

Review

Statistical Modeling of Polarimetric SAR Data: A Survey and Challenges

Xinping Deng ¹, Carlos López-Martínez ², Jinsong Chen ^{1,*} and Pengpeng Han ¹

¹ Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; xp.deng1@siat.ac.cn (X.D.); pp.han@siat.ac.cn (P.H.)

² Remote Sensing Laboratory, Signal Theory and Communications Department, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain; carlos.lopez@tsc.upc.edu

* Correspondence: js.chen@siat.ac.cn; Tel.: +86-755-86392370

Academic Editors: Timo Balz, Uwe Soergel, Mattia Crespi, Batuhan Osmanoglu, Zhenhong Li and Prasad S. Thenkabil

Received: 6 March 2017; Accepted: 2 April 2017; Published: 5 April 2017

Abstract: Knowledge of the exact statistical properties of the signal plays an important role in the applications of Polarimetric Synthetic Aperture Radar (PolSAR) data. In the last three decades, a considerable research effort has been devoted to finding accurate statistical models for PolSAR data, and a number of distributions have been proposed. In order to see the differences of various models and to make a comparison among them, a survey is provided in this paper. Texture models, which could capture the non-Gaussian behavior observed in high resolution data, and yet keep a compact mathematical form, are mainly explained. Probability density functions for the single look data and the multilook data are reviewed, as well as the advantages and applicable context of those models. As a summary, challenges in the area of statistical analysis of PolSAR data are also discussed.

Keywords: statistical modeling; polarimetric SAR; texture models; finite mixture models; copulas

1. Introduction

Synthetic Aperture Radar (SAR) and Polarimetric SAR (PolSAR) are widely used for observation of natural scenes. In most SAR or PolSAR systems, the size of a resolution cell is much larger than the wavelength. The measured signal is then a coherent addition of the echoes from all individual targets within that cell. Depending on the relative phases of each scattered wave, the coherent addition may be constructive or destructive, and it produces a salt-and-pepper appearance known as speckle over SAR images [1]. The target information, therefore, should be extracted through statistical analysis of the data. Hence, an accurate statistical model to describe the data becomes very important for the extraction of ground target properties [2–6].

Gaussian statistics for the radar returns have been frequently assumed when the spatial resolution of PolSAR images is moderate and the speckle is fully developed [1,7,8]. Actually, the number of targets in a resolution cell of low or medium resolution data is large. According to the Central Limit Theorem (CLT), Gaussian statistics could give a proper approximation to the data distribution. The Gaussian distribution is both mathematically tractable and efficient, making it very useful in specific applications. For SAR or PolSAR data, the mean value of the complex echo is generally assumed to be zero, and all the statistical properties are determined by the covariance matrix (or the coherency matrix) under the Gaussian assumption.

As the image resolution increases, analysis of real PolSAR data, however, reveals that non-Gaussian models give a more accurate representation of the data. The change of the observing surface could also give rise to non-Gaussian distributed data. Applications based on such models have better performance [2,4,9,10]. A common way to introduce non-Gaussianity is to divide the

randomness of the radar images into two unrelated factors, texture and speckle. The texture models the natural spatial variation of the radar cross section, whereas the speckle, following a Gaussian distribution, conveys the polarimetric information. The texture and the speckle are incorporated with a product operation which leads to a doubly stochastic model called product model [11]. In the last two decades, a considerable research effort has been dedicated to investigate accurate product models for PolSAR data [12–16].

Another way to model the non-Gaussian behavior of PolSAR data is the so called finite mixture model [17–19], which assumes the data under analysis is a discrete mixture of different targets. This makes sense in certain scenes such as urban areas, which usually consist of coherent targets like houses and roads, as well as distributed targets like trees and grass. The backscattering from the urban area is a combination of different scattering mechanisms. It has been shown that for complex regions with irregular histograms, multimodal or spiky for example, the finite mixture model is more accurate than a single distribution [17].

As summarized in [20], there are many non-Gaussian distributions, including the Weibull distribution, the lognormal distribution, and the α -stable distribution, suggested for the one-dimensional SAR data. However, these distributions are difficult to generalize to the multidimensional PolSAR data. A possible solution to this problem is to consider the idea of copulas [21]. First, we can use various non-Gaussian distributions to model the data of each polarimetric channel separately (called marginal distribution), and then introduce some common multivariate distributions to model the dependence of these marginal distributions. With the copulas, different marginal distributions and simple correlation structure can make up complex distributions for the PolSAR data [22,23].

As we can see, there are many statistical models proposed for the PolSAR data from different aspects. In this paper, a survey of these models is provided. PDFs for the single look data and the multilook data are mainly reviewed, as well as the advantages of those models. Analysis of real PolSAR data are performed using different statistical methods to evaluate the models.

The remainder of this paper is organized as follows. First, a few basic concepts of the polarimetric SAR are introduced, especially the notation employed in this paper. Then, statistics of the fully developed speckle will be discussed. Properties of the single look data and the multilook data are studied under the Gaussian assumption. The introduction of texture is followed, along with the widely studied texture models, including both the scalar texture models and the multi-texture models. Finally, finite mixture models as well as copula based models are detailed. Several experiments to validate applicability of different models are also given. Challenges in statistical modeling is summarized at the end.

2. Polarimetric SAR

PolSAR systems measure the properties of a distant target by detecting the change of the polarization state that the target induces to the incident wave. Let the polarized incident wave \mathbf{E}^i and scattered wave \mathbf{E}^s be expressed as the Jones Vectors [24]

$$\mathbf{E}^i = \begin{bmatrix} E_h^i \\ E_v^i \end{bmatrix} \quad \mathbf{E}^s = \begin{bmatrix} E_h^s \\ E_v^s \end{bmatrix} \quad (1)$$

where h represents the horizontal polarization state and v the vertical polarization state. It is possible to relate the incident and the scattered waves by means of a 2×2 complex matrix [24]

$$\begin{bmatrix} E_h^s \\ E_v^s \end{bmatrix} = \frac{e^{-jkz}}{z} \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \begin{bmatrix} E_h^i \\ E_v^i \end{bmatrix} \quad (2)$$

where z is the distance between the target and the receiving antenna, and k is the wave number of the illuminating wave. The 2×2 transformation matrix is generally referred to as scattering matrix

and denoted by \mathbf{S} . It characterizes the target under observation with four complex-valued scattering coefficients. The diagonal elements of the scattering matrix receive the name “co-pol”, since they relate the same polarization for the incident and the scattered waves. The off-diagonal elements are known as “cross-pol” terms as they relate orthogonal polarization states [24].

The definition of \mathbf{S} depends on the coordinate systems. There are two principal conventions concerning the coordinate systems where the polarimetric scattering process can be considered: Forward Scattering Alignment (FSA) and Back Scattering Alignment (BSA) [24]. The difference lies in the way the coordinate system is selected to describe the polarization state of the scattered wave. The FSA is usually used when the transmitter and the receiver are not placed at the same spatial location, for example, for bistatic radar measurements. In contrast, the BSA is often adopted in monostatic radar measurements, in which the transmitting and receiving antennas are collocated in space. In this paper, we assume that the BSA convention is employed.

The interaction between the electromagnetic waves with a reciprocal medium follows the vector reciprocity theorem, which states that if we transmit a polarization state P_A from position A, then the component polarized in the P_B direction at position B is equal to the P_A component of the scattered radiation when we illuminate the same object from B with polarization P_B [25]. The reciprocity theorem applies to ground targets generally [25]. In the BSA coordinate system, the reciprocity theorem says that the cross-pol channels of the scattering matrix are equal, that is $S_{hv} = S_{vh}$. Therefore, there are only three independent complex coefficients required to characterize the scatterer under observation.

In many cases, it is more flexible to represent the scattering matrix \mathbf{S} as a vector which is known as scattering vector. The vectorization can be performed through [26]

$$\mathbf{k} = \frac{1}{2} \text{Tr}(\mathbf{S}\Psi) \quad (3)$$

where $\text{Tr}(\cdot)$ is the matrix trace and Ψ is a 2×2 complex matrix from a basis set which are constructed as an orthonormal set under the Hermitian inner product. The lexicographic basis and the Pauli basis are the most common ones in the context of radar polarimetry. The selection of the basis to vectorize the scattering matrix depends on the final purpose of the vectorization itself. When studying the statistical behavior of the PolSAR data, the lexicographic basis is more convenient due to its simplicity. The lexicographic basis set consists of the straightforward lexicographic ordering of the elements of the scattering matrix. For a reciprocal target, the scattering vector in this case can be expressed as

$$\mathbf{k} = \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix} \quad (4)$$

Targets under observation are commonly situated in a dynamically changing environment and are subjected to spatial and temporal variations. Despite the radar system transmits a perfectly polarized wave, the wave scattered by the target is partially polarized [25]. Such scatterers are called distributed targets. The analysis of this type of targets can not be performed exactly by one target but a population of targets. More precisely, they are analyzed by introducing the concept of space and time varying stochastic processes, where the targets are described by the second order moments such as the polarimetric coherency or covariance matrices.

The covariance matrix is defined as the expectation of the outer product of the target vector with its transpose conjugate

$$\Sigma = E\{\mathbf{k}\mathbf{k}^\dagger\} = \begin{bmatrix} E\{S_{hh}S_{hh}^*\} & \sqrt{2}E\{S_{hh}S_{hv}^*\} & E\{S_{hh}S_{vv}^*\} \\ \sqrt{2}E\{S_{hv}S_{hh}^*\} & 2E\{S_{hv}S_{hv}^*\} & \sqrt{2}E\{S_{hv}S_{vv}^*\} \\ E\{S_{vv}S_{hh}^*\} & \sqrt{2}E\{S_{vv}S_{hv}^*\} & E\{S_{vv}S_{vv}^*\} \end{bmatrix} \quad (5)$$

where $(\cdot)^\dagger$ and $(\cdot)^*$ denote the transpose conjugate and conjugate, respectively. In practice, the number of scattering vectors used to calculate the expectation is limited. Let L denote the number of pixels to compute the average, the PolSAR data are then represented by the so-called sample covariance matrix

$$\mathbf{C}_L = \frac{1}{L} \sum_{i=1}^L \mathbf{k}_i \mathbf{k}_i^\dagger \quad (6)$$

where \mathbf{k}_i is the i th scattering vector. The averaging is also called multilook processing which can be employed to reduce the speckle of PolSAR data, with L referring to the number of looks.

3. Gaussian Statistics

Under the assumption that the speckle is fully developed, it has been experimentally verified that the Gaussian statistics generally provide a good fit to SAR data, especially in homogeneous natural areas [7,27–30]. The multivariate Gaussian distribution, which is both mathematically tractable and efficient, is proper to model the scattering vectors when the surface roughness is relatively low, the spatial resolution is moderate, and a large number of scatterers are present [1,24]. The Gaussian assumption indicates that the statistical properties of the data are determined by the covariance matrix. The sample covariance matrix in this case follows a complex Wishart distribution, which is widely used in the applications of PolSAR data. There exist also some variations of the Wishart distribution that are shown to be more accurate in certain circumstances.

3.1. Gaussian Distribution

When a radar illuminates an area of a random surface containing many elementary scatterers, the scattering vector, \mathbf{z} , can be modeled as having a d -dimensional complex Gaussian distribution with zero mean. The Probability Density Function (PDF) is given by [31]

$$p(\mathbf{z}; \boldsymbol{\Sigma}) = \frac{1}{\pi^d |\boldsymbol{\Sigma}|} \exp(-\mathbf{z}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{z}) \quad (7)$$

where $|\cdot|$ is the determinant operation. The complex Gaussian distribution is denoted by $\mathbf{z} \sim \mathcal{CN}(0, \boldsymbol{\Sigma})$ for brevity. The real and imaginary parts of any complex element of \mathbf{z} are assumed to follow a circular Gaussian distribution. Consider the i th element $\mathbf{z}_i = x_i + jy_i$ for example, the joint PDF of the real and the imaginary parts can be written as

$$p(x_i, y_i; \sigma_i) = \frac{1}{\pi \sigma_i^2} \exp\left(-\frac{x_i^2 + y_i^2}{\sigma_i^2}\right) \quad (8)$$

where $\sigma_i^2 = \boldsymbol{\Sigma}_{ii}$. Let r_i be the amplitude and θ_i be the phase of a complex value, then the real part of \mathbf{z}_i can be written as $x_i = r_i \cos \theta_i$, and the imaginary part as $y_i = r_i \sin \theta_i$. The Jacobian determinant of the transform from (x_i, y_i) to (r_i, θ_i) is given by

$$\mathbf{J} = \begin{vmatrix} \cos \theta_i & -r_i \sin \theta_i \\ \sin \theta_i & r_i \cos \theta_i \end{vmatrix} = r_i \quad (9)$$

Subsequently, the joint PDF of the amplitude and the phase can be obtained from (8) after changing variables, giving

$$p(r_i, \theta_i; \sigma_i) = \frac{r_i}{\pi \sigma_i^2} \exp\left(-\frac{r_i^2}{\sigma_i^2}\right) \quad (10)$$

The circular Gaussian assumption implies that the phase θ_i is uniformly distributed over $(-\pi, \pi]$, and independent from the amplitude. Averaging over the phase, therefore, gives the PDF of the amplitude

$$p(r_i; \sigma_i) = \frac{2r_i}{\sigma_i^2} \exp\left(-\frac{r_i^2}{\sigma_i^2}\right) \tag{11}$$

Equation (11) is known as the Rayleigh distribution, with mean value $\sigma_i\sqrt{\pi}/2$. The intensity of the i th channel, $I_i = x_i^2 + y_i^2 = r_i^2$, can be easily proved to have a negative exponential distribution

$$p(I_i; \sigma_i) = \frac{1}{\sigma_i^2} \exp\left(-\frac{I_i}{\sigma_i^2}\right) \tag{12}$$

with mean value $E\{I_i\} = \sigma_i^2$ and variance $\text{Var}\{I_i\} = \sigma_i^4$. This distribution shows that the useful information is described by a single degree of freedom, corresponding to the mean intensity.

Besides the intensity, the joint properties of two different polarimetric channels are of great interest. Considering two polarimetric channels $\mathbf{z}_i = x_i + jy_i$ and $\mathbf{z}_k = x_k + jy_k$, the complex correlation coefficient is determined by

$$\rho e^{j\varphi} = \frac{\Sigma_{ik}}{\sqrt{\Sigma_{ii}\Sigma_{kk}}} \tag{13}$$

where ρ and φ are, respectively, the amplitude and the phase of the complex correlation coefficient. The joint PDF of the real part and the imaginary part can be derived from (7), which is given as follows [30,32,33]

$$p(x_i, y_i, x_k, y_k) = \frac{1}{\pi^2 \psi^2 (1 - \rho^2)} \exp\left(-\frac{\sigma_k^2(x_i^2 + y_i^2) + \sigma_i^2(x_k^2 + y_k^2)}{\psi^2(1 - \rho^2)} + \frac{2\psi\rho[(x_i x_k + y_i y_k) \cos \varphi + (x_k y_i - x_i y_k) \sin \varphi]}{\psi^2(1 - \rho^2)}\right) \tag{14}$$

where $\sigma_i^2 = \Sigma_{ii}$, $\sigma_k^2 = \Sigma_{kk}$ and $\psi = \sigma_i \sigma_k$. Write the complex values in the polar form, i.e., $r_i e^{j\theta_i} = x_i + jy_i$ and $r_k e^{j\theta_k} = x_k + jy_k$, by changing variables from (x_i, y_i, x_k, y_k) to $(r_i, \theta_i, r_k, \theta_k)$, the previous distribution becomes

$$p(r_i, \theta_i, r_k, \theta_k) = \frac{r_i r_k}{\pi^2 \psi^2 (1 - \rho^2)} \exp\left(-\frac{\sigma_k^2 r_i^2 + \sigma_i^2 r_k^2 - 2\psi r_i r_k \rho \cos(\theta_i - \theta_k - \varphi)}{\psi^2(1 - \rho^2)}\right) \tag{15}$$

We are interested in the distributions of the product of the two amplitudes $z = r_i r_k$, and the phase difference $\phi = \theta_i - \theta_k$, since their values reflect the correlation between different polarimetric channels. It can be shown that the Jacobian determinant of the transform from $(r_i, r_k, \theta_i, \theta_k)$ to (r_i, z, θ_i, ϕ) is $-1/r_i$. Thus the following distribution can be obtained after changing variables

$$p(r_i, z, \theta_i, \phi) = \frac{z}{\pi^2 \psi^2 (1 - \rho^2)} \frac{1}{r_i} \exp\left(-\frac{\sigma_k^2 r_i^2 + \frac{\sigma_i^2 z^2}{r_i^2} - 2\psi \rho z \cos(\phi - \varphi)}{\psi^2(1 - \rho^2)}\right) \tag{16}$$

from which the joint PDF of z and ϕ can be further derived by integrating over θ_i and r_i and employing the equality (A1)

$$p(z, \phi) = \frac{2z}{\pi \psi^2 (1 - \rho^2)} \exp\left(\frac{2\rho z \cos(\phi - \varphi)}{\psi(1 - \rho^2)}\right) K_0\left(\frac{2z}{\psi(1 - \rho^2)}\right) \tag{17}$$

Here K_v is the modified Bessel function of the second kind of order v [34]. The marginal distribution of the product of the amplitudes, subsequently, is found to be

$$p(z) = \frac{4z}{\psi^2(1-\rho^2)} I_0\left(\frac{2\rho z}{\psi(1-\rho^2)}\right) K_0\left(\frac{2z}{\psi(1-\rho^2)}\right) \tag{18}$$

where $I_0(z)$ is the modified Bessel function of the first kind [34] resulting from the integral identity (A2). Similarly, integrating (17) over the amplitudes and following the identity (A3) gives the marginal distribution of the phase difference

$$p(\phi) = \frac{1-\rho^2}{2\pi(1-\beta^2)} \left\{ \frac{\beta}{\sqrt{\beta^2-1}} \ln(-\beta + \sqrt{\beta^2-1}) + 1 \right\} \tag{19}$$

with $\beta = \rho \cos(\phi - \varphi)$. Note that $-\beta + \sqrt{\beta^2-1}$ is a complex number since β is less than 1. Therefore, it can be represented in the polar form, e.g., $-\beta + \sqrt{\beta^2-1} = \exp(j(\pi - \arccos \beta))$, and as a result, (19) becomes

$$p(\phi) = \frac{1-\rho^2}{2\pi(1-\beta^2)} \left\{ \frac{\beta(\pi - \arccos \beta)}{\sqrt{1-\beta^2}} + 1 \right\} \tag{20}$$

The PDFs shown in (18) and (20) can be also found in [32,33]. The Gaussian assumption implies that the statistics of the PolSAR data is completely determined by the covariance matrix. The properties described by the multivariate distribution (7) can be analyzed separately by the intensity (12), the product of amplitudes (18) and the phase difference (20).

3.2. Wishart Distribution

SAR data are frequently multilook processed for speckle reduction. Under the Gaussian assumption, the sample covariance matrix \mathbf{C}_L follows a complex Wishart distribution, $\mathbf{C}_L \sim \mathcal{CW}(L, \mathbf{\Sigma})$, with PDF given by [31]

$$p(\mathbf{C}_L; L, \mathbf{\Sigma}) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d} \exp(-L \text{Tr}(\mathbf{\Sigma}^{-1} \mathbf{C}_L))}{\Gamma_d(L) |\mathbf{\Sigma}|^L} \tag{21}$$

where the normalization factor $\Gamma_d(L)$ is defined as

$$\Gamma_d(L) = \pi^{\frac{d(d-1)}{2}} \prod_{i=1}^d \Gamma(L - i + 1) \tag{22}$$

with $\Gamma(\cdot)$ referring to the gamma function. The Wishart distribution is valid only if $L \geq d$. The random variables of this distribution are the diagonal terms of \mathbf{C}_L as well as the real and imaginary parts of the upper (or lower) off-diagonal terms. For a d -dimensional radar signal, the total number of independent variables is d^2 .

Considering only one polarimetric channel, from (21), we have the distribution of the intensity as

$$p(I_i; L, \sigma_i) = \frac{1}{\Gamma(L)} \left(\frac{L}{\sigma_i^2}\right)^L I_i^{L-1} \exp\left(-\frac{L}{\sigma_i^2} I_i\right) \tag{23}$$

It is known as the gamma distribution with mean value $E\{I_i\} = \sigma_i^2$ and variance $\text{Var}\{I_i\} = \sigma_i^4/L$ [35]. The number of looks can be estimated using the mean and the variance of the intensity

$$\hat{L} = \frac{E^2\{I_i\}}{\text{Var}\{I_i\}}. \tag{24}$$

When L is equal to 1, the gamma distribution reduces to the exponential distribution (12). The variances of the two different distributions show that the multilook process reduces the speckle by scaling down the fluctuation magnitude with a factor $1/L$.

For two polarimetric channels, saying channel i and channel k , the sample covariance matrix can be written as

$$\mathbf{C}_L = \begin{bmatrix} I_i & R_{ik} + jI_{ik} \\ R_{ik} - jI_{ik} & I_k \end{bmatrix}. \quad (25)$$

Let $\rho e^{j\varphi}$ represent the complex correlation coefficient, the joint distribution of I_i , I_k , R_{ik} and I_{ik} can be derived from (21), giving

$$p(I_i, I_k, R_{ik}, I_{ik}) = \frac{L^{2L}(I_i I_k - R_{ik}^2 - I_{ik}^2)^{L-2}}{\pi \Gamma(L) \Gamma(L-1) \psi^{2L} (1-\rho^2)^L} \times \exp\left(-L \frac{\sigma_i^2 I_k + \sigma_k^2 I_i - 2\rho\psi(R_{ik} \cos\varphi - jI_{ik} \sin\varphi)}{\psi^2(1-\rho^2)}\right) \quad (26)$$

where $\sigma_i^2 = \Sigma_{ii}$, $\sigma_k^2 = \Sigma_{kk}$, and $\psi = \sigma_i \sigma_k$. Write the off-diagonal element in the polar form, $ze^{j\varphi} = R_{ik} + jI_{ik}$, by changing variables from $(I_i, I_k, R_{ik}, I_{ik})$ to (I_i, I_k, z, φ) , the following result can be obtained

$$p(I_i, I_k, z, \varphi) = \frac{zL^{2L}(I_i I_k - z^2)^{L-2}}{\pi \Gamma(L) \Gamma(L-1) \psi^{2L} (1-\rho^2)^L} \exp\left(-L \frac{\sigma_i^2 I_k + \sigma_k^2 I_i - 2z\rho\psi \cos(\varphi - \varphi)}{\psi^2(1-\rho^2)}\right) \quad (27)$$

The determinant of \mathbf{C}_L must be greater than 0, therefore, we have $I_i I_k - z^2 > 0$. Integrating I_i over $(z^2/I_k, \infty)$ using (A4) and then I_k over $(0, \infty)$ using (A1) gives

$$p(z, \varphi) = \frac{2L^{L+1}z^L}{\pi \Gamma(L) \psi^{L+1} (1-\rho^2)} \exp\left(\frac{2Lz\rho \cos(\varphi - \varphi)}{\psi(1-\rho^2)}\right) K_{L-1}\left(\frac{2Lz}{\psi(1-\rho^2)}\right) \quad (28)$$

Subsequently, the marginal distribution of the amplitude can be obtained following the integral identity (A2)

$$p(z) = \frac{4L^{L+1}z^L}{\Gamma(L) \psi^{L+1} (1-\rho^2)} I_0\left(\frac{2Lz\rho}{\psi(1-\rho^2)}\right) K_{L-1}\left(\frac{2Lz}{\psi(1-\rho^2)}\right) \quad (29)$$

and the distribution of the phase difference by identity (A5)

$$p(\varphi) = \frac{(1-\rho^2)^L}{2\sqrt{\pi}(1-\beta)^{2L}} \frac{\Gamma(2L)}{\Gamma(L)\Gamma(L+\frac{3}{2})} {}_2F_1\left(2L, L - \frac{1}{2}, L + \frac{3}{2}, \frac{\beta+1}{\beta-1}\right). \quad (30)$$

where $\beta = \rho \cos(\varphi - \varphi)$, and ${}_2F_1(a, b; c; z)$ is the Gauss hypergeometric function [34]. Again, the statistical properties of the multilook data can be analyzed separately using (23), (29) and (30). The Wishart distribution is widely used in the modeling of PolSAR data [7,36–38], and there are several variations that make the model more accurate or efficient.

3.2.1. Relaxed Wishart Model

Compared with the multivariate complex Gaussian distribution, the Wishart distribution depends on an additional parameter, L , the number of looks. Assume that the multilook processing has different contributions to different types of targets, Anfinsen et al. proposed a refined model called relaxed Wishart distribution [39], in which the number of looks L is treated as a variable shape parameter. In other words, the number of looks is assumed to be distinct in different areas. It is observed that varying L gives a better representation of the data than using a constant L over all regions [39].

3.2.2. Wishart-Kotz Distribution

Another variation of the Wishart distribution is the Wishart-Kotz model [40,41], which exhibits the heavy tails needed to fit the data found in high resolution PolSAR images. In this model, there are no special mathematical functions involved that limit the usefulness by inflicting high computational cost and numerical instability. The sample covariance matrix in the Wishart-Kotz model is assumed to follow a Wishart-Kotz type I distribution with PDF defined as [40]

$$p(\mathbf{C}_L; L, \boldsymbol{\Sigma}, \rho, \beta) = \frac{c |\mathbf{C}_L|^{L-d}}{|\boldsymbol{\Sigma}|^L} (\text{Tr}(\boldsymbol{\Sigma}^{-1} \mathbf{C}_L))^{\beta-1} \exp(-[L \text{Tr}(\boldsymbol{\Sigma}^{-1} \mathbf{C}_L)]^\rho) \quad (31)$$

with additional parameters ρ and β , and a normalization constant factor c

$$c = \frac{\rho L^{\beta+Ld-1} \Gamma(Ld)}{\Gamma_d(L) \Gamma\left(\frac{\beta+Ld-1}{\rho}\right)} \quad (32)$$

Here $\Gamma_d(L)$ is the same as that in Wishart model, see (22). The Wishart-Kotz distribution is a generalization of the Wishart distribution, which reduces to the latter when $\rho = 1$ and $\beta = 1$.

4. Texture Model

The properties of the fully developed speckle are detailed in the previous section. This section illustrates how to model the texture statistically. There are two main manners to manage this: (1) consider the texture as a scalar random variable, or (2) consider it as a vector having the same dimension as the speckle component. They lead to the so called scalar texture model and multi-texture model, respectively. The texture random variable is generally assumed to be positive with unity mean. Therefore, it models the variation of the radar cross section only, leaving the intensities to the speckle component [7,42]. The statistical properties could be described by a certain distribution, or just a stochastic process without a specific PDF.

4.1. Scalar Texture Model

The scalar texture model assumes that the texture component in the product model is a positive scalar random variable. The scattering vector in this case can be written as [7,11,43,44]

$$\mathbf{k} = \sqrt{\tau} \mathbf{z} \quad (33)$$

where τ is the texture parameter with mean value equal to 1, and \mathbf{z} is the speckle vector, following a multivariate Gaussian distribution (7). The scalar texture model is also referred to as scale mixture of Gaussian [4], or Spherically Invariant Random Vector (SIRV) [45–47]. For the multilook data, the sample covariance matrix can be expressed as

$$\mathbf{C}_L = \frac{1}{L} \sum_{i=1}^L \tau_i \mathbf{z}_i \mathbf{z}_i^\dagger = \frac{\tau}{L} \sum_{i=1}^L \mathbf{z}_i \mathbf{z}_i^\dagger \quad (34)$$

under the assumption that the texture has a higher spatial correlation than the speckle and the texture parameter is constant over the multilook processing window [13].

For a known τ , (33) implies that the scattering vector \mathbf{k} follows a complex Gaussian distribution (see Section 3.1) with PDF given by

$$p(\mathbf{k}|\tau; \boldsymbol{\Sigma}) = \frac{1}{\pi^d |\boldsymbol{\Sigma}|} \frac{1}{\tau^d} \exp\left(-\frac{\mathbf{k}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{k}}{\tau}\right) \quad (35)$$

And the distribution of the sample covariance matrix is given by

$$p(\mathbf{C}_L|\tau; L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \frac{1}{\tau^{Ld}} \exp\left(-\frac{L \text{Tr}(\Sigma^{-1} \mathbf{C}_L)}{\tau}\right) \tag{36}$$

which is known as the Wishart distribution detailed in Section 3.2.

If the PDF of the texture random variable is not explicitly specified, τ can be viewed as an unknown deterministic parameter from pixel to pixel [47]. According to the concept of SIRV, an approximate maximum likelihood estimator for the texture parameter of each pixel is found to be [45,47]

$$\begin{aligned} \hat{\tau}_i &= \frac{\mathbf{k}_i^\dagger \hat{\Sigma}^{-1} \mathbf{k}_i}{d} \\ \hat{\Sigma} &= \frac{1}{N} \sum_{i=1}^N \frac{\mathbf{k}_i \mathbf{k}_i^\dagger}{\hat{\tau}_i} \end{aligned} \tag{37}$$

where $\hat{\tau}_i$ is the texture parameter of the i th pixel, d is the dimension of the target vector, and N is the number of pixels in the neighborhood. The estimators of the texture parameter and the covariance matrix depend on each other. They can be decoupled using a recursive process. Inserting $\hat{\tau}_i$ into the second line of the above equation, the covariance matrix in the $(k + 1)$ th iteration can be estimated by [45–47]

$$\hat{\Sigma}_{k+1} = \frac{d}{N} \sum_{i=1}^N \frac{\mathbf{k}_i \mathbf{k}_i^\dagger}{\mathbf{k}_i^\dagger \hat{\Sigma}_k^{-1} \mathbf{k}_i} \tag{38}$$

The process can be initialized by any matrix, even an identity matrix [47], and it is stopped when the Frobenius distance between two consecutive estimated matrices reaches some limit. More details about the existence as well as the convergence can be found in [46]. This estimator is referred to as fixed point estimator [47].

On the contrary, if the texture random variable is specified by a distribution, averaging all possible τ gives the unconditional or marginal PDF of the scattering vector

$$p(\mathbf{k}; \Sigma) = \int_0^\infty p(\mathbf{k}|\tau; \Sigma) p(\tau) d\tau \tag{39}$$

which is analytically solvable for some choices of $p(\tau)$. The PDF of the sample covariance matrix can be obtained similarly by

$$p(\mathbf{C}_L; L, \Sigma) = \int_0^\infty p(\mathbf{C}_L|\tau; L, \Sigma) p(\tau) d\tau \tag{40}$$

A number of models have been proposed in the literature by introducing different distributions for the texture component, including the \mathcal{K} distribution [13], the \mathcal{G}^0 distribution [14,15], the Kummer- \mathcal{U} distribution [16], the \mathcal{W} , and the \mathcal{M} distribution [48], to represent different scenes of PolSAR data. They are explained in the following subsections.

4.1.1. \mathcal{K} Distribution

The \mathcal{K} distribution, assuming that the texture is gamma distributed, is widely used to model forests and the sea surface, and it can be unarguably regarded as one of the most successful radars models [4,10,12,13]. The gamma distribution is given by [35]

$$p(x; \alpha, \theta) = \frac{1}{\Gamma(\alpha) \theta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\theta}\right) \tag{41}$$

with shape parameter α and scale parameter θ . The mean value is $\mu = \alpha\theta$. Let $\tau = \frac{x}{\mu}$ to ensure the mean value of the texture is equal to 1, the distribution can be written as

$$p(\tau; \alpha) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} \exp(-\alpha\tau) \quad (42)$$

The PDF of the scattering vector \mathbf{k} can be obtained by substituting the texture distribution into (39) and employing the integral equality (A1)

$$p(\mathbf{k}; \alpha, \Sigma) = \frac{1}{\pi^d |\Sigma|} \frac{2\alpha^{\frac{\alpha+d}{2}}}{\Gamma(\alpha)} (\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k})^{\frac{\alpha-d}{2}} K_{\alpha-d} \left(2\sqrt{\alpha \mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}} \right) \quad (43)$$

By the same procedure, inserting (42) into (40), we have the PDF of the sample covariance matrix as follows

$$p(\mathbf{C}_L; \alpha, L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \frac{2\alpha^{\frac{\alpha+Ld}{2}}}{\Gamma(\alpha)} \left(L \operatorname{Tr}(\Sigma^{-1} \mathbf{C}_L) \right)^{\frac{\alpha-Ld}{2}} K_{\alpha-Ld} \left(2\sqrt{\alpha L \operatorname{Tr}(\Sigma^{-1} \mathbf{C}_L)} \right) \quad (44)$$

4.1.2. Normal Inverse Gaussian (NIG)

The Normal Inverse Gaussian (NIG) distribution assumes that the texture follows an inverse Gaussian distribution [49,50]. The PDF of the inverse Gaussian distribution is given by

$$p(x; \mu, \gamma) = \left(\frac{\gamma}{2\pi x^3} \right)^{1/2} \exp \left(-\frac{\gamma(x-\mu)^2}{2\mu^2 x} \right) \quad (45)$$

where μ is the mean value. By letting μ equal to 1 and replacing the random variable x with τ , the texture distribution becomes

$$p(\tau; \gamma) = \left(\frac{\gamma}{2\pi} \right)^{1/2} \tau^{-3/2} e^\gamma \exp \left(-\frac{1}{2} \left(\frac{\gamma}{\tau} + \gamma\tau \right) \right) \quad (46)$$

Subsequently, the PDF of the scattering vector and the sample covariance matrix can be obtained by following the integral Equation (A1), giving

$$p(\mathbf{k}; \gamma, \Sigma) = \frac{1}{\pi^d |\Sigma|} \frac{\sqrt{2\gamma} e^\gamma}{\sqrt{\pi}} \left(\frac{\gamma}{2\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k} + \gamma} \right)^{\frac{1+2d}{4}} K_{d+\frac{1}{2}} \left(\sqrt{\gamma(\gamma + 2\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k})} \right) \quad (47)$$

and

$$p(\mathbf{C}_L; \gamma, L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \frac{\sqrt{2\gamma} e^\gamma}{\sqrt{\pi}} \left(\frac{\gamma}{2L \operatorname{Tr}(\Sigma^{-1} \mathbf{C}_L) + \gamma} \right)^{\frac{1+2Ld}{4}} \times K_{Ld+\frac{1}{2}} \left(\sqrt{\gamma(\gamma + 2L \operatorname{Tr}(\Sigma^{-1} \mathbf{C}_L))} \right) \quad (48)$$

The NIG distribution has strong theoretical grounds derived from Brownian motion theory. Experiments demonstrate that it usually gives a better representation of the data than the Wishart distribution or the \mathcal{K} distribution does, because the inverse Gaussian distribution captures larger distribution shape variations than the gamma distribution [50]. In addition, the NIG distribution has less trouble at boundary mixtures than the \mathcal{K} distribution [50].

4.1.3. \mathcal{G} and \mathcal{G}^0 Distributions

It is shown that the \mathcal{G} distribution and the \mathcal{G}^0 distribution have a good representation of extremely heterogeneous regions such as urban areas [15]. Especially, the \mathcal{G}^0 distribution has the same number of parameters as the \mathcal{K} distribution, but without complex special functions like the Bessel function which requires intensive computations [14,15].

The \mathcal{G} distribution assumes that the texture parameter obeys the Generalized Inverse Gaussian (GIG) law which is characterized by the PDF [14,51]

$$p(x; a, b, p) = \frac{1}{2K_p(\sqrt{ab})} \left(\frac{a}{b}\right)^{\frac{p}{2}} x^{p-1} \exp\left(-\frac{1}{2}\left(\frac{b}{x} + ax\right)\right) \tag{49}$$

where $a > 0, b > 0$ and p is a real parameter. The mean value of this distribution is $\mu = \sqrt{\frac{b}{a} \frac{K_{p+1}(\sqrt{ab})}{K_p(\sqrt{ab})}}$. Letting $\tau = \frac{x}{\mu}$ gives

$$p(\tau; a, b, p) = \frac{1}{2} \frac{K_{p+1}^p(\sqrt{ab})}{K_p^{p+1}(\sqrt{ab})} \tau^{p-1} \exp\left(-\frac{\sqrt{ab}}{2} \left(\frac{K_p(\sqrt{ab})}{K_{p+1}(\sqrt{ab})} \frac{1}{\tau} + \frac{K_{p+1}(\sqrt{ab})}{K_p(\sqrt{ab})} \tau\right)\right) \tag{50}$$

which can be further rewritten as follows by replacing \sqrt{ab} with ω to reduce the number of parameters

$$p(\tau; \omega, p) = \frac{1}{2} \frac{K_{p+1}^p(\omega)}{K_p^{p+1}(\omega)} \tau^{p-1} \exp\left(-\frac{\omega}{2} \left(\frac{K_p(\omega)}{K_{p+1}(\omega)} \frac{1}{\tau} + \frac{K_{p+1}(\omega)}{K_p(\omega)} \tau\right)\right) \tag{51}$$

Substituting (51) into (39) and (40), and calculating the integral using (A1) leads to

$$p(\mathbf{k}; \omega, p, \Sigma) = \frac{1}{\pi^d |\Sigma|} \frac{1}{\eta^p K_p(\omega)} \left(\eta^2 + \frac{2\eta}{\omega} \mathbf{k}^t \Sigma^{-1} \mathbf{k}\right)^{\frac{p-d}{2}} K_{p-d} \left(\sqrt{\omega^2 + \frac{2\omega}{\eta} \mathbf{k}^t \Sigma^{-1} \mathbf{k}}\right) \tag{52}$$

and

$$p(\mathbf{C}_L; \omega, p, L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \frac{1}{\eta^p K_p(\omega)} \left(\eta^2 + \frac{2\eta}{\omega} L \text{Tr}(\Sigma^{-1} \mathbf{C}_L)\right)^{\frac{p-Ld}{2}} \times K_{p-Ld} \left(\sqrt{\omega^2 + \frac{2\omega}{\eta} L \text{Tr}(\Sigma^{-1} \mathbf{C}_L)}\right) \tag{53}$$

where $\eta = \frac{K_p(\omega)}{K_{p+1}(\omega)}$. The above expressions are the PDFs of the scattering vector and the sample covariance matrix following \mathcal{G} distributions [14,52].

The \mathcal{G}^0 distribution can be obtained from the \mathcal{G} distribution by letting $a \rightarrow 0$. Representing the modified Bessel function $K_\nu(z)$ using (A6), Equation (49) becomes

$$p(x; a, b, p) = \frac{2^{p-1} \Gamma\left(p + \frac{1}{2}\right)}{b^p \sqrt{\pi}} x^{p-1} \exp\left(-\frac{1}{2}\left(\frac{b}{x} + ax\right)\right) \times \left(\int_1^\infty e^{-\sqrt{abt}t} (t^2 - 1)^{p-\frac{1}{2}} dt\right)^{-1} \tag{54}$$

If $a \rightarrow 0, p = -\lambda, b = 2\beta$, then after calculating the integral via (A7), the PDF of the GIG distribution is reduced to

$$p(x; \lambda, \beta) = \frac{\beta^\lambda}{\Gamma(\lambda)} x^{-\lambda-1} \exp\left(-\frac{\beta}{x}\right) \tag{55}$$

Equation (55) is known as the inverse gamma distribution, or the reciprocal of the gamma distribution, with mean value $\mu = \frac{\beta}{\lambda-1}$. Let $\tau = \frac{x}{\mu}$ to ensure the mean value of the texture τ is equal to 1, the PDF becomes

$$p(\tau; \lambda) = \frac{(\lambda - 1)^\lambda}{\Gamma(\lambda)} \tau^{-\lambda-1} \exp\left(-\frac{\lambda - 1}{\tau}\right) \tag{56}$$

The PDFs of the scattering vector and the sample covariance matrix of the \mathcal{G}^0 distribution can be obtained by plugging the texture distribution into (39) and (40) respectively, and calculating the integral by (A9), giving

$$p(\mathbf{k}; \lambda, \Sigma) = \frac{1}{\pi^d |\Sigma|} \frac{\Gamma(\lambda + d)(\lambda - 1)^\lambda}{\Gamma(\lambda)} \left(\lambda - 1 + \mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}\right)^{-\lambda-d} \tag{57}$$

and

$$p(\mathbf{C}_L; \lambda, L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \frac{\Gamma(\lambda + Ld)(\lambda - 1)^\lambda}{\Gamma(\lambda)} \left(\lambda - 1 + L \text{Tr}(\Sigma^{-1} \mathbf{C}_L)\right)^{-\lambda-Ld} \tag{58}$$

Another extreme case of the GIG distribution is the gamma distribution when $b \rightarrow 0$, which leads to the \mathcal{K} distribution [14].

4.1.4. Kummer- \mathcal{U} Distribution

Assuming that the texture parameter follows a Fisher distribution, also known as the F-distribution or the Fisher-Snedecor distribution, with PDF given by [35]

$$p(x; d_1, d_2) = \frac{1}{B\left(\frac{d_1}{2}, \frac{d_2}{2}\right)} \left(\frac{d_1}{d_2}\right)^{\frac{d_1}{2}} x^{\frac{d_1}{2}-1} \left(1 + \frac{d_1}{d_2} x\right)^{-\frac{d_1+d_2}{2}} \tag{59}$$

where $d_1 > 0$ and $d_2 > 0$, the scattering vector or the sample covariance matrix are Kummer- \mathcal{U} distributed, with the ability to model different types of textures, because the Fisher distribution covers a large range of distributions [16,53]. The mean value of the Fisher distribution is $\mu = \frac{d_2}{d_2-2}$. Let $\tau = \frac{x}{\mu}$ to ensure the mean value of the texture variable equal to 1, we have the distribution for the texture as

$$p(\tau; \xi, \zeta) = \frac{\Gamma(\xi + \zeta)}{\Gamma(\xi)\Gamma(\zeta)} \frac{\xi}{\zeta - 1} \left(\frac{\xi}{\zeta - 1} \tau\right)^{\xi-1} \left(\frac{\xi}{\zeta - 1} \tau + 1\right)^{-\xi-\zeta} \tag{60}$$

Here parameters d_1 and d_2 are replaced by $\xi = d_1/2$ and $\zeta = d_2/2$ to make the expression more concise. Inserting the texture distribution into (39), the PDF of the scattering vector can be calculated by

$$p(\mathbf{k}; \xi, \zeta, \Sigma) = \frac{\Gamma(\xi + \zeta)}{\Gamma(\xi)\Gamma(\zeta)\pi^d |\Sigma|} \left(\frac{\xi}{\zeta - 1}\right)^\xi \times \int_0^\infty \tau^{\xi-1-d} \left(\frac{\xi}{\zeta - 1} \tau + 1\right)^{-\xi-\zeta} \exp\left(-\frac{\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}}{\tau}\right) d\tau \tag{61}$$

Replacing τ by $\frac{\xi-1}{\xi} t^{-1}$, and using (A10) to calculate the integral results into the distribution of the scattering vector

$$p(\mathbf{k}; \xi, \zeta, \Sigma) = \frac{1}{\pi^d |\Sigma|} \frac{\Gamma(\xi + \zeta)\Gamma(\zeta + d)}{\Gamma(\xi)\Gamma(\zeta)} \left(\frac{\xi}{\zeta - 1}\right)^d \times U\left(d + \zeta, d - \zeta + 1, \frac{\xi}{\zeta - 1} \mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}\right) \tag{62}$$

where $U(a, b, z)$ is the hyper-geometric function of the second kind [34]. By the same procedure, the distribution of the sample covariance matrix can be obtained as

$$p(\mathbf{C}_L; \xi, \zeta, L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d} \Gamma(\xi + \zeta) \Gamma(\zeta + Ld)}{\Gamma_d(L) |\Sigma|^L \Gamma(\xi) \Gamma(\zeta)} \left(\frac{\xi}{\zeta - 1} \right)^{Ld} \times U \left(Ld + \zeta, Ld - \xi + 1, \frac{\xi}{\zeta - 1} L \text{Tr}(\Sigma^{-1} \mathbf{C}_L) \right) \tag{63}$$

As a matter of fact, Fisher distributions are the Pearson VI solutions and cover a large range of distributions. It is not only confined to urban scenes, but also fits reasonably in forest and agricultural fields [16,53]. The behavior of the head and tail of the distribution can be controlled by the two parameters ξ and ζ .

4.1.5. \mathcal{W} Distribution

The \mathcal{W} distribution assumes the texture to follow a beta distribution [48], which is given by [35]

$$p(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad x \in [0, 1] \tag{64}$$

The mean value of the beta distribution is $\mu = \frac{\alpha}{\alpha + \beta}$. Let $\tau = \frac{x}{\mu}$, $\xi = \alpha$, $\zeta = \alpha + \beta$, the distribution of the normalized texture can be written as

$$p(\tau; \xi, \zeta) = \frac{\Gamma(\zeta)}{\Gamma(\xi) \Gamma(\zeta - \xi)} \frac{\xi}{\zeta} \left(\frac{\xi}{\zeta} \tau \right)^{\xi-1} \left(1 - \frac{\xi}{\zeta} \tau \right)^{\zeta-\xi-1}, \quad \tau \in \left[0, \frac{\zeta}{\xi} \right] \tag{65}$$

The distribution of the scattering vector in this case can be calculated by

$$p(\mathbf{k}; \xi, \zeta, \Sigma) = \frac{\Gamma(\zeta)}{\Gamma(\xi) \Gamma(\zeta - \xi) \pi^d |\Sigma|} \left(\frac{\xi}{\zeta} \right)^{\zeta-1} \times \int_0^{\frac{\zeta}{\xi}} \tau^{\xi-1-d} \left(\frac{\zeta}{\xi} - \tau \right)^{\zeta-\xi-1} \exp \left(-\frac{\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}}{\tau} \right) d\tau \tag{66}$$

which leads to the following result according to the integral identity (A11)

$$p(\mathbf{k}; \xi, \zeta, \Sigma) = \frac{1}{\pi^d |\Sigma|} \frac{\Gamma(\zeta)}{\Gamma(\xi)} \left(\frac{\xi}{\zeta} \right)^{\frac{\xi+d-1}{2}} \left(\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k} \right)^{\frac{\xi-d-1}{2}} \times \exp \left(-\frac{\xi}{2\zeta} \mathbf{k}^\dagger \Sigma^{-1} \mathbf{k} \right) W_{\frac{d+1+\xi-2\zeta}{2}, \frac{\xi-d}{2}} \left(\frac{\xi}{\zeta} \mathbf{k}^\dagger \Sigma^{-1} \mathbf{k} \right) \tag{67}$$

where $W_{a,b}(z)$ is Whittaker W function [34]. The distribution of the sample covariance matrix can be obtained by the same way

$$p(\mathbf{C}_L; \xi, \zeta, L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d} \Gamma(\zeta)}{\Gamma_d(L) |\Sigma|^L \Gamma(\xi)} \left(\frac{\xi}{\zeta} \right)^{\frac{\xi+Ld-1}{2}} \left(L \text{Tr}(\Sigma^{-1} \mathbf{C}_L) \right)^{\frac{\xi-Ld-1}{2}} \times \exp \left(-\frac{\xi}{2\zeta} L \text{Tr}(\Sigma^{-1} \mathbf{C}_L) \right) W_{\frac{Ld+1+\xi-2\zeta}{2}, \frac{\xi-Ld}{2}} \left(\frac{\xi}{\zeta} L \text{Tr}(\Sigma^{-1} \mathbf{C}_L) \right) \tag{68}$$

4.1.6. \mathcal{M} Distribution

Another possible distribution for the texture is the beta prime distribution, also known as inverted beta distribution, with PDF given by [35]

$$p(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1+x)^{-\alpha-\beta}, \quad x > 0 \tag{69}$$

The mean value can be calculated by $\mu = \frac{\alpha}{\alpha + \beta}$. Again, scale the random variable to ensure the mean value is equal to 1 by letting $\tau = \frac{\beta - 1}{\alpha + \beta - 1}(1 + x)$, the above distribution becomes

$$p(\tau; \xi, \zeta) = \frac{\Gamma(\zeta)}{\Gamma(\xi)\Gamma(\zeta - \xi)} \frac{\zeta - 1}{\xi - 1} \left(\frac{\zeta - 1}{\xi - 1}\tau\right)^{-\zeta} \left(\frac{\zeta - 1}{\xi - 1}\tau - 1\right)^{\zeta - \xi - 1}, \quad \tau > \frac{\xi - 1}{\zeta - 1} \tag{70}$$

where the parameters are changed to $\zeta = \alpha + \beta$, $\xi = \beta$ to make the expression brief. Equation (70) is the texture distribution of the \mathcal{M} distribution [48]. According to the product model, the distribution of the scattering vector can be calculated by

$$p(\mathbf{k}; \xi, \zeta, \Sigma) = \frac{\Gamma(\zeta)}{\Gamma(\xi)\Gamma(\zeta - \xi)\pi^d |\Sigma|} \left(\frac{\xi - 1}{\zeta - 1}\right)^\xi \times \int_{\frac{\xi - 1}{\zeta - 1}}^\infty \tau^{-\zeta - d} \left(\tau - \frac{\xi - 1}{\zeta - 1}\right)^{\zeta - \xi - 1} \exp\left(-\frac{\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}}{\tau}\right) d\tau \tag{71}$$

Employing the integral identity (A12), we have the PDF of the scattering vector as

$$p(\mathbf{k}; \xi, \zeta, \Sigma) = \frac{1}{\pi^d |\Sigma|} \frac{\Gamma(\zeta)\Gamma(\xi + d)}{\Gamma(\xi)\Gamma(\zeta + d)} \left(\frac{\xi - 1}{\zeta - 1}\right)^d M\left(\xi + d, \zeta + d, -\frac{\xi - 1}{\zeta - 1} \mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}\right) \tag{72}$$

and the PDF of the sample covariance matrix as

$$p(\mathbf{C}_L; \xi, \zeta, L, \Sigma) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L) |\Sigma|^L} \frac{\Gamma(\zeta)\Gamma(\xi + Ld)}{\Gamma(\xi)\Gamma(\zeta + Ld)} \left(\frac{\xi - 1}{\zeta - 1}\right)^{Ld} \times M\left(\xi + Ld, \zeta + Ld, -\frac{\xi - 1}{\zeta - 1} L \text{Tr}(\Sigma^{-1} \mathbf{C}_L)\right) \tag{73}$$

Here $M(a, b, z)$ is the confluent hypergeometric function of the first kind, also known as the KummerM function [34]. The \mathcal{W} distribution and the \mathcal{M} distribution are able to model data with low variance but extreme skewness, which is particularly relevant to data with textural variability after a speckle filtering [48].

4.1.7. Wishart-Generalized Gamma Distribution

The Wishart-Generalized Gamma (WGF) distribution employs the generalized gamma distribution to model the texture. The generalized gamma distribution has a more compact form and a larger variety of alternative distributions, with the gamma, the Weibull, the Rayleigh, and the exponential distributions being its special cases. Thus it is of greater flexibility in the statistical modelling [54]. The PDF of the generalized gamma distribution is given by [35]

$$p(x; v, \theta, k) = \frac{v}{\theta \Gamma(k)} \left(\frac{x}{\theta}\right)^{kv-1} \exp\left(-\left(\frac{x}{\theta}\right)^v\right), \quad v > 0, \theta > 0, k > 0 \tag{74}$$

which reduces to the gamma distribution (41) when $v = 1$. The mean value is given by $\mu = \theta \Gamma(k + \frac{1}{v}) / \Gamma(k)$. Scaling the mean value to 1, the PDF for the texture is obtained as

$$p(\tau; v, k) = \frac{v \beta^{kv}}{\Gamma(k)} \tau^{kv-1} e^{-(\beta \tau)^v} \tag{75}$$

where $\beta = \Gamma(k + \frac{1}{v}) / \Gamma(k)$. The distribution of the scattering vector \mathbf{k} then can be calculated by

$$p(\mathbf{k}; v, k, \Sigma) = \frac{v \beta^{kv}}{\Gamma(k) \pi^d |\Sigma|} \int_0^\infty \tau^{kv-d-1} \exp\left(-(\beta \tau)^v - \frac{\mathbf{k}^\dagger \Sigma^{-1} \mathbf{k}}{\tau}\right) d\tau \tag{76}$$

There is no closed form expression for the above equation, but it can be solved numerically [54]. The distribution of the sample covariance matrix can be calculated by

$$p(\mathbf{C}_L; v, k, L, \mathbf{\Sigma}) = \frac{v\beta^{kv}L^{LD}|\mathbf{C}_L|^{L-d}}{\Gamma(k)I(L, d)|\mathbf{\Sigma}|^L} \int_0^\infty \tau^{kv-Ld-1} \exp\left(-(\beta\tau)^v - \frac{L \operatorname{Tr}(\mathbf{\Sigma}^{-1}\mathbf{C}_L)}{\tau}\right) d\tau \quad (77)$$

It is reported that the WGF distribution could provide better fitness than the \mathcal{K} and Kummer- \mathcal{U} distributions for different land cover types of homogeneous, heterogeneous, and extremely heterogeneous terrains [54].

4.1.8. Generalized \mathcal{K} Distribution

The well-known gamma distribution sometimes cannot fit the texture distribution accurately in very heterogeneous areas. In order to improve the flexibility of the model, it is assumed that the texture follows a Laguerre expansion of the gamma distribution [55], with its PDF given by

$$p(\tau; \alpha, \mu) = \frac{\tau^{\alpha-1}}{\Gamma(\alpha)} \left(\frac{\alpha}{\mu}\right) \exp\left(-\frac{\alpha\tau}{\mu}\right) \sum_{u=0}^\infty \xi_u \frac{\Gamma(\alpha)u!}{\Gamma(u+\alpha)} L_u^{\alpha-1}\left(\frac{\alpha\tau}{\mu}\right) \quad (78)$$

where μ , the mean value, is normally assumed to be equal to 1, and

$$\xi_u = \sum_{k=0}^u (-1)^k \binom{u+\alpha-1}{u-k} \frac{1}{k!} \left(\frac{\alpha}{\mu}\right)^k E\{x^k\} \quad (79)$$

The Laguerre polynomial $L_u^{\alpha-1}(x)$ is given by

$$L_u^{\alpha-1}(x) = \sum_{k=0}^u (-1)^k \binom{u+\alpha-1}{u-k} \frac{x^k}{k!} \quad (80)$$

The PDF of the sample covariance matrix in this case can be expressed as [55]

$$p(\mathbf{C}_L; \alpha, \mu, L, \mathbf{\Sigma}) = \frac{L^{Ld}|\mathbf{C}_L|^{L-d}}{\Gamma_d(L)|\mathbf{\Sigma}|^L} \frac{\alpha^\alpha}{\Gamma(\alpha)\mu^\alpha} \times \sum_{u=0}^\infty \xi_u \frac{\Gamma(\alpha)u!}{\Gamma(u+\alpha)} \sum_{k=0}^u (-1)^k \frac{2}{k!} \frac{(u+\alpha-1)!}{(u-k)!(\alpha-1+k)!} \left(\frac{\alpha}{\mu}\right)^k \left(\frac{L\mu \operatorname{Tr}(\mathbf{\Sigma}^{-1}\mathbf{C}_L)}{\alpha}\right)^{\frac{\alpha+k-Ld}{2}} K_{\alpha+k-Ld} \left(2\sqrt{\frac{\alpha}{\mu}} L \operatorname{Tr}(\mathbf{\Sigma}^{-1}\mathbf{C}_L)\right) \quad (81)$$

which is a weighted combination of a series of \mathcal{K} distributions based on a Laguerre polynomial expansion. It shows that the generalized \mathcal{K} distribution gives a better approximation than the \mathcal{K} distribution when there exist strong scatterers in the scene [55].

4.2. Multi-Texture Model

In the scalar texture model, different polarimetric channels are assumed to have a common texture variable. However, if the electromagnetic wave sees different geometrical or dielectric properties of the target, and if those properties are spatially modulated, then the texture of each channel should be different [56]. For example, in scattering from forest areas, volume scattering will affect the cross-pol component stronger than the co-pol channels, whereas surface scattering will have the opposite effect [57]. The scalar texture model must, therefore, be extended to take into consideration the different radar cross section modulations in polarimetric channels. One solution is to allow for a vector component of the texture in the product model. This type of models are called multi-texture models.

Under the assumption of reciprocity, there are only three independent complex coefficients required to characterize the scatterer under observation. The multi-texture model then can be formulated as [57–60]

$$\mathbf{k} = \mathbf{\Lambda}^{1/2} \mathbf{z} \tag{82}$$

where \mathbf{z} represents the speckle, following a multivariate Gaussian distribution (see Section 3.1), and $\mathbf{\Lambda}$ is a diagonal matrix containing texture variables for each channel

$$\mathbf{\Lambda} = \begin{bmatrix} \tau_{hh} & 0 & 0 \\ 0 & \tau_{hv} & 0 \\ 0 & 0 & \tau_{vv} \end{bmatrix} \tag{83}$$

The texture parameters are assumed to be positive, and we have $E\{\mathbf{\Lambda}\}$ equal to \mathbf{I} , the identity matrix. Assuming that the texture variables are constant on the scale of the multilook processing window, the sample covariance matrix can be written as

$$\mathbf{C}_L = \frac{1}{L} \sum_{i=1}^L \mathbf{k}_i \mathbf{k}_i^T = \mathbf{\Lambda}^{1/2} \mathbf{W} \mathbf{\Lambda}^{1/2} \tag{84}$$

where \mathbf{W} is Wishart distributed, see Section 3.2.

Provided that the distributions of the texture variables are known, the PDF of the scattering vector can be calculated using

$$p(\mathbf{k}; \mathbf{\Sigma}) = \int_{\Omega^+} p(\mathbf{k} | \mathbf{\Lambda}; \mathbf{\Sigma}) p(\mathbf{\Lambda}) d\mathbf{\Lambda} \tag{85}$$

where Ω^+ is the set of all diagonal matrices with non-negative entries. After changing variable by $\mathbf{z} = \mathbf{\Lambda}^{-1/2} \mathbf{k}$, the conditional distribution of \mathbf{k} on $\mathbf{\Lambda}$ can be obtained from (7), giving

$$p(\mathbf{k} | \mathbf{\Lambda}; \mathbf{\Sigma}) = \frac{1}{\pi^d |\mathbf{\Sigma}| |\mathbf{\Lambda}|} \exp \left(-\mathbf{k}^T \mathbf{\Lambda}^{-1/2} \mathbf{\Sigma}^{-1} \mathbf{\Lambda}^{-1/2} \mathbf{k} \right) \tag{86}$$

By the similar way, we have the distribution of the sample covariance matrix as [57,59]

$$p(\mathbf{C}_L; L, \mathbf{\Sigma}) = \int_{\Omega^+} p(\mathbf{C}_L | \mathbf{\Lambda}; L, \mathbf{\Sigma}) p(\mathbf{\Lambda}) d\mathbf{\Lambda} \tag{87}$$

where

$$p(\mathbf{C}_L | \mathbf{\Lambda}; L, \mathbf{\Sigma}) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L) |\mathbf{\Sigma}|^L |\mathbf{\Lambda}|^L} \exp \left(-L \text{Tr}(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}^{-1/2} \mathbf{C}_L \mathbf{\Lambda}^{-1/2}) \right) \tag{88}$$

Different texture variables for the multi-texture model can be: (1) totally dependent, in which case it reduces to the scalar texture model, (2) independent from each other, that is, texture variables follow different distributions with different parameters, or (3) partially correlated [58,61]. In many cases, it is reasonable to assume co-pol channels have the same texture but different from that of the cross-pol channels. This type of models is usually referred to as dual-texture model [57,59,62]. For reciprocal media with reflection symmetry for example, the PDF of the sample covariance matrix can be expanded as [59]

$$p(\mathbf{C}_L; L, \mathbf{\Sigma}) = \frac{L^{3L} |\mathbf{C}_L|^{L-3}}{I(L, 3) |\mathbf{\Sigma}|^L} \int_0^\infty \exp \left(-L \frac{q_{22} c_{22}}{T_x} \right) \frac{p(T_x)}{T_x^L} dT_x \times \int_0^\infty \exp \left(-L \frac{q_{11} c_{11} + q_{13} c_{31} + q_{31} c_{13} + q_{33} c_{33}}{T_{co}} \right) \frac{p(T_{co})}{T_{co}^{2L}} dT_{co} \tag{89}$$

where q_{ij} and c_{ij} denote the (i, j) th entry of matrix \mathbf{C}_L and $\mathbf{\Sigma}$ respectively. The texture of the co-pol channels is represented by T_{co} and that of the cross-pol channel by T_x .

4.2.1. Correlated \mathcal{K} Distribution

The correlated \mathcal{K} distribution assumes that the texture variables of different polarimetric channels are partially correlated, each following a gamma distribution [58,61]. Unfortunately, there is no explicit expression for the joint distribution of the texture variables, or the correlated gamma distribution. In this model, the texture of polarimetric channel i , specified by the PDF (42) with parameter α , is given by [61]

$$\tau_i = \frac{1}{2\alpha} \sum_{k=1}^{2\alpha} [g_i^{(k)}]^2 \quad (90)$$

where $g_i^{(k)}$ is the i th element of the vector $\mathbf{g}^{(k)}$, $k = 1, \dots, 2\alpha$, which is Gaussian distributed with zero mean, variance one, and correlation matrix \mathbf{T} . The correlation properties of the texture variables is also specified by \mathbf{T} . The characteristic function of the vector containing all texture variables is [61]

$$C(\boldsymbol{\omega}) = \frac{1}{|\mathbf{I} + j(1/\alpha)\mathbf{TW}|^\alpha} \quad (91)$$

where \mathbf{W} is a diagonal matrix having the entry (i, i) equal to the i th element of the characteristic function variable $\boldsymbol{\omega}$. This model requires that all polarimetric channels have the same half-integer distribution parameter α , e.g., 0.5, 1.5, 2.5 and so on.

4.2.2. Dual-Texture \mathcal{G} Distribution

The dual-texture \mathcal{G} distribution is derived by considering different texture variables for co-pol and cross-pol channels. Both the co-pol and the cross-pol texture variables are modelled by the GIG distributions (49), which yields a more flexible multivariate distribution [62]. Under the assumption of reciprocity and reflection symmetry, the statistical properties of the single look complex data is characterized by the distribution [62]

$$p(\mathbf{k}; \boldsymbol{\Sigma}, \boldsymbol{\theta}) = \frac{1}{\pi^d |\boldsymbol{\Sigma}|} \prod_{i=1}^2 \frac{(\eta_i^2 + 2\eta_i s_i / \omega_i)^{\frac{p_i - d + i}{2}}}{\eta_i^{p_i} K_{p_i}(\omega_i)} K_{p_i - d + i} \left(\sqrt{\omega_i^2 + 2\omega_i s_i / \eta_i} \right) \quad (92)$$

where $\boldsymbol{\theta} = \{\omega, p_i, \eta_i\}$ consists of all parameters for the GIG texture distributions (see Section 4.1.3), $s_1 = z_{11}c_{11} + z_{13}c_{31} + z_{31}c_{13} + z_{33}c_{33}$, and $s_2 = z_{22}c_{22}$, with z_{ij} and c_{ij} indexing entries of $\mathbf{Z} = \mathbf{k}\mathbf{k}^\dagger$ and $\boldsymbol{\Sigma}$ respectively.

5. Other Models

To model a complex scene using texture models, we often need to introduce complicated distributions with many parameters to describe the statistical behavior of the texture component. However, having more parameters requires a more complicated estimation process by considering higher order statistics. In addition, higher order moment estimators are known to have higher variance. With the limited sample sizes used in the modelling, such complicated modelling may be very inefficient [50]. To overcome this problem, some researchers try to divide a complex model into multiple simple components and then find a way to combine these components together. The finite mixture model and copula based model detailed as follows are based on this idea.

5.1. Finite Mixture Model

The heterogeneity that appears in PolSAR data may result from the mixture of different targets. For instance, from an urban area which usually consists of different objects like houses, trees and roads,

the backscattering is a combination of different scattering mechanisms. The forest areas sometimes can be treated as a mixture of bright clutters and dark ones, corresponding to the strong returns from the crowns of trees and the shadows behind them. To represent this type of data, a simple model would be inappropriate. Finite mixture models, instead, could achieve reasonable level of accuracy [17–19,63,64].

Assume that the region under analysis can be modeled by a mixture of K components, then the overall PDF of the data can be written as a weighted sum of the probabilities of each component [65]

$$p(\mathbf{x}; \boldsymbol{\theta}) = \sum_{k=1}^K w_k p_k(\mathbf{x}; \boldsymbol{\theta}_k) \quad (93)$$

where $\boldsymbol{\theta}$ is a vector containing all the parameters of the distribution and the mixing proportions obey

$$\sum_{k=1}^K w_k = 1, \quad w_k \geq 0 \quad (94)$$

It has been shown that for complicated regions with more irregular histograms (multimodal, spiky), the finite mixture model is more accurate than a single distribution [17–19].

There are many options for the distributions of the mixing components, but here we mainly focus on the mixture of Wishart distributed components. For different mixing components, the number of looks are the same. The PDF, therefore, can be written as

$$p(\mathbf{C}_L; L, \boldsymbol{\theta}) = \frac{L^{Ld} |\mathbf{C}_L|^{L-d}}{\Gamma_d(L)} \sum_{k=1}^K \frac{w_k \exp(-L \operatorname{Tr}(\boldsymbol{\Sigma}_k^{-1} \mathbf{C}_L))}{|\boldsymbol{\Sigma}_k|^L} \quad (95)$$

where $\boldsymbol{\theta} = \{\boldsymbol{\Sigma}_k, k = 1, \dots, K\}$ and $\Gamma_d(L)$ is given by (22). The PDF of the i th channel intensity, which is also a finite mixture, is found to be

$$p(I_i; L, \boldsymbol{\theta}) = \frac{I_i^{L-1}}{\Gamma(L)} \sum_{k=1}^K \left(\frac{L}{\sigma_{k,i}^2} \right)^L \exp \left(-\frac{L}{\sigma_{k,i}^2} I_i \right) \quad (96)$$

where $\sigma_{k,i}^2 = [\boldsymbol{\Sigma}_k]_{ii}$. The most interesting property of a mixture density is that the shape of the density is extremely flexible. A mixture density may be multimodal, or even if it is unimodal, may exhibit considerable skewness or additional humps. For this reason, finite mixture distributions offer a flexible way to describe rather heterogeneous data by summarizing the characteristics of the data in terms of the number and the spread of the mixture components [65].

5.2. Copula Based Model

Copulas are popular in high-dimensional statistical applications as they allow one to easily model and estimate the distribution of random vectors by estimating marginals and dependence separately [21]. They are of great interest for two main reasons: (1) as a way to study scale-free measures of dependence; and (2) as a starting point for constructing families of multivariate distributions [21]. For the PolSAR data, we often have a much better idea about the marginal behaviour of individual polarimetric channels than we do about their dependence structure. The copula approach allows us to combine our more developed marginal models with a variety of possible dependence models to investigate the statistical behavior of the data.

A d -dimensional copula, denoted by $C(\mathbf{u}) = C(u_1, \dots, u_d)$, is a joint Cumulative Distribution Function (CDF) of a d -dimensional random vector on the unit hypercube $[0, 1]^d$ with uniform marginals. More specifically, a copula is a function C from $[0, 1]^d$ to $[0, 1]$ with the following properties [21,66]:

1. $C(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0$, the copula is equal to 0 if at least one parameter is 0.
2. $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$, the copula is equal to u_i if all parameters are 1 except u_i .
3. For each hyperrectangle $B = \prod_{i=1}^d [x_i, y_i] \subseteq [0, 1]^d$ where $x_i \leq y_i$, the C -volume of B is non-negative

$$\sum_{\mathbf{z} \in \times_{i=1}^d \{x_i, y_i\}} (-1)^{N(\mathbf{z})} C(\mathbf{z}) \geq 0 \quad (97)$$

where \mathbf{z} represents the corners of the hyperrectangle, and $N(\mathbf{z}) = \#\{k : z_k = x_k\}$ is the number of elements in \mathbf{z} reaching the lower bound of the hyperrectangle.

According to Sklar's Theorem, any multivariate joint distribution can be written in terms of univariate marginal distribution functions and a copula which describes the dependence structure between the variables [21]

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (98)$$

where F_i is the continuous marginal CDF $F_i(x) = P(X_i \leq x)$. The copula C contains all information about the dependence structure whereas the marginal cumulative distribution functions F_i contains all information about an individual random variable.

There are many parametric copula families available, which usually have one or more parameters controlling the strength of dependence. The most popular ones include the elliptical copulas (such as the Gaussian copula and the student t copula), and the Archimedian copulas. In the context of PolSAR data modeling, the Ali-Mikhail-Haq copula which belongs to the Archimedian family is demonstrated to be appropriate [22,23,67]. The Gaussian copula is also found to be proper to model the wavelet coefficients [68]. Though it is a hot topic, the study of copulas and the role they play in statistics and stochastic processes is a subject still in its infancy. There are many open problems and much work to be done.

6. Model Analysis

In the previous sections, the statistical models proposed for the PolSAR data are reviewed, with an emphasis on the derivation of PDFs for the scattering vector and the sample covariance matrix. The models are categorized into three groups: (1) Gaussian Models, (2) Texture Models, and (3) The Others. Table 1 shows a summary of all these models. As we can see, texture models are still the main focus in statistical modeling of PolSAR data. Several examples of the texture distributions with different distribution parameters are plotted in Figure 1.

In the remaining of this section, we will show some experimental results on the applicability of different statistical models.

First of all, two homogeneous Regions Of Interest (ROI) over the farmland of a RADARSAT-2 image are analyzed, as shown in Figure 2. The data, in single look complex format, has a spatial resolution of 11.1 m \times 7.6 m (Range \times Azimuth). It was acquired over Flevoland (The Netherlands) with the Fine Quad-Pol mode during the ESA-led AgriSAR 2009 campaign. Statistical properties are analyzed separately by the histograms of the intensity, the product of amplitudes and the phase difference between two polarimetric channels. To tell whether Gaussian distributions are proper or not, the histograms are compared with the PDFs defined by (12), (18) and (20). The covariance matrices of the Gaussian distributions are estimated using the simple mean estimator.

Figure 2 shows the fit of the HH intensity, and the fit of the product of HH Channel and HV channel. It demonstrates that the histograms conform to the corresponding PDFs, implying that Gaussian distributions are suitable for these crops areas. Though it could work, the comparison of histogram and PDF is not visually effective, see Figure 2f,g for example. So in the next experiments we will try different methods to validate the applicability of statistical models.

As mentioned in the previous sections, the spatial resolution of PolSAR images is one of the most important factors that have strong impact on the data statistics. To demonstrate this, real SAR data including a RADARSAT-2 Fine Quad-Pol data (RST2) as well as a F-SAR X-band full-pol data (FSAR)

are analyzed. The two data have quite different spatial resolutions, $11.1 \text{ m} \times 7.6 \text{ m}$ for the RST2 data, and $0.25 \text{ m} \times 0.25 \text{ m}$ for the FSAR data. Three ROIs over the crops area from each data are tested, see Figure 3a,b. For the RST2 data, each ROI covers 50×50 pixels. The ROIs in the FSAR data are much larger thanks to a higher spatial resolution, each covering 200×200 pixels. The Pauli decomposition shows that the ROIs in both images are very homogeneous, no appreciable texture is observed.

Table 1. Summary of statistical PolSAR data models.

Category	Model	PDF	References	Summary
Gaussian	Gaussian	(7)	[31,33]	Simple, high mathematical tractability, suitable for data of low or moderate spatial resolution.
	Wishart	(21)	[31–33]	
	Relaxed Wishart	(21)	[39]	More flexible than the Wishart distribution, but assigning different values to the number of looks L is not so convincing.
	Wishart-Kotz	(31)	[40,41]	With ability to model heavy tail behaviors, computationally efficient and numerically stable, but at the expense of adding two more parameters.
Texture Models	\mathcal{K}	(43), (44)	[4,7,10]	Suitable for non-Gaussian data, widely used to model forest, ocean and so on, strong physical background.
	NIG	(47), (48)	[49,50]	Large shape variations, strong theoretical grounds derived from Brownian motion.
	\mathcal{G}	(52), (53)	[14,15,52]	Able to model different types of texture, but requires more parameters (two parameters).
	\mathcal{G}^0	(57), (58)	[14,15]	Suitable for extremely heterogeneous data, no complex special function involved.
	Kummer- \mathcal{U}	(62), (63)	[16,53]	Able to model different types of texture, but requires more parameters (two parameters), texture distribution belongs to Pearson family.
	\mathcal{W}	(67), (68)	[5]	Able to model data with low variance but extreme skewness, e.g., textured data after speckle filtering.
	\mathcal{M}	(72), (73)	[5]	Of great flexibility (generalization of many other distributions), but the PDF needs to be calculated numerically.
	WGT	No Explicit	[54]	
	Generalized \mathcal{K}	(81)	[55]	Good approximation of data when there exist strong scatterers, very complex PDF with polynomial expansions.
	Correlated \mathcal{K}	No Explicit	[58,61]	Able to model texture correlations of different channels, no explicit expression for the texture variables, distribution parameters are limited to specific values.
	Dual-Texture \mathcal{G}	(92)	[62]	Different texture distributions for the co-pol and the cross-pol channels.
	Others	Finite Mixture	(93)	[17–19]
Copula Based		No Explicit	[22,67]	Divides complex multivariate distributions into marginal distributions and dependence structure, and analyze them separately, but it is not very straightforward to choose the best copulas.

Normalized Intensity Moments (NIM) are employed to determine whether the Gaussian distributions are suitable for the test ROIs or not [9,69]. Let I denote the intensity of a polarimetric channel, the NIM of the v th order is defined as

$$NIM_v = \frac{E\{I^v\}}{E^v\{I\}} \quad (99)$$

For Gaussian distributed data, the intensity will follow an exponential distribution as defined in (12). The NIMs in this case are independent of the data, which can be calculated by

$$NIM_v^* = \Gamma(v + 1) \tag{100}$$

By comparing the estimated values from the data with those of Gaussian distributions, we can easily make conclusions on the applicability of Gaussian distributions.

The HH channel is analyzed for both the RST2 data and the FSAR data. Results are shown in Figure 3c,d, where black lines represent theoretical values of the exponential distribution and different markers represent values estimated from the test ROIs. As it can be seen, the NIMs estimated from the RST2 data fit those calculated from the exponential distribution very well. Same results can be obtained for the HV channel and the VV channel. It is rational to conclude that these ROIs can be modeled by Gaussian distributions. In contrast, the result on the FSAR data shows different behaviors. There are large discrepancies between the estimated values and the theoretical values for all ROIs. Apparently, Gaussian distributions are not accurate any more.

A further validation on the FSAR data is performed. Assuming that the intensity of each ROI can be modeled by a Weibull distribution, then the distribution parameter, denoted by γ , can be estimated using the first order moment. Furthermore, the NIM of the v th order can be computed by

$$NIM_v^\dagger = \frac{\Gamma\left(1 + \frac{v}{\gamma}\right)}{\Gamma^v\left(1 + \frac{1}{\gamma}\right)} \tag{101}$$

In Figure 3e, the estimated NIMs (markers) and those calculated using the above equation (lines) are plotted for each ROI. The Weibull distribution seems to be applicable in ROI 2 and ROI 3. Compared with the exponential distribution, the Weibull distribution could capture larger variance by introducing an additional distribution parameter. However, even the Weibull distribution could not give an accurate representation for ROI 1. Complex distributions with more parameters may achieve reasonable fit.

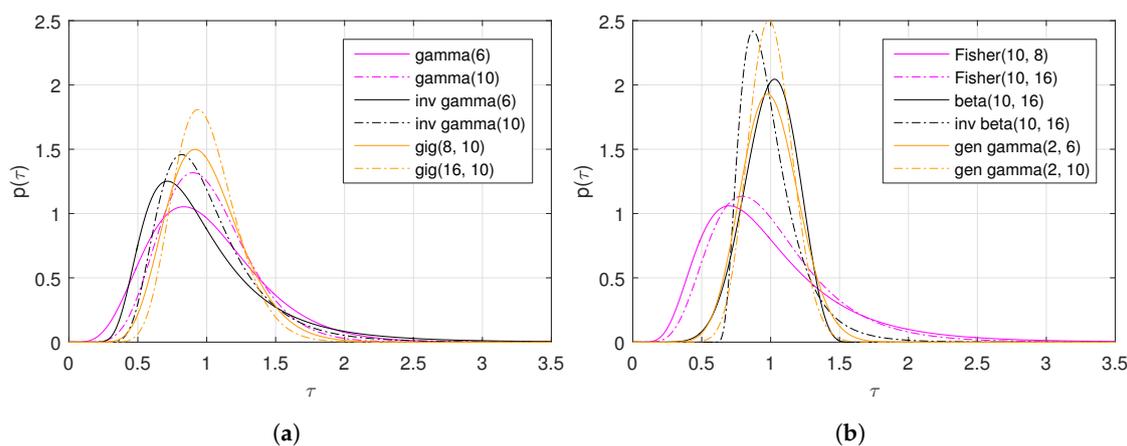


Figure 1. Examples of different texture distributions. (a) PDFs of the gamma (gamma(6) and gamma(10)), the inverse gamma (inv gamma(6) and inv gamma(10)) and the GIG (gig(8, 10) and gig(16, 10)) distributions. (b) PDFs of the Fisher (Fisher(10, 8) and Fisher(10, 16)), the beta (beta(10, 16)), the inverted beta (inv beta(10, 16)), and the generalized gamma (gen gamma(2, 6) and gen gamma(2, 10)) distributions.

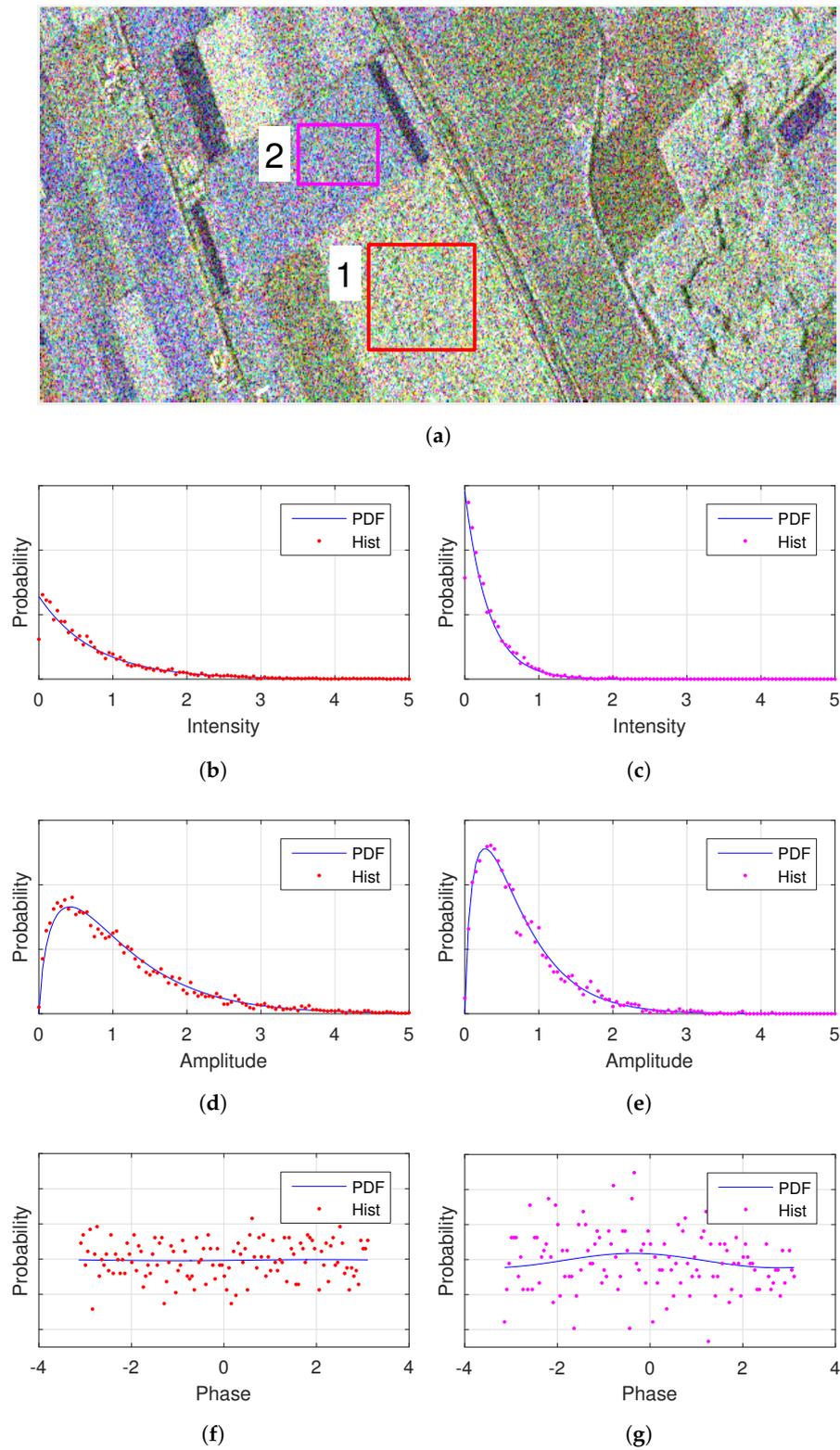


Figure 2. Histograms of two homogeneous areas in a RADARSAT-2 image and the PDFs under Gaussian assumption. Parameters of the Gaussian distributions are estimated using moments. (a) Pauli decomposition of the RADARSAT-2 data and two ROIs. (b,c) Intensity of the S_{hh} . (d,e) Amplitude of $S_{hh}S_{hv}$. (f,g) Phase of $S_{hh}S_{hv}$.

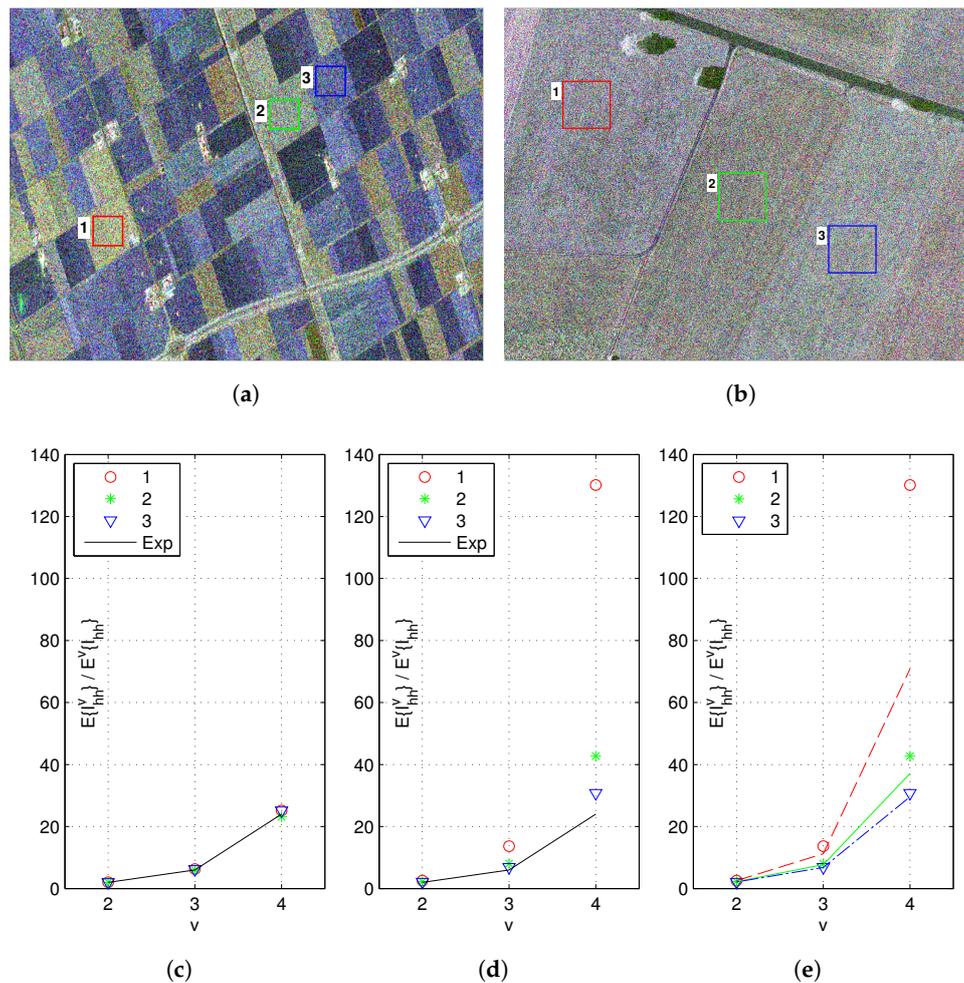


Figure 3. NIMs of the 2nd, the 3rd and the 4th order estimated from three crops areas in a RADARSAT-2 data and a F-SAR data. (a) Pauli decomposition and ROIs of the RADARSAT-2 image. (b) Pauli decomposition and ROIs of the F-SAR image. (c) NIMs of the ROIs in the RADARSAT-2 data and the exponential distribution. (d) NIMs of the ROIs in the F-SAR data and the exponential distribution. (e) NIMs of the ROIs in the F-SAR data and Weibull distributions.

In general, the Weibull distribution is advisable to model the intensity of high resolution single channel data. However, for PolSAR data, the correlations between different polarimetric channels convey useful information, besides the intensities. In order to describe the statistical behavior correctly, copulas (introduced in Section 5.2) can be adopted. By modeling the dependence structure between polarimetric channels using copulas, and the intensities by Weibull distributions, a good approximation of the data could be expected. However, how to choose the proper copulas needs to be investigated intensively. We haven't found a copula capturing the dependency properly for the testing ROIs.

Another aspect that causes non-gaussianity in PolSAR data is the fluctuation of radar cross section due to the change of surface properties, e.g., height of the observing surface. Usually, this type of data should be modeled using texture distributions. To validate the applicability of texture models, two PolSAR images, the RADARSAT-2 Fine Quad-Pol data (RST2) and the ALOS-2 level 1.1 Full-Pol data (ALOS2), are analyzed. Both images were acquired over Barcelona (Spain) with similar incidence angles. The spatial resolution is different, $11.1 \text{ m} \times 7.6 \text{ m}$ (Range \times Azimuth) for the RST2 data and $3.49 \text{ m} \times 3.84 \text{ m}$ for the ALOS2 data, respectively. Original data are in the single look complex format, from which the sample covariance matrices are obtained after applying a multilook processing.

We have selected ROIs locating at similar positions in the urban area, in the agriculture area, in the sea and the forest areas. Pauli decomposition of the test data and ROIs are shown in Figure 4.

Matrix variate log-cumulants [42] are used to examine the suitability of texture models. The matrix variate log-cumulants are of great value for the analysis of sample covariance matrix, and that they can be employed to derive estimators for the distribution parameters with low bias and variance [42]. Different from the NIMs, there is no need to study each polarimetric channel separately with matrix variate log-cumulants. Define the Mellin kind matrix-variate characteristic function as the Mellin Transform of the PDF

$$\phi(s) = \int_{\Omega_+} |\mathbf{Z}|^{s-d} p(\mathbf{Z}) d\mathbf{Z} \quad (102)$$

then, the v th-order log-cumulant, or Mellin kind cumulant, is given by

$$\kappa_v = \left. \frac{d^v}{ds^v} \ln \phi(s) \right|_{s=d} \quad (103)$$

Meanwhile, the sample log-cumulants can be estimated from the data using

$$\hat{\kappa}_v = \hat{\mu}_v - \sum_{i=1}^{v-1} \binom{v-1}{i-1} \hat{\kappa}_i \hat{\mu}_{v-i} \quad (104)$$

where $\hat{\mu}_v$ is the estimated log-moments

$$\hat{\mu}_v = \frac{1}{M} \sum_{i=1}^M (\ln |\mathbf{C}_i|)^v \quad (105)$$

with M denoting the number of samples and \mathbf{C}_i the i th sample covariance matrix.

To see if a texture model is applicable, we can compare the log-cumulants calculated from the PDF (κ_v) and those estimated from the sample data ($\hat{\kappa}_v$). In [42], a diagram is proposed to visualize the comparison by plotting the second order log-cumulants κ_2 against the third order log-cumulants κ_3 in a plane, where different distributions place in different regions, as shown in Figure 4c,d. In this diagram, estimated log-cumulants are represented by the "+" markers (values from different ROIs are distinguished by various colors), and theoretical values of different texture distributions are represented by curves (the \mathcal{K} and the \mathcal{G}^0 distributions) as well as regions (the Kummer- \mathcal{U} , the \mathcal{M} and the \mathcal{W} distributions).

From Figure 4, we can see that the urban areas (red and green rectangles) can be modeled by the \mathcal{G}^0 or the Kummer- \mathcal{U} distributions, which have the capability to model heterogeneous areas. The two ROIs in urban area represent two different urban structures, one is of tall and densely distributed apartments, the other is of short and sparse houses. This may be an explanation as to why different statistics, the \mathcal{G}^0 vs the Kummer- \mathcal{U} , are obtained. In agriculture areas (cyan and yellow rectangles), \mathcal{K} distribution is shown to be the most suitable model. The forest area (black rectangle) shows weak texture in the RST2 data. In the ALOS2 data, there is a strong fluctuation in the backscattering due to the radar foreshortening. To eliminate the effect of radar image distortions, another forest region (purple rectangle) is analyzed, which is found to follow a \mathcal{K} distribution. In most cases, texture is not observed in the sea areas.

As explained in Section 5.1, the finite mixtures could also give rise to non-gaussian statistics. To further distinguish textures from mixtures, higher order log-cumulants are required [70]. A large number of samples are demanded in order to estimate the higher order log-cumulants correctly. There are only 20×20 pixels in each of the previous ROIs, not enough to obtain a satisfying estimation of the fourth order log-cumulants. So another experiment is carried out on an airborne SAR data, a UAVSAR image.

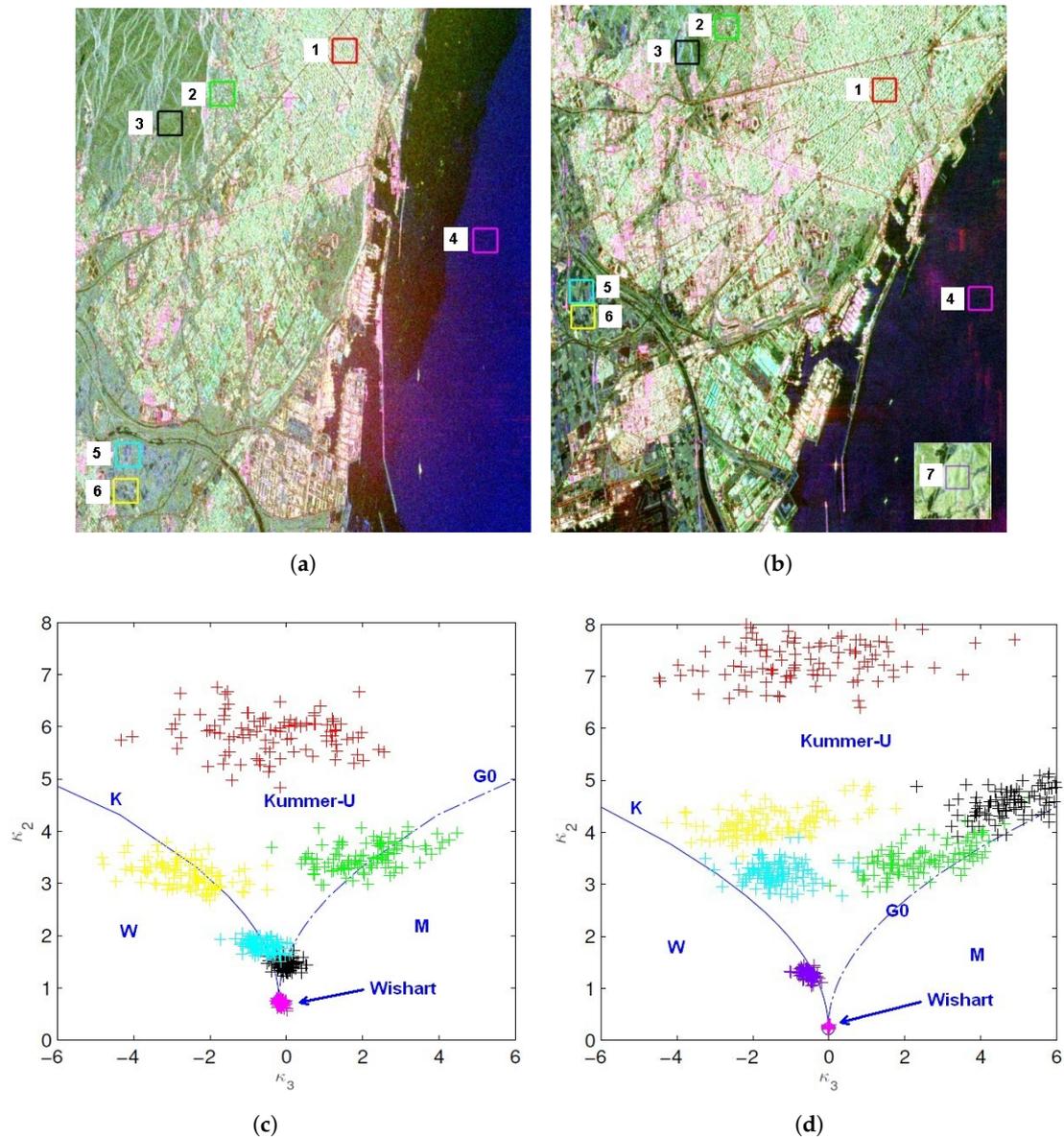


Figure 4. Matrix variate log-cumulants of the 2nd and the 3rd order estimated from a RADARSAT-2 data and an ALOS-2 data. Theoretical values calculated from the \mathcal{K} , the \mathcal{G}^0 , the Kummer- \mathcal{U} , the \mathcal{W} , the \mathcal{M} and the Wishart distributions are also plotted as references. (a) ROIs of the RADARSAT-2 data. (b) ROIs of the ALOS-2 data. (c) Matrix variate log-cumulants of the RADARSAT-2 data. (d) Matrix variate log-cumulants of the ALOS-2 data.

The test site is in the West Panhandle of Florida (USA), and the data is in the multilook cross-product slant range format, with number of looks in the range dimension and azimuth dimension equal to 3 and 12 respectively. The ENL is estimated as 12.73 over a homogeneous sea area. Four ROIs covering land types of ocean area (ROI 1), forest (ROI 2), wetland (ROI 3), and urban area (ROI 4), are tested, see Figure 5. Thanks to a higher spatial resolution, $1.67 \text{ m} \times 0.8 \text{ m}$ (Range \times Azimuth), each ROI contains 90×70 pixels, much more samples than those in the previous experiment.

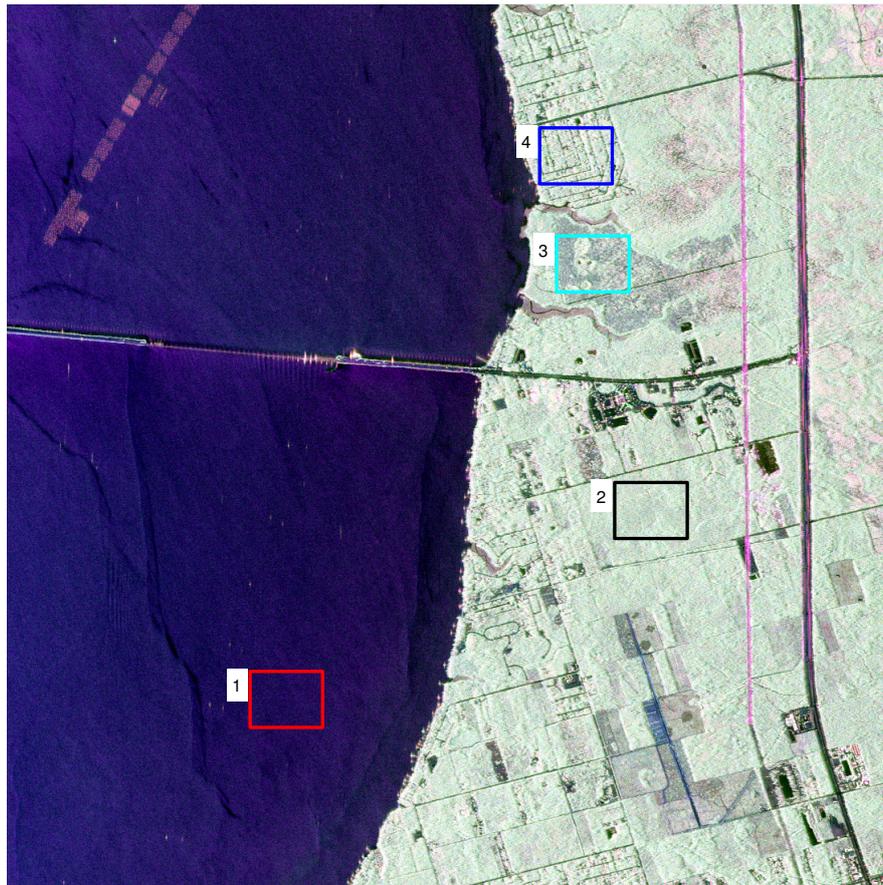


Figure 5. Test regions on the UAVSAR data. Four ROIs over different land types are tested, including the sea, the forest, the wetland and the urban area.

The log-cumulants of the second order, the third order and the fourth order are calculated. From the log-cumulant diagrams (Figure 6a,b), we can see that different ROIs show different statistical behaviors. The ocean area can be modeled by a Wishart distribution, and the forest by a \mathcal{K} distribution. The wetland and the urban area are very heterogeneous, especially the urban area, which has a very small κ_3 . The point clouds representing estimated statistics are less widely spread than those in Figure 4c,d. This is because more samples are used to estimate the values.

The fourth order log-cumulant is considered to further discriminate the texture from mixture. As shown in Figure 6c,d, the log-cumulants of major texture models can construct a smooth surface, while those of the finite mixture model will lie below this surface. The results show that texture models are proper for the sea area and forest area, while a finite mixture model make a better representation than a texture model for the wetland area and the urban area, because the point clouds estimated from ROI 1 and ROI 2 are on the product model surface, whereas those from ROI 3 and ROI 4 are below it. Actually, the Pauli decomposition in Figure 5 shows that the first two ROIs are very homogeneous and ROI 3 consists of different targets. Urban area, made up of distributed targets and point targets usually, has very large variance. This can be verified by the log-cumulant cube in Figure 6d, where both of the absolute values of κ_3 and κ_4 are very large. The estimation of the number as well as the weights of mixing components needs to be further studied.

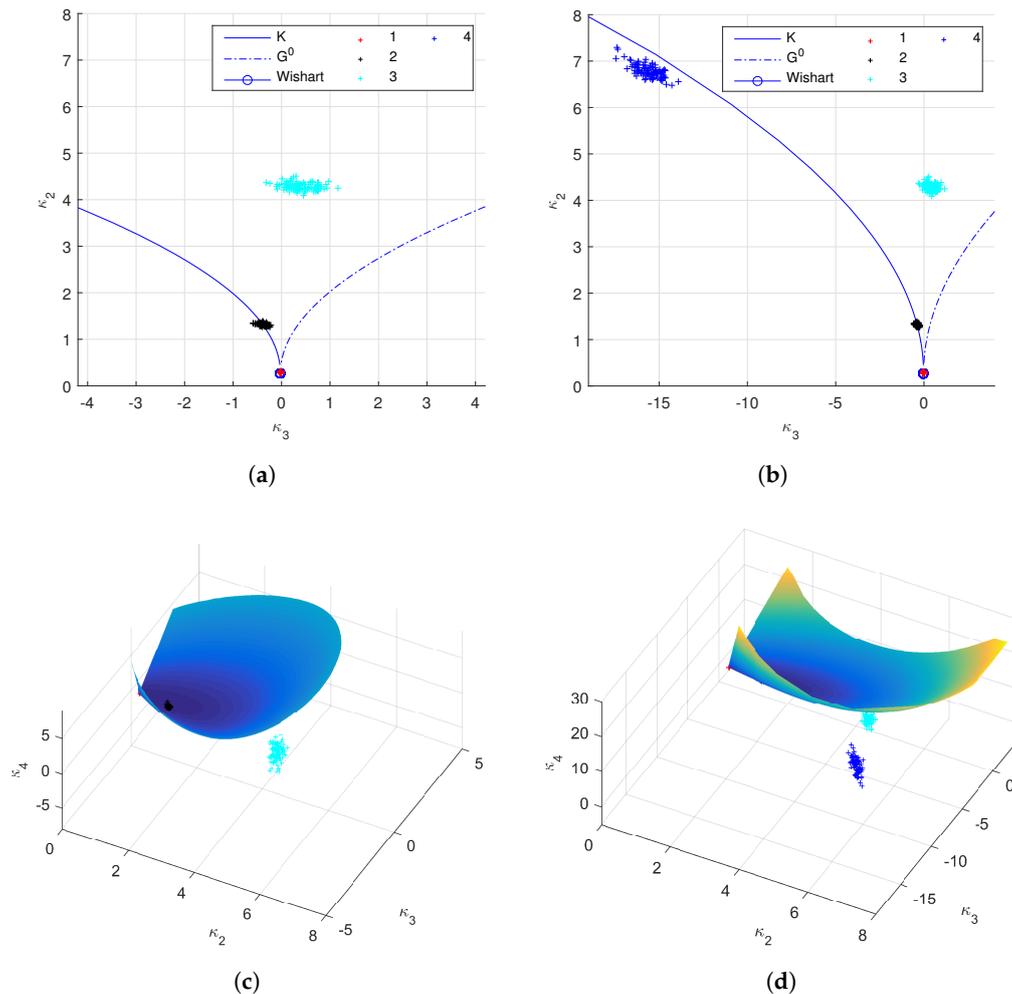


Figure 6. Log-cumulants of the 2nd, the 3rd and the 4th orders on the UAVSAR data. The right column and the left column are the same results but with different axes limits. (a,b) Log-cumulants of the 2nd and the 3rd order. ROIs over different ground targets show different statistics. (c,d) Log-cumulant up to the 4th order. It shows some ROIs can be modeled by texture models, while others should be represented using finite mixture model.

At last, statistical properties of the sea area at two different conditions are examined. One is with smooth surface, and the other with waves, as shown in Figure 7. Both data are acquired by RADARSAT-2 at C-band, and they have similar spatial resolutions. To study the textures of different polarimetric channels, the intensity of each channel is checked separately using the scalar log-cumulants [42,71]. The second order and the third order log-cumulants are employed as before. From Figure 7, we can see that in the first case, no texture is observed in all polarimetric channels. This can be also validated by the Pauli decomposition. When there exist sea waves, however, the log-cumulants are quite different. The HH channel and the VV channel have similar statistics, but different from those of the HV channel. In other words, multi-texture is observed in the test area. The result supposes that we can model the data using a dual-texture model in which the co-pol channels share a same texture distribution and the cross-pol channel with another one. The correlation between different textures needs to be further investigated to see if a partially correlated texture distribution is required.

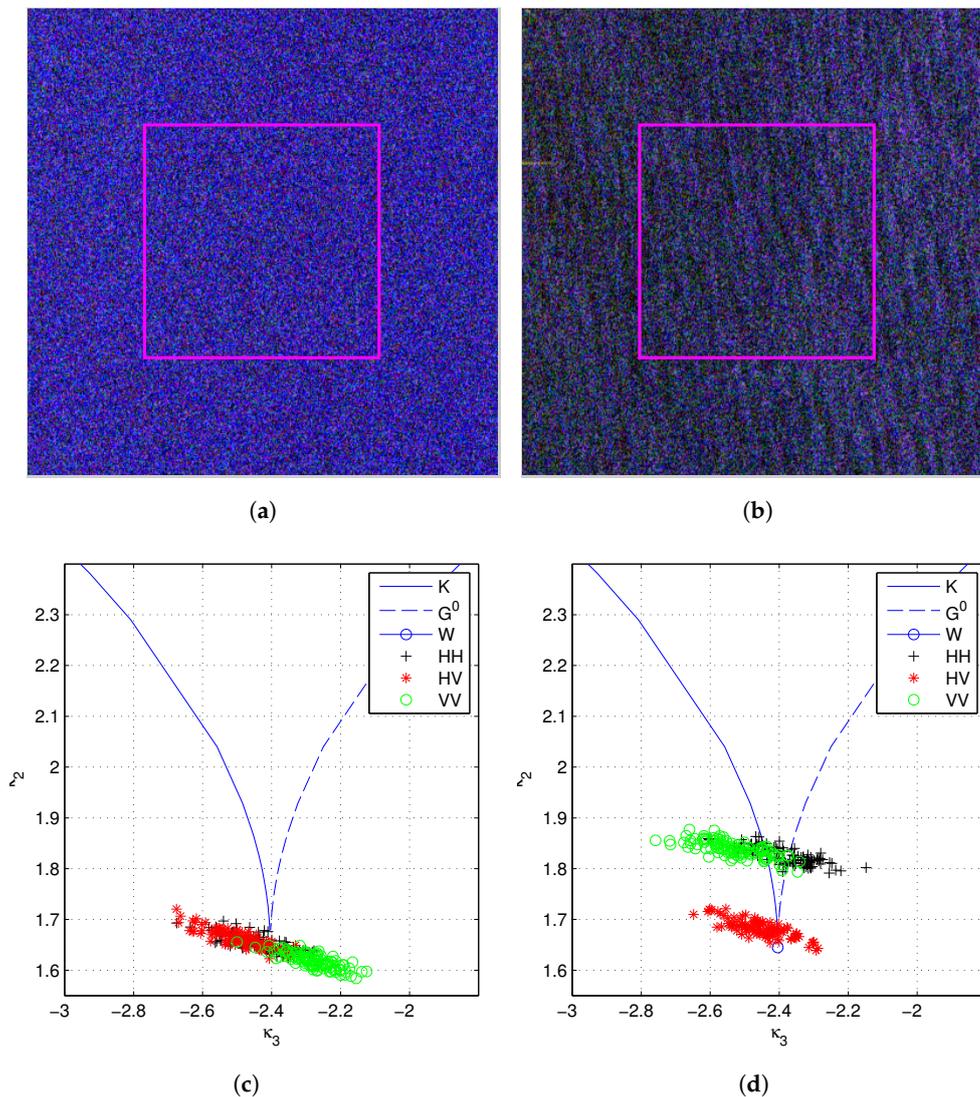


Figure 7. Scalar log-cumulants of different polarimetric channels (the HH, the HV and the VV channels) over sea areas. (a) Test region over smooth sea surface. (b) Test region over sea surface with waves. (c) Log-cumulants of the smooth sea surface. (d) Log-cumulants of the sea surface with waves.

7. Challenges

When the spatial resolution is not very high and the data is very homogeneous, the Gaussian distributions could provide a good representation of the data. As the spatial resolution increases, PolSAR data usually show non-Gaussian behaviour, e.g., exhibiting heavy tails. The texture models, which adopt additional random variables to model the spatial variation of the radar cross section, are found to be accurate for this kind of data. Texture models could model the non-Gaussian behavior observed in high resolution data, and yet keep a compact mathematical form. However, to model textures over complex scenes, sophisticated distributions are generally required. In addition, they are known to present problems in estimating parameters accurately. General distributions that cover a wide range of other distributions are suggested by many researchers. However, they usually have a complex form. Using several simple distributions from a certain family, Pearson family for example, could be a better idea. The distributions from a same family often have similar behaviors, but can be further distinguished by statistics like higher order log-cumulants [70].

In the product model as shown in (33) or (34), positive random variable following any distribution can be employed to model the texture. Additionally, PDFs of the scattering vector or the sample

covariance matrix can be obtained subsequently after mathematical calculations. The PDFs, however, give no information about why the data following a specific distribution is obtained. Most of the texture models lack a physical explanation of the underlying scattering process. A possible way to solve this problem could be the random walk model, which treats the received signal as an addition of responses from all the scatterers in the same resolution cell [9,12,72]. The random walk model can be interpreted as a discrete analog of the SAR focusing process.

Texture information has been used in optical image processing for a long time. In SAR or PolSAR images, it is also found to be useful to distinguish different target types. For example, trees of different heights can be distinguished by texture information [73]. However, currently the most common way to make use of texture models is to design probability based algorithms (e.g., classification and segmentation) by replacing the Gaussian distribution or the Wishart distribution [4–6]. How to extract texture information and let PolSAR applications benefit from it is not involved. Apparently, combining polarimetric information and texture information could improve the performance of applications since more knowledge is introduced. Therefore, a further study in this aspect will be of great value.

Besides texture models, there are non-Gaussian models subdividing complicated distributions into components, each with a simple distribution. For example, the finite mixture model treats a distribution as a weighted sum of those of different target types. In addition, the copula based model divides a multivariate distribution into marginal distributions and a general dependence structure. In the finite mixture model, robust algorithms to estimate the mixing number and the mixing weights are in urgent need. For the copula based method, we have many options for the marginal distributions. However, for the dependence structure, not many experiments were implemented to show which copula is the best. Additionally, it is a big challenge to extend the bivariate copulas which are intensively investigated in the field of statistical analysis to multivariate ones to fit the PolSAR data.

The statistical properties of PolSAR data are characterized totally by the PDFs of the scattering vectors or the sample covariance matrices. However, it is difficult to use these PDFs directly because they are multivariate ones. Normally, the statistics of each polarimetric channel are studied separately, and the correlation between different polarimetric channels are neglected. Another way is to analyze the determinants of sample covariance matrices. The widely used matrix variate log-cumulant is an example. However, we need to filter the data (the multilook process) to obtain the sample covariance matrices, which could change the actual statistical properties of the data. To overcome these problems, the l_2 -norms of the scattering vectors can be employed, and they are found to be a useful tool for texture analysis of PolSAR data [73]. However, there are also limitations, e.g., the difference between models are not very large.

In summary, statistical modeling and texture analysis of PolSAR data covers a wide range of topics. To make a better understanding of texture and to make good use of it, there is still have a lot of work to do.

Acknowledgments: This work has been financed by: the Shenzhen Science & Technology Program with code JSGG20150512145714247, the Spanish Science, Research and Innovation Plan (Ministerio de Economía y Competitividad) with project code TIN2014-55413-C2-1-P, and the National Key Research Plan with code 2016YFC0500201-07. In addition, the authors would like to acknowledge ESA (European Space Agency) for the RADARSAT-2 data provided in the frame of the AgriSAR 2009 campaign, JAXA (Japan Aerospace Exploration Agency) for the ALOS-2 data provided in the framework of the 4th ALOS Research Announcement for ALOS-2, NASA (National Aeronautics and Space Administration) for the UAVSAR data, and Andreas Reigber from DLR for the F-SAR data.

Author Contributions: Xinpeng Deng designed and performed the experiments, analyzed the data, and wrote the paper; Carlos López-Martínez and Jinsong Chen provided valuable suggestions to write the paper, and revised the paper; Pengpeng Han prepared some of the data used in the experiments.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

BSA	Back Scattering Alignment
CDF	Cumulative Distribution Function
CLT	Central Limit Theorem
FSA	Forward Scattering Alignment
GIG	Generalized Inverse Gaussian
NIG	Normal Inverse Gaussian
NIM	Normalized Intensity Gaussian
PDF	Probability Density Function
PolSAR	Polarimetric SAR
ROI	Region Of Interest
SAR	Synthetic Aperture Radar
SIRV	Spherically Invariant Random Vector

Appendix A

Some integral identities used in this paper are listed out here.

1. ([74] p. 368, Equation (3.471-9))

$$\int_0^{\infty} x^{v-1} \exp\left(-\frac{\beta}{x} - \alpha x\right) dx = 2 \left(\frac{\beta}{\alpha}\right)^{v/2} K_v\left(2\sqrt{\beta\alpha}\right) \quad (\text{A1})$$

$\text{Re } \beta > 0, \text{Re } \alpha > 0$

K_v is the modified Bessel function of the second kind of order v .

2. ([74] p. 340, Equation (3.339))

$$\int_0^{\pi} \exp(z \cos x) dx = \pi I_0(z) \quad (\text{A2})$$

$I_0(z)$ is the modified Bessel function of the first kind.

3. ([74] p. 702, Equation (6.624-1))

$$\int_0^{\infty} x e^{-\alpha x} K_0(\beta x) dx = \frac{1}{\alpha^2 - \beta^2} \times \left\{ \frac{\alpha}{\sqrt{\alpha^2 - \beta^2}} \ln \left(\frac{\alpha}{\beta} + \sqrt{\left(\frac{\alpha}{\beta}\right)^2 - 1} \right) - 1 \right\} \quad (\text{A3})$$

4. ([74] p. 347, Equation (3.382-2))

$$\int_u^{\infty} (x-u)^v e^{-\mu x} dx = \mu^{-v-1} e^{-u\mu} \Gamma(v+1), \quad u > 1, \text{Re } v > -1, \text{Re } \mu > 0 \quad (\text{A4})$$

5. ([74] p. 700, Equation (6.621-3))

$$\int_0^{\infty} x^{\mu-1} e^{-\alpha x} K_v(\beta x) dx = \frac{\sqrt{\pi}(2\beta)^v}{(\alpha + \beta)^{\mu+v}} \frac{\Gamma(\mu + v)\Gamma(\mu - v)}{\Gamma(\mu + 1/2)} \times {}_2F_1\left(\mu + v, v + \frac{1}{2}; \mu + \frac{1}{2}; \frac{\alpha - \beta}{\alpha + \beta}\right) \quad (\text{A5})$$

$\text{Re } \mu > |\text{Re } v|, \text{Re } (\alpha + \beta) > 0$

${}_2F_1(a, b; c; z)$ is the Gauss hypergeometric function.

6. ([74] p. 917, Equation (8.432-3))

$$K_v(z) = \frac{\left(\frac{z}{2}\right)^v \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(v + \frac{1}{2}\right)} \int_1^\infty e^{-zt}(t^2 - 1)^{v-\frac{1}{2}} dt, \quad \operatorname{Re}\left(v + \frac{1}{2}\right) > 0, |\arg z| < \frac{\pi}{2} \quad (\text{A6})$$

7. ([74] p. 325, Equation (3.252-3))

$$\int_1^\infty x^{\mu-1}(x^p - 1)^{v-1} = \frac{1}{p} B\left(1 - v - \frac{\mu}{p}, v\right) \quad (\text{A7})$$

$$p > 0, \operatorname{Re} v > 0, \operatorname{Re} \mu < p(1 - \operatorname{Re} v)$$

8. The gamma function is defined as

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx. \quad (\text{A8})$$

Let $x = \frac{\beta}{y}$ where $\beta > 0$, we have the following equation after changing variables

$$\int_0^\infty y^{-t-1} \exp\left(-\frac{\beta}{y}\right) dy = \Gamma(t) \beta^{-t} \quad (\text{A9})$$

9. ([34] p. 505, Equation (13.2.5))

$$\int_0^\infty e^{-zt} t^{a-1} (1+t)^{b-a-1} dt = \Gamma(a) U(a, b, z) \quad (\text{A10})$$

U is the confluent hypergeometric function of the second kind, or KummerU function.

10. ([74] p. 367, Equation (3.471-2))

$$\int_0^u x^{v-1} (u-x)^{\mu-1} \exp\left(-\frac{\beta}{x}\right) dx = \beta^{\frac{v-1}{2}} u^{\frac{2\mu+v-1}{2}} \Gamma(\mu) \times \exp\left(-\frac{\beta}{2u}\right) W_{\frac{1-2\mu-v}{2}, \frac{v}{2}}\left(\frac{\beta}{u}\right) \quad (\text{A11})$$

$$\operatorname{Re} \mu > 0, \operatorname{Re} \beta > 0, \mu > 0$$

W is Whittaker W function.

11. ([74] p. 368, Equation (3.471-5))

$$\int_u^\infty x^{v-1} (x-u)^{\mu-1} \exp\left(\frac{\beta}{x}\right) dx = B(1-\mu-v, \mu) u^{\mu+v-1} \times M\left(1-\mu-v, 1-v, \frac{\beta}{u}\right) \quad (\text{A12})$$

$$0 < \operatorname{Re} \mu < \operatorname{Re}(1-v), u > 0$$

M is the confluent hypergeometric function of the first kind, also known as the KummerM function.

References

1. Goodman, J.W. Some fundamental properties of speckle. *JOSA* **1976**, *66*, 1145–1150.
2. Lee, J.S.; Grunes, M.R.; De Grandi, G. Polarimetric SAR speckle filtering and its implication for classification. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2363–2373.
3. Kersten, P.R.; Lee, J.S.; Ainsworth, T.L. Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 519–527.

4. Doulgeris, A.P.; Anfinson, S.N.; Eltoft, T. Classification with a non-Gaussian model for PolSAR data. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2999–3009.
5. Bombrun, L.; Vasile, G.; Gay, M.; Totir, F. Hierarchical segmentation of polarimetric SAR images using heterogeneous clutter models. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 726–737.
6. Doulgeris, A.P.; Akbari, V.; Eltoft, T. Automatic PolSAR segmentation with the U-distribution and Markov random fields. In Proceedings of the 2012 EUSAR—9th European Conference on Synthetic Aperture Radar, Nuremberg, Germany, 23–26 April 2012; pp. 183–186.
7. Lee, J.S.; Mitchell, R.G.; Kwok, R. Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution. *Int. J. Remote Sens.* **1994**, *15*, 2299–2311.
8. Lee, J.S.; Grunes, M.; Ainsworth, T.; Du, L.J.; Schuler, D.; Cloude, S. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2249–2258.
9. Jakeman, E.; Pusey, P.N. A model for non-Rayleigh sea echo. *IEEE Trans. Antennas Propag.* **1976**, *24*, 806–814.
10. Jakeman, E.; Pusey, P. Significance of K distributions in scattering experiments. *Phys. Rev. Lett.* **1978**, *40*, 546–550.
11. Ward, K. Compound representation of high resolution sea clutter. *Electron. Lett.* **1981**, *17*, 561–563.
12. Yueh, S.H.; Kong, J.A.; Jao, J.K.; Shin, R.T.; Novak, L.M. K-distribution and polarimetric terrain radar clutter. *J. Electromagn. Waves Appl.* **1989**, *3*, 747–768.
13. Lee, J.S.; Schuler, D.L.; Lang, R.H.; Ranson, K.J. K-distribution for multi-look processed polarimetric SAR imagery. In Proceedings of the International Geoscience and Remote Sensing Symposium, Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation, Pasadena, CA, USA, 8–12 August 1994; Volume 4, pp. 2179–2181.
14. Frery, A.C.; Muller, H.J.; Yanasse, C.C.F.; Sant’Anna, S.J.S. A model for extremely heterogeneous clutter. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 648–659.
15. Freitas, C.C.; Frery, A.C.; Correia, A.H. The polarimetric G distribution for SAR data analysis. *Environmetrics* **2005**, *16*, 13–31.
16. Bombrun, L.; Beaulieu, J.M. Fisher distribution for texture modeling of polarimetric SAR data. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 512–516.
17. Moser, G.; Zerubia, J.; Serpico, S.B. Dictionary-based stochastic expectation-maximization for SAR amplitude probability density function estimation. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 188–200.
18. Krylov, V.; Moser, G.; Serpico, S.B.; Zerubia, J. *Modeling the Statistics of High Resolution SAR Images*; Technical Report; Institut National de Recherche en Informatique et en Automatique, INRIA: Rocquencourt, France, 2008.
19. Wang, Y.; Ainsworth, T.L.; Lee, J. On Characterizing High-Resolution SAR Imagery Using Kernel-Based Mixture Speckle Models. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 968–972.
20. Gao, G. Statistical modeling of SAR images: A survey. *Sensors* **2010**, *10*, 775–795.
21. Nelsen, R.B. *An Introduction to Copulas*; Springer Science & Business Media: New York, NY, USA, 2007.
22. Mercier, G.; Bouchemakh, L.; Smara, Y. The use of multidimensional copulas to describe amplitude distribution of polarimetric SAR data. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 2236–2239.
23. Voisin, A.; Krylov, V.A.; Moser, G.; Serpico, S.B.; Zerubia, J. Supervised Classification of Multisensor and Multiresolution Remote Sensing Images With a Hierarchical Copula-Based Approach. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 3346–3358.
24. Lee, J.S.; Pottier, E. *Polarimetric Radar Imaging: From Basics to Applications*; CRC Press: Boca Raton, FL, USA, 2009.
25. Cloude, S. *Polarisation: Applications in Remote Sensing*; Oxford University Press: Oxford, UK, 2009.
26. Cloude, S.R.; Pottier, E. A review of target decomposition theorems in radar polarimetry. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 498–518.
27. Frost, V.S.; Stiles, J.A.; Shanmugan, K.S.; Holtzman, J.C. A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise. *IEEE Trans. Pattern Anal. Mach. Intell.* **1982**, *PAMI-4*, 157–166.
28. Kong, J.; Yueh, S.; Lim, H.; Shin, R.; Van Zyl, J. Classification of earth terrain using polarimetric synthetic aperture radar images. *Prog. Electromagn. Res.* **1990**, *3*, 327–370.

29. Sarabandi, K. Derivation of phase statistics from the Mueller matrix. *Radio Sci.* **1992**, *27*, 553.
30. Ulaby, F.; Sarabandi, K.; Nashashibi, A. Statistical properties of the Mueller matrix of distributed targets. In Proceedings of the IEE Proceedings F—Radar and Signal Processing, Barcelona, Spain, 6 August 1992; Volume 139, pp. 136–146.
31. Goodman, N.R. Statistical analysis based on a certain multivariate complex Gaussian distribution (An Introduction). *Ann. Math. Stat.* **1963**, *34*, 152–177.
32. Lee, J.S.; Hoppel, K.W.; Mango, S.A.; Miller, A.R. Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 1017–1028.
33. Tough, R.J.A.; Blacknell, D.; Quegan, S. A statistical description of polarimetric and interferometric synthetic aperture radar data. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **1995**, *449*, 567–589.
34. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*; Courier Corporation: North Chelmsford, MA, USA, 1964; Volume 55.
35. Walck, C. *Hand-Book on Statistical Distributions for Experimentalists*; University of Stockholm: Stockholm, Sweden, 2007.
36. López-Martínez, C.; Fabregas, X. Polarimetric SAR speckle noise model. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2232–2242.
37. López-Martínez, C.; Pottier, E.; Cloude, S.R. Statistical assessment of eigenvector-based target decomposition theorems in radar polarimetry. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2058–2074.
38. Alonso-González, A.; López-Martínez, C.; Salembier, P. Filtering and segmentation of polarimetric SAR data based on binary partition trees. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 593–605.
39. Anfinson, S.N.; Eltoft, T.; Doulgeris, A.P. A relaxed Wishart model for polarimetric SAR data. In *Proceedings of the Fourth International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry PolInSAR*; European Space Agency: Noordwijk, The Netherlands, 2009; pp. 26–30.
40. Kersten, P.R.; Anfinson, S.N. A flexible and computationally efficient density model for the multilook polarimetric covariance matrix. In Proceedings of the 9th European Conference on Synthetic Aperture Radar, Nuremberg, Germany, 23–26 April 2012; pp. 760–763.
41. Kersten, P.R.; Anfinson, S.N.; Doulgeris, A.P. The Wishart-Kotz classifier for multilook polarimetric SAR data. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 3146–3149.
42. Anfinson, S.N.; Eltoft, T. Application of the matrix-variate Mellin transform to analysis of polarimetric radar images. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2281–2295.
43. Novak, L.M.; Burl, M.C. Optimal speckle reduction in polarimetric SAR imagery. *IEEE Trans. Aerosp. Electron. Syst.* **1990**, *26*, 293–305.
44. Sheen, D.R.; Johnston, L.P. Statistical and spatial properties of forest clutter measured with polarimetric synthetic aperture radar (SAR). *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 578–588.
45. Gini, F.; Greco, M. Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter. *Signal Process.* **2002**, *82*, 1847–1859.
46. Pascal, F.; Forster, P.; Ovarlez, J.; Larzabal, P. Performance analysis of covariance matrix estimates in impulsive noise. *IEEE Trans. Signal Process.* **2008**, *56*, 2206–2217.
47. Vasile, G.; Ovarlez, J.P.; Pascal, F.; Tison, C. Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1809–1826.
48. Bombrun, L.; Anfinson, S.N.; Harant, O. A complete coverage of log-cumulant space in terms of distributions for Polarimetric SAR data. In Proceedings of the 5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry (POLInSAR 2011), Friscati, Italy, 24–28 January 2011; pp. 1–8.
49. Doulgeris, A.; Anfinson, S.N.; Eltoft, T. Analysis of non-Gaussian POLSAR data. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007, pp. 160–163.
50. Doulgeris, A.P.; Eltoft, T. Scale Mixture of Gaussian Modelling of Polarimetric SAR Data. *EURASIP J. Adv. Signal Process.* **2009**, *2010*, 874592.
51. Koudou, A.E.; Ley, C. Characterizations of GIG laws: A survey. *Probab. Surv.* **2014**, *11*, 161–176.
52. Khan, S.; Guida, R. Application of Mellin-Kind Statistics to Polarimetric G Distribution for SAR Data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 3513–3528.

53. Tison, C.; Nicolas, J.M.; Tupin, F.; Maître, H. A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2046–2057.
54. Song, W.; Li, M.; Zhang, P.; Wu, Y.; Jia, L.; An, L. The WGF Distribution for Multilook Polarimetric SAR Data and Its Application. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2056–2060.
55. Bian, Y.; Mercer, B. Multilook polarimetric SAR data probability density function estimation using a generalized form of multivariate K-distribution. *Remote Sens. Lett.* **2014**, *5*, 682–691.
56. De Grandi, G.; Lee, J.S.; Schuler, D.; Nezry, E. Texture and speckle statistics in polarimetric SAR synthesized images. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2070–2088.
57. Eltoft, T.; Anfinson, S.N.; Doulgeris, A.P. A multitexture model for multilook polarimetric radar data. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 1048–1051.
58. Yu, Y. Textural-partially correlated polarimetric K-distribution. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Seattle, WA, USA, 6–10 July 1998; Volume 1, pp. 60–62.
59. Eltoft, T.; Anfinson, S.N.; Doulgeris, A.P. A multitexture model for multilook polarimetric synthetic aperture radar data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2910–2919.
60. Doulgeris, A.P.; Anfinson, S.N.; Eltoft, T. Segmentation of polarimetric SAR data with a multi-texture product model. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Nuremberg, Germany, 22–27 July 2012; pp. 1437–1440.
61. Lombardo, P.; Farina, A. Coherent radar detection against K-distributed clutter with partially correlated texture. *Signal Process.* **1996**, *48*, 1–15.
62. Khan, S.; Guida, R. The new dual-texture G distribution for single-look PolSAR data. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Nuremberg, Germany, 22–27 July 2012; pp. 1469–1472.
63. Anfinson, S.N. *Statistical Unmixing of SAR Images*; Technical Report; Munin Open Research Archive, University of Tromsø—The Arctic University of Norway: Tromsø, Norway, 2016.
64. Wang, Y.; Ainsworth, T.L.; Lee, J.S. Application of Mixture Regression for Improved Polarimetric SAR Speckle Filtering. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 453–467.
65. Frühwirth-Schnatter, S. *Finite Mixture and Markov Switching Models*; Springer: Berlin, Germany, 2006.
66. McNeil, A.J.; Frey, R.; Embrechts, P. *Quantitative Risk Management: Concepts, Techniques and Tools*; Princeton University Press: Princeton, NJ, USA, 2005.
67. Mercier, G.; Moser, G.; Serpico, S.B. Conditional copulas for change detection in heterogeneous remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1428–1441.
68. Regniers, O.; Bombrun, L.; Guyon, D.; Samalens, J.C.; Germain, C. Wavelet-based texture features for the classification of age classes in a maritime pine forest. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 621–625.
69. Oliver, C. Fundamental properties of high-resolution sideways-looking radar. *IEEE F-Commun. Radar Sig.* **1982**, *129*, 385–402.
70. Deng, X.; López-Martínez, C. Higher Order Statistics for Texture Analysis and Physical Interpretation of Polarimetric SAR Data. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 912–916.
71. Nicolas, J.M. Introduction aux statistiques de deuxième espèce: Applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar. *TS. Traitement du Signal* **2002**, *19*, 139–167.
72. Deng, X.; López-Martínez, C.; Varona, E.M. A Physical Analysis of Polarimetric SAR Data Statistical Models. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3035–3048.
73. Deng, X.; López-Martínez, C. On the Use of the l_2 -Norm for Texture Analysis of Polarimetric SAR Data. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6385–6398.
74. Jeffrey, A.; Zwillinger, D. *Table of Integrals, Series, and Products*, 7th ed.; Academic Press: Boston, MA, USA, 2007.

