To appear in *Remote Sensing Letters*
Vol. 00, No. 00, Month 201X, 1–13

Scalar and Representative Observables for Polarimetric SAR Data

Thanh-Hai Le*†, and Ian McLoughlin‡, and Chan-Hua Vun†

†School of Computer Engineering, Nanyang Technological University, Singapore.
‡ School of Information Science and Technology, University of Science and Technology of China.

(*November 2013*)

This paper proposes a scalar and representative observable for multi-dimensional POLSAR data, from which statistically consistent discrimination measures can be derived. Specifically, the statistical behaviour of the POLSAR covariance matrix determinant is used to derive a scalar and generic statistical model for multi-dimensional POLSAR data, which is specifically applicable to the two and three dimensional versions of partial and full monostatic polarimetric SAR data. As the POLSAR covariance matrix determinant generalizes the SAR intensity towards multiple dimensions, the proposed model is able to subsume the traditional SAR intensity model under the umbrella of a unified model. Consequently, the main beneficial implication of the proposed approach is that it provides a consistent theory unifying the currently disconnected proposals for SAR and POLSAR discrimination measures which simplifies the adaptation of existing SAR data processing techniques for POLSAR data.

1. Introduction

During the past decades, exponential growth in computing power has allowed the once computationally-demanding Synthetic Aperture Radar (SAR) technology to become a feasible and preferred technique for earth observation applications. SAR technology has since been extended in a few directions, one of which is polarimetric SAR (POLSAR). POLSAR extends SAR by exploiting the natural polarization property of Electro-Magnetic (EM) waves, leading to the availability of multi-channel POLSAR data, compared to traditional one-channel SAR data.

Like SAR data, POLSAR data is stochastic. Moreover, it is multi-dimensional, making it even harder to interpret. Under this context, it is therefore important to establish a simple and intuitive understanding of the data. Statistical models are undoubtedly crucial in understanding its stochastic nature. While several models have been proposed to describe POLSAR data, they unfortunately tend to be complex and not very intuitive due to the multidimensional nature of the data. Practical POLSAR data processing however, makes heavy use of scalar discrimination measures, which should be based on statistically consistent models for scalar and representative observables of the multidimensional data. It is thus important to establish a scalar and representative observable for this multi-dimensional POLSAR data.

*Corresponding author. Email: le_thanh_hai@hotmail.com

There has been a few scalar observables with accompanying statistical models proposal for POLSAR (Conradsen *et al.* 2003, Alberga *et al.* 2008, Joughin *et al.* 1994, Lee *et al.* 1994b, Touzi and Lopes 1996, Lopez-Martinez and Fabregas 2003, Erten 2012), but none of them is able to provide meaningful scalar discrimination measures. As such no observable has been widely accepted as being highly representative of this multi-dimensional data, which severely limits their applicability in practical data processing applications. From an alternative approach, a few POLSAR discrimination measures have been proposed, but all are based on the likelihood ratio statistics. This statistical test should be based on an exact and consistent statistical distribution but so far has only shown to be based on an asymptotic distribution.

This article hence presents a scalar and representative observable, and its associated generic statistical model to describe multi-dimensional POLSAR data and provide a consistent foundation for the derivation of POLSAR discrimination measures. POLSAR data have been statistically modelled as following the complex Wishart distribution. Consequently, the generic statistical model for the covariance matrix determinant is presented as being just a scalar projection of the multidimensional data. This model is then used to derive several scalar and consistent statistical descriptions suggesting that their associated observables are capable of being used as discrimination measures for POLSAR data.

Interestingly the paper also shows that the specific one dimensional versions of the proposed models match perfectly with the traditional statistical model used for SAR intensity. This effectively incorporates the common one dimensional SAR theory under the unbrella of the proposed scalar approach for multi dimensional POLSAR. With this new insight the different discrimination proposals for SAR and POLSAR are reviewed and it is shown that the approach proposed in this paper provides a strong, unifying and consistent foundation relating them all together. The applicability of these theoretical models will be illustrated by experiments where the specific one, two and three dimensional versions of the proposed generic models are validated against practical data.

The remainder of this article is organized as follows. After the next section reviews existing discrimination measures as well as scalar observables models for POLSAR, section 3. derives the generic statistical model for the determinant of the POLSAR covariance matrix and proposes several new discrimination measures for the multi-dimensional data. After section 4. illustrates the match between the specific model of one dimensional and the traditional model for SAR intensity, section 5. links the disparate discrimination measure proposals for the traditional SAR and the more recent POLSAR data. The applicability of these models against practical data is illustrated in section 6.. Section 7. presents a high-level discussion of the results presented before section 8. finally concludes the paper.

2. Related Work in Literature

Section 2.1 reviews various statistical models used for different scalar observables for POLSAR. where it is shown that none are able to provide statistically consistent discrimination measures. Section 2.2 further strengthens those findings by discussing discrimination measures that have been proposed for POLSAR, and demonstrates that almost all of them are based on the likelihood statistical test for complex Wishart distribution. While an exact statistical distribution is expected to be necessary for a test, only an asymptotic distribution is used in the underlying approach (Conradsen *et al.* 2003).

2.1 *Scalar Observables for POLSAR Data and their Statistical Models*

Different target decomposition theorems have identified many possible scalar observables for complex POLSAR data. In (Alberga *et al.* 2008), the performance of different scalar POLSAR observables is evaluated for classification purposes. While many scalar observables for POLSAR were presented, their corresponding statistical models and classifiers were not available. Furthermore, at its conclusion, the paper indicated that it is impossible to identify a single best representation. Although, to be fair, the observables were identified for describing a decomposed portion of the complex POLSAR data, rather than providing a unified representation of the POLSAR data.

Using a different approach, given that the joint distribution for POLSAR is known to be the multi-variate complex Wishart distribution, it is possible to derive the scalar statistical models for some univariate POLSAR observables. However, such derivations are not trivial tasks, and so far, only a handful of such statistical models have been proposed, including the following:

(i) cross-pol ratio $r_{HV/HH} = |S_{HV}|^2/|S_{HH}|^2$ (Joughin *et al.* 1994),
(ii) co-pol ratio $r_{VV/HH} = |S_{VV}|^2/|S_{HH}|^2$ (Joughin *et al.* 1994),
(iii) co-pol phase difference $\phi_{VV/HH} = arg(S_{VV}S^*_{HH})$ (Joughin *et al.* 1994) (Lee *et al.* 1994b),
(iv) magnitude $g = |avg(S_{pq}S^*_{rs})|$ (Lee *et al.* 1994b),
(v) normalized magnitude $\xi = \frac{|avg(S_{pq}S^*_{rs})|}{\sqrt{avg(|S_{pq}|^2)avg(|S_{rs}|^2)}}$ (Lee *et al.* 1994b),
(vi) intensity ratio $w = avg(|S_{pq}|^2)/avg(|S_{rs}|^2)$ (Lee *et al.* 1994b),
(vii) and the Stokes parameters $S_i, 0 \leq i \leq 3$ (Touzi and Lopes 1996).

More recently, statistical models for each element of the POLSAR covariance matrix, i.e. $S_{pq}S^*_{rs}$, (Lopez-Martinez and Fabregas 2003) as well as for the largest eigen-value of the covariance matrix λ_1 (Erten 2012) have been proposed.

While these models undoubtedly help to further our understandings of the POLSAR data, none of the underlying observables have been shown to meet the dual criteria of 1) resulting in statistically consistent discrimination measures and thus 2) being representative of the complex POLSAR data.

2.2 *POLSAR Discrimination Measures*

The commonly used measure of distance for matrices are either the Euclidean or Manhattan distances, defined respectively as:

$$d(C_x, C_y) = \sum_{i,j} |\mathbb{R}(C_x - C_y)_{i,j}| + \sum_{i,j} |\mathbb{I}(C_x - C_y)_{i,j}| \quad (1)$$

$$d(C_x, C_y) = \sqrt{\sum_{i,j} |C_x - C_y|^2_{i,j}} \quad (2)$$

where $C_{i,j}$ denotes the (i, j) elements of the POLSAR covariance matrix C, $||$ denotes absolute values and $\mathbb{R}, \mathbb{I}$ denote the real and imaginary parts respectively. However, in the context of POLSAR, these dissimilarity measures are not widely used mainly due to the multiplicative nature of the noisy data.

In the field of POLSAR, the Wishart distance is probably the most widely used, as part of the well-known Wishart classifier (Lee *et al.* 1999). It is defined (Lee *et al.* 1994a) as:

$$d(C_x, C_y) = \ln|C_y| + tr(C_x C_y^{-1}) \tag{3}$$

, where $tr(C)$ denotes the trace of the matrix C. As a measure of distance, its main disadvantage is that $d(C_y, C_y) = \ln|C_y| \neq 0$.

Recent works have suggested alternative dissimilarity measures including the symmetric and asymmetric refined Wishart distance (Anfinsen *et al.* 2007),

$$d(C_x, C_y) = \frac{1}{2} tr(C_x^{-1} C_y + C_y^{-1} C_x) - d \tag{4}$$

$$d(C_x, C_y) = \ln|C_x| - \ln|C_y| + tr(C_x C_y^{-1}) - d \tag{5}$$

the Bartlett distance (Kersten *et al.* 2005),

$$d(C_x, C_y) = 2\ln|C_{x+y}| - \ln|C_x| - \ln|C_y| - 2d\ln 2 \tag{6}$$

the Bhattacharyya distance (Lee and Bretschneider 2011),

$$r(C_x, C_y) = \frac{|C_x|^{1/2}|C_y|^{1/2}}{|(C_x + C_y)/2|} \tag{7}$$

and the Wishart Statistical test distance (Cao *et al.* 2007)

$$d(C_x, C_y) = (L_x + L_y)\ln|C| - L_x \ln|C_x| - L_y \ln|C_y| \tag{8}$$

Closer examination of these dissimilarity measures reveals that most of them are related in some ways. The Bhattacharyya distance is easily shown to be related to the Barlett distance. At the same time the Barlett distance can be considered a special case of the Wishart Statistical Test distance, when the two data sets have the same number of looks, i.e. $L_x = L_y$. The close relation among the measures is further supported by the fact that all of their publications refered the same statistical model developed in (Conradsen *et al.* 2003) as the foundation. In (Conradsen *et al.* 2003), to determine if the two scaled multi-look POLSAR covariance matrixes Z_x and Z_y, which have L_x and L_y as the corresponding number of looks, come from the same underlying stochastic process, the likelihood ratio statistics for POLSAR covariance matrix is considered:

$$Q = \frac{(L_x + L_y)^{d\cdot(L_x+L_y)}}{L_x^{d\cdot L_x} L_y^{d\cdot L_y}} \frac{|Z_x|^{L_x}|Z_y|^{L_y}}{|Z_x + Z_y|^{(L_x+L_y)}}$$

Taking the log-transformation of the above equation, and denoting $C_{vx} = Z_x/L_x$, $C_{vy} = Z_y/L_y$ and $C_{vxy} = (Z_x + Z_y)/(L_x + L_y)$ then:

$$Q = \frac{|C_{vx}|^{L_x} \cdot |C_{vy}|^{L_y}}{|C_{vxy}|^{L_x+L_y}} \tag{9}$$

$$\ln Q = L_x \ln|C_{vx}| + L_y \ln|C_{vy}| - (L_x + L_y)\ln|C_{vxy}| \tag{10}$$

To detect changes, a test statistic is developed for this discrimination measure. This means a distribution is to be derived for the dissimilarity measure. However, in the original work (Conradsen *et al.* 2003), only an asymptotic distribution was used. This paper proposes a statistical model for the determinant of the POLSAR covariance matrix $|C_v|$ which is capable of providing an exact distribution for the test.

3. The Generic Scalar Statistical Model for POLSAR

In this section, after the basic foundations of POLSAR statistical analysis are introduced, the generic scalar statistical model for the multi-dimensional POLSAR will be presented.

In this paper, the POLSAR scattering vector is denoted as s. For partial polarimetric SAR (single polarization in transmit and dual polarization in receipt), the vector is two-dimensional ($d = 2$) and is normally written as:

$$s_{part} = \begin{bmatrix} S_h \\ S_v \end{bmatrix} \tag{11}$$

For full and monostatic POLSAR data, the vector is three-dimensional ($d = 3$) and is presented as:

$$s_{full} = \begin{bmatrix} S_{hh} \\ \sqrt{2}S_{hv} \\ S_{vv} \end{bmatrix} \tag{12}$$

Let $\Sigma = E[ss^{*T}]$ denote the population expected value of the POLSAR covariance matrix, where s^{*T} is the complex conjugate transpose of s. Assuming s is jointly circular complex Gaussian with the expected covariance matrix Σ, then the probability density function (PDF) of s can be written as:

$$pdf(s;\Sigma) = \frac{1}{\pi^d|\Sigma|}e^{-s^{*T}\Sigma^{-1}s} \tag{13}$$

where $||$ denotes the matrix determinant.

The sample POLSAR covariance matrix is formed as the mean of Hermitian outer product of independent single-look scattering vectors,

$$C_v = \langle ss^{*T}\rangle = \frac{1}{L}\sum_{i=1}^{L} s_i s_i^{*T} \tag{14}$$

where L is the number of looks and s_i denotes the single-look scattering vector, which equals s_{part} for the case of partial POLSAR or s_{full} for the case of full polarimetry.

Complex Wishart distribution statistics are normally used for the scaled covariance matrix $Z = LC_v$, whose PDF is given as:

$$pdf(Z;d,\Sigma,L) = \frac{|Z|^{L-d}}{|\Sigma^L|\Gamma_d(L)}e^{-tr(\Sigma^{-1}Z)} \tag{15}$$

with $\Gamma_d(L) = \pi^{d(d-1)/2}\prod_{i=0}^{d-1}\Gamma(L-i)$ and d is the dimensional number of the POLSAR covariance matrix.

The approach taken in this paper differs by applying the homoskedastic log transformation on a less-than-well-known result for the determinant of the covariance matrix. Goodman (Goodman 1963) proved that the ratio between the observable and expected values of the sample covariance matrix determinants behaves like a product of d chi-squared random variables with different degrees of freedom:

$$\chi_L^d = (2L)^d \frac{|C_v|}{|\Sigma_v|} \sim \prod_{i=0}^{d-1} \chi(2L-2i) \tag{16}$$

This result is used here to develop the generic scalar statistical model for POLSAR. From Eqn. 16 we have:

$$|C_v| \sim |\Sigma_v| \cdot \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi(2L-2i) \tag{17}$$

In a given homogeneous POLSAR area, the parameters Σ_v, d and L can be considered constant. Thus Eqn. 17 indicates that in the original POLSAR domain, a multiplicative speckle noise pattern is present.

Moreover, since the average and variance of these chi-squared distributions are known to be constant, i.e. $avg\left[\chi(2L)\right] = 2L$ and $var\left[\chi(2L)\right] = 4L$, the product and summation of these random variables also have fixed summary statistics. Specifically:

$$\begin{aligned}
avg\left[\prod_{i=0}^{d-1} \chi(2L-2i)\right] &= 2^d \cdot \prod_{i=0}^{d-1} (L-i), \\
var\left[\prod_{i=0}^{d-1} \chi(2L-2i)\right] &= \prod_{i=0}^{d-1} 4(L-i)(L-i+1) - \prod_{i=0}^{d-1} 4(L-i)^2,
\end{aligned}$$

Combining these results with Eqn. 17 , we have:

$$avg\left[|C_v|\right] = \frac{|\Sigma_v|}{L^d} \prod_{i=0}^{d-1} (L-i) \tag{18}$$

$$var\left[|C_v|\right] = \frac{|\Sigma_v|^2 \left[\prod_{i=0}^{d-1} (L-i)(L-i+1) - \prod_{i=0}^{d-1} (L-i)^2\right]}{L^{2d}} \tag{19}$$

For a real world captured image, while the parameters d and L do not change for the whole image, the underlying Σ_v is expected to differ from one region to the next. Thus over a heterogeneous scene, the stochastic process for $|C_v|$ and $\ln |C_v|$ vary depending on the underlying signal Σ_v. In such context, Eqn. 19 implies that the variance of $|C_v|$ also differs depending on the underlying signal Σ_v (i.e. it is heteroskedastic).

Similar to the way intensity-ratio is proposed as the discrimination measure for the multiplicative and heteroskedastic SAR intensity (Rignot and van Zyl 1993), in this paper, the determinant-ratio and the change-ratio are also proposed as a discrimination measure for the POLSAR data. which is shown above to also suffer from the multiplicative and heteroskedastic phenomena.

For cases where the true value of the underlying signal Σ_v is known *a priori*, then the determinant-ratio of the signal random variable ($\mathbb{R}_\Sigma$) is defined as:

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \tag{20}$$

For cases where the POLSAR data is known to have come from a homogeneous area, but the true value of the underlying signal Σ_v is *unknown*, then a random variable called the change-ratio ($\mathbb{R}_C$) is defined as:

$$\mathbb{R}_C = \frac{|C_1|}{|C_2|} \tag{21}$$

where C_1 and C_2 are samples of the covariance matrix determinant in an assumed homogeneous area.

Using the results from Eqn. 17, we have

$$\mathbb{R}_\Sigma \sim \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi(2L-2i) \tag{22}$$

$$\mathbb{R}_C \sim \prod_{i=0}^{d-1} \frac{\chi(2L-2i)}{\chi(2L-2i)} \tag{23}$$

Since each elementary component follows fixed distributions (i.e. $\chi^2(2L)$), it is natural that this variable also follows fixed distributions. Moreover, it is independent of the underlying signal Σ_v, indicating its statistically consistent properties and hence its applicability as a POLSAR discrimination measure. This claim is extrapolated from the widely-used intensity-ratio as a SAR discrimination measure and will be further discussed in the next sections.

4. SAR as a one-dimensional case of POLSAR

This section shows that the proposed generic model is equally applicable to the 1-dimensional ($d = 1$) case, which is physically equivalent to collapsing the multi-dimensional POLSAR dataset into conventional single dimensional SAR data. Mathematically, the sample covariance matrix C_v is reduced to the sample variance while the determinant $|C_v|$ is equivalent to the scalar variance value. As variance is equal to intensity I for SAR data, our result is consistent with previous results for SAR intensity data. Thus the proposed generic model for POLSAR, collapsed into one dimension will be shown to incorporate the traditional SAR intensity as a natural case of the unified model.

The results for our models can be summarized using the following equations:

$$|C_v| \sim |\Sigma_v| \cdot \frac{1}{(2L)^d} \cdot \prod_{i=0}^{d-1} \chi(2L-2i)$$

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \sim \frac{1}{(2L)^d} \prod_{i=0}^{d-1} \chi(2L-2i)$$

$$\mathbb{R}_C = \frac{C_1}{C_2} \sim \prod_{i=0}^{d-1} \frac{\chi(2L-2i)}{\chi(2L-2i)}$$

Upon setting $d = 1$ into the above equations, the equations become:

$$|C_v| \sim \frac{|\Sigma_v|}{(2L)} \cdot \chi(2L)$$

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \sim \frac{1}{(2L)} \cdot \chi(2L)$$

$$\mathbb{R}_C = \frac{C_1}{C_2} \sim \prod_{i=0}^{d-1} \frac{\chi(2L-2i)}{\chi(2L-2i)}$$

Since the PDF of chi-squared distribution can be written as:

$$\chi(2L) \sim pdf \left[\frac{\chi^{L-1} e^{-\chi/2}}{2^L \Gamma(L)} \right]$$

Applying variable change theorem into the above equations results in:

$$|C_v| \sim pdf \left[\frac{L^L x^{L-1} e^{-Lx/|\Sigma_v|}}{\Gamma(L)|\Sigma_v|^L} \right]$$

$$\mathbb{R}_\Sigma = \frac{|C_v|}{|\Sigma_v|} \sim pdf \left[\frac{L^L x^{L-1} e^{-Lx}}{\Gamma(L)} \right]$$

$$\mathbb{R}_C = \frac{C_1}{C_2} \sim pdf \left[\frac{\Gamma(2L-1) x^{L-1}}{\Gamma^2(L-1)(1+x)^{2L}} \right]$$

These equations match exactly with the following traditional model for multi-look SAR intensity:

$$I \sim pdf \left[\frac{L^L x^{L-1} e^{-Lx/\bar{I}}}{\Gamma(L)\bar{I}^L} \right] \tag{24}$$

$$\mathbb{R}_{\bar{I}} = \frac{I}{\bar{I}} \sim pdf \left[\frac{L^L x^{L-1} e^{-Lx}}{\Gamma(L)} \right] \tag{25}$$

$$\mathbb{R}_I = \frac{I_1}{I_2} \sim pdf \left[\frac{\Gamma(2L-1) x^{L-1}}{\Gamma^2(L-1)(1+x)^{2L}} \right] \tag{26}$$

considering that $|C_v| \mapsto I$ and $|\Sigma_v| \mapsto \bar{I}$ as multi-dimensional POLSAR collapses into single-dimensional SAR.

5. Unifying the different discrimination measures proposals for both SAR and POLSAR

Statistical models have always been the foundation for deriving discrimination measures. This is true for both fields of SAR and POLSAR. For SAR, where the field is more or less settled on the issue, the statistical model for SAR intensity has been used to derive the most widely used intensity-ratio discrimination measure (Rignot and van Zyl 1993). For POLSAR, where the research field is much less mature, the same case should apply, except that so far only asymptotic distributions have been derived for the most common foundation, i.e. the likelihood test statistics (Conradsen *et al.* 2003).

With the insight gained from section 4., this section presents a few results. First, similar to the way that the statistical model for SAR intensity has been used to provide the foundation for SAR discrimination measures, e.g. intensity-ratio (Rignot and van Zyl 1993), the statistical model proposed for POLSAR covariance matrix determinant is to be reviewed as providing the foundation for POLSAR discrimination measures, i.e. the likelihood test statistics. And second, new discrimination measures for POLSAR may be derived by learning from the existing SAR discrimination measures.

As for the first matter, in view of the models given in Eqn 17, the likelihood test statistics presented in (Conradsen *et al.* 2003) and rewritten in Eqns 9 & 10 can be expressed as:

$$\ln Q \sim k + L_x \Lambda^d_{L_x} + L_y \Lambda^d_{L_y} - (L_x + L_y)\Lambda^d_{(L_x+L_y)}$$

$$Q \sim e^k \frac{(\chi^d_{L_x})^{L_x} \cdot (\chi^d_{L_y})^{L_y}}{(\chi^d_{L_x+L_y})^{L_x+L_y}}$$

where $k = d\left[(L_x + L_y)\ln(L_x + L_y) - L_x \ln L_x - L_y \ln L_y\right]$. This, in essence, derives an exact statistical distribution for the likelihood test statistics, as opposed to the asymptotic distribution derived in (Conradsen *et al.* 2003).

As a by-product of this exact derivation, several discrimination measures for the common case of $L_x = L_y$ are further proposed. They are the determinant-ratio and the change-ratio presented in Section 3.. Compared to existing discrimination measures for POLSAR reviewed in Section 2., the proposed dissimilarity measures are simpler in both concept and computation. They, at the same time, are multi-dimensional extensions of the intensity-ratio discrimination measure widely used in SAR.

In short, with the newly gained insight, the approach proposed in this paper provides a new bridge between the two fields of SAR and POLSAR. This invites the adaptation of many established SAR data processing techniques towards POLSAR data, where this Section has just described a few examples. However, before further discussion ensues, it is necessary to validate the presented theoretical model with real-life practical data.

6. Validating the proposed models against real-life data

This section aims to verify the theoretical models given in Eqns. 17, 22 and 23 against practical data. All of these models require the estimation of two parameters from the captured data. These are (i) the dimensional number d and (ii) the look number L. The dimensional number d is related to the type of (POL)SAR data captured, with $d = 1, 2, 3$ corresponding to the cases of SAR, partial and full polarimetric SAR data, respectively. The look-number L can be taken either as nominally stated by the data provider, or it can be estimated using the technique proposed by (Anfinsen *et al.* 2009).

To show the robustness of the proposed models, their validations are also put to test on different types of POLSAR sensors. Specifically in this paper, the air-borne four-look ($L = 4$) AIRSAR image of Flevoland and a fine-quad single-look complex RADARSAT2 image ($L = 1$) are used. Since the determinant of the covariance matrix is only significant on multi-look data, nine-look processing is first applied on the single-look RADARSAT2 data ($L = 9$).

6.1 *The Traditional case of SAR* ($d = 1$)

Since it has been shown in Section 4. that the model for the case of $d = 1$ matches exactly with the traditional model for SAR intensity, the validation of the proposed $d = 1$ model is quite straightforward. Fig. 1 presents the results of a test performed for the stated purpose. In the test, the intensity of single-channel SAR data (HH) for sample homogeneous areas in both datasets is extracted. The histograms for both the intensity and intensity ratio are then plotted against the theoretical PDF given above. The plots are obtained having set ENL to an estimated value obtained using the technique presented in (Anfinsen *et al.* 2009). In all cases, the good visual match between the actual data and model distribution tends to validate the proposed model.

6.2 *The Multi-dimensional case of POLSAR* ($d = 2, 3$)

The remainder of this section now focuses on validating the models for both partial ($d = 2$) and full ($d = 3$) POLSAR. Similar tests are carried out for both types of polarimetric SAR on both AIRSAR and RADARSAT2 datasets. The look-number is estimated for each type of data and for each image using the technique described in (Anfinsen *et al.* 2009). The practical and predicted distributions are again plotted to determine whether their visual match is close in each case. Fig. 2 shows the $d = 2$ plots for determinant, determinant ratio and change ratio for both datasets. It is very clear that the model histogram matches the data well, in fact appearing similar to a smoothed data response. Evidently the match is good for $d = 2$.

Similarly, Fig. 3 explores the $d = 3$ case for the same data, same model parameters. Again, although the histogram is much tighter, the match is visually obvious in all cases.

7. Discussion

To begin, a few theoretical properties of the proposed statistical models are discussed. First, the use of covariance matrix log-determinant may be related to the standard eigen-decomposition method of the POLSAR covariance matrices. In fact, the log-determinant can also be computed as the sum of log-eigenvalues. Specifically $\ln |M| = \sum \ln \lambda_M$ where λ_M denotes all the eigenvalues of M. Thus, similar to other eigenvalue based approaches (e.g. entropy/anisotropy, ...), the models presented here are invariant to polarization basis transformations.

Second, the models are developed for the POLSAR covariance matrix. However, since the POLSAR coherency matrix is related to the covariance matrix via a unitary transformation which preserves the determinant, the model is also applicable to the coherency matrix.

It should be noted that despite these advantages, the model descriptions are far from complete. While it is desirable to reduce the multi-dimensional POLSAR data to a scalar value for many applications, such a reduction is unlikely to be lossless. Ideally the use of this technique could be complemented by some high-dimensional POLSAR target-decomposition techniques, such as the Freeman Durden decomposition (Freeman and Durden 1998) or the entropy/anisotropy decomposition (Cloude and Pottier 1997) or similar.

Nevertheless the proposed models are promising. Even though initially developed for partial and monostatic POLSAR data, it was shown to be applicable to traditional SAR data as well. Since the model assumptions are quite minimal, they may potentially be applicable to bi-static and interferometric data, although that would require further study.

Intuitively, the concept of the determinant being representative of the POLSAR data can be thought of as being similar to the concept of using the magnitude to represent a complex number (in a sense this is also why intensity is widely considered representative of the complex SAR data). Similar to the determinant, the magnitude observable is also scalar and is not lossless, but it is undeniably useful. To fully represent the complex number, another observable would needed besides the magnitude, for example a polar representation. Magnitude is, of course, widely used as it can be shown to be invariant to a change in the reference frame. Similarly, the determinant of the POLSAR covariance matrix is also scalar and invariant to a change of polarization basis. Thus while it may not be *fully* representative of the POLSAR data, it can be said to be *highly* representative of it. The important point is that it can provide a scalar comparison of the multi-dimensional data.

Besides the above properties, the theoretical models may also provide an alternative derivation for the likelihood test statistics widely used in POLSAR. Similar to the way that different measures of distance can be used to derive POLSAR classifiers (Lee *et al.* 1999), change detectors (Conradsen *et al.* 2003), edge detectors (Schou *et al.* 2003) or other clustering and speckle filtering techniques (Le *et al.* 2010) (Le and McLoughlin 2011), new detection, classification, clustering or speckle filtering algorithms can be derived using the models presented in this paper. These applications however are not included in this paper, but rather left as a suggestion for future investigations.

8. Conclusion

This paper has proposed a generic statistical model for the POLSAR covariance matrix determinant based on the complex Wishart POLSAR target vector distribution.This model has been validated for specific ($d = 2$) and ($d = 3$) using real-life partial and full polarimetric SAR data. The paper also establishes a few POLSAR discrimination measures: the determinant-ratio and the change-ratio, which essentially are the generic version of the SAR intensity-ratio. At the same time, we have also shown a near-perfect match between the specific $d = 1$ model and that for traditional SAR intensity, effectively bringing the existing SAR theories under the umbrella of this new model.

The main emphasis of this paper is the proposal to consider the POLSAR covariance matrix determinant as a scalar and representative observable for the multi-dimensional POLSAR data. Compared to other published scalar observables, the determinant is highly representative of the multi-dimensional data. Its representative power is justified for the following reasons. Firstly, the covariance matrix determinant, when collapsed into one-dimensional case, transforms neatly into the representative SAR intensity. Secondly, the statistical model for this observable is shown to be the generic multi-dimensional extension of the traditional model for the one dimensional SAR intensity. Finally, this observable also leads to the establishment of the determinant-ratio discrimination measure for POLSAR data, which should increase the usability of this proposal.

There are several beneficial implications of the proposed approach presented in this paper. Currently while the field of SAR is much more developed than the field of POLSAR, the two fields remain quite separated. Hence many techniques applicable to SAR can not be directly translated to POLSAR. Since the scalar and representative model for POLSAR generalizes the traditional model for SAR intensity, its main benefit is that it enables the convenient adaptation of many existing SAR data processing techniques for POLSAR data. One example has already been presented in this paper, where existing SAR discrimination measures are extended towards POLSAR. At the same time, it provides a consistent theory unifying the seemingly disparate discrimination measure proposals for SAR and POLSAR data.

Of course, the models proposed in this article also have their own limitations. Firstly, they are based on the complex Wishart distribution which is only guaranteed to work for homogeneous areas. Secondly, while it is desirable for a large number of applications to reduce the multi-dimensional POLSAR data to a scalar value, such a reduction is unlikely to be lossless in the general case.

In this paper, the theoretical model has only been validated for partial ($d = 2$) and full ($d = 3$) monostatic POLSAR data, leaving the test of its validity for other datasets such as bistatic POLSAR data ($d = 4$) or interferrometric POLSAR data ($d = 6$) as possible future work. Other possible future work includes an investigation on the applicability of this observables into POLSAR heterogeneous areas, as well as further studies to explore the use of these generic models in conjunction with other target decomposition techniques such as the Freeman-Durden decomposition (Freeman and Durden 1998) or the entropy/anisotropy decomposition (Cloude and Pottier 1997).

References

Alberga, V., Satalino, G. and Staykova, D.K., 2008, Comparison of polarimetric SAR observables in terms of classification performance. *International Journal of Remote Sensing*, **29**, pp. 4129–4150.

Anfinsen, S., Doulgeris, A. and Eltoft, T., 2009, Estimation of the Equivalent Number of Looks in Polarimetric Synthetic Aperture Radar Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **47**, pp. 3795 –3809.

Anfinsen, S.N., Jenssen, R. and Eltoft, T., 2007, Spectral Clustering of Polarimetric SAR Data With Wishart-Derived Distance Measures. In *Proceedings of the 3rd International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry*, 3, Jan.

Cao, F., Hong, W., Wu, Y. and Pottier, E., 2007, An Unsupervised Segmentation With an Adaptive Number of Clusters Using the SPAN/H/α/A Space and the Complex Wishart Clustering for Fully Polarimetric SAR Data Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, pp. 3454 –3467.

Cloude, S. and Pottier, E., 1997, An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, pp. 68 –78.

Conradsen, K., Nielsen, A., Schou, J. and Skriver, H., 2003, A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 4 – 19.

Erten, E., 2012, The performance analysis based on SAR sample covariance matrix. *Sensors (Basel)*, **12**, pp. 2766–2786.

Freeman, A. and Durden, S., 1998, A three-component scattering model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, pp. 963 –973.

Goodman, N.R., 1963, The Distribution of the Determinant of a Complex Wishart Distributed Matrix. *Annals of Mathematical Statistics*, **34**, pp. 178–180.

Joughin, I., Winebrenner, D. and Percival, D., 1994, Probability density functions for multilook polarimetric signatures. *Geoscience and Remote Sensing, IEEE Transactions on*, **32**, pp. 562 –574.

Kersten, P., Lee, J.S. and Ainsworth, T., 2005, Unsupervised classification of polarimetric synthetic aperture Radar images using fuzzy clustering and EM clustering. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, pp. 519 – 527.

Le, T.H. and McLoughlin, I.V., 2011, SAR Fuzzy-MLE Speckle Filter using the Distance Consistency Property in Homoskedastic Log-Transformed Domain. In *Proceedings of the Proceedings of the 32nd Asian Conference on Remote Sensing (ACRS)*, Nov, Taipei, Taiwan.

Le, T.H., McLoughlin, I.V., Lee, K.Y. and Brestchneider, T., 2010, SLC SAR speckle filtering using homoskedastic features of logarithmic transformation. In *Proceedings of the Proceedings of the 31st Asian Conference on Remote Sensing (ACRS)*, Nov, Hanoi, Vietnam.

Lee, J.S., Grunes, M.R. and Kwok, R., 1994a, Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution. *International Journal of Remote Sensing*, **15**, pp. 2299–2311.

Lee, J.S., Grunes, M. and de Grandi, G., 1999, Polarimetric SAR speckle filtering and its implication for classification. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, pp. 2363 –2373.

Lee, J.S., Hoppel, K., Mango, S. and Miller, A., 1994b, Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **32**, pp. 1017 –1028.

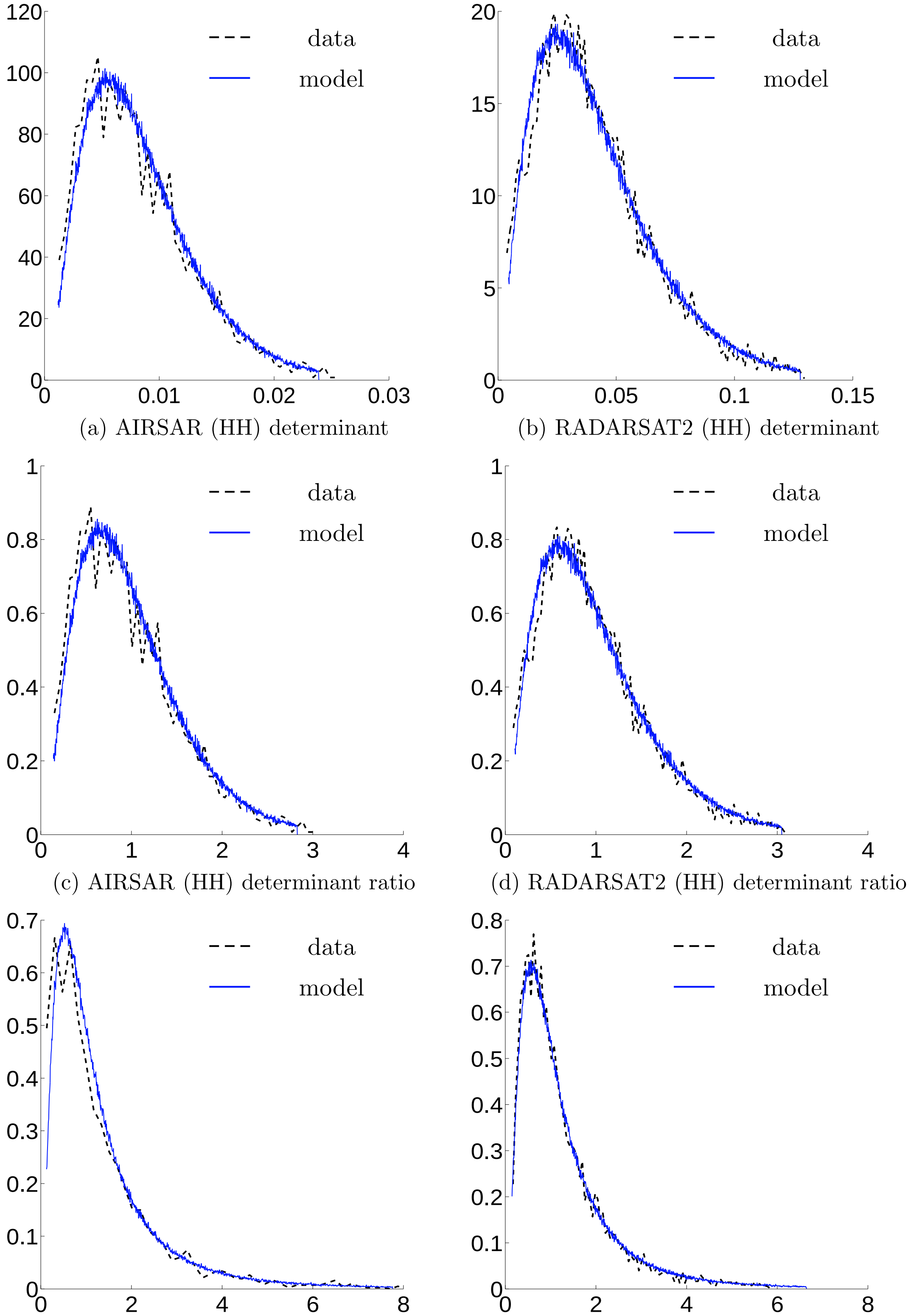
Lee, K.Y. and Bretschneider, T., 2011, Derivation of separability measures based on central complex Gaussian and Wishart distributions. In *Proceedings of the IEEE Internationaal Geoscience and Remote Sensing Symposium (IGARSS), 2011*, Jul, pp. 3740 –3743.

Lopez-Martinez, C. and Fabregas, X., 2003, Polarimetric SAR speckle noise model. *Geoscience and Remote Sensing, IEEE Transactions on*, **41**, pp. 2232–2242.

Rignot, E. and van Zyl, J., 1993, Change detection techniques for ERS-1 SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **31**, pp. 896 –906.

Schou, J., Skriver, H., Nielsen, A. and Conradsen, K., 2003, CFAR edge detector for polarimetric SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 20 – 32.

Touzi, R. and Lopes, A., 1996, Statistics of the Stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields. *Geoscience and Remote Sensing, IEEE Transactions on*, **34**, pp. 519 –531.



(a) AIRSAR (HH) determinant

(b) RADARSAT2 (HH) determinant

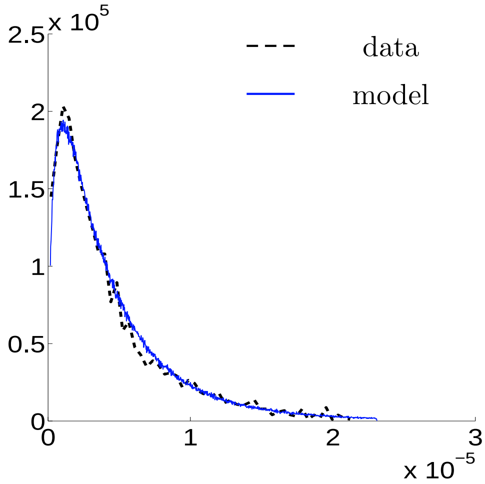
(c) AIRSAR (HH) determinant ratio

(d) RADARSAT2 (HH) determinant ratio

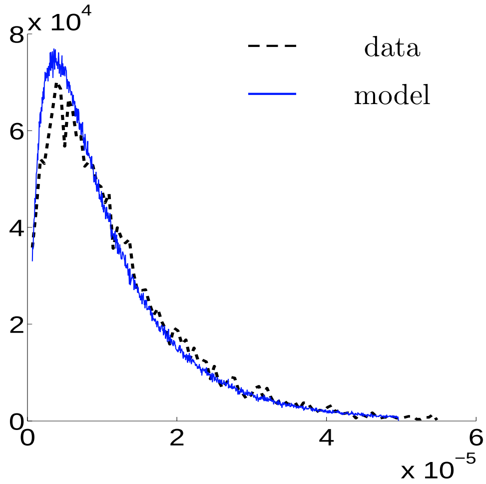
(e) AIRSAR (HH) change ratio

(f) RADARSAT2 (HH) change ratio

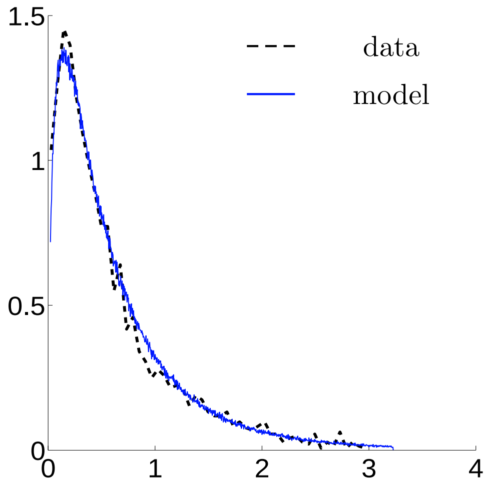
Figure 1.: The specific d=1 models are validated for both RADARSAT2 and AIRSAR datasets. Each histogram plots the signal under investigation (along the x-axis) against its probabilistic distribution across the patch (on the y-axis).



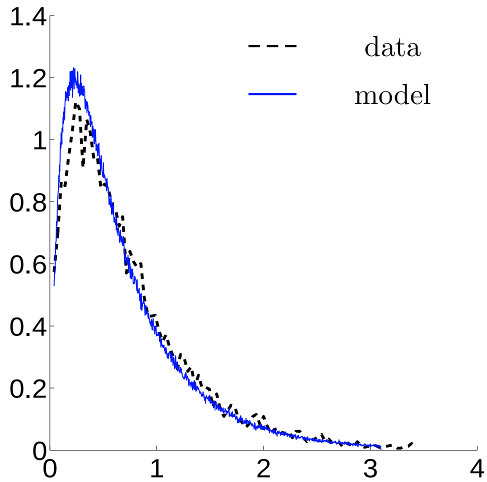
(a) AIRSAR (HH-HV) determinant



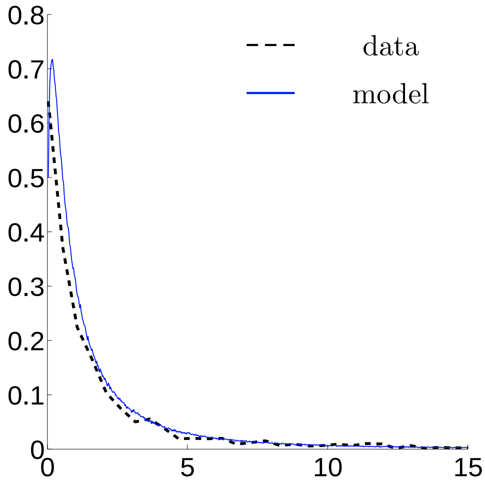
(b) RADARSAT2 (HH-HV) determinant



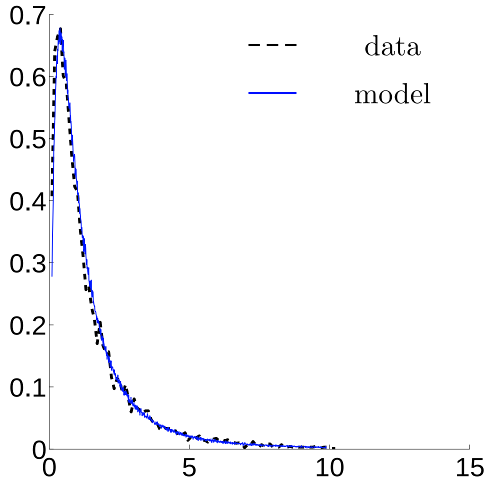
(c) AIRSAR (HH-HV) determinant ratio



(d) RADARSAT2 (HH-HV) determinant ratio

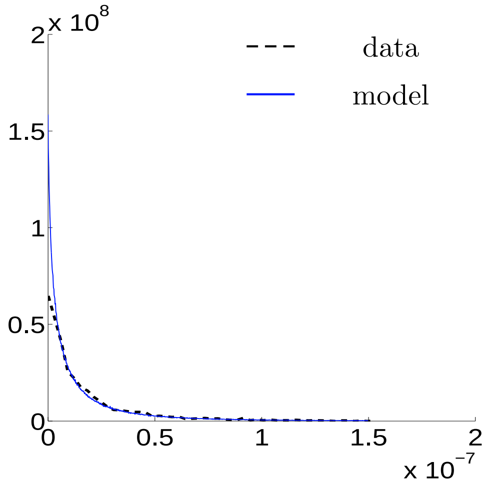


(e) AIRSAR (HH-HV) change ratio

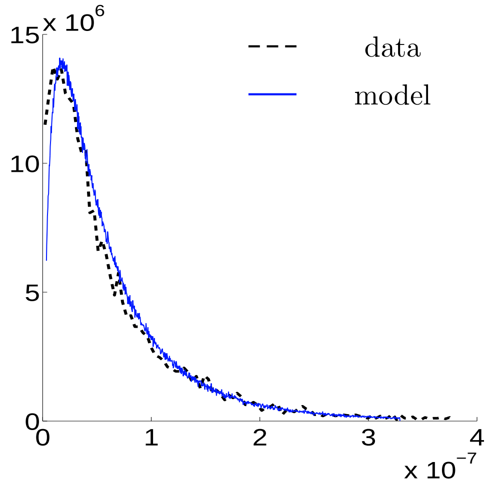


(f) RADARSAT2 (HH-HV) change ratio

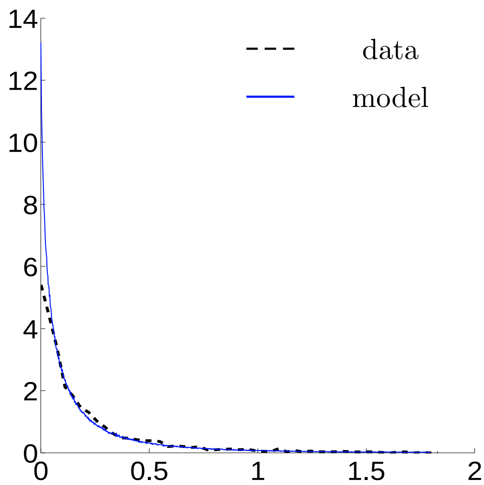
Figure 2.: The specific $d = 2$ models are validated on both RADARSAT2 and AIRSAR datasets. Each histogram plots the signal under investigation (along the x-axis) against its probabilistic distribution across the patch (on the y-axis).



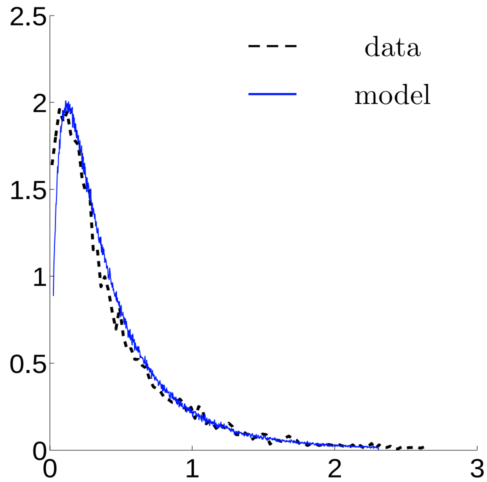
(a) AIRSAR (HH-HV-VV) determinant



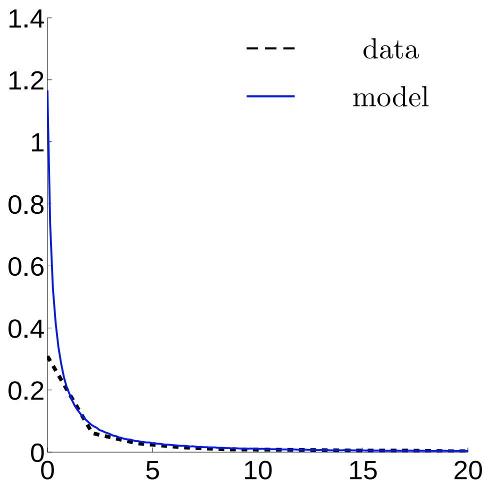
(b) RADARSAT2 (HH-HV-VV) determinant



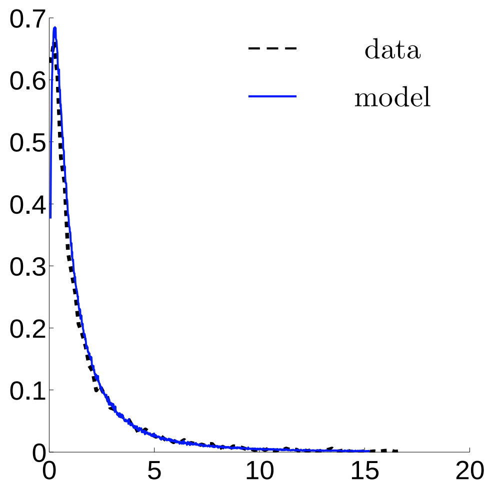
(c) AIRSAR (HH-HV-VV) determinant ratio



(d) RADARSAT2 (HH-HV-VV) determinant ratio



(e) AIRSAR (HH-HV-VV) change ratio



(f) RADARSAT2 (HH-HV-VV) change ratio

Figure 3.: The specific $d = 3$ models are validated on both RADARSAT2 and AIRSAR datasets. Each histogram plots the signal under investigation (along the x-axis) against its probabilistic distribution across the patch (on the y-axis).