gumbel for maximum CDF for a reference window of 16×16 pixels, in a high-density forest.

form because it requires the knowledge of the joint distribution of the variable $X_i'(m, n)$ and of the location and scale estimators. More easily, the value of the quantile γ_i can be computed via Monte Carlo simulation where N_p realizations of the test statistic are generated and the threshold multiplier level is estimated from the empirical CDF [16]. In order to improve the estimators reliability, a suitable number of trials is needed.

In Fig. 2, plots show the value of the threshold multiplier versus the false alarm probability for a Log-Weibull distribution with $N = 16$, $M = 256$ and various values of the censoring depth r. As seen in Fig. 2 the value of threshold multiplier increases by reducing the false alarm probability.

6) Finally, the decision rule

$$X_i'(m, n) \underset{H_{B|D_i}}{\overset{H_{T|D_i}}{\gtrless}} \hat{T}\left(\gamma_i, \hat{\theta}_{L,i}, \hat{\theta}_{S,i}\right) \quad (9)$$

compares each data pixel with the adaptive threshold for CFAR detection. The decision rule (9) is applied using the threshold value $\hat{T}\left(\gamma_i, \hat{\theta}_{L,i}, \hat{\theta}_{S,i}\right)$ associated to the specific distribution.

The algorithm adapts itself with respect to the distribution that fits better the real data in a specific reference window, with no *a priori* knowledge of trees density.

As can be seen from (8), the false alarm probability is conditionally CFAR. Nevertheless, if the same false alarm probability is kept for all single models, the multimodel approach keeps the CFAR property.[2]

PROOF. Considering $P_{FA}|D_i = \tilde{P}_{FA} \ \forall i = 1 \dots K$, for the law of total probability

$$P_{FA} = \sum_{i=1}^{K} Pr(D_i) \cdot P_{FA}|D_i = \tilde{P}_{FA} \cdot \sum_{i=1}^{K} Pr(D_i) = \tilde{P}_{FA} \quad (10)$$

---

[2]CFAR property ensured an estimated P_{FA} values in agree with the designed value when the correct model is chosen during the statistical characterization of the background stage.
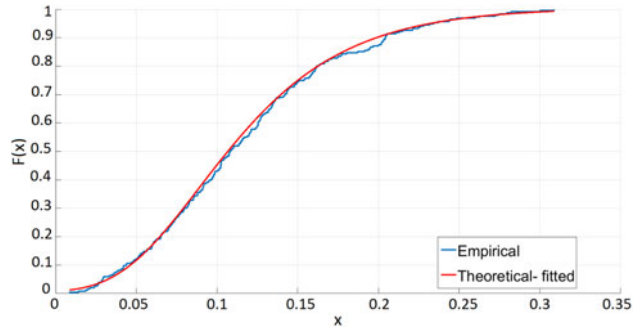
where $\sum_{i=1}^{K} Pr(D_i) = 1$, and P_{FA} is the total false alarm probability of the multimodel algorithm. ∎

III. STATISTICAL CHARACTERIZATION OF BACKGROUND

In this section, we present the results obtained in terms of goodness of fit for the selection of the statistical distributions in the H_B hypothesis. The dataset was acquired using the Swedish low-frequency SAR system *CARABAS-II VHF SAR* [17]. The system transmits HH-polarized radio waves between 20 and 90 MHz, corresponding to wavelengths between 3.3 and 15 m. In the imaged areas, 25 military vehicles are concealed by the forest, in four deployments (for reader's convenience see [17]). Due to the presence of trees in the scenario, we deal with extremely inhomogeneous data. During setup, several light and heavy tailed distributions were analyzed, e.g., the Rayleigh, Normal, Log-Normal, Gamma, Weibull, and Extreme Value Distributions. The Log-Normal distribution in low-density forest is rejected in the 42.33% of the analyzed cells. Weibull Distribution and the Gumbel for maximum distribution provide the best results for low- and high-density forests, respectively. For this reason, and for shortness, these two distributions have been selected and results will be discussed in this section.

A. High-Density Forest

In order to analyze the goodness of fit of the data to the distributions under test, we consider the area shown in Fig. 4(a). Within this image, we selected the subimage enclosed in the red frame in Fig. 4(b), that contains an homogeneous area without targets, with high density of trees. The selected area has the same characteristics of the surrounding area, where the targets are placed. For this area, a set of 29×29 reference windows, indexed through the respective quadrant in Fig. 4(b), represents the dataset for the statistical validation. Each reference window is composed of 16×16 pixels producing a set of 256 samples. For each reference window, a Lilliefors test is carried out to verify the compatibility of the data with the design distribution. To this aim, the empirical CDF has been compared to the theoretical CDF, considering the Gumbel for maximum
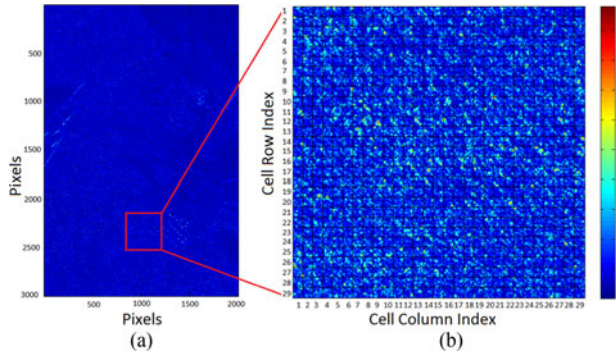
Fig. 4. (a) SAR image of 2000 × 3000 samples, related to flight pass number 1 of Fredrik mission. (b) Homogeneous area under test with pixel coordinates [2101–2565, 701–1165], within Forest 1 [17].
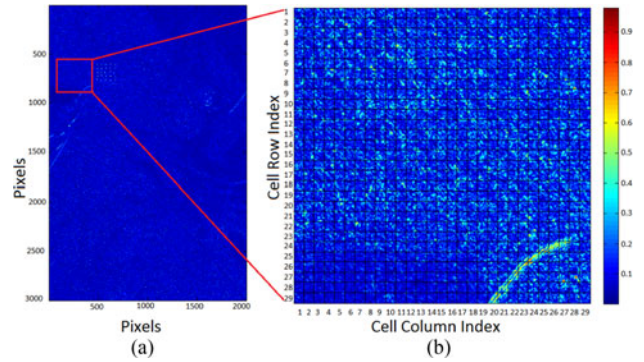

Fig. 5. (a) SAR image of flight pass number 1 of Sigismund mission, showing 2000 × 3000 acquired samples. (b) Homogeneous area under test with coordinates [451–915, 1–465] in Forest 2.

TABLE I
Lilliefors Test Outcomes for High Density and Low Density Forests: Percentages Where the H_0 Hypothesis Using Gumbel Maximum Distribution and Weibull Distribution is Accepted

Fredrik Mission	Gumbel Maximum Distribution
16 × 16	**96.91%**
24 × 24	**93.35%**
32 × 32	**89.80%**
Sigismund Mission	Weibull Distribution
16 × 16	**97.50%**
24 × 24	**94.46%**
32 × 32	**92.86%**

Fig. 6. Empirical CDF compared to the theoretical Weibull CDF for a reference window of 16 × 16 pixels, in a low-density forest.

distribution for high-density forest data. This comparison is shown in Fig. 3.

The p-values for each reference window in the homogeneous area of Fig. 4(b), are evaluated for both H_0 hypothesis, where the Gumbel for maximum distribution fits well the real data, and H_1 hypothesis, where the Gumbel for maximum distribution is rejected (with a significance level set to 0.1%).

Varying the reference window size, as shown in the Table I, best results are obtained with a reference window of 16 × 16 pixels. In this case, it is found that H_0 hypothesis was rejected 26 times, corresponding to a percentage of acceptance of the Gumbel for maximum distribution of 96.91%.

B. Low-Density Forests

In a low-density forest, the homogeneous area into Forest 2 dataset [17] is considered. The statistical characterization was carried out by considering an image of the flight pass 1 of Sigismund mission. In Fig. 5, a set of 29 × 29 reference windows is used for test, with reference window size equal to 16 × 16 pixel.

For low-density forest data, the Weibull distribution is considered. For each reference window, a Lilliefors test is carried out to verify the compatibility of the data with the design distribution. The comparison between the empirical CDF and the theoretical Weibull is shown in Fig. 6.
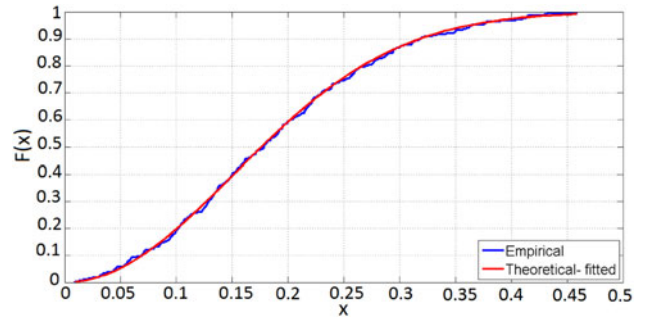
From Table I, we note that the Weibull distribution shows a good fit in 97.50% of cases (H_0 hypothesis was rejected 21 times). By changing the reference window dimensions a slight performance degradation was found for the low-density forest scenario as already seen in high-density forest case; 16 × 16 reference window size, provides the best fitting, for both Weibull and Gumbel for Maximum distributions.

C. Multimodel Fitting

Best performance are achieved through the implementation of a multimodel fitting approach as described in Section I. In order to evaluate the advantage of this approach in both low- and high-density forests, the same homogeneous areas of Forest 1 in Fig. 4 and Forest 2 in Fig. 5 were considered.

Differently from the single-model approach, whereby either the Gumbel or Weibull distributions have been exploited, with the multimodel approach the Lilliefors test is applied considering both the distributions and choosing the one that provides the highest p-value, greater than the significance level. The H_0 percentage of Lilliefors test outcomes obtained using the multimodel fitting approach, for homogeneous area of Figs. 4 and 5, increases in both scenarios.

For high-density forest, from a 96.91% of H_0 acceptance percentage, associated to the Gumbel for maximum single model, a 98.57% has been achieved using the multimodel algorithm.

TABLE II
Simulated Data From Gumbel Distribution: False Alarm
Probabilities Using Both the Correct Single-Model, the Wrong
Single-Model, and Multimodel CFAR Detector for Various
Values of the Censoring Depth, With a Designed $P_{\text{FA}} = 10^{-4}$

Censoring Depth	P_{FA}_GUM	P_{FA}_WBL	P_{FA}_MM
0	1.19×10^{-4}	1.3×10^{-3}	2.21×10^{-4}
32	1.11×10^{-4}	2.22×10^{-3}	2.48×10^{-4}
64	9.28×10^{-5}	2.29×10^{-3}	2.49×10^{-4}
96	1.04×10^{-4}	3.30×10^{-3}	2.67×10^{-4}
115	9.49×10^{-5}	3.13×10^{-3}	2.61×10^{-4}
128	1.01×10^{-4}	2.90×10^{-3}	2.68×10^{-4}

TABLE III
Simulated Data From Weibull Distribution: False Alarm
Probabilities Using Both the Correct Single-Model, the Wrong
Single-Model, and Multimodel CFAR Detector for Various
Values of the Censoring Depth, With a Designed $P_{\text{FA}} = 10^{-4}$

Censoring Depth	P_{FA}_WBL	P_{FA}_GUM	P_{FA}_MM
0	9.88×10^{-5}	3.9×10^{-6}	7.67×10^{-5}
32	8.86×10^{-5}	3.72×10^{-6}	6.15×10^{-5}
64	9.98×10^{-5}	3.38×10^{-6}	5.83×10^{-5}
96	1.08×10^{-4}	5.71×10^{-6}	5.38×10^{-5}
115	1.02×10^{-4}	7.55×10^{-6}	4.69×10^{-5}
128	1.04×10^{-4}	1.01×10^{-5}	4.81×10^{-5}

For low-density forest, by comparing the Weibull single-model approach with the multimodel approach, we find that the percentage of H_0 hypothesis is increased from 97.50%, in the Weibull single model, to 99.16% in the multimodel approach (H_0 is rejected only seven times in the second case).

It is worth to remark that the H_0 hypothesis is rejected only for the sake of statistical validation, but in the detection scheme, a distribution is always selected according to the maximum p-value, even if it is lower than the significance level.

IV. PERFORMANCE ANALYSIS OF THE CFAR DETECTOR

In order to demonstrate the CFAR property of the single- and multimodel approaches, simulated data have been considered. In particular, a $P_{\text{FA}} = 10^{-4}$ was set like designed false alarm rate and $100/P_{\text{FA}}$ realizations of simulated data composed by N^2 samples are drawn exactly from both Weibull and Gumbel distributions. Successively, sorting and censoring block is applied, accounting for the number of valid samples available for each censoring depth r.

The P_{FA} probability are reported in the Tables II and III, choosing the locations θ_L and scale θ_S parameters from a uniform distribution $\mathcal{U}(a, b)$ with a and b values tuned on the real data. In particular for the Gumbel case, $(a_l = 0.051, b_l = 0.275)$ and $(a_s = 0.029, b_s = 0.157)$ are considered for the location and scale parameters, respectively, while $(a_s = 0.770, b_s = 0.397)$ and $(a_c = 1.450, b_c = 2.824)$ are used for the scale and shape parameter of the Weibull distribution. For the single model,

TABLE IV
Real Data From Fredrik Deployment: False Alarm Probabilities
Using Both the Correct Single-Model (Gumbel Max), the Wrong
Single-Model (Weibull), and Multimodel CFAR Detector for
Various Values of the Censoring Depth, With a Designed
$P_{\text{FA}} = 10^{-4}$

Censoring Depth r	P_{FA}_GUM	P_{FA}_WBL	P_{FA}_MM
0	5.55×10^{-5}	2.54×10^{-4}	7.86×10^{-5}
32	0.97×10^{-4}	1.11×10^{-3}	1.76×10^{-4}
64	1.20×10^{-4}	1.17×10^{-3}	1.80×10^{-4}
96	1.66×10^{-4}	2.00×10^{-3}	2.45×10^{-4}
115	1.99×10^{-4}	2.01×10^{-3}	3.05×10^{-4}
128	2.40×10^{-4}	2.01×10^{-3}	3.56×10^{-4}

how clearly reported in the first column of the Tables II and III, the P_{FA} values oscillate around the designed P_{FA} ensuring the CFAR property.

In addiction, a model miss-match analysis has been done. In this case, simulated data from a chosen distribution have been evaluated, using a single-model CFAR detector based on the other distribution. How clearly reported in the second column of the Tables II and III, the P_{FA} values do not respect the designed P_{FA}. The same procedure has been done with the multimodel approach. The outcomes reported in the Tables II and III, confirm the effectiveness of the proposed approach ensuring robustness of mixed scenario. The discrepancies between the single- and multimodel results are due mainly to cases in which the Lilliefors test is not able to discriminate correctly the true data distribution. Furthermore, in this section, we investigate the performance of the single-model and multimodel CFAR detector for different forest environments, using real data. The CFAR detector performance are optimized by choosing the best size for the reference window and for the depth of censoring. A reasonable rule is to take the number of samples in the reference window much greater than the maximum expected object size, and at the same time, to discard a number of samples that is at least equal to the overall size of the objects. The efficiency of the algorithm has been tested for different values of the reference window size, censoring depth, and false alarm rate. The attention will focus on the flight pass number 1 of each mission/target deployment [17] with the following parameters: $P_{\text{FA}} = 10^{-4}$; $N = 16$; $r = [0 \quad 32 \quad 64 \quad 96 \quad 115 \quad 128]$. Single-model, model mismatch, and multimodel analysis for high- and low-density forests are considered and the results are discussed in the next sections.

A. False Alarm Rate

The first analysis is performed for an area with high density within Forest 1 scenario, in the absence of targets (see Fig. 4). The single-model CFAR detector is applied using a Gumbel for maximum distribution and the false alarm probabilities for the single- and multimodel CFAR detector are reported in Table IV, for a design false alarm rate equal to 10^{-4} and for various values of the censoring depth. The detectors exhibit P_{FA} values that are compatible

TABLE V

Real Data From Sigismund Deployment: False Alarm Probabilities Using Both the Correct Single-Model (Weibull), the Wrong Single-Model (Gumbel Max), and Multimodel CFAR Detector for Various Values of the Censoring Depth, With a Designed $P_{FA} = 10^{-4}$

Censoring Depth r	P_{FA}_WBL	P_{FA}_GUM	P_{FA}_MM
0	1.11×10^{-4}	0	4.16×10^{-05}
32	3.00×10^{-4}	1.37×10^{-5}	6.94×10^{-05}
64	4.53×10^{-4}	1.38×10^{-5}	4.42×10^{-05}
96	5.87×10^{-4}	1.39×10^{-5}	4.16×10^{-05}
115	4.81×10^{-4}	1.39×10^{-5}	5.55×10^{-05}
128	5.97×10^{-4}	1.39×10^{-5}	7.86×10^{-05}

TABLE VI

Detection Probabilities for a Single-Model CFAR Gumbel for Maximum Detector versus Multimodel CFAR Detector in Fredrik Deployment

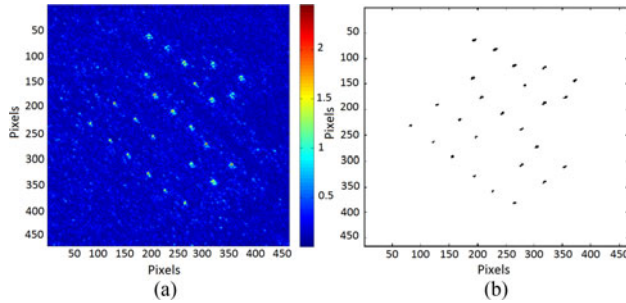| Censoring Depth r	P_D_MM	P_D_GUM
0	0.076	0.076
32	0.125	0.124
64	0.172	0.170
96	0.211	0.209
115	0.222	0.220
128	0.237	0.233

Fig. 7. (a) SAR image of the 5 by 5 targets array within Forest 1 characteristic of Fredrik deployment. (b) Targets ground truth for Fredrik deployment.
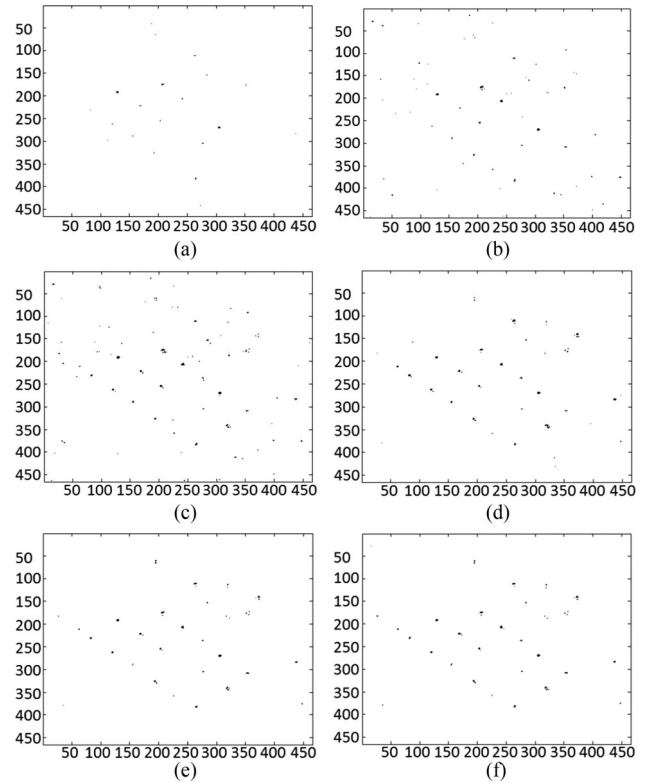

Fig. 8. Detection maps for multimodel CFAR detector, referred to Fredrik deployment (Forest 1) and varying censoring depths; $P_{FA} = 10^{-4}$ and cell size 16×16. (a) $r = 0$. (b) $r = 32$. (c) $r = 64$. (d) $r = 96$. (e) $r = 115$. (f) $r = 128$.

with the design rate, however, the multimodel approach ensures higher reliability thanks to its capability to adapt to the statistical model.

A similar analysis is considered for the Forest 2 dataset (low forest density, in the absence of targets) shown in Fig. 5. Both the single-model CFAR detector for the low-density forest (using a Weibull distribution) and the multimodel approach have been considered. The resulting false alarm probabilities are reported in Table V, for a design false alarm rate equal to 10^{-4} and various values of the censoring depth. Again, the false alarm probabilities are compatible with the design value.

B. Detection Probability

The detection capability of the single- and multi-model CFAR detector is first evaluated by considering the area of Forest 1. Fig. 7(a) and (b) shows the SAR image of 5×5 array of targets and the targets ground truth.[3]

The single-model CFAR detector has been applied using the design values $P_{FA} = 10^{-4}$ and various censoring depths. In Table VI, the probability of detection for a single-model CFAR with Gumbel for Maximum detector with P_{FA} set to 10^{-4} are reported along with the probabilities of detection of a multimodel CFAR detector referred to the same area under test.

When no censoring is applied the targets are not correctly detected, due to the self-masking effect from the targets. More correct detections are obtained when r is increased up to 128 samples. Further increase of the depth of censoring does not introduce additional features to the targets detection but, otherwise, generates underestimation of the distribution parameters.

Detection maps for the multimodel CFAR detector are reported in Fig. 8, for various censoring depths and $P_{FA} = 10^{-4}$. Comparing Fig. 8(c) with (f), where an increase of the depth of censoring is applied, we can clearly note the performance optimization in terms of false alarms.

The multimodel CFAR algorithm achieves equivalent performance in terms of detection probabilities as the single-model CFAR algorithm (see Table VI). The most

---

[3]It has to be considered that the targets ground truth has been estimated affecting the detection probability performance.
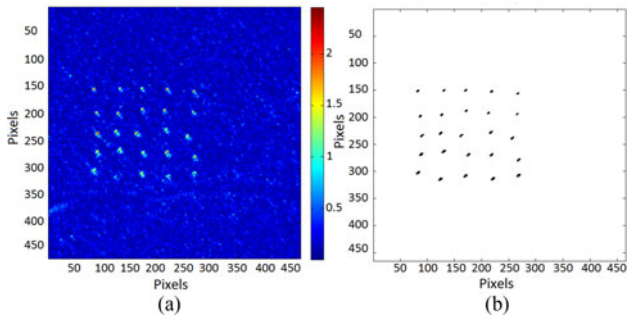
Fig. 9. (a) SAR image representing the 5 by 5 targets array within forest 2 characteristic of Sigismund deployment. (b) Targets ground truth for Sigismund deployment.

TABLE VII
Detection Probabilities for a Single-Model
CFAR Weibull Detector versus Multimodel
CFAR Detector for Sigismund Deployment

Censoring Depth r	P_D_MM	P_D_WBL
0	0.105	0.045
32	0.167	0.120
64	0.257	0.277
96	0.355	0.396
115	0.373	0.434
128	0.394	0.468

important remark is the higher accuracy of the multimodel approach in the estimate of the background parameters, hence a higher reliability of the obtained detections. To complete the performance analysis, the low-density area in Forest 2 shown in Fig. 9(a) is considered. The targets empirical ground truth in Fig. 9(b), has been considered to quantify the capability of the algorithm to detect targets under foliage.

In Table VII, the probability of detection for a single-model CFAR Weibull detector with P_{FA} set to 10^{-4} is reported along with the probability of detection of a multi-model CFAR detector referred to the same area under the test.

The detection maps for the multimodel detection cases reported in Table VII are shown in Fig. 10.

C. Clustering Strategy

We now assess the capability to detect extended targets using the proposed algorithm. In particular, by increasing the design false alarm probability, a higher detection probability is achievable. By introducing some domain knowledge like the size of the targets of interest, an a *posteriori* analysis can be performed to decrease the false alarm rate (preserving the high detection probability). Precisely, by using a *clustering strategy*[4] and target size selection, the

---

[4]Clustering is a division of data into groups of similar objects. Each group, called cluster, consists of objects that are similar between themselves and dissimilar to objects of other groups [18]. In this paper, the "aggregation" method is exploited, an algorithm based on the aggregation of adjoining pixels.
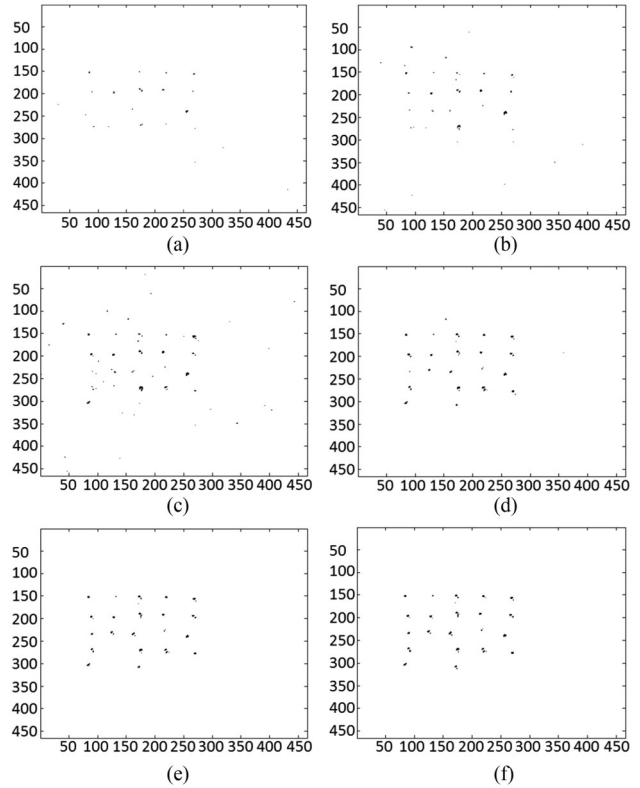

Fig. 10. Detection maps for multi-model CFAR detector, referred to Sigismund deployment (Forest 2) and varying censoring depths; $P_{FA} = 10^{-4}$ and cell size 16×16. (a) $r = 0$. (b) $r = 32$. (c) $r = 64$. (d) $r = 96$. (e) $r = 115$. (f) $r = 128$.

extended targets can be extracted from the detection map. The clustering algorithm classifies the detections according to a pixels aggregation procedure and assigns a value to every cluster equal to the number of pixels forming it. Then, using a *priori* knowledge about targets size, clusters smaller than a fixed threshold can be discarded. This operation leads to a strong decrement of the false alarms. Setting the design P_{FA} of a multimodel CFAR algorithm to 10^{-2} an higher number of detected pixels is achieved. An area with targets within Forest 1 with Fredrik deployment (as shown in Fig. 7) is considered. The clustering algorithm has been applied to the area under test. Following the clustering, the distribution of the detected target sizes can be obtained, as shown in Fig. 11. Exploiting the *a priori* domain knowledge about the size of the smaller target ($\approx 14m^2$), the clusters smaller than that value have been discarded (red bars in Fig. 11). The resulting detection map after the clustering and the size threshold is shown in Fig. 12, where the red detections can be discarded.

In Table VIII, the performance in terms of false alarms and detection probabilities on a pixel basis before and after the clustering and target size threshold are reported. In this case, the proposed approach provides a bounded false alarm rate. The clustering strategy applied to the multi-model CFAR algorithm provides a decrement of the false alarm probability, keeping the same detection probability. Hence, the clustering algorithm can be used if an higher capability of detecting extended targets is required. For
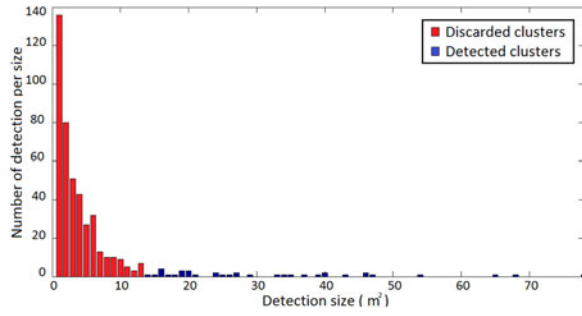
Fig. 11. Histogram of number of detections per size versus detection size. In red are shown the target with size less than 14 m², while in blue are shown the targets with size greater or equal than 14 m².
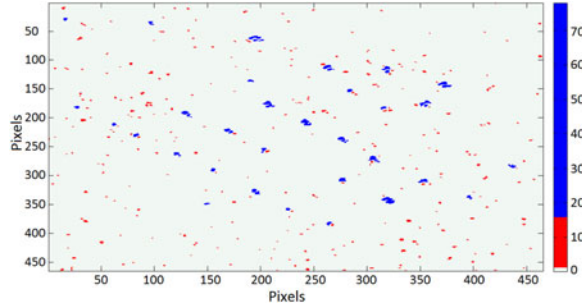

Fig. 12. Clustering strategy: Detection map of a multimodel CFAR detector with $P_{FA} = 10^{-2}$ and $r = 128$. The red color represents the discarded pixels, instead the blue color highlights the potential targets.

TABLE VIII
Multimodel CFAR Detector versus the Same
Detector With Clustering Strategy:
Performance Analysis

$P_{FA} = 10^{-2}, r = 128$	P_D	P_{FA}
Multimodel CFAR	0.5804	1×10^{-2}
Multimodel CFAR with Clustering	0.5804	3×10^{-3}

example, for the case in hand, the 96% of targets were detected (one vehicle only was missed).

V. CONCLUSION

In this paper, a novel framework for CFAR detection of extended target in FOPEN SAR images has been proposed using a multimodel approach. The novel framework exploits a CFAR detection algorithm based on LS distributions. Due to the peculiar problem in hand, the LS property was not the only constraint on the selection of the statistical distribution to be considered but, due to the presence of trees in the scene, distributions having light and heavy tails were considered as well. The proposed framework is able to control the false alarm probability in the FOPEN SAR scenario, adapting the best LS distribution fitting the background data. The effectiveness of the proposed approach has been demonstrated on real data, showing that the framework is able to adapt both the model and the threshold, and provides a higher level of reliability. Moreover, the

capability to detect extended targets has been demonstrated by exploiting domain knowledge. The usefulness of the multimodel approach, derived by generalization of the single model, is reinforced by the possibility of using it also for other kind of clutter, since that it could be used for any types of LS distribution. In this paper, it has been used for FOPEN clutter, but it does not exclude the possibility of using this approach for SAR data collected in other conditions. For example sea clutter at low grazing angle exhibits sea spikes that can lead to a behavior of the radar returns similar to those of large trunks in forests. Future work will deal with the generalization of the statistical model and to the application to 3-D FOPEN SAR.

APPENDIX A

Letting $G_i^{(m)}$ be the mth order derivative of the inverse cumulative distribution function $G_i(\cdot) = F_i^{-1}(\cdot)$, the approximate expressions of mean and covariance matrix of the ordered samples are [11]

$$\mu_{0(k),i} \approx G_{k,i} + \frac{p_k q_k}{2(n+2)} G_{k,i}^{(2)} + \frac{p_k q_k}{(n+2)^2}$$
$$\times \left(\frac{1}{3}(q_k - p_k)G_{k,i}^{(3)} + \frac{1}{8}p_k q_k G_{k,i}^{(4)} \right)$$

$$C_{0(k,h),i} \approx \frac{p_k q_h}{(n+2)} G_{k,i}^{(1)} G_{h,i}^{(1)} + \frac{p_k q_h}{(n+2)^2}$$
$$\times [(q_k - p_k)G_{k,i}^{(2)}G_{h,i}^{(1)} + (q_h - p_h)G_{h,i}^{(2)}G_{k,i}^{(1)}$$
$$+ \frac{1}{2}p_k q_k G_{k,i}^{(3)}G_{h,i}^{(1)} + \frac{1}{2}p_h q_h G_{k,i}^{(1)}G_{h,i}^{(3)}$$
$$+ \frac{1}{2}p_k q_h G_{k,i}^{(2)}G_{h,i}^{(2)}], \qquad k \leq h \qquad (11)$$

with $p_k = k/(n+1)$ and $q_k = (1 - p_k)$ and $G_{k,i} = G(p_k)$. The moments in (11) can be evaluated with reference to a specific reduced distribution by direct substitution of the derivatives (up to the fourth order) of its quantile function (inverse CDF). Let $G_{i=1}(x)$ be the quantile function associated with the reduced extreme value distribution of type I for maximum (Gumbel for Maximum), the set of equations to substitute in (11), and containing the four derivatives under test are

$$G_1(x) = -ln(-ln(x))$$
$$G_1^{(1)}(x) = -\left(\frac{1}{x\,ln(x)} \right)$$
$$G_1^{(2)}(x) = \left(\frac{ln(x) + 1}{x^2\,ln^2(x)} \right)$$
$$G_1^{(3)}(x) = -\left(\frac{2\,ln^2(x) + 3\,ln(x) + 2}{x^3\,ln^3(x)} \right)$$
$$G_1^{(4)}(x) = \left(\frac{6\,ln^3(x) + 11\,ln^2(x) + 12\,ln(x) + 6}{x^4\,ln^4(x)} \right). \quad (12)$$

The same procedure can be applied to obtain the quantile function associated to the reduced extreme value distribution of type I for minimum (Log-Weibull).

In this case, we obtain

$$G_2(x) = ln\left(-ln(1-x)\right)$$

$$G_2^{(1)}(x) = \left(\frac{1}{(x-1)\,ln(1-x)}\right)$$

$$G_2^{(2)}(x) = -\left(\frac{ln(1-x)+1}{(x-1)^2\,ln^2(1-x)}\right)$$

$$G_2^{(3)}(x) = \left(\frac{2\,ln^2(1-x)+3\,ln(1-x)+2}{(x-1)^3\,ln^3(1-x)}\right)$$

$$G_2^{(4)}(x) =$$
$$-\left(\frac{6\,ln^3(1-x)+11\,ln^2(1-x)+12\,ln(1-x)+6}{(x-1)^4\,ln^4(1-x)}\right).$$
(13)

REFERENCES

[1] M. E. Davis
Foliage Penetration Radar. Detection and Characterization of Objects Under Trees. Raleigh, NC, USA: Scitech Publishing, 2012.

[2] J. Fransson, F. Walter, and L. Ulander
Estimation of forest parameters using CARABAS-II VHF SAR data
IEEE Trans. Geosci. Remote Sens., vol. 38, no. 2, pp. 720–727, Mar. 2000.

[3] G. Smith and L. Ulander
A model relating VHF-band backscatter to stem volume of coniferous boreal forest
IEEE Trans. Geosci. Remote Sens., vol. 38, no. 2, pp. 728–740, Mar. 2000.

[4] M. Williams *et al.*,
Modeling the SAR response of pine forest in Southern Finland
In *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 2, Jul. 2003, pp. 1350–1352.

[5] A. Kononov and M.-H. Ka
Model-associated forest parameter retrieval using VHF SAR data at the individual tree level
IEEE Trans. Geosci. Remote Sens., vol. 46, no. 1, pp. 69–84, Jan. 2008.

[6] J. A. Jackson and R. L. Moses
A model for generating synthetic VHF SAR forest clutter images
IEEE Trans. Aerosp. Electron. Syst., vol. 45, no. 3, pp. 1138–1152, Jul. 2009.

[7] J. Nanis, S. Halversen, G. Owirka, and L. Novak
Adaptive filters for detection of targets in foliage
IEEE Aerosp. Electron. Syst. Mag., vol. 10, no. 8, pp. 34–36, Aug. 1995.

[8] A. Mitra, T. Lewis, and A. Shaw
Rank-order filters for FOPEN target detection
IEEE Signal Process. Lett., vol. 11, no. 2, pp. 93–96, Feb. 2004.

[9] R. Kapfer and M. Davis
Along track interferometry for foliage penetration moving target indication
In *Proc. IEEE Radar Conf.,* May 2008, pp. 1–6.

[10] Y. Norouzi, F. Gini, M. Nayebi, and M. Greco
Non-coherent radar CFAR detection based on goodness-of-fit tests
IET Radar, Sonar Navigat., vol. 1, no. 2, pp. 98–105, Apr. 2007.

[11] M. D. Bisceglie and C. Galdi
CFAR detection of extended objects in high-resolution SAR images
IEEE Trans. Geosci. Remote Sens., vol. 43, no. 4, pp. 833–843, Apr. 2005.

[12] M. Longo, M. Guida, and M. Lops
Biparametric linear estimation for CFAR against Weibull clutter
IEEE Trans. Aerosp. Electron. Syst., vol. 28, no. 1, pp. 138–151, Jan. 1992.

[13] E. Conte, M. Lops, and A. Tulino
Hybrid procedure for CFAR in non-Gaussian clutter
Proc. IEE, vol. 144, no. 6, pp. 361–369, Dec. 1997.

[14] W. Conover
Practical Nonparametric Statistics (Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics). New York, NY, USA: Wiley, 1980. [Online]. Available: https://books.google.co.uk/books?id=m54s2puW_5AC

[15] M. Richards, J. Scheer, J. Scheer, and W. Holm
Principles of Modern Radar (Principles of Modern Radar), vol. 1. Edison, NJ, USA: SciTech Publishing, 2010. [Online]. Available: https://books.google.it/books?id=nD7tGAAACAAJ

[16] S. M. Kay
Fundamentals of Statistical Signal Processing Volume II: Detection Theory. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[17] H. Hellsten, L. Ulander, A. Gustavsson, and B. Larsson
Development of VHF CARABAS II SAR
Proc. SPIE, vol. 2747, vol. 2747, pp. 48–60, Apr. 8–9, 1996.

[18] L. Kaufman and P. Rousseeuw
Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken, NJ, USA: Wiley, 1990.
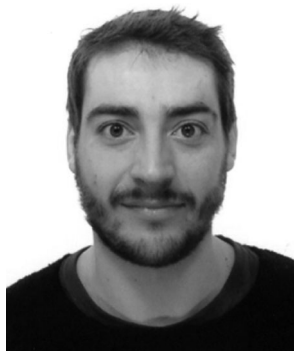
Alessio Izzo received the Laurea (B.Sc.) and Laurea Specialistica (M.Sc.) degrees in telecommunication engineering and electronic engineering for automation and telecommunications from Universita' degli Studi del Sannio, Benevento, Italy, in 2012 and 2015, respectively. He is currently working toward the Ph.D. degree from the University of Strathclyde, Glasgow, U.K, working on adaptive filtering, audio signal processing, radar signal processing, micro-Doppler signature analysis, and classification.

During his master degree, he has been visiting student with the Centre for Excellence in Signal and Image Processing Group, University of Strathclyde.

Marco Liguori was born in Benevento, Italy, on December 21, 1987. He received the Bachelor's degree in telecommunication engineering in 2012 and the Master's degree in electronic engineering for automation and telecommunications in 2015, from the Universita' degli Studi Del Sannio, Benevento, Italy.

In 2014, he joined a internship with the Centre for Excellence in Signal and Image Processing, University of Strathclyde, Glasgow, U.K. He is currently a System Analyst consultant within Telecom Italia Company, Roma, Italy.

Carmine Clemente (S'09–M'13) received the Laurea cum laude (B.Sc.) and Laurea Specialistica cum laude (M.Sc.) degrees in telecommunications engineering from Universita' degli Studi del Sannio, Benevento, Italy, in 2006 and 2009, respectively, and the Ph.D. degree from the University of Strathclyde, Glasgow, U.K., in 2012.

He is currently a Lecturer with the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, U.K., working on advanced radar signal processing algorithm, MIMO radar systems, and micro-Doppler analysis. His research interests include synthetic aperture radar (SAR) focusing and bistatic SAR focusing algorithms development, micro-Doppler signature analysis and extraction from multistatic radar platforms, micro-Doppler classification, and statistical signal processing.

Carmela Galdi (M'00) received the Dr.Eng. and Ph.D. degrees in electronic engineering from the Universita' degli Studi di Napoli "Federico II," Naples, Italy, in 1992 and 1997, respectively.

In 1995, she spent a period for study and research in the Signal Processing Division, University of Strathclyde, Glasgow, U.K. In 1997, she was a Visiting Scientist with the University College of London, London, U.K., and also with the Defence Evaluation and Research Agency, Malvern, U.K. From 1997 to 2000, she was with the Universita' degli Studi di Napoli "Federico II." In 2000, she joined the Universita' degli Studi del Sannio, Benevento, Italy, where she is currently a Professor in telecommunications. Her research interests include the field of statistical signal processing, non-Gaussian models of radar backscattering, and global navigation satellite system reflectometry. She is currently associated with the NASA Cyclone Global Navigation Satellite System project as an external science team member.

Maurizio Di Bisceglie (M'91) received the Ph.D. degree in electronic and communications engineering from Universita' degli Studi di Napoli "Federico II," Naples, Italy.

Since 1998, he has been with Universita' degli Studi del Sannio, Benevento, Italy, as an Associate Professor of telecommunications. He was a Visiting Scientist with the University College of London, London, U.K. and with the Defence Evaluation and Research Agency, Malvern, U.K. He was in the organizing committee of the Italian phase of European AQUA Thermodynamic Experiment mission, in 2004, a Cochair of the NASA Direct Readout Conference, in 2005, and in the organizing committee of the IEEE 2008 Radar Conference. Since 2016, he has been a member of the NASA Cyclone Global Navigation Satellite System External Science Team. His research interests include the field of statistical signal processing with applications to stochastic modelling and global navigation satellite system reflectometry.

John J. Soraghan (S'83–M'84–SM'96) received the B.Eng. (Hons.) and M.Eng.Sc. degrees in electronic engineering from University College Dublin, Dublin, Ireland, in 1978 and 1983, respectively, and the Ph.D. degree in electronic engineering from the University of Southampton, Southampton, U.K., in 1989.

His doctoral research focused on synthetic aperture radar processing on the distributed array processor. After graduating, he worked with the Electricity Supply Board in Ireland and with Westinghouse Electric Corporation in the United States. In 1986, he joined the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, U.K., as a Lecturer and became a Senior Lecturer in 1990, a Reader in 2000, and a Professor of signal processing in September 2003, within the Institute for Communications and Signal Processing (ICSP). In December 2005, he became the Head of the ICSP. He currently holds the Texas Instruments Chair in Signal Processing with the University of Strathclyde. He was a Manager of the Scottish Transputer Centre from 1988 to 1991, and a Manager of the DTI Parallel Signal Processing Centre from 1991 to 1995. His main research interests include signal processing theories, algorithms and architectures with applications to remote sensing, telecommunications, biomedicine, and condition monitoring.

Prof. Soraghan is a member of the Institution of Engineering and Technology.