Statistical Analysis of VHF-Band Tree Backscattering Using Forest Ground Truth Data and PO Scattering Model

Anatolii Kononov, *Member*, *IEEE*, Annelie Wyholt, Gustaf Sandberg, and Lars M. H. Ulander, *Senior Member*, *IEEE*

***Abstract*— The paper analyzes statistical properties of the VHF-band radar backscattering from coniferous trees by incorporating forest ground truth data into a physical-optics model that assumes horizontally transmit and receive polarizations and dominant double-bounce scattering from vertical stems standing on undulating ground surface. The analysis shows that a statistically adequate model for the tree backscattering amplitude can be presented as a mixture of generalized gamma or lognormal distribution and the mixture model can be reduced to a single density model if the trees with trunk volumes exceeding an appropriate threshold are to be taken into account. The generalized gamma density is shown to provide appreciably better fit to the exceedance functions associated with the PO-model data than that for the lognormal density. The results can be used to design statistically adequate models of forest clutter for VHF SAR systems.**

***Index Terms*— Backscattering, coniferous forest, finite mixture model, generalized gamma distribution, lognormal distribution, physical-optics (PO), radar cross section (RCS), synthetic aperture radar (SAR), Very High Frequency (VHF)**

I. INTRODUCTION

Over the past decade, much attention has been drawn toward the studies of low-frequency SAR systems for accurate mapping of forest biomass [1]–[9]. The physical basis for using these systems is that low-frequency radar signals are better able to penetrate the forest canopy and the strength of radar backscatter in low-frequency band is strongly correlated with the forest biomass in tree trunks and large branches. Recent works [10]–[12] on change detection using data from low-frequency SAR such as the VHF-band (20–90 MHz) CARABAS II [13], [14] and P-band (200–500 MHz) LORA [15], [16] have also shown great potential of these systems in detecting targets concealed by tree canopies.

A need for design and performance characterization of algorithms for detection of targets hidden under foliage has demanded the development of accurate models for low-frequency radar backscattering from forests. The availability of such models also allows one to assess target detection performance over a wide range of environmental conditions by using synthetic data generated by means of appropriate models. VHF-band radar backscattering from forests (forest clutter), consists of two components: distributed responses and discrete scattering terms of typically larger magnitude, such as tree trunk responses. This paper focuses only on the latter.

A new approach to the synthesis of statistical forest clutter models has recently been proposed in [17]. The procedure suggested for the synthesis includes both a statistically relevant model for clutter scene and a model of SAR image formation process. As has been shown, the fit error for the clutter model synthesized by using the suggested approach is considerably less than that achieved without including the image formation process.

The lognormal distribution is chosen in [17] to model the backscattering from tree trunks. As is also pointed out in [17], the lognormal distribution could be replaced with a new one that captures the variations of tree backscattering to a great degree. In particular, an important feature to be captured by the model is the effect of ground topography, which is not taken into account in [17]. As has been shown in an early study [18] and later in [19] – [23], the HH-polarized scattering from mature forests at low-frequency is dominated by a ground-trunk double-bounce mechanism when vertical trees are standing on horizontal flat ground. The results also indicate that due to the dominant scattering mechanism there is strong dependence of the tree backscatter on ground surface topography.

Two distribution models are used in our analysis: the generalized gamma (GGD) and lognormal distribution (LOGN). The choice of the GGD is motivated by the fact that it includes as special cases [24]: gamma, exponential, Weibull, chi-squared, as well as Rayleigh distributions and tends to a lognormal distribution under some conditions. Hence, one should expect that GGD will be able to capture the specific statistical features of tree backscattering resulted from diverse physical phenomena. To find out how a model, that has demonstrated a good fit in case of clutter scene on a horizontal flat terrain, is able to work in case of topographical ground, the LOGN is also used in the analysis.

The purpose of this paper is to ascertain a statistically adequate model of the VHF-band tree backscattering taking into account the effect of trunk volume distribution and ground topography. To achieve this purpose we perform statistical analysis of several tree backscattering

Manuscript received XXX XX, 2009; revised XXX XX, 2010. This work was supported by the Swedish National Space Board.

A. Kononov, A. Wyholt, and G. Sandberg are with the Department of Radio and Space Science, Chalmers University of Technology, 412 89 Göteborg, Sweden.

Lars M. H. Ulander is with the Department of Radio and Space Science, Chalmers University of Technology, 41289 Göteborg, Sweden, and also with the Swedish Defence Research Agency (FOI), 581 11 Linköping, Sweden.

sets generated by incorporating forest ground truth data into a recently proposed physical-optics (PO) model [23] for low-frequency radar backscatter from coniferous forests. This model has been validated in [23] against SAR image data captured with CARABAS-II and LORA SAR systems (both systems measure essentially HH-polarized backscatter) over a coniferous forest stand with well-characterized forest and ground topography. The model allows one to generate radar backscatter for a single tree at the input of a SAR system, i.e. before the image formation. We believe that using tree backscattering data in their "clean" form, when the data are not distorted by the process of SAR image formation, is required in selecting statistical adequate models. Once these models are selected, the procedure suggested in [17] can be used to find out which of them provides better fit to data from measured SAR image, i.e. to data that include the effect of image formation process.

The PO model used here calculates the strength of the double-bounce scattering from vertical stems standing on undulating ground surface and assumes deterministic ground topography [23]. According to this model, a single tree radar cross section (RCS) σ for the horizontally polarized components of the transmitted and received electric fields can be computed from

$$\sigma = 4\pi \left| \frac{1}{mn} \sum_{i=1}^{m} \sum_{k=1}^{n} S_{hh}(i,k) \right|^2 , \qquad (1)$$

where m and n is the number of model sampling points in the system frequency band and Doppler cone angles, respectively, and $S_{hh}(i,k)$ is the scattering amplitude predicted by the PO scattering model at the (i,k)-th sampling point. To compute $S_{hh}(i,k)$ several datasets are required; see for details [23, pp. 2611-2612]. These include a dataset of the SAR system parameters, parameters related to the SAR imaging geometries, parameters characterizing the electrical properties of soil and trees, forest ground truth data at a single tree level and measurements of ground topography. The first three datasets used in this paper for the computations are exactly the same that have been used in [23] to perform the model verification for the SAR system CARABAS-II. The last two datasets are different and briefly described in Section II.

II. Ground Truth Data

A. Forest Description

The forest ground truth data used in this paper are from the Remningstorp forest estate area, located in the province of Västergötland, southern Sweden (lat. 58˚30´ N, lon. 13˚40´ E). The forests in this area are dominated by Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies L. Karst*), and birch (*Betula spp.*), which are typical for Scandinavian forests. The soil type is till with a field layer consisting of blueberry (*Vaccinium myrtillus*) and cowberry (*Vaccinium vitis-idaea*).

The measurements at the single tree level were performed in the Remningstorp forests during fall 2006 and spring 2007 [25]. Due to a severe storm on 14 January 2007, i.e. between the two inventories, the data from 2006 have been updated based on a field visit during the fall of 2007. All trees with a diameter larger than 5 cm at breast height (1.3 m above ground) have been recorded together with tree species and stem positions. The stem positions were measured with a positioning error on the order of 5–10 cm. In total, 17 areas have been inventoried and the database consists in total of 4358 trees. Ten largest forest stands corresponding to 80 m × 80 m square plots are analyzed in this paper. For each of these stands only trees with valid measurements of diameter, trunk volume and position are selected for statistical analysis. Some ground truth data including the number of test trees for the selected forest stands are listed in Table I. The test trees present a subset of trees, for which the height was also measured. The tree height was measured by an ultrasound ranger together with an electronic angle decoder. The errors in the height measurements did not exceed a few decimeters. When the number of test trees was less than the total number of selected trees a third-order polynomial approximation was used to estimate tree heights from measurements of stem diameter. Trunk volumes were computed based on the tree measurements using species-dependent functions [25].

B. Topographical Data

Topographical data used for the computations with the PO model are in the form of digital elevation models (DEM). The DEMs were generated over all the selected forest stands using the TopEye helicopter-borne laser scanner [26]. The acquisition was done in April 2007. The system used a wavelength of 1064 nm with a beam divergence of 1.0 mrad. The beam was constantly scanned in cross-track direction, producing an approximately z-shaped trace on the ground. First and last return were recorded, and for each pulse, the positions of the corresponding reflection points were translated into Cartesian coordinates using the measured distance and data from inertial navigation system and differential global positioning system with a nearby reference station. The estimated positioning error was on the order of 10–30 cm, both horizontally and vertically. The flight altitude was 130 m above ground. The average laser-spot density was 30–50 m^{-2} on ground. This enables a 0.25m×0.25m horizontal grid for laser measurements. Ground-surface height was extracted from the laser data points using the progressively irregular triangular-network densification method [27], [28]. To speedup computations with the PO model, the surface was resampled to a 5m×5m horizontal grid.

Table I summarizes the ground topography for the investigated forest stands in terms of root mean square (RMS) slope computed from the corresponding DEMs. The RMS slope is a highly generalized characteristic. The real ground topography shows high complexity, which can not be described by a single parameter. Clearly, the DEMs

contain accurate and complete description of ground topography.

III. Statistical Analysis

A. Methodology

This section addresses the statistical analysis of the tree backscatter for the forest stands listed in Table I. The analysis focuses on identifying to what degree a particular theoretical distribution fits to the distribution of "observed" data, which are represented by a set of backscattering amplitudes computed at the individual tree level for each of the selected forest stands. The backscattering amplitude A (in meters) corresponding to a single tree is given by

$$A = \sqrt{\sigma}\,, \tag{2}$$

where σ is the single tree RCS given by (1).

To perform analysis we apply a fitting procedure based on the goodness-of-fit tests. The goodness-of-fit tests evaluate the quality of fit between a specified theoretical (hypothesized) distribution and the distribution of the observed data. A test-specific fit statistic can be used as a quantitative measure for the quality of fit. In our analysis the problem involves a comparison of the empirical distribution function computed for a set of tree backscattering amplitudes against a particular hypothesized distribution with unknown parameters. A conventional approach to implementing a goodness-of-fit test in this case is using a two-step procedure. First, one should estimate the unknown parameters from observed data by using the method of maximum likelihood, and then apply a goodness-of-fit test for known parameters using instead of the unknown parameters their maximum likelihood estimates. A significant disadvantage of this approach is that it does not utilize the goodness-of-fit test to its highest potential because no optimality criteria for the quality of fit are involved in this two-step procedure.

The modified goodness-of-fit tests are used in this paper for the analysis. These modified tests are based on the approach suggested in [29]. The principal idea of that approach is to combine the estimation of the hypothesized distribution parameters and a goodness-of-fit test procedure in such a manner as to find the estimates of parameters, which maximize the p -value of the goodness-of-fit test. Thus, such a modified goodness-of-fit test strives to optimally (in the sense of the p -value maximization) fit the hypothesized distribution to the distribution of the observed data.

We start the analysis from estimating the probability density function (pdf) of the tree backscatter by using a Matlab implementation of the kernel density (KD) estimation [30]–[32].To help in understanding of the effect of ground topography on the tree backscatter we also compute the KD-estimates of the pdf for the corresponding trunk volumes. Fig. 1 (a) and (b) plots some examples of these estimates for the tree backscatter and trunk volume pdfs, respectively. These examples indicate that statistically adequate description for the VHF tree backscatter and trunk volume distributions can be specified in terms of a finite mixture model. By visual inspecting the pdf plots for all the selected forest stands we have approximately estimated the number of terms (see Table II) in the corresponding mixture distributions.

Next, a statistical procedure based on a simple heuristic approach is performed to find out if a 2-term mixture distribution can be statistically adequate model for the tree backscattering by using data sets associated with the studied forest stands. As a criterion for the quality of model fit we use the criterion of maximum p -value of a modified Kolmogorov-Smirnov (MKS) goodness-of-fit test. A modified chi-squared (MCHI2) goodness-of-fit test proposed in [29] is also used in our analysis with the following addition. After fitting a model by using the chi-squared goodness-of-fit test from [29], a conventional Kolmogorov-Smirnov (KS) goodness-of-fit test is run in order to compute the p -value and test statistic for the MCHI2-fitted model in terms of the quality metric of the KS goodness-of-fit test.

The pdf $q_m(x)$ associated with a 2-term mixture distribution has the following form [33], [34]

$$q_m(x) = w_1 q(x \mid \theta_1) + (1 - w_1) q(x \mid \theta_2)\,, \tag{3}$$

where w_1 and $1 - w_1$ are the mixing weights, and θ_1 and θ_2 are the vectors of parameters for the first and second terms of the mixture, respectively. Distribution models considered in the analysis are the generalized gamma and lognormal distributions. When the hypothesized model is assumed to be a generalized gamma distribution the pdf $q(x \mid \theta)$ in (3) is given by [24]

$$q(x \mid a, b, c) = \frac{c}{b^{ac}\Gamma(a)} x^{ac-1} e^{-\left(\frac{x}{b}\right)^c}\,, \tag{4}$$

where $a > 0$ and $c > 0$ are the shape parameters, $b > 0$ is the scale parameters, so that the vector of parameters is $\theta = [a, b, c]$, and $\Gamma(a)$ is the gamma function. For the lognormal model the pdf $q(x \mid \theta)$ is [24]

$$q(x \mid \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{[\ln(x) - \mu]^2}{2\sigma^2}}\,, \tag{5}$$

where μ and σ are the mean and standard deviation, respectively, of the associated normal distribution and the vector of parameters is $\theta = [\mu, \sigma]$.

The method proposed in [29] for the estimation of parameters in the distribution (4) employs a simple relation between the generalized and gamma variates. If a random variable X has a generalized gamma distribution with parameters a, b, c, then the random variable $Y = X^c$ has a gamma distribution with parameters a, b^c, i.e. the pdf of Y is given by [24]

$$q(x \mid \alpha, \beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)}, \qquad (6)$$

where $\alpha = a$ and $\beta = b^c$ are the parameters of distribution. Using this relation is reasonable because the estimation of the parameters α, β in the gamma density (6) is much simpler than that for the parameters a, b, c in the generalized density (4). As is suggested in [29], estimates of the parameters in (6) can be constructed by using simple equations [24]

$$\alpha = \left(E[Y]\right)^2 / VAR[Y], \ \beta = VAR[Y] / E[Y],$$

where $E[Y]$ and $VAR[Y]$ are the mean and variance of the random variable Y, respectively.

We now turn to the description of the MCHI2 (MKS) goodness-of-fit test procedure when the null hypothesis distribution is assumed to be a generalized gamma one. Let ξ be the observed dataset for which we want to optimally fit the generalized gamma distribution with parameters a, b, c using the MCHI2 or MKS goodness-of-fit test. The algorithm for the MCHI2 tests can be sketched as follows (see [29] for details).

1. Loop on the parameter c, where c takes on values from the specified interval ($c \in [0.5, 50]$ in our analysis).

2. Compute $\zeta = \xi^c$, where c is a fixed value from step 1.

3. Compute the estimates $\hat{\alpha}$, $\hat{\beta}$ as $\hat{\alpha} = \left(\overline{\zeta} / s_{\zeta}\right)^2$, $\hat{\beta} = s_{\zeta}^2 / \overline{\zeta}$, where $\overline{\zeta}$ and s_{ζ} are the sample mean and sample standard deviation computed for ζ.

4. Run conventional chi-squared goodness-of-fit test (at a specified significance level α) for the sample ζ assuming the pdf (6) in the form $q(x \mid \hat{\alpha}, \hat{\beta})$ for the null hypothesis.

5. Keep the p-value and test statistic from step 4 and the associated estimates for the parameters in the generalized gamma distribution as $[\hat{a}, \hat{b}, \hat{c}] = [\hat{\alpha}, \hat{\beta}^{1/c}, c]$.

6. At the end of the loop on c, return the vector $[\hat{a}, \hat{b}, \hat{c}]$ for which the p-value of the performed goodness-of-fit test is maximum.

7. Perform conventional KS goodness-of-fit test with the pdf (4) in the form $q(x \mid \hat{a}, \hat{b}, \hat{c})$ in order to compute the p-value and test statistic for the MCHI2-fitted GGD model in terms of the quality metric of the KS goodness-of-fit test.

Note: To implement the MKS test the algorithm should be changed: at step 4, the conventional chi-squared test has to be replaced with the conventional KS test and step 7 may be excluded.

Below we sketch the algorithm that implements the MCHI2 (MKS) goodness-of-fit test in case when the null hypothesis distribution is assumed to be a lognormal one.

1. Compute the maximum likelihood estimates $\hat{\mu}_{ML}$ and $\hat{\sigma}_{ML}$ of the parameters in the distribution (5) for the dataset ξ.

2. Loop on the parameter μ, where μ takes on values from the interval $[0.5\hat{\mu}_{ML}, 1.5\hat{\mu}_{ML}]$.

3. Solve the following maximization problem. For a fixed value of μ specified at step 2, find such a value of the parameter σ, $\sigma \in [0.5\hat{\sigma}_{ML}, 1.5\hat{\sigma}_{ML}]$, which maximizes the p-value for the conventional chi-squared test (at a specified significance level α) for the sample ξ assuming the pdf (5) $q(x \mid \mu, \sigma)$ for the null hypothesis. (7)

4. Keep $p_{\max}$, which is the maximum p-value, and associated values of μ and σ from step 3 as $[\hat{\mu}, \hat{\sigma}] = [\mu, \sigma]$.

5. At the end of the loop on μ, return the vector $[\hat{\mu}, \hat{\sigma}]$ for which $p_{\max}$ is maximum $\max(p_{\max})$.

6. Perform conventional KS test with the pdf (5) in the form $q(x \mid \hat{\mu}, \hat{\sigma})$ in order to compute the p-value and test statistic for the MCHI2-fitted LOGN model in terms of the quality metric of the KS test. Set $\max(p_{\max}) = p$, where p is the p-value computed at this step.

Note: To implement the MKS test this algorithm should be changed: at step 3, the conventional chi-squared test must be replaced with the conventional KS test, and step 6 should be excluded.

Fig. 2 shows the flow chart of the statistical procedure that is performed for optimum 2-term mixture fitting by using the MCHI2 and MKS goodness-of-fit tests. As can be seen from this figure, the procedure starts from dividing the full sample of the tree backscattering amplitudes into two subsets by using a threshold A_t. Then, these subsets are inputs for the MCHI2 and MKS goodness-of-fit tests, which generate such estimates $\hat{\theta}_1^{\chi^2}$, $\hat{\theta}_2^{\chi^2}$ and $\hat{\theta}_1^{KS}$, $\hat{\theta}_2^{KS}$ (subscripts stand for the numbers of subsets) that maximize the p-values $p_1^{\chi^2}$, $p_2^{\chi^2}$ and p_1^{KS}, p_2^{KS}, respectively, with the pdf (4) or (5) for the null hypothesis. These estimates are the vectors $\hat{\theta}_1 = [\hat{a}_1, \hat{b}_1, \hat{c}_1]$, $\hat{\theta}_2 = [\hat{a}_2, \hat{b}_2, \hat{c}_2]$ for the generalized gamma density, and the vectors $\hat{\theta}_1 = [\hat{\mu}_1, \hat{\sigma}_1]$, $\hat{\theta}_2 = [\hat{\mu}_2, \hat{\sigma}_2]$ for the lognormal density (upper indices KS and χ^2 are omitted).

At the next step the algorithm searches for the optimum estimates $\hat{w}_1^{\chi^2}$ and $\hat{w}_1^{KS}$ of the mixing weight w_1 that maximize the p-values $p_m^{\chi^2}$ and p_m^{KS} for the KS goodness-of-fit test with the pdf given by the expressions

$w_1 q(x \mid \hat{\theta}_1^{\chi^2}) + (1-w_1) q(x \mid \hat{\theta}_2^{\chi^2})$ and $w_1 q(x \mid \hat{\theta}_1^{KS}) + (1-w_1) q(x \mid \hat{\theta}_2^{KS})$, respectively, for the null hypothesis. Hence, $p_m^{\chi^2}$ (p_m^{KS}) is the corresponding p -value for a 2-term mixture distribution. Then a value of p_m is calculated as $p_m = \max(p_m^{\chi^2}, p_m^{KS})$ and stored along with the associated estimates of parameters, which are also stored as $\hat{\theta}_1$ and $\hat{\theta}_2$. The sequence of these steps is being iterated in the threshold A_t until the maximum of p_m is reached. Then, the Monte Carlo test is performed to statistically verify the optimum p -value and the significance level α.

Since even a 2-term mixture distribution is not an easily tractable statistical model (it includes at least 5 parameters) it is reasonable to find out the suitability of a single pdf model specified in terms of the generalized gamma or lognormal distributions. The single pdf model is a special case of the model (3) with w_1=1. It can be shown, by using [2, equation (20)] for typical parameters of the CARABAS system, soil, and trees, that the average normalized backscattering amplitude ($\overline{A}_n$) from trees with trunk volume $V < 0.2\,\mathrm{m}^3$ is on the order of –22 dB. Hence, $\overline{A}_n$ does not exceed the noise-equivalent backscattering level, which is about –20 dB for normal CARABAS images. Thus, one can refine a set of the tree backscattering amplitudes by selecting for the analysis only those above the noise floor, i.e. which correspond to the trees with $V \geq 0.2$ m^3. Then the modified goodness-of-fit tests with a single pdf model can be applied to a selected subset of the backscattering amplitudes as well as to a corresponding subset of the trees.

Fig. 3 shows the flow chart of statistical procedure used in the paper to study the statistical adequacy of a single pdf model for the tree backscattering corresponding to the trees with trunk volumes exceeding 0.2 m^3.

B. Numerical Results

Table III presents the results of 2-term mixture model fitting obtained by using the statistical procedure in Fig. 2 for the generalized gamma and lognormal distributions. The table contains the optimum values of the normalized threshold ($A_t / 1\,\mathrm{m}$), the numbers of trees (in the form $N_1 + N_2$) corresponding to the subsets of backscattering amplitudes after optimum thresholding, the optimum mixing weight $\hat{w}_1$ (below the values for $N_1 + N_2$), the estimates for the parameters of the first and second terms (these estimates are denoted by $\hat{a}_i, \hat{b}_i, \hat{c}_i, i = 1,2$ and $\hat{\mu}_i, \hat{\sigma}_i, i = 1,2$ for the general gamma and lognormal distribution, respectively), the maximum p -values and corresponding fit statistic for the general gamma and lognormal distribution (columns marked as GGD and LOGN, respectively). The fit statistic D in terms of the KS test quality metric is given by

$$D = \max_x | F_{PO}(x) - F_{TM}(x) |, \qquad (9)$$

where $F_{PO}(x)$ is the empirical cumulative distribution function (empirical cdf) computed for the data generated by means of the PO model and $F_{TM}(x)$ is the theoretical cdf associated with a probability density model specified for the null hypothesis.

For the purpose of the false alarm rate prediction one can rewrite (9) as

$$D = \max_x | [1 - F_{TM}(x)] - [1 - F_{PO}(x)] |, \qquad (10)$$

where $1 - F_{TM}(x)$ and $1 - F_{PO}(x)$ are, by the definition, the theoretical and empirical exceedance functions, respectively. The exceedance function, corresponding to the forest clutter cdf, is equal to the probability of false alarm $P_{FA}(\tau)$ at any given threshold τ [17]: $P_{FA}(\tau) = 1 - F(\tau)$. As follows from (10), the fit statistic D can be also interpreted as a fit quality measure that determines the maximum absolute difference between the "measured" and "theoretical" (computed for a hypothesized forest clutter model) false alarm probabilities. It should be noted, that maximizing the p -value for a KS test also leads to minimizing the statistic D, other conditions being equal. A substantial advantage of this statistic is that, in contrast to the squared-error metric suggested in [17], it is a *direct* measure of the difference between the "measured" and "theoretical" false alarm probabilities at any threshold. Hence, using the statistic D in the procedure proposed in [17] for forest clutter synthesis will lead to minimizing that maximum difference over any specified interval for the threshold. In particular, such an interval can be specified by appending (9) with the following weighting function: $f(x) = 1$, $x \in [\tau_1, \tau_2]$, zero elsewhere. The left endpoint τ_1 and right endpoint τ_2 should match the expected variations in target distributions [17].

The exceedance plots for the data generated by using the PO model and optimally fitted 2-term mixture models based on the GGD and LOGN pdfs are shown in Fig. 4.

Table IV provides the results of fitting by using the single density model for tree backscattering data sets corresponding to truncated samples of trees (these are the samples of trees with trunk volume $V \geq 0.2$ m^3). Fig. 5 shows the exceedance plots for optimally fitted single density models and those for the subsets of data extracted from the corresponding full sets of tree backscattering amplitude computed from (2). Fig. 6 displays the KD-estimates of pdf for the truncated samples of (a) tree

backscattering and (b) trunk volume, respectively. The estimates for the number of terms in the corresponding distributions are given in Table V.

The Monte-Carlo (MC) estimates (5000 of independent trials are performed) for the maximum p -values and for the significance level are given in Table VI and Table VII for 2-term mixture density model and single density model, respectively. For ease of comparison the maximum p -values from Table III and Table IV are also presented in Table VI and Table VII. The results in Table III and Table IV are obtained at a significance level of $\alpha = 0.05$. Analyzing the data in Table VI and Table VII shows that the MC estimates are in full conformity with those obtained by fitting.

IV. Discussion

The results of the preliminary statistical analysis (Section III, Table II and Fig. 1) show that a statistically adequate model for the tree backscattering amplitude should be specified in terms of a finite mixture distribution. As can be seen from Table II, this assertion is also valid for the trunk volume distribution. Hence, the multimodality of VHF-band tree backscattering distributions should not exclusively be considered as the effect of ground topography, rather it is a combined effect of the primary mixture distribution for the tree trunk volume and topography of ground surface, other conditions being equal.

Analyzing Table II also shows that the ground topography modifies the number of components in the tree backscattering distribution with respect to the associated trunk volume distribution in two ways. Increasing or decreasing this number depends, probably, on a specific combination of the spatial distribution of trees with small or large trunk volumes and topographic features over an area of interest.

Inspecting Table III yields that a 2-term mixture model provides good fits (for both the GGD and LOGN): the maximum p -values are higher than 0.8 and the associated fit statistics do not exceed 0.0333 even when the number of terms in the respective mixture of distributions exceeds 2. The explanation is that in all those cases there are only two significant terms, i.e., terms with relatively large mixing weights, as Fig. 1 (a) illustrates for the stand 10. The quality of fitting is mainly determined by these significant terms since the minor terms make a relatively low contribution to the total distribution function. Nevertheless, we do not know any likely reasons that may preclude the existence of such forest stands and/or ground topographies when a 2-term model is not able to provide such high quality fits like those in Table III.

The exceedance plots in Fig. 4 (a) and (b) for the 2-term mixture model show that the GGD-based model outperforms the LOGN-based one over the region of large backscattering amplitudes despite the fit statistic for the latter is less than that for the former; as can be seen, the plot for the GGD exceedance function goes markedly closer to the plot of exceedance function associated with the PO-model data. It should be noted that this result is observed for all the studied forest stands. The region of large backscattering amplitudes is that domain for the exceedance function where accurate prediction of the false alarm rate is of most importance [17].

As can be seen from Table IV, the single density model applied to the tree backscattering samples associated with the samples of truncated trunk volumes provides good fits for both the GGD and LOGN except the forest stand ID = 18. Yet again, Fig. 5 shows that within the domain of large tree backscattering amplitudes the exceedance plots for the GGD-based single density model provides a markedly better fit than the LOGN-based model does. Moreover, the former model has demonstrated this advantage over the latter one for all the investigated forest stands.

Table V shows that for several stands the trunk volume thresholding reduces the number of terms in the trunk volume distributions though it does not definitely lead to decreasing the number of terms in the associated tree backscattering distribution. But even if the thresholding does not reduce the number of terms in the tree backscattering distribution it may lead to such a mixture of distribution (for the tree backscattering amplitude) that consists of one or two significant terms, which have closely spaced distributions, for which the difference of mean values is small enough with respect to the standard deviations of the distributions, as is illustrated in Fig. 6 (a). This explains why a single density model can provide good fits. A poor quality fit for the stand ID=18 can be explained by that the difference between the mean values associated with the significant terms is so large that a single density model is not able to capture well the statistical variation of tree backscattering for such a pronounced mixture distribution. At the same time the 2-term model provides a very good fit for the tree backscattering associated with this forest stand (see Table III).

It should also be noted that the results of this paper supplement those published in numerous works on statistical models of SAR image data, see e.g., some recent works [35]–[41], in particular [40], which has shown that the generalized gamma distribution is the best-fit model compared to gamma, Weibull and lognormal distributions for estimating speckle noise in SAR images. The present paper suggests that the generalized gamma distribution, finite mixture model structure, and modified KS goodness-of-fit test can be also relevant for building models that adequately describe backscattering statistics in SAR imagery.

V. Conclusion

The results reported in this paper have shown that an adequate probabilistic model that is able to capture the essential statistical variations of the VHF-band tree backscattering should generally be specified in terms of finite mixture distributions. The reason for this is a

combined effect of the natural multimodality of the trunk volume distribution and ground surface topography. Specifically, a 2-term mixture distribution specified in terms of generalized gamma or lognormal pdf has been shown to be a quite adequate statistical model for the VHF-band tree backscattering amplitude.

It has also been shown that under a reasonable assumption the general mixture model can be reduced to a single density model with a generalized gamma or lognormal distribution. In particular, such a single density model is suitable for the VHF-band SAR system CARABAS-II in order to model forest clutter from trees with trunk volume above 0.2 m^3 that corresponds to the noise-equivalent backscattering for normal CARABAS II images.

From this paper, one can draw the following inferences for the VHF-band forest clutter modelling. First, in designing a forest clutter model one should compare at least two choices for the density function: lognormal and generalized gamma pdf. The latter is a more complex, it includes 3 parameters, but it has been shown to be able to provide more accurate false alarm rate prediction than the former. This should not be surprising since the extra degree of freedom allows for more flexibility in the model fit that, in turn, results in improving the fit quality. Second, the fit statistic defined by formula (10) should be considered as a reasonable alternative to the weighted squared-error metric defined in [17]. Using the statistic (10) in the clutter synthesis procedure proposed in [17] will guarantee that the maximum difference between the measured and predicted false alarm probabilities (taken over a specified interval of threshold values) does not exceed a certain level, which is to be minimized during the procedure of synthesis.

It is important to comment on the implication of this paper for the forest stem volume (biomass) retrieval by inversion of low-frequency SAR data. The information on the forest biomass is "encrypted" into SAR image data in a very complicated manner. First, the natural statistical distribution of trunk volume exhibits features that are specific for multi-term finite mixture distributions. Second, this natural distribution is modified by the ground topography, when electromagnetic waves interact with the forest scene, and then by the SAR image formation process. Thus, biomass retrieval from inversion of SAR image data requires taking into account the natural complexity of trunk volume distribution and all the distortions of the biomass related information in SAR images.

Additional information and results related to the present paper will be included in the doctoral thesis being prepared by A. Wyholt and G. Sandberg.

ACKNOWLEDGMENT

The ground truth data used in this work were obtained with the financial sponsorship of Hildur and Sven Wingquist's Foundation for Forest Research, Umeå, Sweden. The authors are sincerely grateful to Dr. Johan Holmgren and Associate Professor Johan E.S. Fransson from the Swedish University of Agricultural Sciences (SLU) in Umeå, Sweden for their great efforts at managing the field inventories. The authors would also like to thank the three anonymous reviewers for their constructive comments and suggestions to improve the manuscript.

REFERENCES

[1] H. Israelsson, L. M. H. Ulander, J. I. H. Askne, J. E. S. Fransson, P.-O. Frölind, A. Gustavsson, and H. Hellsten, "Retrieval of forest stem volume using VHF SAR", *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 36-40, Jan. 1997.

[2] G. Smith and L. M. H. Ulander, "A model relating VHF-band backscatter to stem volume of coniferous boreal forest", *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 728-740, Mar. 2000.

[3] P. Melon, J.-M. Martinez, T. Le Toan, L. M. H. Ulander, and A. Beaudoin, "On the retrieval of forest stem volume from VHF SAR data: Observation and modeling", *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 2364-2372, Nov. 2001.

[4] J. Fransson, G. Smith, F. Walter, A. Gustavsson, and L. Ulander, "Estimation of forest stem volume in sloping terrain using CARABAS-II VHF SAR data*", The Canadian Journal of Remote Sensing*, vol. 30, no. 4, pp. 651-660, Aug. 2004.

[5] B. Hallberg, G. Smith-Jonforsen, and L. Ulander, "Measurements on individual trees using multiple VHF SAR images", *IEEE Trans. Geosci. Remote Sensing*, vol. 43, pp. 2261-2269, Oct. 2005.

[6] H. Nguyen, H. Roussel and W. Tabbara, "A coherent model of forest scattering and SAR imaging in the VHF and UHF-band", *IEEE Trans. Geosci. Remote Sensing*, vol. 44, pp. 838-848, Apr. 2006.

[7] G. Smith-Jonforsen, K. Folkesson, B. Hallberg and L.M.H Ulander, "Effects of Forest Biomass and Stand Consolidation on P-Band Backscatter", *IEEE Geosci. Remote Sens. Letters*, vol. 4, no. 4, pp. 669 - 673, Oct. 2007.

[8] A. Kononov and M.-H. Ka, "Model-Associated Forest Parameter Retrieval Using VHF SAR Data at the Individual Tree Level", *IEEE Trans. on Geosci. Remote Sens.*, vol. 46, no. 1, pp. 69-83, Jan. 2008.

[9] K. Folkesson, G. Smith-Jonforsen, L.M.H Ulander, "Model-Based Compensation of Topographic Effects for Improved Stem-Volume Retrieval from CARABAS-II VHF-Band SAR Images", *IEEE Trans. on Geosci. Remote Sens.*, vol. 47, no. 41, pp. 1045 - 1055, Apr. 2009.

[10] L. Ulander, M. Lundberg, W. Pierson and A. Gustavsson, "Change detection for low-frequency SAR ground surveillance", *Radar and Signal Processing, IEE Proceedings*, vol. 152, No. 6, pp. 413-420, Dec. 2005.

[11] L. Ulander, W. Pierson, M. Lundberg, et al., "Performance of VHF-band SAR change detection for wide-area surveillance of concealed ground targets", in *Proc. Algorithms for Synthetic Aperture Radar Imagery XI* (ser. Proceedings of SPIE, E. G. Zelnio and F. D. Garber (Eds.)), vol. 5427, pp. 259—270, 2004.

[12] L. Ulander, P.-O. Frölind, A. Gustavsson, H. Hellsten and B. Larsson, "Detection of concealed ground targets in CARABAS SAR images using change detection", in *Proc. Algorithms for Synthetic Aperture Radar Imagery XI* (ser. Proceedings of SPIE, E. G. Zelnio and F. D. Garber (Eds.)), vol. 3721, pp. 243-252, 1999.

[13] H. Hellsten, L. M. H. Ulander, A. Gustavsson, and B. Larsson, "Development of VHF CARABAS II SAR," in *Proc. SPIE—Radar Sensor Technology*, Orlando, FL, Apr. 8–9, 1996, vol. 2747, pp. 48–60.

[14] H. Hellsten, L. Ulander, and J. D. Taylor, "The CARABAS II VHF synthetic aperture radar," in *Ultra-Wideband Radar Technology*, J. D. Taylor, Ed. London, U.K.: CRC Press, 2000, pp. 329–343.

[15] L. Ulander, M. Blom, B. Flood, P. Follo, P.-O. Frölind, A. Gustavsson, T. Jonsson, B. Larsson, D. Murdin, M. Pettersson, U. Rääf, and G. Stenström, "Development of the ultra-wideband LORA SAR operating in the VHF/UHF-band," in *Proc. IGARSS*, Toulouse, France, Jul. 21–25, 2003, vol. 7, pp. 4268–4270.

[16] L. Ulander, M. Blom, B. Flood, P. Follo, P.-O. Frölind, A. Gustavsson, T. Jonsson, B. Larsson, D. Murdin, M. Pettersson, U.

Rääf, and G. Stenström, "The VHF/UHF-band LORA SAR and GMTI system," in *Proc. SPIE—Algorithm for Synthetic Aperture Radar Imagery X*, Orlando, FL, Apr. 21–23, 2003, vol. 5095, pp. 206–215.

[17] J. Jackson and R. Moses, "A Model for Generating Synthetic VHF SAR Forest Clutter Images", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1138-1152, July 2009.

[18] J. van Zyl, "The effect of topography on radar scattering from vegetated areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 1, pp. 153–160, Jan. 1993.

[19] G. Smith-Jonforsen, L. Ulander, and X. Luo, "Low VHF-band backscatter from coniferous forests on sloping terrain", *IEEE Trans. Geosci. Remote Sensing*, vol. 43, pp. 2246-2260, Oct. 2005.

[20] J. Fransson, G. Smith, F. Walter, A. Gustavsson, and L. Ulander, "Estimation of forest stem volume in sloping terrain using CARABAS-II VHF SAR data*", The Canadian Journal of Remote Sensing*, vol. 30, no. 4, pp. 651-660, Aug. 2004.

[21] J. He, N. Geng; L. Nguyen, L. Carin, "Rigorous modeling of ultrawideband VHF scattering from tree trunks over flat and sloped terrain", *IEEE Trans. on Geosci. Remote Sens.*, vol. 39, no. 10, pp. 2182-2193, Oct. 2001.

[22] L. Yi-Cheng and K. Sarabandi, "Electromagnetic scattering model for a tree trunk above a tilted ground plane", *IEEE Trans. on Geosci. Remote Sens.*, vol. 33, no. 4, pp. 1063 - 1070, July, 1995.

[23] B. Hallberg, G. Smith-Jonforsen, L. M. H. Ulander, G. Sandberg, "A Physical Optics Model for Double-Bounce Scattering From Tree Stems Standing on an Undulating Ground Surface", *IEEE Trans. on Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2607-2621, Sep. 2008.

[24] M. Evans, N. Hastings and B. Peacock, *Statistical Distributions*. New York: Wiley, 1993.

[25] I. Hajnsek, R. Scheiber, L. Ulander, et al., *BioSAR 2007 technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2007 experiment: Final report without synthesis*, Contract no.: 20755/07/nl/cb, European Space Agency,2008

[26] H. Sterner, "Helicopter aerial laser ranging," in *Proc. 3rd EARSeL Workshop LIDAR Remote Sens. Land Sea*, Tallinn, Estonia, Jul. 17–19, 1997, pp. 113–118.

[27] P. Axelsson, "Processing of laser scanner data—Algorithms and applications," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, no. 2/3, pp. 138–147, Jul. 1999.

[28] P. Axelsson, "DEM generation from laser scanner data using adaptive TIN models," *Int. Arch. Photogramm. Remote Sens.*, vol. 33(B), pp. 110–118, 2000.

[29] O. Gomes, C. Combes and A. Dussauchoy, "Parameter estimation of the generalized gamma distribution", *Mathematics and Computer in simulation*, vol. 79, pp. 955-963, 2008.

[30] M.P. Wand, M.C. Jones, *Kernel Smoothing*. Chapman & Hall, London, 1995.

[31] F.M. Dekking C. Kraaikamp, H. P. Lopuhaa and L.E. Meester, *A Modern Introduction to Probability and Statistics*. Springer, 2005.

[32] W. L. Martinez and A. R. Martinez, *Computational Statistics Handbook with MATLAB*. Chapman & Hall, 2002.

[33] D. Titterington, A. Smith, and U. Makov, *Statistical Analysis of Finite Mixture Distributions*. Wiley, 1985.

[34] G.J. McLachlan and D. Peel, *Finite Mixture Models.* New York: Wiley, 2000.

[35] A. Nascimento, R. Cintra and A. Frery, "Hypothesis Testing in Speckled Data With Stochastic Distances", *IEEE Trans. on Geosci. Remote Sens.*, vol. 48, no. 1, pp. 373–385, Jan. 2010.

[36] M. Collins and J. Allan, "Modeling and Simulation of SAR Image Texture", *IEEE Trans. on Geosci. Remote Sens.*, vol. 47, no. 10, pp. 3530–3546, Oct. 2009.

[37] S. Anfinsen, A. Doulgeris and T. Eltoft, "Estimation of the Equivalent Number of Looks in Polarimetric Synthetic Aperture Radar Imagery", *IEEE Trans. on Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3795–3809, Nov. 2009.

[38] H. Wang and K. Ouchi, "Accuracy of the K-Distribution Regression Model for Forest Biomass Estimation by High-Resolution Polarimetric SAR: Comparison of Model Estimation and Field Data", *IEEE Trans. on Geosci. Remote Sens.*, vol. 46, no. 4, pp. 1058–1064, Apr. 2008.

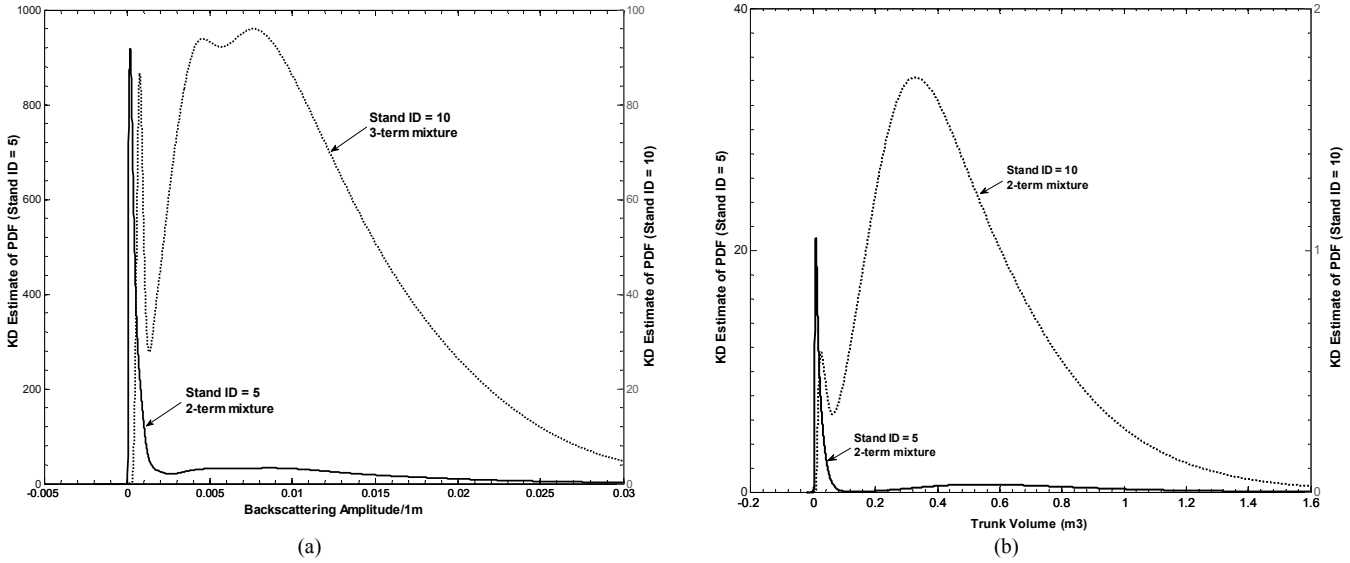
[39] S. Intajag and S. Chitwong, "Speckle noise estimation with generalized gamma distribution", in *Proc. Int. Conf. SICE-ICASE*, Busan, Korea, 18-21 Oct. 2006, pp. 1164 – 1167.

[40] H. Wang et al., "In Search of the Statistical Properties of High-Resolution Polarimetric SAR Data for the Measurements of Forest Biomass Beyond the RCS Saturation Limits", *IEEE Geosci. Remote Sens. Letters*, vol. 3, no. 4, pp. 495–499, Oct. 2006.

[41] E. Kuruoğlu and J. Zerubia, "Modeling SAR Images with a Generalization of the Rayleigh Distribution", *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 527–533, Apr. 2004.

TABLE I
SOME FOREST GROUND TRUTH DATA

Forest Stand ID	Forest Stand Inventory Number	Number of Selected Trees	Number of Test Trees	Average Diameter (cm)	Average Height (m)	Average Trunk Volume (m^3)	RMS slope (deg)
1	3926	262	262	29.14	22.16	1.125	4.30
5	4626	530	530	19.07	16.82	0.418	3.21
9	4327	332	32	29.25	23.29	0.838	4.13
10	5932	389	30	24.05	20.21	0.478	4.34
12	4430	339	43	32.73	27.35	1.050	3.37
14	2728	333	333	16.03	14.16	0.147	2.75
15	5721	370	370	19.20	15.91	0.289	7.46
16	4430	401	38	29.87	25.32	0.844	4.06
17	5729	360	30	23.73	21.25	0.496	4.14
18	5332	305	33	32.50	24.52	1.079	4.35



(a) (b)

Fig. 1. KD-estimates for probability density function of (a) tree backscattering amplitude and (b) trunk volume

TABLE II
ESTIMATES FOR NUMBER OF TERMS IN MIXTURE DISTRIBUTIONS

	Stand ID	1	5	9	10	12	14	15	16	17	18
Number of terms	Tree Volume	3	2	4	2	4	2	2	3	2	5
	Backscatter	2	2	4	3	5	2	2	4	2	5

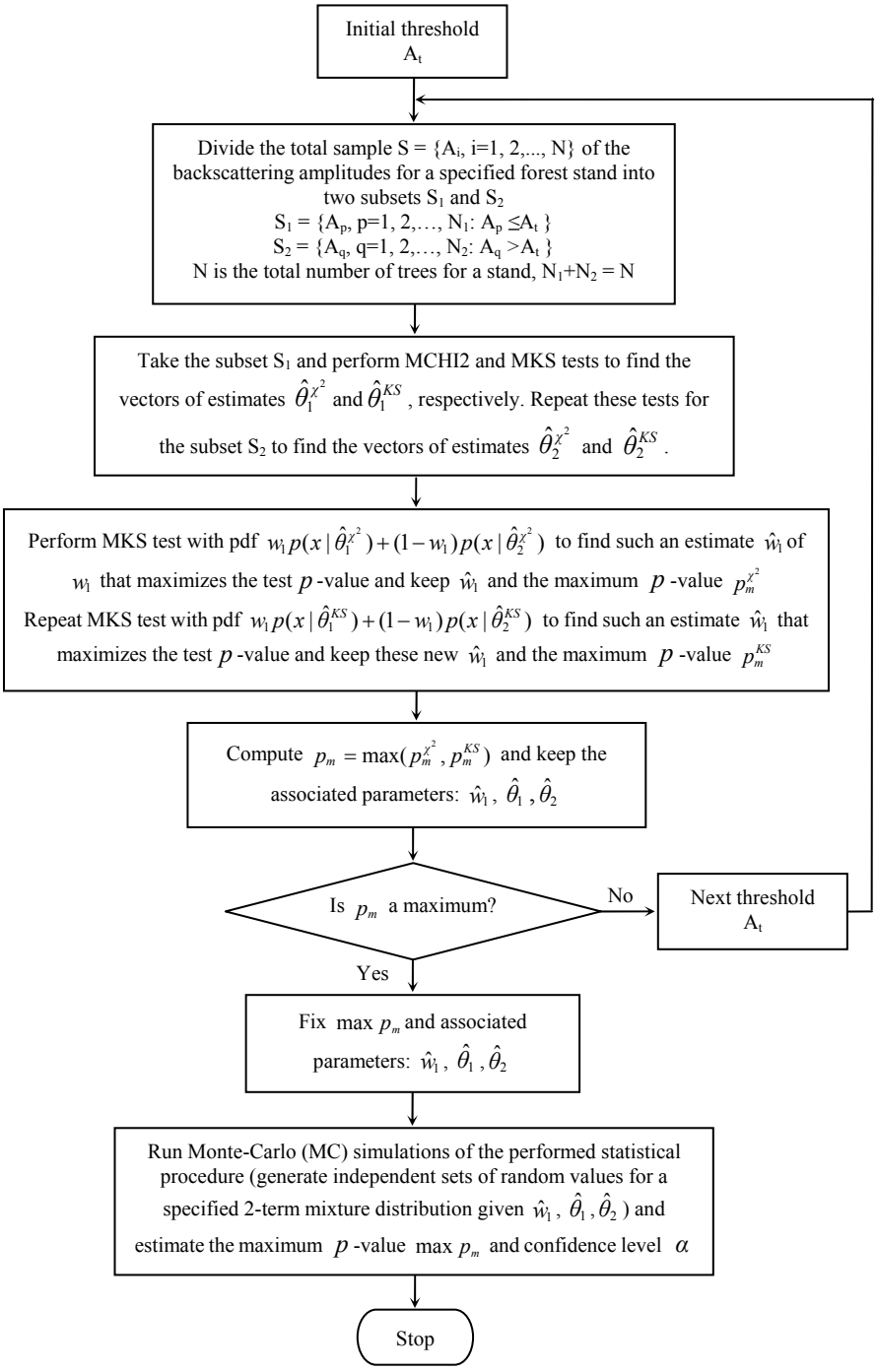


Fig. 2 Statistical procedure for optimum 2-term mixture model fitting by using modified goodness-of-fit tests

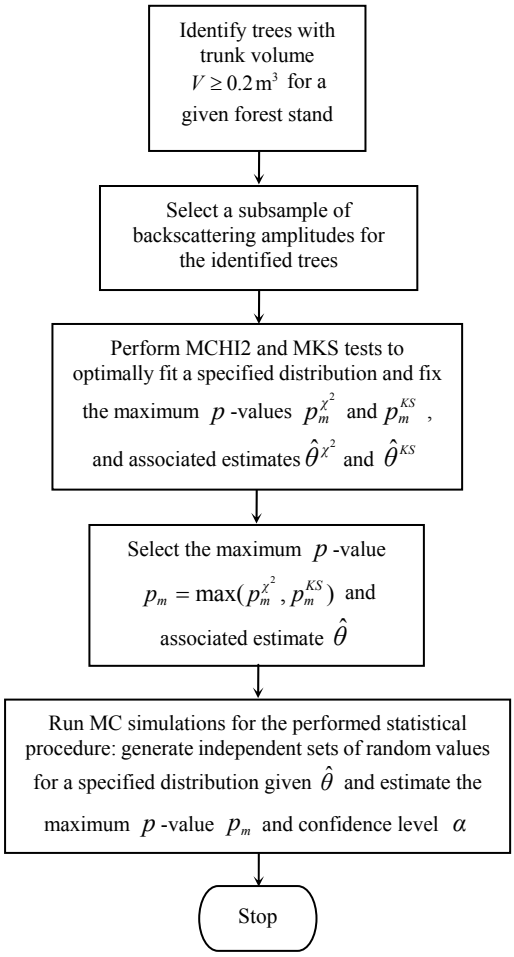
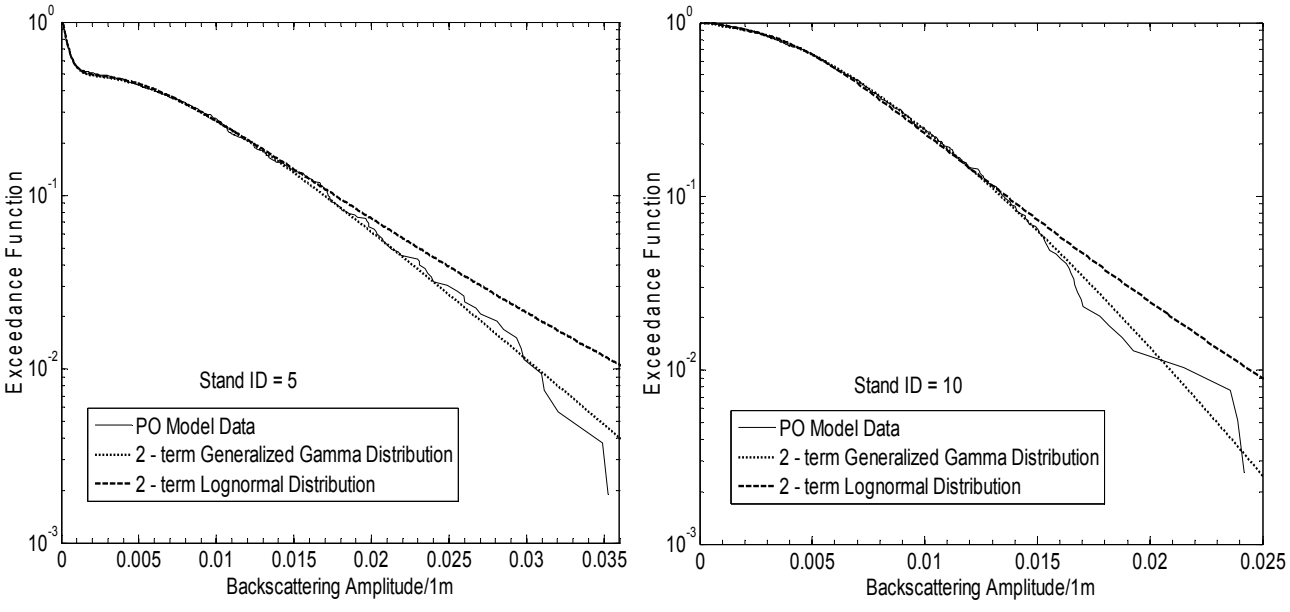


Fig. 3 Statistical procedure for optimum single density model fitting by using modified goodness-of-fit tests

TABLE III
RESULTS OF 2-TERM MIXTURE MODEL FITTING BY USING MODIFIED GOODNESS-OF-FIT TESTS

Forest Area		Generalized Gamma Distribution (GGD)					Lognormal Distribution (LOGN)				Maximum p-value		Fit Statistic	
Stand ID	Number of Trees	$\frac{A_t}{1\text{m}}$	N_1+N_2 $\hat{w}_1$	$\hat{a}_1$ $\hat{a}_2$	$\hat{b}_1$ $\hat{b}_2$	$\hat{c}_1$ $\hat{c}_2$	$\frac{A_t}{1\text{m}}$	N_1+N_2 $\hat{w}_1$	$\hat{\mu}_1$ $\hat{\mu}_2$	$\hat{\sigma}_1$ $\hat{\sigma}_2$	GGD	LOGN	GGD	LOGN
1	262	0.0012	83+179 0.3197	14.1900 1.6579	1.7674e-6 0.0170	0.5102 1.2555	0.0031	90+172 0.3435	-8.0484 -3.8304	0.6306 0.6304	0.9999	0.9981	0.0190	0.0235
5	530	0.0024	266+264 0.5083	6.3799 6.4569	1.1172e-5 7.7150e-4	0.5000 0.6884	0.0032	272+258 0.5131	-7.8196 -4.5260	0.8169 0.5956	0.9869	0.9999	0.0193	0.0136
9	332	0.0042	72+260 0.2149	3.4594 3.2793	4.5399e-4 0.0054	0.8488 0.9480	0.0058	93+239 0.2792	-6.0875 -4.0312	0.7364 0.5730	0.8907	0.9409	0.0313	0.0286
10	389	0.0012	16+373 0.0411	0.9292 2.1730	7.9458e-4 0.0040	2.7084 1.1756	0.0024	42+347 0.1107	-6.6238 -4.9592	0.5615 0.5473	0.9964	0.9612	0.0202	0.0252
12	339	0.0015	3+336 0.0088	1.8246 1.2703	6.0412e-4 0.0172	1.3963 1.5465	0.0064	43+296 0.1269	-5.4014 -4.0134	0.2773 0.5284	0.9802	0.9457	0.0250	0.0280
14	333	0.0010	67+266 0.2012	3.0401 12.3897	1.2027e-4 3.8826e-5	0.9541 0.5673	0.0010	67+266 0.2012	-8.0268 -5.7787	0.6061 0.5235	0.9363	0.9971	0.0289	0.0215
15	370	0.0039	193+177 0.5215	0.9105 2.4864	0.0012 0.0068	1.2706 1.3229	0.0062	218+152 0.5892	-6.9965 -4.3262	1.1143 0.4111	0.8116	0.9737	0.0327	0.0247
16	401	0.0011	5+396 0.0165	0.4596 0.8087	6.0235e-4 0.0184	1.5302 2.5376	0.0062	42+359 0.1066	-5.5726 -4.2308	0.5240 0.4384	0.8582	0.9356	0.0298	0.0264
17	360	0.00176	40+320 0.1117	2.4186 7.6182	8.8740e-4 1.9464e-4	2.7841 0.5000	0.0051	97+263 0.2737	-6.1491 -4.4137	0.8151 0.6045	0.9999	0.9874	0.0173	0.0233
18	305	0.00855	76+229 0.2449	0.0354 5.7090	0.0084 0.0036	36.0334 0.9707	0.0115	105+200 0.3444	-5.2155 -3.8146	0.6130 0.3726	0.9498	0.8770	0.0292	0.0333

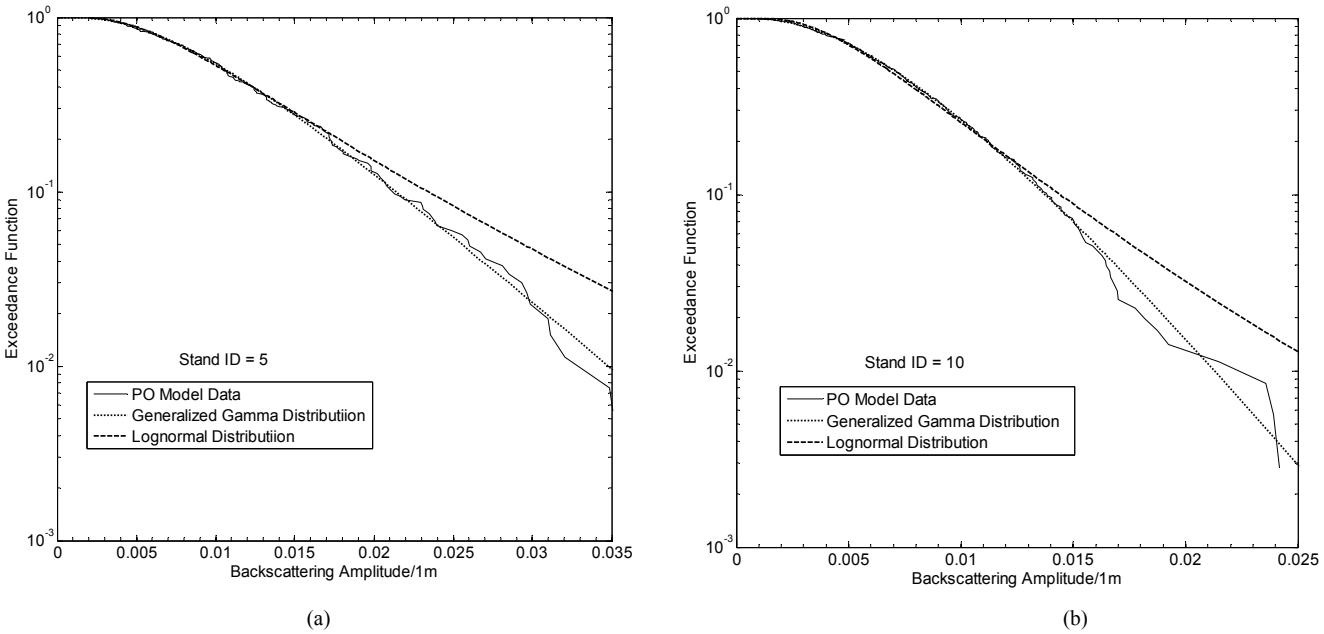


(a) (b)

Fig. 4. Exceedance plots for PO model data and optimally fitted 2-term mixture model for generalized gamma and lognormal distribution: (a) forest stand ID = 5 and (b) forest stand ID = 10

TABLE IV
RESULTS OF SINGLE DENSITY MODEL FITTING BY USING MODIFIED GOODNESS-OF-FIT TESTS

Forest Area		Generalized Gamma Distribution (GGD)			Lognormal Distribution (LOGN)		Maximum p-value		Fit Statistic	
Stand ID	Number of Trees in Sample	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{\mu}$	$\hat{\sigma}$	GGD	LOGN	GGD	LOGN
1	172	3.3747	0.0066	0.9148	-3.8303	0.6304	0.9918	0.9747	0.0322	0.0358
5	266	5.4125	0.0012	0.7422	-4.5560	0.6254	0.9764	0.9506	0.0287	0.0312
9	277	0.5420	0.0288	2.3100	-4.2263	0.7415	0.7655	0.5052	0.0395	0.0489
10	354	2.8945	0.0029	1.0583	-4.9832	0.5797	0.9995	0.8258	0.0187	0.0329
12	333	1.3519	0.0165	1.5113	-4.1260	0.6311	0.9745	0.5761	0.0259	0.0423
14	85	59.0763	1.5317e-6	0.5000	-5.2589	0.2330	0.7543	0.9815	0.0712	0.0488
15	173	2.3604	0.0075	1.3882	-4.4011	0.5027	0.9914	0.8777	0.0322	0.0439
16	389	0.9091	0.0176	2.4289	-4.2978	0.4859	0.7379	0.7908	0.0343	0.0326
17	302	8.4572	0.0002	0.5000	-4.5285	0.6836	0.9761	0.9995	0.0270	0.0201
18	279	1.4771	0.0158	1.4943	-4.0675	0.6060	0.3015	0.1274	0.0576	0.0696



(a) (b)

Fig. 5. Exceedance plots for PO model data and optimally fitted single density model for generalized gamma and lognormal distribution: (a) forest stand ID = 5 and (b) forest stand ID = 10

TABLE V
ESTIMATES FOR NUMBER OF TERMS FOR TRUNCATED SAMPLES

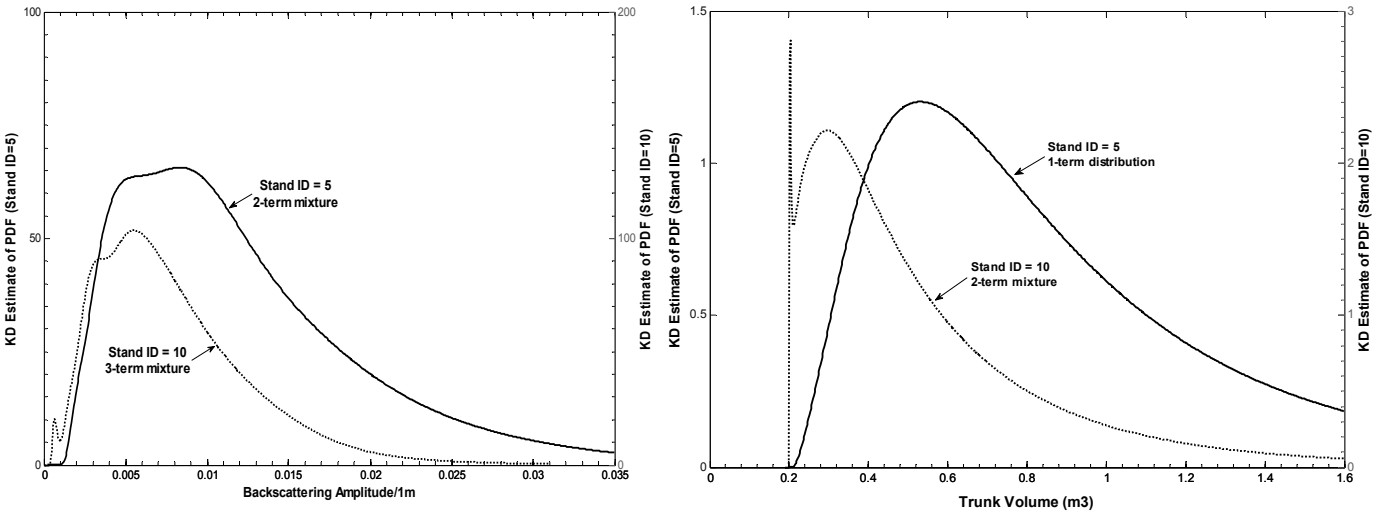
	Stand ID	1	5	9	10	12	14	15	16	17	18
Number of terms	Tree Volume	1	1	3	2	2	1	2	2	2	5
	Backscatter	2	2	3	3	2	2	2	1	2	3

TABLE VI
MONTE-CARLO ESTIMATES OF p -VALUE AND CONFIDENCE LEVEL FOR 2-TERM MIXTURE MODEL

Forest Stand ID			1	5	9	10	12	14	15	16	17	18
p -value	GGD	Result of fit	0.9999	0.9869	0.8907	0.9964	0.9802	0.9363	0.8116	0.8582	0.9999	0.9498
		MC estimate	0.9998	0.9850	0.8838	0.9949	0.9779	0.9299	0.7726	0.8462	0.9999	0.9427
	LOGN	Result of fit	0.9981	0.9999	0.9409	0.9612	0.9457	0.9971	0.9737	0.9356	0.9874	0.8770
		MC estimate	0.9973	0.9999	0.9390	0.9549	0.9458	0.9964	0.9613	0.9313	0.7382	0.8095
Confidence level	GGD	MC estimate	0.0526	0.0532	0.0488	0.0484	0.0534	0.0488	0.0536	0.0506	0.0468	0.0528
	LOGN	MC estimate	0.0508	0.0516	0.0504	0.0494	0.0516	0.0546	0.0514	0.0518	0.0520	0.0478

TABLE VII
MONTE-CARLO ESTIMATES OF p -VALUE AND CONFIDENCE LEVEL FOR SINGLE DENSITY MODEL

Forest Stand ID			1	5	9	10	12	14	15	16	17	18
p -value	GGD	Result of fit	0.9918	0.9764	0.7655	0.9995	0.9745	0.7543	0.9914	0.7379	0.9761	0.3015
		MC estimate	0.9618	0.9349	0.7188	0.9901	0.9682	0.7496	0.9895	0.7288	0.9724	0.3057
	LOGN	Result of fit	0.9747	0.9506	0.5052	0.8258	0.5761	0.9815	0.8777	0.7908	0.9995	0.1274
		MC estimate	0.9728	0.9456	0.5025	0.8292	0.5773	0.9776	0.8765	0.7847	0.9993	0.1287
Confidence level	GGD	MC estimate	0.0536	0.0534	0.0554	0.0498	0.0508	0.0476	0.0494	0.0428	0.0486	0.0538
	LOGN	MC estimate	0.0493	0.0489	0.0516	0.0521	0.0482	0.0499	0.0524	0.0497	0.0504	0.0457



(a) (b)

Fig. 6. KD-estimates for probability density function for truncated sets of trees: (a) tree backscattering amplitude and (b) trunk volume