Determining the Points of Change in Time Series of Polarimetric SAR Data

Knut Conradsen, Allan Aasbjerg Nielsen, and Henning Skriver, *Member, IEEE*

***Abstract*—We present the likelihood ratio test statistic for the homogeneity of several complex variance–covariance matrices that may be used in order to assess whether at least one change has taken place in a time series of SAR data. Furthermore, we give a factorization of this test statistic into a product of test statistics that each tests simpler hypotheses of homogeneity up to a certain point and that are independent if the hypothesis of total homogeneity is true. This factorization is used in determining the (pixelwise) time points of change in a series of six L-band EMISAR polarimetric SAR data. The pixelwise analyses are applied on homogeneous subareas covered with different vegetation types using the distribution of the observed p-values.**

***Index Terms*—Complex covariance matrix test statistic, complex Wishart distribution, dual polarization, EMISAR, full polarization, multitemporal synthetic aperture radar (SAR) data, omnibus test statistic, quad polarization, remote sensing change detection.**

I. Introduction

CHANGE detection is a very important method for many applications of remotely sensed data from satellites. In particular, synthetic aperture radar (SAR) data are useful due to their all-weather capabilities; hence, planned acquisitions for change detection are normally secured. A number of studies have applied SAR data to change detection applications in single-channel SAR images applying different methods using, e.g., the classical ratio detection [1], [2], the Kittler–Illingworth threshold selection criterion [3]–[6], hidden Markov chains for thresholding [7], wavelets [8]–[11], linear features [12], Kullback–Leibler divergence [13], multivariate gamma distributions [14], neural networks [15], fusion of multisimilarity measures [16], and Markov random fields [17], where most of the methods are based on the classic ratio detector and improvements thereof. In addition, methods for change detection using multichannel SAR data (e.g., polarimetric) have been studied using, e.g., polarimetric parameters [18], [19], Markov random fields for multichannel SAR data [20], [21], a generalized maximum likelihood test for covariance matrices [22] and the same test statistics for classification [23], partial vectors for the suppression of the backscatter coefficient influence [24], the Hotelling–Lawley trace statistic [25], and a non-Wishart change detector [26], where a characteristic of the multichannel change detectors is that they are used for change detection between bitemporal acquisitions.

In change detection between two polarimetric SAR images, tests comparing two complex variance–covariance matrices have turned out to be very efficient. For example, this is used in [22], where the likelihood ratio test statistic is derived, and an approximate expression for the distribution of the statistic under the hypothesis that no changes have occurred is found. In radar literature, the term "variance–covariance matrix" is not commonly used. We use the term to indicate that we here deal with quadratic positive-definite (dispersion) matrices and not cross-covariance matrices between different multivariate observations. In this paper, we apply the test statistics developed to multilook SAR data in the so-called complex covariance formulation. In this paper, we will use the usual radar term "covariance matrix."

When comparing several images, one may apply the simple approach making pairwise comparisons. However, this approach makes it virtually impossible to control the rates of false positives (postulating a change when there actually is none) and of false negatives (missing an actual change). In general, a better approach for comparing several distributions is to perform a simultaneous test of the hypothesis of homogeneity of the said distributions, i.e., a so-called omnibus test (e.g., see [27]). In this paper, we enable this by deriving the likelihood ratio test statistic for the equality of several, say, k, complex variance–covariance matrices and by finding an approximation for the distribution of this test statistic under the hypothesis of equality. If the conclusion of such an analysis is that the parameters in the underlying Wishart distributions are not constant, i.e., we have a nonstationary time series, then the following question naturally arises: When do the changes actually occur? In this paper, we present a factorization of the likelihood ratio statistic into a product of test statistics that each tests simpler hypotheses of homogeneity up to a certain point and that are independent if the hypothesis of total homogeneity is true. We show how this may be used in setting up a change detector for solving the said problem.

The first results in this paper are direct generalizations of the $k = 2$ case reported in [22] and further described in [28]. As mentioned above, some other studies have been reported on change detection between two polarimetric SAR images, but the approach presented in this paper of detecting changes in a series of polarimetric SAR data in the covariance matrix representation is new. The launches of a number of satellite SAR systems during the latest ten years and more have made it more common to use SAR data and, particularly, the time series

Manuscript received January 28, 2015; revised July 10, 2015 and October 17, 2015; accepted December 14, 2015. Date of publication February 3, 2016; date of current version March 25, 2016.

K. Conradsen and A. A. Nielsen are with the Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kongens Lyngby, Denmark (e-mail: alan@dtu.dk).

H. Skriver is with the National Space Institute, Technical University of Denmark, 2800 Kongens Lyngby, Denmark.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2015.2510160

of SAR data for different applications. The SAR systems include, e.g., ALOS, the Radarsat-2, the TerraSAR-X, COSMO-SkyMed satellites, and the Sentinel-1 mission. Some of these systems have a full-polarimetric mode, whereas others have single and/or dual polarization. The time series of SAR data enable a number of applications where a core method for the utilization of the data is the detection of changes in the time series, e.g., the detection of a specific event for an agricultural crop (e.g., sowing or harvesting), the detection of a flooding event in a monitoring system, or the detection of changes in an urban area.

Section II includes a description of the covariance representation of polarimetric SAR data and describes the data set we have used in the demonstration of the test statistics. Section III gives the basic results on the omnibus test statistic and on its decomposition into a product of test statistics that test simpler hypotheses in the general complex Wishart case. Since the real Gamma distribution is a univariate special case, the theory is exemplified on this distribution. In addition, a scheme for using the test statistics in a change detection method to detect the changes in a time series is shown. Section IV shows the results of change detection in multitemporal polarimetric data over five time points for the omnibus test and six time points for the decomposition. Section V contains the conclusions. The Appendix gives details on the new test statistics and their distributions.

II. Data

This section first describes the covariance representation of multilook polarimetric SAR data, followed by a description of the SAR data used in this paper.

A. Polarimetric SAR

A fully polarimetric SAR measures the 2×2 complex so-called scattering matrix at each resolution cell on the ground. The scattering matrix relates the incident and scattered electric fields (see [29]). If S_{rt} denotes the complex scattering amplitude for receive and transmit polarizations ($r, t \in \{h, v\}$ for horizontal and vertical polarizations), then reciprocity, which normally applies to natural targets, gives $S_{hv} = S_{vh}$ (in the backscattering direction using the backscattering alignment convention) [29]. Assuming reciprocity, the scattering matrix is represented by the three-component complex target vector $\boldsymbol{s} = [S_{hh}\ S_{hv}\ S_{vv}]^T$, where superscript T denotes the matrix transpose.

The inherent speckle in the SAR data can be reduced by spatial averaging at the expense of spatial resolution. In this so-called multilook case (in the following, n is the number of looks), a more appropriate representation of the backscattered signal is the covariance matrix in which the average properties of a group of resolution cells can be expressed in a single matrix formed by the outer products of the averaged target vectors. The sample covariance matrix is defined as [29]

$$\langle \boldsymbol{C} \rangle_{\text{full}} = \left\langle \boldsymbol{s}(i)\boldsymbol{s}(i)^H \right\rangle = \begin{bmatrix} \langle S_{hh}S_{hh}^* \rangle & \langle S_{hh}S_{hv}^* \rangle & \langle S_{hh}S_{vv}^* \rangle \\ \langle S_{hv}S_{hh}^* \rangle & \langle S_{hv}S_{hv}^* \rangle & \langle S_{hv}S_{vv}^* \rangle \\ \langle S_{vv}S_{hh}^* \rangle & \langle S_{vv}S_{hv}^* \rangle & \langle S_{vv}S_{vv}^* \rangle \end{bmatrix} \quad (1)$$

where $\langle \cdot \rangle$ denotes ensemble averaging, * denotes complex conjugation, and superscript H denotes the complex conjugate transpose. Reciprocity results in a covariance matrix with rank 3. $n\langle \boldsymbol{C} \rangle$ follows a complex Wishart distribution.

Spaceborne instruments often only transmit one polarization, e.g., horizontal, and receive both polarizations, giving rise to dual-polarization data, e.g., S_{hh} and S_{hv}. In this case, we have components $\langle S_{hh}S_{hh}^* \rangle$, $\langle S_{hh}S_{hv}^* \rangle$, and $\langle S_{hv}S_{hv}^* \rangle$ only. The resulting covariance matrix, i.e.,

$$\langle \boldsymbol{C} \rangle_{\text{dual}} = \begin{bmatrix} \langle S_{hh}S_{hh}^* \rangle & \langle S_{hh}S_{hv}^* \rangle \\ \langle S_{hv}S_{hh}^* \rangle & \langle S_{hv}S_{hv}^* \rangle \end{bmatrix} \quad (2)$$

has rank 2. The availability of full-polarimetric data allows us to extract dual-polarimetric subsets.

B. Multitemporal EMISAR Data Set

The SAR data used in this paper have been acquired by the fully polarimetric Danish airborne SAR system, i.e., EMISAR, which operates at two frequencies, i.e., C-band (5.3 GHz/5.7-cm wavelength) and L-band (1.25 GHz/24-cm wavelength) [30]. The nominal one-look spatial resolution is 2 m $\times$ 2 m, the ground range swath is approximately 12 km, and typical incidence angles range from 35° to 60°. The processed data from this system are fully calibrated by using an advanced internal calibration system [31]. In 1998, L-band data were acquired over a Danish agricultural test site on March 21, April 17, May 20, June 16, July 15, and August 16.

All acquisitions have been coregistered by identifying ground control points in the images and using an interferometric digital elevation model acquired by the EMISAR system [18]. Before the resampling, the original one-look scattering matrix data have been transformed to covariance matrix data, and these data have been averaged to reduce the speckle by a cosine-squared weighted 9×9 filter. The new pixel spacing in the images is 5 m by 5 m, and the effective spatial resolution is approximately 8 m by 8 m at midrange. After the averaging, the equivalent number of looks is approximately 13.

Fig. 1 (rowwise) shows RGB combinations of the diagonal elements of the full-polarimetry covariance matrix at L-band for March, April, and May (top row, from left to right) and for June, July, and August (bottom row, from left to right). $\langle S_{hv}S_{hv}^* \rangle$ (red) is linearly stretched between -36 and -6 dB; $\langle S_{hh}S_{hh}^* \rangle$ (green), between -30 and 0 dB; and $\langle S_{vv}S_{vv}^* \rangle$ (blue), between -24 and 0 dB. The darker areas in the March and April images are bare surfaces corresponding to spring crops, and the very bright areas in all images are forest areas, which are primarily coniferous forest. The development of the crops during the growing season is clearly seen in the series of images from March to August.

The changes we are looking for are changes in the average properties of the pixels/fields, i.e., the speckle patterns are not correlated. The images used are acquired with about a one-month interval, and the areas used are vegetated areas, forests, and agricultural fields. The speckle patterns of such areas are very likely uncorrelated after one month. We are dealing with incoherent change detection, where it seems realistic to assume that the measurements on the same pixel taken a month apart in scenes with natural vegetation are temporally independent.

Fig. 1. RGB images of diagonal elements of the L-band data. (Top row, from left to right) March, April, and May. (Bottom row, from left to right) June, July, and August.

III. Theory

In this section, we first discuss the challenges of the multiple testing problem. We then give the test statistic for the equality of several complex-Wishart-distributed matrices and the associated probability measure. Following this, the test statistic is factorized into independent test statistics for partial hypotheses. These two results are used in setting up the pixelwise change detector, and this is finally used in defining the fieldwise change detector.

A. The Multiple Testing Problem

In our setting, we consider pixels from a series of images taken at time points $t_1 < \cdots < t_k$, with distributions characterized by often multivariate parameters $\mu_1, \ldots, \mu_k$. We are concerned with detecting changes in those parameters, i.e., in assessing situations such as

$$\mu_1 = \cdots = \mu_{i-1} \neq \mu_i = \cdots = \mu_{\ell-1} \neq \mu_\ell = \cdots = \mu_k \quad (3)$$

which states that we have only changed after time point $i-1$ and again after time point $\ell - 1$. A simple approach would be sequentially testing hypotheses $\mu_j = \mu_{j-1}$ against $\mu_j \neq \mu_{j-1}$. However, this may give tests with a large false negative rate. For example, small gradual changes may not be detected, even when there is a clear trend throughout the time series. Therefore, we suggest to apply an omnibus test of hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ against all alternatives (see [27]). If this test is accepted, we conclude that no changes have occurred in the time interval $[t_1, t_k]$. If we reject the hypothesis, we may determine the occurrence of changes by suitable *post hoc* analysis using the results in Section III-C.

To clarify some concepts, let us briefly summarize some definitions in the hypothesis testing theory. In general, if we test a statistical hypothesis H_0, i.e., the item studied has no effect (here, corresponding to no change over time), against the alternative H_1, i.e., the item has an effect (here, corresponding to change over time), we may commit two different types of errors

- *Type-I error: Rejecting a true hypothesis (false positive, false alarm).* The *significance level* α of the test or the *false positive rate* is the probability of committing a type-I error, or, if we have a composite hypothesis, the maximum of the possible probabilities.
- *Type-II error: Accepting a false hypothesis (false negative).* The *type-II error rate, miss rate* or the *false negative rate* β is the probability of committing a type-II error. The *power* of the test is $1 - \beta$.

If a decision procedure is based on multiple (say, n) independent tests, each with significance level α_c, then the combined significance level $\alpha = \alpha_{\mathrm{FWER}}$, i.e., the so-called *familywise error rate* (FWER), is determined as $\alpha = \alpha_{\mathrm{FWER}} = 1 - [1 - \alpha_c]^n$. This error rate increases with n. For $\alpha_c = 0.05$ and $n = 5$, 10, and 15, we obtain the values 0.2262, 0.4013, and 0.5367, respectively, thus giving rather substantial false positive rates. If the individual tests are not independent, we can still state that $\alpha = \alpha_{\mathrm{FWER}} \leq n\alpha_c$. This is used in the so-called *Bonferroni correction*, where we put the per-comparison significance level $\alpha = \alpha_{\mathrm{FWER}}/n$ in order to control the familywise error (e.g., see [32]).

Example: We shall briefly outline some of the problems of multiple comparisons by a very simple example. We consider independent random variables $X_1, \ldots, X_k$ with means $\mu_1, \ldots, \mu_k$ and (known) standard deviation σ. Furthermore, we consider two different types of changes corresponding to two simple mean value structures as follows: 1) a signal that is linearly increasing with time (a constant increase of 2σ from one time period to the next); or 2) a piecewise constant signal with a jump/step at a given time point (a constant value until time point t_j, where it jumps $s\sigma$ and is again constant from then on), cf. Fig. 2.

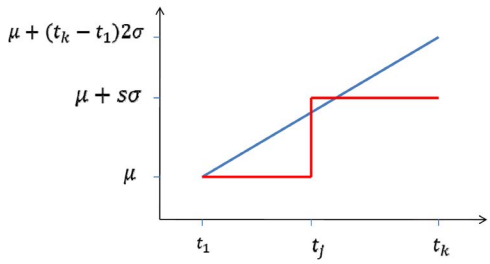


Fig. 2. Two mean value situations 1 and 2 considered in the text. The blue line corresponds to a steady increase, and the red corresponds to a discontinuous jump.

In situation 1, it is difficult to detect the changes by comparing successive observations. If we are using a standard statistical test using test statistic $U_m = (X_m - X_{m-j+1})/(\sqrt{2}\sigma)$ and a significance level of 5%, the false negative rate is as high as 70.70% for $j = 2$. However, if we consider comparisons between measurements corresponding to $j = 3,\ 4,$ and 5, the false negative rates will decrease to 19.26%, 1.12%, and 0.01%, respectively, but at the cost of an increased FWER (cf. the earlier discussion).

If we try to compensate for this increased FWER by using the Bonferroni correction for ten comparisons, i.e., using $\alpha = 0.5\%$, giving $\alpha_{\mathrm{FWER}} \leq 5\%$, the false negative rates will be 91.82%, 49.15%, 7.56%, and 0.22% for $j = 2, 3, 4,$ and 5, respectively, i.e., considerably higher than above. Thus, it may be difficult to control the FWER, while avoiding an unnecessarily large false negative rate.

If we instead use the omnibus test in the case with true means $\mu, \mu + 2\sigma, \mu + 4\sigma, \mu + 6\sigma$, and $\mu + 8\sigma$ (situation 1), the (likelihood ratio) test statistic is $\Sigma(X_i - \overline{X})^2/\sigma^2$, which is noncentrally chi-squared distributed with $k - 1$ degrees of freedom and noncentrality parameter $\Sigma(\mu_i - \overline{\mu})^2/\sigma^2$. If we use a significance level of 5%, the false negative rate becomes 0.02%. If the same overall change in the means instantaneously happens after time period 3 (situation 2), the true means are $\mu, \mu, \mu, \mu + 8\sigma$, and $\mu + 8\sigma$ (corresponding to situation 2), and the false negative rate becomes 0.00%. This example enhances the advantages of the omnibus test with respect to limiting the error rates.

B. Test for Equality of Several Complex Covariance Matrices

As stated in Section II-A, the sample covariance matrix of multilook fully polarimetric SAR data when multiplied by the number of looks will follow a complex Wishart distribution. In order to test for possible changes between several, say, k, time points, we must therefore investigate whether we may assume that several sample covariance matrices have the same expected value, say, $\boldsymbol{\Sigma}$, or whether we must assume that the expected values are different, say, equal to, say, $\boldsymbol{\Sigma}_i, i = 1, \ldots, k$, where at least two $\boldsymbol{\Sigma}_i$ differ. In a general setting, we therefore consider independent random variables $\boldsymbol{X}_i, i = 1, \ldots, k$, that follow complex Wishart distributions

$$\boldsymbol{X}_i \sim W_C(p, n, \boldsymbol{\Sigma}_i), \qquad i = 1, \ldots, k \tag{4}$$

where $E\{\boldsymbol{X}_i/n\} = \boldsymbol{\Sigma}_i$. To test the null hypothesis

$$H_0 : \ \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_k \tag{5}$$

against all alternatives we use the following test statistic (see the Appendix for the derivation, the work in [33] for the real case, and the work in [22] for the case with two complex matrices)

$$Q = k^{pnk} \frac{\prod_{i=1}^{k} |\boldsymbol{X}_i|^n}{|\boldsymbol{X}|^{nk}} = \left\{ k^{pk} \frac{\prod_{i=1}^{k} |\boldsymbol{X}_i|}{|\boldsymbol{X}|^k} \right\}^n. \tag{6}$$

Here, $|\cdot|$ denotes the determinant, the independent $\boldsymbol{X}_i = n\langle \boldsymbol{C} \rangle_i$ follow the complex Wishart distribution, i.e., $\boldsymbol{X}_i \sim W_{\mathbb{C}}(p, n, \boldsymbol{\Sigma}_i)$, and $\boldsymbol{X} = \sum_{i=1}^{k} \boldsymbol{X}_i \sim W_C(p, nk, \boldsymbol{\Sigma})$, where n is the number of looks. In addition, under H_0, $\hat{\boldsymbol{\Sigma}} = \boldsymbol{X}/(kn)$. $Q \in [0, 1]$, with $Q = 1$ for equality. For the logarithm of the test statistic, we get

$$\ln Q = n \left\{ pk \ln k + \sum_{i=1}^{k} \ln |\boldsymbol{X}_i| - k \ln |\boldsymbol{X}| \right\}. \tag{7}$$

If

$$f = (k-1)p^2 \tag{8}$$

$$\rho = 1 - \frac{(2p^2 - 1)}{6(k-1)p} \left(\frac{k}{n} - \frac{1}{nk} \right) \tag{9}$$

$$\omega_2 = \frac{p^2(p^2-1)}{24\rho^2} \left(\frac{k}{n^2} - \frac{1}{(nk)^2} \right) - \frac{p^2(k-1)}{4} \left(1 - \frac{1}{\rho} \right)^2 \tag{10}$$

then the probability of finding a smaller value of $-2\rho \ln Q$ is

$$P\{-2\rho \ln Q \leq z\} \simeq P\left\{\chi^2(f) \leq z\right\} + \omega_2 \left[P\left\{\chi^2(f+4) \leq z\right\} - P\left\{\chi^2(f) \leq z\right\} \right] \tag{11}$$

where $z = -2\rho \ln q$, and q is a particular realization (an observed value) of the stochastic variable Q. Instead of q, we may write q_{obs}; q and q_{obs} are used interchangeably in the following. See also the Appendix.

For full-polarimetry data, $p = 3$, and for dual polarimetry, $p = 2$; for single-band (HH, HV, or VV) data, $p = 1$. In the latter case, $\boldsymbol{X}_i$ and $\boldsymbol{X}$ are Gamma-distributed scalar random variables X_i and X, respectively, and Q becomes

$$Q = \left\{ k^k \frac{\prod_{i=1}^{k} X_i}{X^k} \right\}^n. \tag{12}$$

For two time points, i.e., $k = 2$, this is equivalent to the classical ratio detector [1], [2].

C. Test for Equality of First $j < k$ Complex Covariance Matrices

If the aforementioned test shows that we cannot reject the hypothesis of equality, no change has occurred over the time span covered by the data. If we, on the other hand, can reject the hypothesis, change has occurred at some time point. In order to establish at which time a change has occurred, we shall use the fact that the likelihood ratio test statistic may be decomposed into a product of test statistics that test simpler hypotheses and that are independent if H_0 is true. To test whether the first $j,\ 1 < j < k$, complex covariance matrices $\boldsymbol{\Sigma}_i$ (p by p) are equal, i.e., given that

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_{j-1} \tag{13}$$

then likelihood ratio test statistic R_j for testing the hypothesis

$$H_{0,j} : \boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}_{j-1} \quad \text{against} \quad H_{1,j} : \boldsymbol{\Sigma}_j \neq \boldsymbol{\Sigma}_{j-1} \tag{14}$$

is

$$\begin{aligned} R_j &= \frac{j^{jpn}}{(j-1)^{(j-1)pn}} \frac{|\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_{j-1}|^{(j-1)n} |\boldsymbol{X}_j|^n}{|\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_j|^{jn}} \\ &= \left\{ \frac{j^{jp}}{(j-1)^{(j-1)p}} \frac{|\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_{j-1}|^{(j-1)} |\boldsymbol{X}_j|}{|\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_j|^j} \right\}^n \end{aligned} \tag{15}$$

or

$$\begin{aligned} \ln R_j = n \Bigg\{ & p(j \ln j - (j-1)\ln(j-1)) \\ & + (j-1) \ln \left| \sum_{i=1}^{j-1} \boldsymbol{X}_i \right| + \ln |\boldsymbol{X}_j| - j \ln \left| \sum_{i=1}^{j} \boldsymbol{X}_i \right| \Bigg\}. \end{aligned} \tag{16}$$

Furthermore, we have

$$Q = \prod_{j=2}^{k} R_j \tag{17}$$

and if H_0 is true, then the random variables $R_2, \ldots, R_k$ are independent.

Finally, letting

$$f = p^2 \tag{18}$$

$$\rho_j = 1 - \frac{2p^2 - 1}{6pn} \left(1 + \frac{1}{j(j-1)} \right) \tag{19}$$

$$\begin{aligned} \omega_{2j} = & -\frac{p^2}{4} \left(1 - \frac{1}{\rho_j} \right)^2 \\ & + \frac{1}{24n^2} p^2 (p^2 - 1) \left(1 + \frac{2j-1}{j^2(j-1)^2} \right) \frac{1}{\rho_j^2} \end{aligned} \tag{20}$$

we get

$$\begin{aligned} P\{-2\rho_j \ln R_j \leq z\} \simeq & P\left\{\chi^2(f) \leq z\right\} \\ & + \omega_{2j} \left[P\left\{\chi^2(f+4) \leq z\right\} - P\left\{\chi^2(f) \leq z\right\} \right] \end{aligned} \tag{21}$$

where $z = -2\rho_j \ln r_j$, and r_j is a particular realization (an observed value) of the stochastic variable R_j. Instead of r_j, we may write $r_{j,\text{obs}}$; r_j and $r_{j,\text{obs}}$ are used interchangeably in the following. See the Appendix that also gives the resulting formulas for the Gamma-distributed scalar case.

D. Pixelwise Change Detection

We start by looking at the Gamma-distributed case (cf. the Appendix). This will allow a more intuitive presentation, e.g., by plotting the power of the SAR signal as a function of time. Furthermore, the notation becomes somewhat simpler. We have time points $t_1, \ldots, t_k$ corresponding to parameters $\beta_1, \ldots, \beta_k$, respectively, and we introduce the global hypotheses

$$H_0^{(\ell)} : \beta_\ell = \beta_{\ell+1} = \cdots = \beta_k, \qquad \ell = 1, \ldots, k-1 \tag{22}$$

i.e., the last $k - \ell + 1$ of all the parameters are equal. Furthermore, we consider the marginal hypotheses

$$\begin{aligned} H_{0,j}^{(\ell)} : \beta_{\ell+j-1} = \beta_{\ell+j-2} (= \beta_{\ell+j-3} = \cdots = \beta_\ell), \\ j = 2, \ldots, k-\ell+1 \end{aligned} \tag{23}$$

i.e., the first j of the parameters in the global hypotheses $H_0^{(\ell)}$ are equal.

TABLE I
(HYPOTHETICAL) DATA CONSIDERED FOR THE GAMMA DISTRIBUTION EXAMPLE

j	1	2	3	4	5	6	7	8
x_j	1.3338	2.0683	1.3494	1.3858	0.0806	1.6302	1.5201	1.9932

The omnibus test statistic for testing $H_0^{(\ell)}$ against all alternatives based on $X_\ell, X_{\ell+1}, \ldots, X_k$ is

$$Q^{(\ell)} = \left\{ (k-\ell+1)^{k-\ell+1} \frac{X_\ell \ldots X_k}{(X_\ell + \cdots + X_k)^{k-\ell+1}} \right\}^n. \tag{24}$$

Let us assume that $H_{0,j-1}^{(\ell)}$ is true, i.e., the first $j-1$ parameters are equal or $\beta_\ell = \beta_{\ell+1} = \cdots = \beta_{\ell+j-2}$. Then, the test statistic for testing $H_{0,j}^{(\ell)}$ against the alternative $\beta_{\ell+j-1} \neq \beta_{\ell+j-2}$, i.e., for testing that the jth parameter is equal to the $j-1$ preceding parameters, becomes

$$\begin{aligned} R_j^{(\ell)} = & \left\{ \frac{j^j}{(j-1)^{j-1}} \frac{(X_\ell + \cdots + X_{\ell+j-2})^{j-1} \, X_{\ell+j-1}}{(X_\ell + \cdots + X_{\ell+j-1})^j} \right\}^n, \\ & j = 2, \ldots, k-\ell+1 \end{aligned} \tag{25}$$

$$Q^{(\ell)} = R_2^{(\ell)} \cdots R_{k-\ell+1}^{(\ell)} \tag{26}$$

with independence if $H_0^{(\ell)}$ is true.

In broad terms, the algorithm becomes

1) Set $\ell = 1$.
2) Test $H_0^{(\ell)}$ against all alternatives.
   If accepted, conclude that there are no changes in the interval $[t_\ell, t_k]$. Go to 5.
   If rejected, conclude that there is at least one change in the interval $[t_\ell, t_k]$, and go to 3.
3) Test marginal hypotheses $H_{0,j}^{(\ell)}$, and let the first significant hypothesis be $H_{0,r+1}^{(\ell)}$.
   Conclude that we have a change in $[t_{\ell+r-1}, t_{\ell+r}]$.
4) Set $\ell = \ell + r$, and go to 2.
5) Finish.

The algorithm is illustrated in the next example.

Example: We now consider (hypothetical) $\langle S_{hh} S_{hh}^* \rangle$ values (corresponding to 13-look EMISAR data) from eight time points. The observations are given in Table I. It is assumed that they represent independent realizations of Gamma-distributed random variables $X_i \sim G(13, \beta_i)$, $i = 1, \ldots, 8$.

By direct computation, we get that $-2 \ln Q = 54.2510$, and comparing this to quantiles in a $\chi^2(7)$-distribution ($(k-1)p^2 = 7$ for $k = 8$ and $p = 1$) shows that $P\{Q^{(1)} < q^{(1)}\} \simeq 0$, i.e., this value is significant at all reasonable levels. Therefore, we conclude that we have (at least) one change in the time period considered, i.e., $i = 1, \ldots, 8$. We want to determine the time point for the first change. Therefore, we successively compute the quantities $R_j^{(1)}$ and $P\{R_j^{(1)} < r_j^{(1)}\}$, $j = 2, \ldots, 8$, where the upper index (1) indicates that we are looking at the first global hypothesis presented in Table II, column 2.

In the decomposition of the likelihood ratio test statistic Q, we see that $R_5^{(1)}$ is the first significant component corresponding to the rejection of the hypothesis $\beta_5 = \beta_4$. Thus, the conclusion is that, so far, we may assume that $\beta_1 = \beta_2 = \beta_3 = \beta_4$, and we must investigate whether there are changes in period $i = 5, \ldots, 8$, i.e., test global hypothesis $H_0^{(5)} : \beta_5 = \cdots = \beta_8$. Based on observations $X_5, \ldots, X_8$, we compute the likelihood ratio test statistic for $H_0^{(5)}$ and find that $P\{Q^{(5)} < q^{(5)}\} \simeq 0$. We conclude that

TABLE II
p-VALUES OF THE DIFFERENT TEST STATISTICS FOR THE GAMMA-DISTRIBUTED RANDOM VARIABLES. THE CHANGE INDEXES ARE HIGHLIGHTED IN YELLOW, AND THE p-VALUES FOR THE ASSOCIATED GLOBAL TESTS, IN GREEN. WE NEED TO CALCULATE THE HIGHLIGHTED NUMBERS ONLY, THE OTHER NUMBERS ARE SHOWN FOR ILLUSTRATIVE PURPOSES

Marg. Hyp.	Global Hypotheses $H_0^{(\ell)}$, $\ell = 1,\ldots,7$						
	$\beta_1 = \cdots = \beta_8$	$\beta_2 = \cdots = \beta_8$	$\beta_3 = \cdots = \beta_8$	$\beta_4 = \cdots = \beta_8$	$\beta_5 = \cdots = \beta_8$	$\beta_6 = \cdots = \beta_8$	$\beta_7 = \beta_8$
$H_{0,j}^{(\ell)}$	$P\{R_j^{(1)} < r_j^{(1)}\}$	$P\{R_j^{(2)} < r_j^{(2)}\}$	$P\{R_j^{(3)} < r_j^{(3)}\}$	$P\{R_j^{(4)} < r_j^{(4)}\}$	$P\{R_j^{(5)} < r_j^{(5)}\}$	$P\{R_j^{(6)} < r_j^{(6)}\}$	$P\{R_j^{(7)} < r_j^{(7)}\}$
$\beta_2 = \beta_1$	0.2653						
$\beta_3 = \beta_2$	0.5013	0.2780					
$\beta_4 = \beta_3$	0.6801	0.5423	0.9459				
$\beta_5 = \beta_4$	0.0000	0.0000	0.0000	0.0000			
$\beta_6 = \beta_5$	0.3587	0.3378	0.0723	0.0151	0.0000		
$\beta_7 = \beta_6$	0.6096	0.6057	0.2980	0.2129	0.0824	0.8585	
$\beta_8 = \beta_7$	0.1581	0.1642	0.0744	0.0636	0.0442	0.4831	0.4903
$P\{Q^{(\ell)} < q^{(\ell)}\}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.7696	0.4903

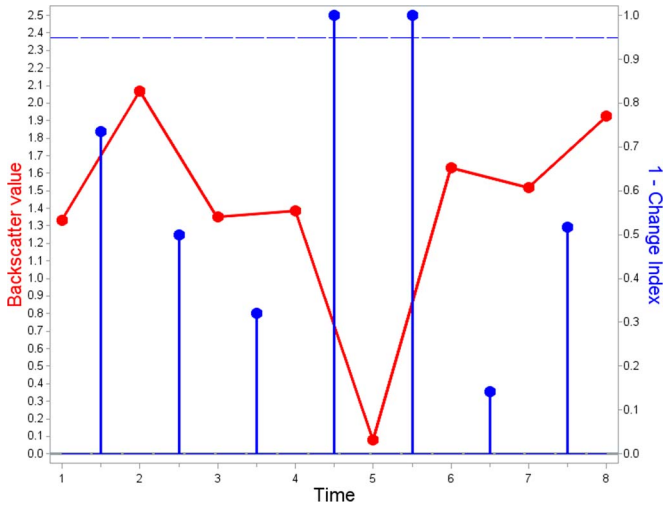


Fig. 3. (Hypothetical) observations and the change indexes for the Gamma distribution example.

there is at least one change in the period considered and decompose the likelihood ratio statistic. We see that $R_2^{(5)}$ is the first significant component corresponding to the rejection of the hypothesis $\beta_6 = \beta_5$. We thus have a change between time points 5 and 6. Then, we must investigate whether there is a change between the last three time points, i.e., we consider global hypothesis $H_0^{(6)} : \beta_6 = \beta_7 = \beta_8$. We find that $P\{Q^{(6)} < q^{(6)}\} = 0.7696$, i.e., we assume that we have no changes in this period.

Thus, the conclusion is that we observe significant changes between

- β_4 and β_5;
- β_5 and β_6.

Therefore, we conclude that we have the following distinct populations

- $\beta_1 = \beta_2 = \beta_3 = \beta_4$.
- β_5.
- $\beta_6 = \beta_7 = \beta_8$.

Remarks: If we introduce the term change index for the relevant quantities $P\{R_j^{(\ell)} < r_j^{(\ell)}\}$, we have that large values of one minus the change index correspond to changes. From a statistical point of view, a threshold of 0.95 seems natural. In the Gamma case, we may plot the observations and the quantities one minus the change index in the same coordinate system. This is done in Fig. 3, which illustrates the outcome of the change detection algorithm. ■

The above description of the change detection algorithm may be immediately generalized to the complex Wishart distribution by simply replacing parameter β with $\mathbf{\Sigma}$ and by using (6) and (15) in setting up the global and marginal test statistics based on the last $k - \ell + 1$ observations. A detailed description of the algorithm is given in Fig. 4.

E. Fieldwise Change Detection

Before defining the change index for a field, let us initially state some facts from statistical testing theory. The p-value of a statistical test is the probability of getting a test statistic that is at least as extreme as the test statistic observed, assuming that the null hypothesis is actually true (and that the assumptions of the analysis are met). If the p-value is smaller than the prescribed significance level α (e.g., $\alpha = 0.05$), we reject the hypothesis since the discrepancy between the data and the hypothesis is too large.

Furthermore, if the sampling distribution is continuous, the distribution of the p-values will be uniform over the interval [0, 1] if the null hypothesis is true. This implies, of course, that if the same experiment is replicated many times, if the hypothesis is true, and if we test on a significance level of 5%, then 5% of the p-values will fall in the interval [0, 0.05], and the remaining p-values will be larger than 0.05. If the null hypothesis is not true, then the fraction of p-values falling in the interval [0, 0.05] will be, in general, much larger than 5%. How large it is depends on the power of the test.

The change index for a field may now be defined by first considering the empirical distributions of the different p-values obtained (i.e., the $P\{Q < q\} = 1 - P\{-2\rho \ln Q \leq -2\rho \ln q\}$ and $P\{R_j^{(\ell)} < r_j^{(\ell)}\} = 1 - P\{-2\rho \ln R_j^{(\ell)} \leq -2\rho \ln r_j^{(\ell)}\}$ values) for each pixel observed over the time span considered and then by applying the one-pixel definition on suitable measures of location for those distributions, such as the mean or the median.

IV. RESULTS

In this section, the data set described in Section II-B is used to illustrate different aspects of the test statistics, and in that process, we utilize different parts of the time series to provide illustrative examples. We first show the strength of the omnibus test statistic Q when applied to three different areas covering a forest, a rye field, and a grass field using data over five time points from March to July only. We then proceed to show the power of the factorization of Q into R_j, this time using data from six time points from March to August.

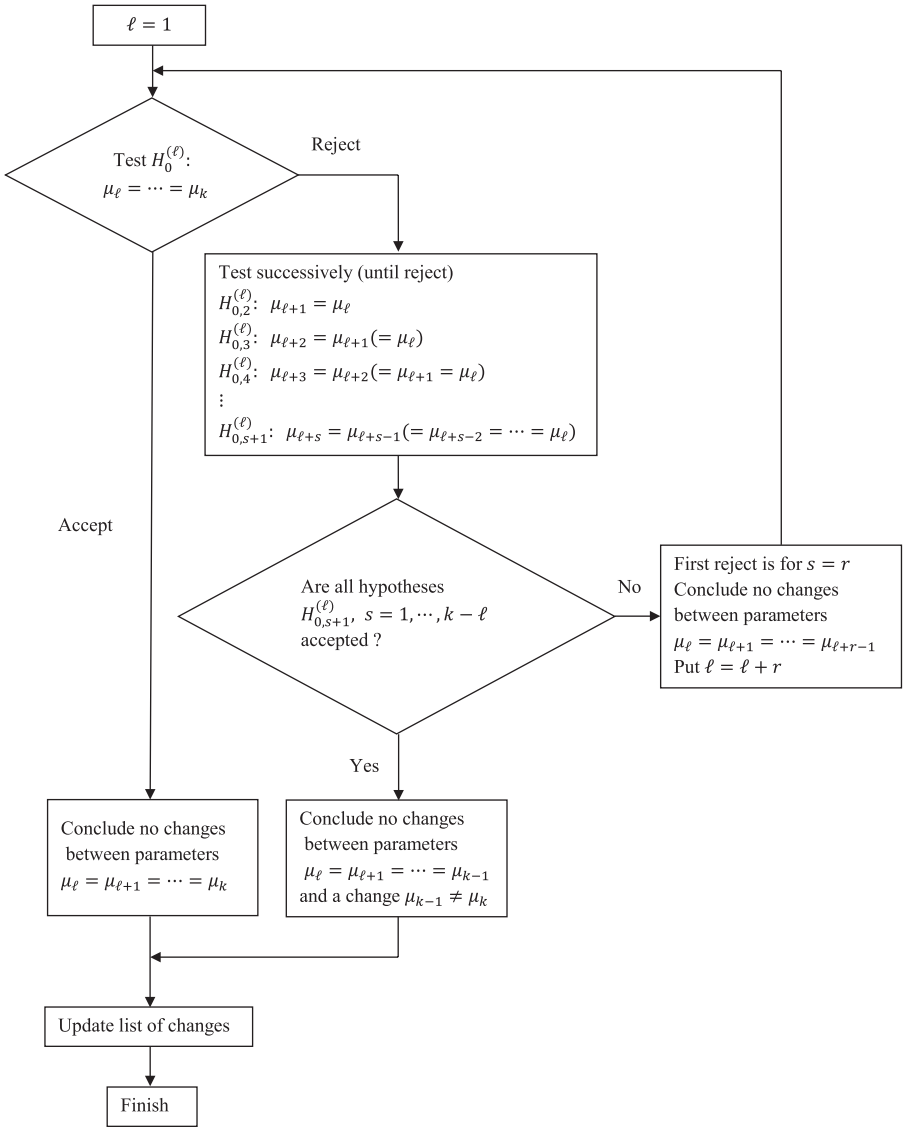
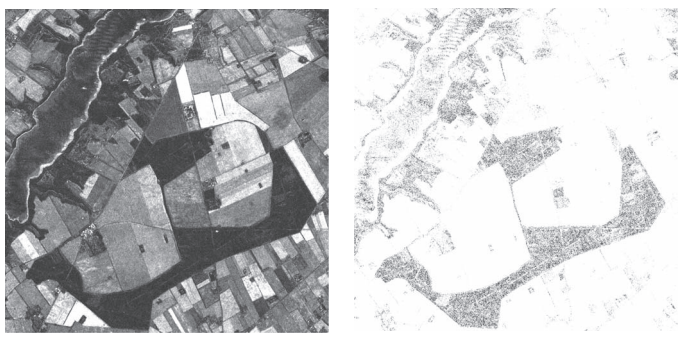


Fig. 4. Flowchart for the change detection method. μ is a generic name for the unknown parameters considered. In the leftmost part, we are investigating whether there are changes in the interval $[t_\ell, t_k]$ using Q-test statistic (6). If the answer is no, the analysis is finished. If the answer is yes, i.e., we have at least one change in $[t_\ell, t_k]$, we go to the column in the middle. Based on observations at $t_\ell, \ldots, t_{\ell+s}$, we (for $s = 1, \ldots, k-1$) successively investigate whether there are changes between time points $t_{\ell+s-1}$ and $t_{\ell+s}$. This is done by testing hypothesis $H_{0,j+1}^{\ell} : \mu_{\ell+j} = \mu_{\ell+j-1}$ corresponding to $H_{0,j}$ from (14) using test statistic $R_j^{(\ell)}$ corresponding to R_j from (15). If we do not identify any changes before time point t_{k-1}, we conclude that the change in the interval $[t_\ell, t_k]$ falls in the interval $[t_{k-1}, t_k]$. If the first change we identify occurs in $[t_{\ell+r-1}, t_{\ell+r}]$, we conclude that there are no changes in the interval $[t_\ell, t_{\ell+r-1}]$ and that there is a change in $[t_{\ell+r-1}, t_{\ell+r}]$. We then update ℓ to $\ell + r$ and start again in the leftmost column.

A. Change Indexes for Three Different Cases: Forest, Rye, and Grass

Fig. 5 shows $-2\rho \ln q_{\mathrm{obs}}$ for full polarimetry (left image) and $P\{Q \geq q_{\mathrm{obs}}\}$ (right image). Here, $\rho = 0.91282$, and $\omega_2 = 0.023577$. The left-hand image shows low values for the forest areas, indicating that no changes have occurred during the five time points, which is a reasonable result for the coniferous forest areas.

In order to provide explanations of the results for the test statistics, the following analysis will be based on results from areas such as fields with a given crop, where we have *in situ* information on ground usage. In this section, we show the results for three cases with different change patterns over time, i.e., a forest area with no changes, a rye field where changes



(a) (b)

Fig. 5. (a) $-2\rho \ln q_{\mathrm{obs}}$ for full polarimetry (linearly stretched between 0 and 300). (b) $P\{Q \geq q_{\mathrm{obs}}\}$. Both are based on data from March to July. In both images, low values (dark tones) correspond to no change.

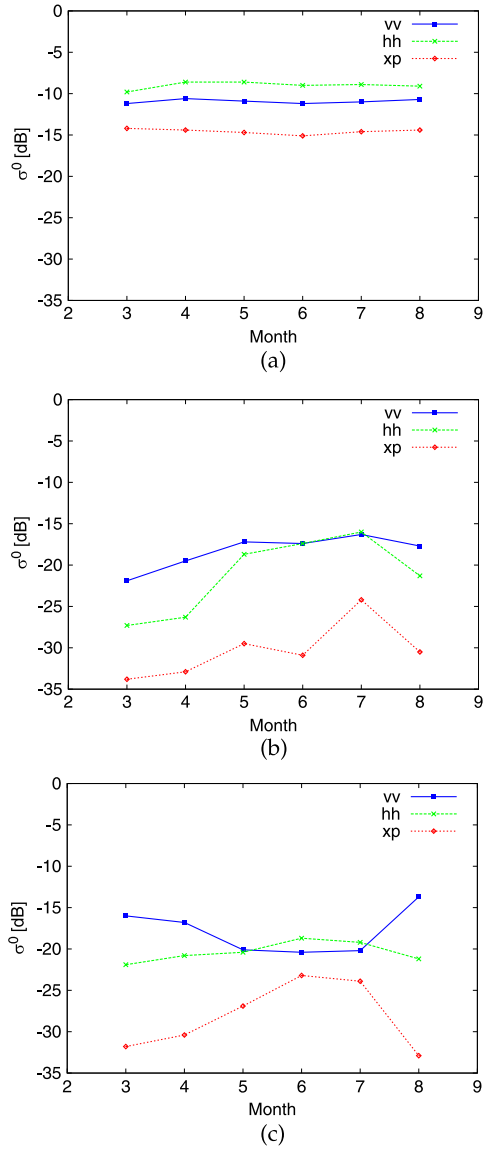


Fig. 6. Backscatter coefficients for (a) forest, (b) rye, and (c) grass. In the legend, vv denotes $\langle S_{vv}S_{vv}^{*}\rangle$, hh denotes $\langle S_{hh}S_{hh}^{*}\rangle$, and xp denotes $\langle S_{hh}S_{vv}^{*}\rangle$.

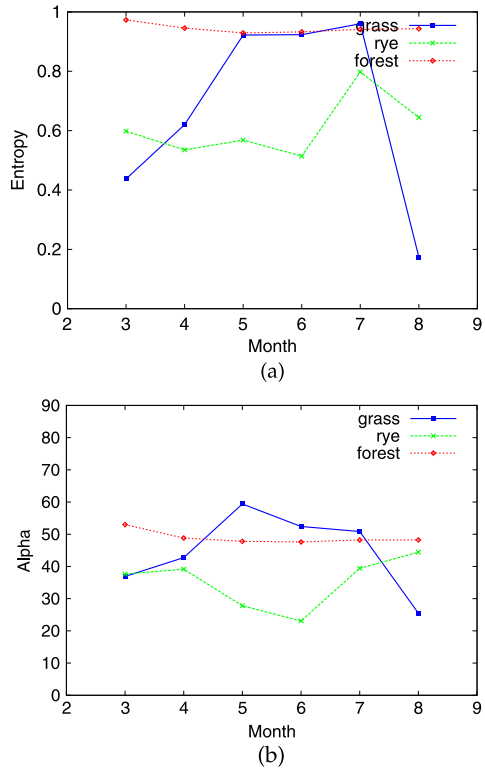


Fig. 7. (a) Entropy and (b) alpha angle for grass, rye, and forest. Entropy has no unit, and the alpha angle is in degrees.

may be detected by both the omnibus test and the pairwise tests, and a grass area where changes may be detected only by the omnibus test and not the pairwise tests.

Backscatter coefficients for these areas are shown in Fig. 6, and the polarimetric entropy and alpha angle parameters from the Cloude–Pottier eigenanalysis decomposition [34] are shown in Fig. 7. The forest area has very constant backscatter through all the images. The entropy for the forest area is very close to 1 for all acquisitions, and the alpha angle is approximately 50°, which indicates, as expected, that the backscatter from the forest is dominated by volume scattering. It is clearly seen in these results that no changes occur for the forest area through all the acquisitions.

For the grass and rye areas, on the other hand, the backscatter coefficients and the entropy and alpha angle parameters change through the time series. For the rye field, the entropy and alpha angle values show for the first four acquisitions that the backscattering mechanism is the medium entropy surface scattering type [34], which indicates rough surface scattering with canopy penetration effects. Rye is a relatively sparse crop, which corresponds to the results that the backscatter is dominated by surface scattering affected by the vegetation layer. Between April and May, we see an increased influence of the vegetation layer by the increase in the HH backscatter [cf. Fig. 6(b)]. For the July acquisition, the vegetation layer has become very dense that we now see volume scattering, i.e., increased cross-polarized backscatter [cf. Fig. 6(b)], and increased entropy [cf. Fig. 7(a)]. Finally, for the August acquisition, the crop is now dried out; thus, less volume scattering occurs, hence we have decreasing cross-polarized backscatter and smaller entropy.

The scattering mechanism for the grass area changes from surface scattering affected by the vegetation layer in March and April with the entropy and alpha angle values similar to the rye field (cf. Fig. 7), over volume scattering in May with some contribution of double-bounce scattering with an alpha angle of about 60° [cf. Fig. 7(b)], to full volume scattering in June and July with the entropy close to 1 [cf. Fig. 7(a)]. These changes correspond to the growth and conditions of the grass. In March, April, and May, the radar wave can penetrate the not-so-dense grass vegetation, and in May, the vegetation is, however, sufficiently dense to result in volume scattering from the vegetation. At the same time the vegetation is still very sparse such that some double-bounce scattering can occur, and in the June and July acquisitions, the vegetation is very dense such that volume scattering occurs. In August, the grass is cut; hence, the entropy and the alpha angle now show surface scattering behavior (cf. Fig. 7), and the VV backscatter and the cross-polarized backscatter increases and decreases, respectively [cf. Fig. 6(c)].

TABLE III
AVERAGE PROBABILITIES OF OBTAINING MORE EXTREME VALUES OF THE TEST STATISTIC FOR THE FOREST, RYE, AND GRASS AREAS FOR PAIRWISE COMPARISONS AND FOR THE OMNIBUS TEST, WITH THE DATA FROM MARCH TO JULY. VALUES BELOW 0.05 INDICATE STATISTICALLY SIGNIFICANT CHANGES

Average $P\{Q < q_{obs}\}$	Pairwise comparisons				Omnibus test
Hypothesis $H_0^{(\ell)}$	Mar=Apr	Apr=May	May=Jun	Jun=Jul	Mar=Apr=May=Jun=Jul
Forest	0.3925	0.6235	0.4913	0.4567	0.3823
Rye	0.4249	0.0131	0.3344	0.0242	0.0000
Grass	0.2808	0.0638	0.1244	0.4311	0.0001

TABLE IV
p-VALUES OF THE DIFFERENT TEST STATISTICS FOR THE FOREST AREA. THE PATH LEADING TO THE CHANGE INDEXES IS INDICATED

Forest	Global Hypothesis				
Marg. Hyp.	Mar= $\cdots$ =Aug	Apr= $\cdots$ =Aug	May= $\cdots$ =Aug	Jun= $\cdots$ =Aug	Jul=Aug
$H_{0,j}^{(\ell)}$	$P\{R_j^{(Mar)} < r_j^{(Mar)}\}$	$P\{R_j^{(Apr)} < r_j^{(Apr)}\}$	$P\{R_j^{(May)} < r_j^{(May)}\}$	$P\{R_j^{(Jun)} < r_j^{(Jun)}\}$	$P\{R_j^{(Jul)} < r_j^{(Jul)}\}$
Apr = Mar	0.3925				
May = Apr	0.5469	0.6235			
Jun = May	0.4121	0.4085	0.4913		
Jul = Jun	0.4229	0.4289	0.4629	0.4567	
Aug = Jul	0.4116	0.4099	0.4095	0.4154	0.4523
$P\{Q < q\}$	0.3494	0.4218	0.4230	0.4108	0.4523

TABLE V
p-VALUES OF THE DIFFERENT TEST STATISTICS FOR THE RYE AREA. THE PATH LEADING TO THE CHANGE INDEXES IS INDICATED

Rye	Global Hypothesis				
Marg. Hyp.	Mar= $\cdots$ =Aug	Apr= $\cdots$ =Aug	May= $\cdots$ =Aug	Jun= $\cdots$ =Aug	Jul=Aug
$H_{0,j}^{(\ell)}$	$P\{R_j^{(Mar)} < r_j^{(Mar)}\}$	$P\{R_j^{(Apr)} < r_j^{(Apr)}\}$	$P\{R_j^{(May)} < r_j^{(May)}\}$	$P\{R_j^{(Jun)} < r_j^{(Jun)}\}$	$P\{R_j^{(Jul)} < r_j^{(Jul)}\}$
Apr = Mar	0.4249				
May = Apr	0.0026	0.0131			
Jun = May	0.1400	0.2599	0.3344		
Jul = Jun	0.0002	0.0023	0.0237	0.0242	
Aug = Jul	0.1758	0.1292	0.0669	0.0468	0.0208
$P\{Q < q\}$	0.0000	0.0001	0.0011	0.0080	0.0208

Table III shows the average probabilities of the test statistics for the forest, rye, and grass areas using the first five acquisitions. The values shown are the average values over all pixels for each area of the observed significance levels $\alpha_{\rm obs}$ (i.e., the probabilities $P\{Q < q_{\rm obs}\}$ of getting the values of the test statistic that are more extreme than the observed values under the null hypothesis). Table III shows both the results of performing consecutive pairwise comparisons between the acquisitions (corresponding to the test statistic in [22]) and the results of using the omnibus test statistic presented in this paper [cf. (11)].

It is clear that no change can be detected for the forest area in the pairwise comparisons and in the omnibus test. For the rye area, the omnibus test clearly detects change in the time series. From the pairwise comparisons, it is seen that changes can be detected at a 5% significance level between the April and May acquisitions and between the June and July acquisitions, corresponding to the observed changes in the polarimetric parameters described above.

The omnibus test for the grass area clearly shows that the parameters over the five acquisitions have not been constant, i.e., changes have occurred during this period. The pairwise comparisons, however, show no changes at a 5% significance level. Although the polarimetric analysis above describes some relatively significant changes during the five acquisitions, the pairwise comparisons do not pick up these changes. This result shows the power of the omnibus test, where the changes between the consecutive images are too small to be detected, but the overall change during all the acquisitions is significant.

B. Points of Change for Three Different Cases: Forest, Rye, and Grass

In Table IV, we have shown the p-values for the forest area and the corresponding path leading to the change index for the forest area computed by looking at the averages described earlier. In Tables V and VI, we have shown the equivalent results for a rye field and a grass field, respectively.

In Tables IV–VI, the first number in each column corresponds to tests for the relevant pairwise comparisons, see (6), where $k = 2$ in the pairwise case.

For the forest results in Table IV, it is seen in the $P\{Q < q\}$ value in the first column that no change has occurred for all acquisitions (with a significance level of 0.05). This is also seen from the individual $P\{R_j^{(\rm Mar)} < r_j^{(\rm Mar)}\}$ values in the first column.

For the rye results in Table V, the $P\{Q < q\}$ value for the first column shows that a change has occurred, and according to the $P\{R_j^{(\rm Mar)} < r_j^{(\rm Mar)}\}$ values in the first column, this change has occurred between April and May. Now, moving to the third column in Table V, the $P\{Q < q\}$ value here shows that a change has occurred between May and August. The $P\{R_j^{(\rm May)} < r_j^{(\rm May)}\}$ values in the third column show that the change has occurred between June and July. Finally, moving to the fifth column, it is seen that a change has occurred between

TABLE VI
p-VALUES OF THE DIFFERENT TEST STATISTICS FOR THE GRASS AREA. THE PATH LEADING TO THE CHANGE INDEXES IS INDICATED

Grass	Global Hypothesis				
Marg. Hyp.	Mar= $\cdots$ =Aug	Apr= $\cdots$ =Aug	May= $\cdots$ =Aug	Jun= $\cdots$ =Aug	Jul=Aug
$H_{0,j}^{(\ell)}$	$P\{R_j^{(Mar)} < r_j^{(Mar)}\}$	$P\{R_j^{(Apr)} < r_j^{(Apr)}\}$	$P\{R_j^{(May)} < r_j^{(May)}\}$	$P\{R_j^{(Jun)} < r_j^{(Jun)}\}$	$P\{R_j^{(Jul)} < r_j^{(Jul)}\}$
Apr = Mar	0.2808				
May = Apr	0.0112	0.0638			
Jun = May	0.0056	0.0319	0.1244		
Jul = Jun	0.1094	0.2386	0.3879	0.4311	
Aug = Jul	0.0000	0.0000	0.0000	0.0000	0.0000
$P\{Q < q\}$	0.0000	0.0000	0.0000	0.0000	0.0000

July and August. We thus conclude that we have the following distinct populations for rye

1) March = April.
2) May = June.
3) July.
4) August.

Using the same methodology for grass in Table VI, we find that changes have occurred between April and May and between July and August. In this case, the distinct populations are

1) March = April.
2) May = June = July.
3) August.

These populations are consistent with the analysis of the polarimetric parameters described in Section IV-A. The forest areas clearly show no change for all six acquisitions. The rye field has four distinct populations, i.e., March–April, May–June, July, and August, as shown above. For March and April, surface scattering is dominating, and in May and June, an increased contribution from volume scattering is seen. In July, more volume scattering is seen, and finally, in August, less volume scattering is seen, which is due to the crop drying out. For the grass field, three distinct populations are identified, i.e., March–April, May–June–July, and August. From the backscatter coefficients in Fig. 6(c) and the entropy and the alpha angle in Fig. 7, groupings of the parameters according to these populations are clearly seen, and it also corresponds to the analysis of the scattering mechanisms given in Section IV-A.

C. Distribution of the p-Values for Forest and Grass

The question naturally arises whether the chosen mean values are proper descriptors of the underlying distributions.

In Fig. 8, the histograms (presented rowwise) behind the six p-values that constitute the change index for forest are shown. For the forested area, no changes are seen, and the means of the p-values ($\bar{p}$) for the omnibus test $P\{Q<q\}$ ($\bar{p} = 0.3494$) and for the marginal tests $P\{R_j^{(\ell)} < r_j^{(\ell)}\}$ ($\bar{p} = 0.3925, 0.5469, 0.4121, 0.4229,$ and 0.4116) are fairly representative for the histograms. We only have minor deviations between the histograms and the theoretical uniform distribution that we would get, assuming that no changes have occurred. The reason for these minor deviations is probably twofold: First, there may very well be changes in small patches, and second, some of the assumptions behind the sampling theory (independence of observations and distributional properties) may not be fulfilled.

The grass field shows a slightly more complicated scenario. The first three histograms in Fig. 9 show the distribution of

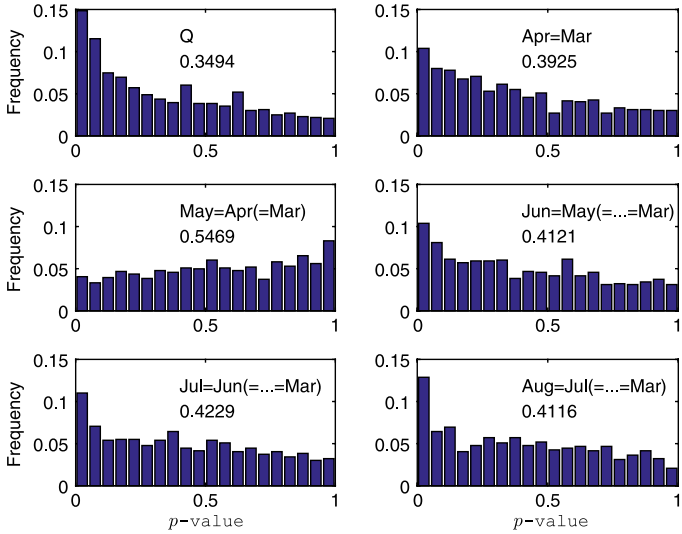


Fig. 8. Histograms of the p-values for testing the hypotheses Mar = Apr = May = Jun = Jul = Aug, i.e., no changes in the entire period, and for testing Apr = Mar, May = Apr(= Mar), Jun = May(= Apr = Mar), Jul = Jun (= May = Apr = Mar), and Aug = Jul(= Jun = May = Apr = Mar) for the forest area. These histograms present the distribution of the pixelwise change indexes for the forest area. The averages are found in Table IV.

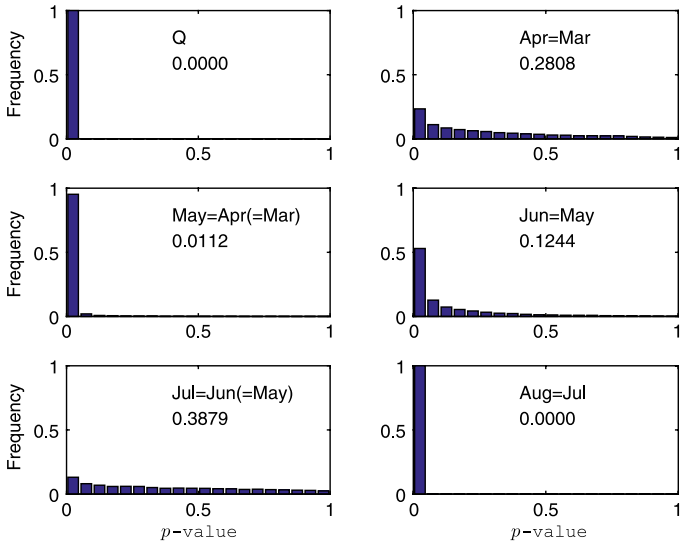


Fig. 9. Histograms of the p-values for testing hypotheses Mar = Apr = May = Jun = Jul = Aug, i.e., no changes in the entire period, and for testing Apr = Mar, May = Apr(= Mar), Jun = May, Jul = Jun(= May), and Aug = Jul for the grass area. These histograms present the distribution of the pixelwise change indexes for the grass area. The averages are found in Table VI.

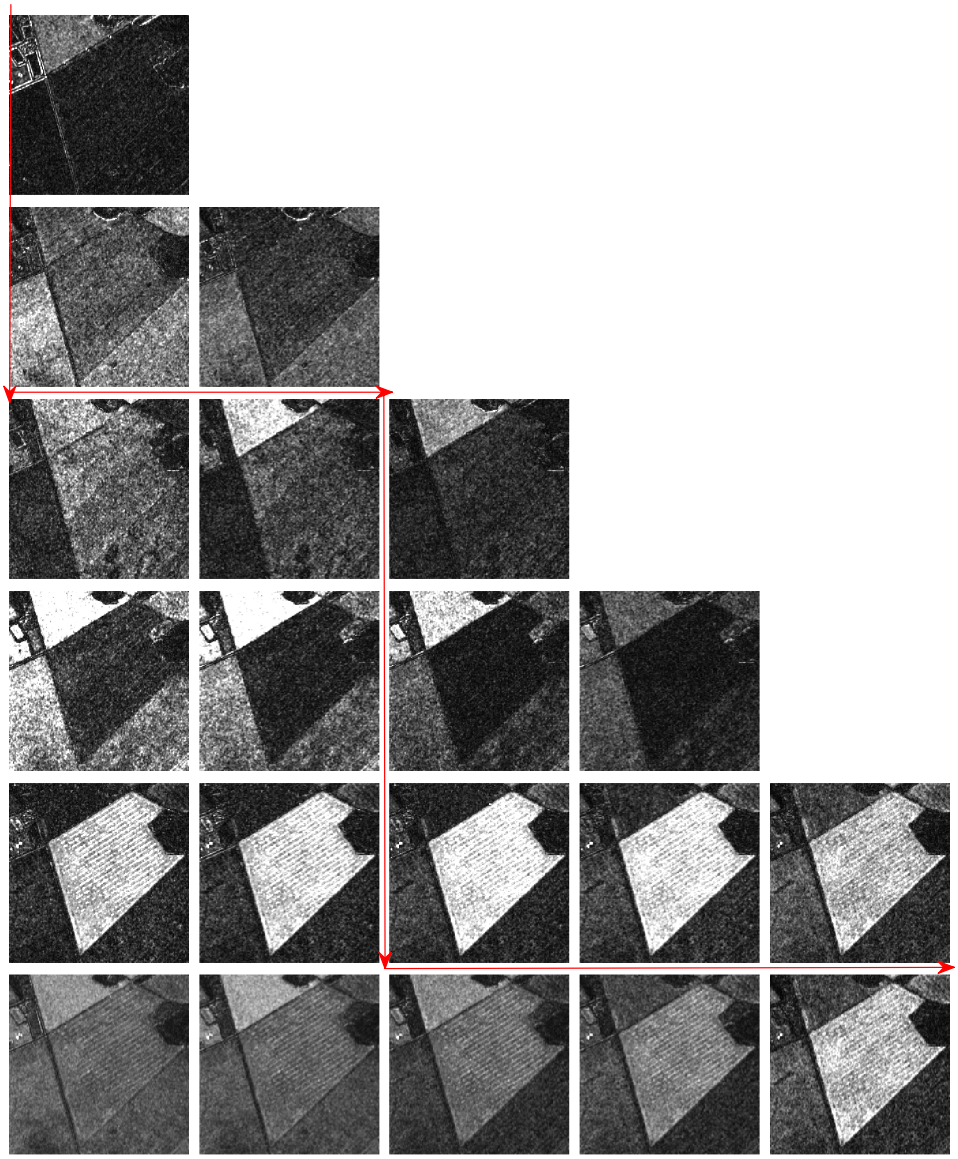


Fig. 10. Zoom on the grass field reported on in Table VI. The p-value corresponding to each subimage has the same relative position as in Table VI. Moreover, the path leading to the change indexes is indicated both in this figure and in Table VI. Rows 1–5: $-2\rho_j \ln r_{j,\mathrm{obs}}$ for full-polarimetry data stretched linearly between 0 and 100. The first column is for March, April, May, June, July, and August. The second column is for April, May, June, July, and August. The third column is for May, June, July, and August. The fourth column is for June, July, and August. The fifth column is for July and August. The last row corresponds to $-2\rho \ln q_{\mathrm{obs}}$ (the first column is stretched between 0 and 500; second column, between 0 and 400; third column, between 0 and 300; fourth column, between 0 and 200; and the last column, between 0 and 100). Dark areas correspond to no change.

the p-values for testing the hypotheses, i.e., total homogeneity (rejected, $\bar{p} = 0.0000$), Apr = Mar (accepted, $\bar{p} = 0.2808$), and May = Apr(= Mar) (rejected, $\bar{p} = 0.0112$). Having a change between May and April, we perform an omnibus test on the remaining months, i.e., we test May = Jun = Jul = Aug. Here, $\bar{p} = 0.0000$ (histogram not shown), and we reject and test Jun = May (accept, $\bar{p} = 0.1244$), Jul = Jun(= May) (accept, $\bar{p} = 0.3879$), and Aug = Jul(= Jun = May) (reject, $\bar{p} = 0.0000$). For the three p-values that are smaller than 0.05, it is seen in the histograms that almost all values (actually between 97.8% and 100%) lie in the interval [0, 0.05]; thus, a small mean p-value is consistent with all pixels in the area showing changes. For the values above 0.05, it follows that a substantial fraction of the observed p-values is larger than 0.05, i.e., the majority of the corresponding pixels show no changes. However, e.g., in the case of June = May (second row, second column), we have an overrepresentation of low p-values when comparing with the uniform distribution. This might be due to changes in smaller patches, which are possibly a sign of local delays of the changes we saw between May and April.

Thus, in both cases, there is a good agreement between the conclusion based on the averages and what the distribution of the p-values indicates for the area considered.

D. Imaging the Decomposed Test Statistics

To give a visual impression of the results, Figs. 10 and 11 show zoom images of the grass field considered in Table VI. Fig. 10 shows images of the values of the different decomposed test statistics for the grass field. The first five rows show $-2\rho_j \ln r_{j,\mathrm{obs}}$ for full-polarimetry data linearly stretched between 0 and 100. The first column is for March, April, May,

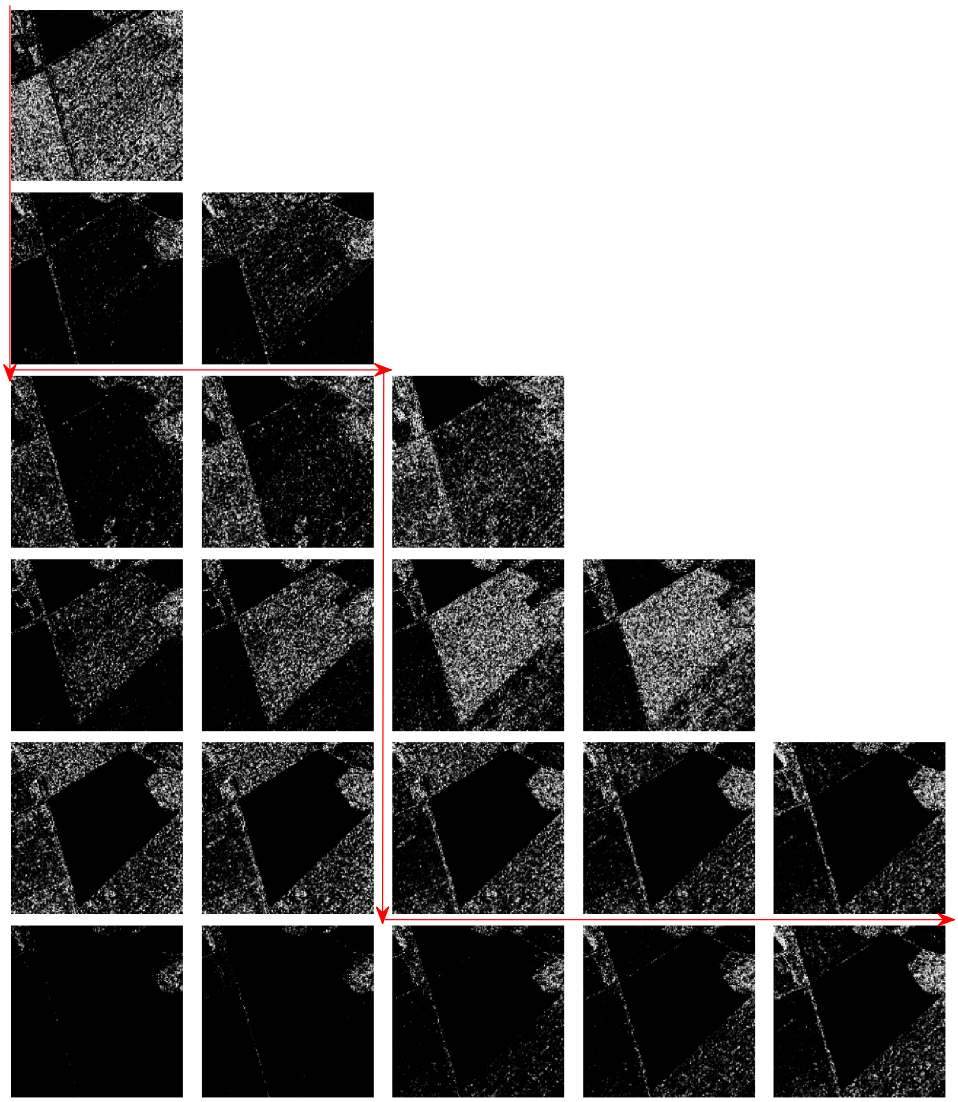


Fig. 11. Zoom on the grass field reported on in Table VI. The p-value corresponding to each subimage has the same relative position as in Table VI. Moreover, the path leading to the change indexes is indicated both in this figure and in Table VI. The p-values, i.e., the no-change probabilities, are linearly stretched between 0 and 1. Dark areas correspond to change.

June, July, and August; the second column is for April, May, June, July, and August; the third column is for May, June, July, and August; the fourth column is for June, July, and August; and the fifth column is for July and August. Row six shows the corresponding $-2\rho \ln q_{\text{obs}}$. The first images in all columns correspond to $-2\rho \ln q_{\text{obs}}$ for the pairwise differences, i.e., the image in the first row, first column is $-2\rho \ln q_{\text{obs}}$ in a $k = 2$ analysis for the pair March and April; the image in the second row, first column is $-2\rho \ln q_{\text{obs}}$ in a $k = 2$ analysis for the pair April and May; the image in the third row, second column is $-2\rho \ln q_{\text{obs}}$ in a $k = 2$ analysis for the pair May and June; the image in the fourth row, third column is $-2\rho \ln q_{\text{obs}}$ in a $k = 2$ analysis for the pair June and July; and the image in the fifth row, fourth column is $-2\rho \ln q_{\text{obs}}$ in a $k = 2$ analysis for the pair July and August. Remember that, if a change is detected at some point, the prerequisite for tests of later changes is no longer valid.

Fig. 11 shows images of the p-values, i.e., the no-change probabilities corresponding to the test statistics in Fig. 10, for the grass field. The values in the grass field in the center of the zoom images clearly show the same trend indicated for the average values in Table VI. The arrows in both figures show the same change patterns, as shown in Table VI and discussed in Section IV-A.

V. Conclusion

A test statistic for the equality of several covariance matrices following the complex Wishart distribution with an associated p-value has been presented. The test statistics are a direct generalization of previously defined test statistics for pairwise comparison [22]. Using data from the airborne EMISAR system at L-band, it is clearly shown how this test statistic is able to detect changes in a series of images, where the pairwise comparison fails to detect the changes. After having detected that at least one change has occurred in a series of polarimetric SAR data, we have shown how one may decompose the likelihood ratio test statistic and thus obtain a procedure for determining the time points of change. Using data from the airborne Danish EMISAR system at L-band clearly shows that this procedure is able to identify the time points of change.

The procedure may be extended to cope with other modes of SAR operations, such as the block-diagonal case, including azimuthally symmetric and diagonal-only data (see [22] for the situation with two time points).

APPENDIX

This Appendix deals with the details of the likelihood ratio test statistics for comparing k Gamma and k complex Wishart distributions. Section A describes the Gamma case. The likelihood ratio test statistic is presented in Theorem 1, and Theorem 2 gives the expected value of Q^h that is used in providing the approximation formula in Theorem 3 using Box's method, [22], [35]. The lemma is a well known result on the independence of the sum and ratio of certain Gamma-distributed random variables. This result is used in the proof of the (other) main result, i.e., the decomposition of the likelihood ratio test into the product of independent (under H_0) random variables R_j. R_j provide tests for when a change actually occurs. Their distributions may be written as a function of independent Beta-distributed random variables, and consequently, the distribution of Q may be written as a product of such variables (Theorem 5), see [22] and [35]. In Theorem 6, it is shown how we may use the F-distribution in testing instead of R_j. This enables two-sided testing. The initial theorems are well known, and the later theorems are not (readily) available in literature.

Section B addresses the same general problem as Section A but considers complex-Wishart-distributed random variables. Theorems 8–10 give the main results on the likelihood ratio test statistic, its expected values, and the large sample distribution. The second main result is the decomposition of Q into independent components R_j presented in Theorems 13 and 14, and the approximative distribution given in Theorem 15. We may apply the same philosophy with respect to mapping change patterns, as was done in Section A. The results in the corollary in Section B are similar to the result in the corollary in Section A and may be used in finding other approximations to the distribution of Q.

A. Comparing k Gamma-Distributed Random Variables

Theorem 1: We consider the independent Gamma-distributed random variables

$$X_i \sim G(n, \beta_i), \quad i = 1, \ldots, k \tag{27}$$

where n is a shape parameter, and β_i is a scale parameter; $E\{X_i\} = n\beta_i$. We want to test the hypothesis H_0 against the alternative H_1, where

$$H_0 : \beta_1 = \cdots = \beta_k \qquad H_1 : \exists i, j : \beta_i \neq \beta_j. \tag{28}$$

Then, the likelihood ratio test statistic for testing H_0 against the alternative H_1 is

$$Q = \left\{ k^k \frac{\prod_{i=1}^k X_i}{X^k} \right\}^n. \tag{29}$$

Proof: Obtained by direct calculations. Q.E.D.

Remarks: In a one-sided analysis of variance (ANOVA, e.g., see [36]) with the same number of observations $f + 1$ in each group, the unbiased estimators s_i^2 of the within-group variances σ_i^2 follow chi-squared distributions with f degrees of freedom, i.e.,

$$s_i^2 \sim \sigma_i^2 \chi^2(f)/f, \qquad i = 1, \ldots, k \tag{30}$$

i.e.,

$$f s_i^2 \sim \sigma_i^2 \chi^2(f) = G\left(\frac{f}{2}, 2\sigma_i^2\right). \tag{31}$$

It now follows that the likelihood ratio test given above is equivalent to what is denoted as *Bartlett's test for homogeneity of variances* in an ANOVA situation. ■

There is no simple closed form for the distribution of Q in Theorem 1, but we may find large sample approximations to the distribution. First, we state the following.

Theorem 2: We consider the likelihood ratio test statistic Q from Theorem 1. Then, we have $(h = 1, 2, 3, \ldots)$

$$\mathrm{E}\{Q^h\} = k^{knh} \left(\frac{\Gamma\left(n(h+1)\right)}{\Gamma(n)} \right)^k \frac{\Gamma(kn)}{\Gamma\left(kn(h+1)\right)}. \tag{32}$$

Proof: Obtained by direct calculations. Q.E.D.

In most cases, it will be sufficient to approximate the distribution of $-2\ln Q$ with a $\chi^2(f)$-distribution (under H_0). However, we may use Box's approximation, e.g., see [22] and [35], to obtain the large sample distribution of $-\ln Q$. We have the following.

Theorem 3: Let the situation be as in Theorem 1. Then, we define

$$f = k - 1 \tag{33}$$

$$\rho = 1 - \frac{k+1}{6kn} \tag{34}$$

$$\omega_2 = -\frac{1}{4}(k-1)\left(1 - \frac{1}{\rho}\right)^2 \tag{35}$$

and have

$$P\{-2\rho \ln Q \leq z\} \simeq P\left\{\chi^2(f) \leq z\right\} + \omega_2 \left[P\left\{\chi^2(f+4) \leq z\right\} - P\left\{\chi^2(f) \leq z\right\}\right]. \tag{36}$$

Proof: Follows by straightforward calculations from Box's theorem by letting:

$$\xi_\ell = 0 \quad \eta_j = 0 \quad a = k \quad b = 1 \quad x_\ell = n \quad y_j = nk. \tag{37}$$

Q.E.D.

We now want to write the likelihood ratio test statistic Q as a product of stochastically independent random variables (if H_0 is true). We start by introducing the following.

Lemma: Let X and Y be independent and Gamma-distributed $G(n, \beta)$ and $G(m, \beta)$, respectively. Then

$$S = X + Y \qquad U = \frac{X}{X+Y} \tag{38}$$

are independent random variables, S follows a $G(n+m, \beta)$-distribution, and U follows a Beta-distribution, i.e., $\mathrm{Be}(n, m)$.

Proof: Straightforward.

We now return to the Gamma-distributed random variables given in (27). We then have the following.

Theorem 4: Let $\beta_1 = \cdots = \beta_{j-1}$. Then, the likelihood ratio test statistic for testing the hypothesis

$$H_{0,j} : \beta_j = \beta_1 \quad \text{against} \quad H_{1,j} : \beta_j \neq \beta_1 \tag{39}$$

is

$$R_j = \frac{j^{jn}}{(j-1)^{(j-1)n}} \frac{(X_1 + \cdots + X_{j-1})^{(j-1)n} X_j^n}{(X_1 + \cdots + X_j)^{jn}}$$
$$= \frac{j^{jn}}{(j-1)^{(j-1)n}} U_j^{(j-1)n} (1-U_j)^n, \quad j=2,\ldots,k \tag{40}$$

where $U_j = S_{j-1}/S_j$ for $S_j = X_1 + \cdots + X_j$. If all β are equal, then $R_2, \ldots, R_k$ will be independent random variables.

Proof: The result on R_j being the likelihood ratio test statistic immediately follows from Theorem 1. According to Theorem 2, S_j and U_j are independent. Therefore, U_j and $S_{j+1} = S_j + X_{j+1}$ will be independent; consequently, U_j and $U_{j+1} = S_j/S_{j+1}$ will be independent, and the theorem follows. Q.E.D.

Theorem 5: Let $\beta_1 = \cdots = \beta_k$. Then, we have

$$Q = R_2 \ldots R_k = k^{kn} \prod_{j=2}^{k} U_j^{(j-1)n} (1-U_j)^n \tag{41}$$

where the U_j are independent and Beta distributed

$$U_j \sim \text{Be}\left((j-1)n, n\right). \tag{42}$$

Proof: Follows by direct computation and from Theorem 2. Q.E.D.

Theorem 6: The critical region for testing $H_{0,j}$ against $H_{1,j}$ is of the following form:

$$R_j \leq c \Leftrightarrow \frac{X_j}{S_{j-1}} \leq c_1 \quad \text{or} \quad \frac{X_j}{S_{j-1}} \geq c_2 \tag{43}$$

where c_1 and c_2 may be determined by realizing that

$$(j-1)\frac{X_j}{S_{j-1}} \sim F\left(2n, 2(j-1)n\right) \tag{44}$$

if $H_{0,j}$ is true, and where F stands for Fisher's F-distribution.

Proof: Straightforward.

B. Comparing k Complex-Wishart-Distributed Random Variables

We start by stating three basic theorems (7–9) on the likelihood ratio test statistic and its distribution.

Theorem 7: We consider the following independent random variables:

$$\boldsymbol{X}_i \sim W_{\mathbb{C}}(p, n, \boldsymbol{\Sigma}_i), \quad i = 1, \ldots, k. \tag{45}$$

We wish to test the hypothesis

$$H_0 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k \quad \text{against} \quad H_1 : \exists i, j : \boldsymbol{\Sigma}_i \neq \boldsymbol{\Sigma}_j. \tag{46}$$

Then, the likelihood ratio test statistic is

$$Q = k^{pnk} \frac{\prod_{i=1}^{k} |\boldsymbol{X}_i|^n}{|\boldsymbol{X}|^{nk}} = \left\{ k^{pk} \frac{\prod_{i=1}^{k} |\boldsymbol{X}_i|}{|\boldsymbol{X}|^k} \right\}^n. \tag{47}$$

where $\boldsymbol{X} = \boldsymbol{X}_1 + \cdots + \boldsymbol{X}_k$.

Proof: Obtained by direct calculations. Q.E.D.

Theorem 8: For the criterion Q, we have

$$\mathrm{E}\{Q^h\} = k^{pknh} \frac{\prod_{j=1}^{p} \Gamma(nk - j + 1)}{\prod_{j=1}^{p} \Gamma\left(nk(1+h) - j + 1\right)} \times \left\{ \frac{\prod_{j=1}^{p} \Gamma\left(n(1+h) - j + 1\right)}{\prod_{j=1}^{p} \Gamma(n - j + 1)} \right\}^k. \tag{48}$$

Using the multivariate Gamma function of the complex kind (e.g., see [37]), i.e.,

$$\Gamma_p(n) = \pi^{p(p-1)/2} \prod_{j=1}^{p} \Gamma(n - j + 1) \tag{49}$$

we get

$$\mathrm{E}\{Q^h\} = k^{pknh} \frac{\Gamma_p(nk)}{\Gamma_p\left(nk(1+h)\right)} \left\{ \frac{\Gamma_p\left(n(1+h)\right)}{\Gamma_p(n)} \right\}^k. \tag{50}$$

Proof: We consider independent p-dimensional random variables

$$Y_{ij} \sim N_{\mathbb{C}}(\mu_i, \boldsymbol{\Sigma}_i), \qquad i=1,\ldots,k; \;\; j=1,\ldots,n+1. \tag{51}$$

For $i = 1, \ldots, k$, we introduce

$$\boldsymbol{X}_i = \sum_{j=1}^{n+1} \left(Y_{ij} - \text{ave}_j(Y_{ij})\right) \left(Y_{ij} - \text{ave}_j(Y_{ij})\right)^H \tag{52}$$

(superscript H denotes the complex conjugate transpose, and ave_j means the average over index j) and have the maximum likelihood estimators

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{n+1} \boldsymbol{X}_i \tag{53}$$

with

$$\boldsymbol{X}_i \sim W_{\mathbb{C}}(p, n, \boldsymbol{\Sigma}_i). \tag{54}$$

Setting $\boldsymbol{X} = \boldsymbol{X}_1 + \cdots + \boldsymbol{X}_k$, the test statistic for testing the equality of the $\boldsymbol{\Sigma}_i$ becomes

$$\Lambda_5 = \left\{ k^{pk} \frac{\prod_{i=1}^{k} \left|\frac{1}{n+1}\boldsymbol{X}_i\right|}{\left|\frac{1}{n+1}\boldsymbol{X}\right|^k} \right\}^{n+1} = \left\{ k^{pk} \frac{\prod_{i=1}^{k} |\boldsymbol{X}_i|}{|\boldsymbol{X}|^k} \right\}^{n+1} \tag{55}$$

(e.g., see [38]). Comparing this with the likelihood ratio test criterion Q in Theorem 4, we see that

$$Q = \Lambda_5^{\frac{n}{n+1}}. \tag{56}$$

$$\begin{aligned}\mathrm{E}\{Q^h\} &= k^{pknh}\frac{\Gamma_p(nk)}{\Gamma_p(n)^k}\frac{\Gamma_p\left(n(1+h)\right)^k}{\Gamma_p\left(nk(1+h)\right)}\\ &= k^{pknh}\frac{\Gamma_p(nk)}{\Gamma_p(n)^k}\frac{\left\{\prod_{j=1}^{p}\Gamma\left(n(1+h)-j+1\right)\right\}\cdots\left\{\prod_{j=1}^{p}\Gamma\left(n(1+h)-j+1\right)\right\}}{\prod_{j=1}^{p}\Gamma\left(nk(1+h)-j+1\right)}\end{aligned} \tag{63}$$

Again, following the work in [38], we have:

$$\begin{aligned}\mathrm{E}\{\Lambda_5^t\} &= k^{pk(n+1)t}\prod_{j=1}^{p}\frac{\Gamma(nk+1-j)}{\Gamma\left(nk+tj+(n+1)kt\right)}\\ &\quad\times\prod_{j=1}^{k}\frac{\Gamma\left(n+1-j+(n+1)t\right)}{\Gamma(n+1-j)}\\ &= k^{pk(n+1)t}\frac{\Gamma_p(nk)}{\Gamma_p\left(nk+(n+1)kt\right)}\\ &\quad\times\left\{\frac{\Gamma_p\left(n+(n+1)t\right)}{\Gamma_p(n)}\right\}^k.\end{aligned} \tag{57}$$

Therefore

$$\begin{aligned}\mathrm{E}\{Q^h\} &= \mathrm{E}\left\{\Lambda_5^{h\frac{n}{n+1}}\right\}\\ &= k^{pknh}\frac{\Gamma_p(nk)}{\Gamma_p(nk+nkh)}\left\{\frac{\Gamma_p(n+nh)}{\Gamma_p(n)}\right\}^k.\end{aligned} \tag{58}$$

and the theorem follows. Q.E.D.

Theorem 9: For Q as in Theorems 4 and 5, we define

$$f = (k-1)p^2 \tag{59}$$

$$\rho = 1-\frac{(2p^2-1)}{6(k-1)p}\left(\frac{k}{n}-\frac{1}{nk}\right) \tag{60}$$

$$\omega_2 = \frac{p^2(p^2-1)}{24\rho^2}\left(\frac{k}{n^2}-\frac{1}{(nk)^2}\right)-\frac{p^2(k-1)}{4}\left(1-\frac{1}{\rho}\right)^2 \tag{61}$$

and have

$$\begin{aligned}P\{-2\rho\ln Q\le z\} &\simeq P\left\{\chi^2(f)\le z\right\}\\ &\quad+\omega_2\left[P\left\{\chi^2(f+4)\le z\right\}-P\left\{\chi^2(f)\le z\right\}\right].\end{aligned} \tag{62}$$

Proof: We rearrange the terms in Q and obtain (63), shown at the top of the page. Letting

$$a = kp, \quad b = p \tag{64}$$

$$x_\ell = n, \quad \ell = 1,\ldots,kp \tag{65}$$

$$y_j = nk, \quad j = 1,\ldots,p \tag{66}$$

$$K = \frac{\Gamma_p(nk)}{\Gamma_p(n)^k} \tag{67}$$

we have

$$\frac{\left\{\prod_{j=1}^{p}y_j^{y_j}\right\}}{\left\{\prod_{\ell=1}^{kp}x_\ell^{x_\ell}\right\}} = \frac{(nk)^{pnk}}{n^{nkp}} = k^{nkp}. \tag{68}$$

We define

$$\eta_j = -j+1, \qquad j = 1,\ldots,p \tag{69}$$

and we let the ξ_ℓ be k versions of those numbers, i.e.,

$$\xi_\ell = -\ell+1, \qquad \ell = 1,\ldots,p \tag{70}$$

$$\vdots$$

$$\xi_\ell = -\ell+(k-1)p+1, \qquad \ell=(k-1)p+1,\ldots,kp. \tag{71}$$

It is now obvious that we may use the result of Box [35], e.g., in the form given in [22], in approximating the distribution of Q. We have

$$\begin{aligned}f &= -2\left\{\sum_\ell\xi_\ell-\sum_j\eta_j-\frac{1}{2}(a-b)\right\}\\ &= -2\left\{k\sum_{\ell=1}^{p}(-\ell+1)-\sum_{j=1}^{p}(-j+1)-\frac{1}{2}(kp-p)\right\}\\ &= -2\left\{-\frac{1}{2}(k-1)p(p-1)-\frac{1}{2}(k-1)p\right\}\\ &= (k-1)p^2.\end{aligned} \tag{72}$$

Furthermore, for $\beta_\ell = (1-\rho)n$ and $\varepsilon_j = (1-\rho)nh$, we have

$$\begin{aligned}\omega_1 &= \frac{1}{2}\left\{\sum_{\ell=1}^{p}\frac{k}{n\rho}B_2\left((1-\rho)n-\ell+1\right)\right.\\ &\qquad\left.-\sum_{j=1}^{p}\frac{1}{\rho nk}B_2\left((1-\rho)nk-j+1\right)\right\}\\ &= \frac{1}{2\rho}\left\{k\sum_{\ell=1}^{p}\frac{1}{n}B_2\left((1-\rho)n-\ell+1\right)\right.\\ &\qquad\left.-\sum_{j=1}^{p}\frac{1}{nk}B_2\left((1-\rho)nk-j+1\right)\right\}.\end{aligned} \tag{73}$$

Now, for $s = n$ or $s = nk$, we have (straightforward calculations)

$$\begin{aligned}\frac{1}{s}B_2\left((1-\rho)s-j+1\right) &= (1-\rho)^2 s\\ &\quad-2(1-\rho)B_1(j)+\frac{1}{s}B_2(j).\end{aligned} \tag{74}$$

The sums of the Bernoulli polynomials are

$$\sum_{j=1}^{p} B_1(j) = \frac{1}{2}p^2$$

$$\sum_{j=1}^{p} B_2(j) = \frac{1}{6}p(2p^2-1)$$

$$\sum_{j=1}^{p} B_3(j) = \frac{1}{4}p^2(p^2-1). \tag{75}$$

Therefore

$$\begin{aligned} 2\rho\omega_1 =& k\left\{p(1-\rho)^2 n - 2(1-\rho)\frac{1}{2}p^2 + \frac{1}{6n}p(2p^2-1)\right\} \\ &- \left\{p(1-\rho)^2 nk - 2(1-\rho)\frac{1}{2}p^2 + \frac{1}{6nk}p(2p^2-1)\right\} \\ =& -(k-1)(1-\rho)p^2 + \frac{1}{6}p(2p^2-1)\left(\frac{k}{n} - \frac{1}{nk}\right). \end{aligned} \tag{76}$$

If we put $\omega_1 = 0$, we obtain

$$(k-1)(1-\rho)p^2 = \frac{1}{6}p(2p^2-1)\left(\frac{k}{n} - \frac{1}{nk}\right) \tag{77}$$

or

$$(1-\rho) = \frac{1}{6(k-1)p}(2p^2-1)\left(\frac{k}{n} - \frac{1}{nk}\right). \tag{78}$$

For this value of ρ, we want to determine ω_2. We have

$$\begin{aligned} -6\rho^2\omega_2 =& k\sum_{\ell=1}^{p}\frac{1}{n^2}B_3\left((1-\rho)n - \ell + 1\right) \\ &- \sum_{j=1}^{p}\frac{1}{(nk)^2}B_3\left((1-\rho)nk - j + 1\right). \end{aligned} \tag{79}$$

Since

$$\begin{aligned} \frac{1}{s^2}B_3\left((1-\rho)s - j + 1\right) =& s(1-\rho)^3 \\ &- 3(1-\rho)^2 B_1(j) + \frac{1}{s}(1-\rho)3B_2(j) - \frac{1}{s^2}B_3(j) \end{aligned} \tag{80}$$

we get

$$-6\rho^2\omega_2 = \frac{3}{2}(k-1)p^2(1-\rho)^2 - \frac{1}{4}p^2(p^2-1)\left(\frac{k}{n^2} - \frac{1}{(nk)^2}\right) \tag{81}$$

or

$$\omega_2 = -(k-1)\frac{p^2}{4}\left(1-\frac{1}{\rho}\right)^2 + \frac{1}{24\rho^2}p^2(p^2-1)\left(\frac{k}{n^2} - \frac{1}{(nk)^2}\right) \tag{82}$$

and the theorem follows. Q.E.D.

Coelho *et al.* [38] give other approximations to the distribution of expressions such as Q.

In order to decompose the likelihood ratio criterion into independent components showing where possible changes may take place, we need some auxiliary results on the distributions of complex matrices.

Lemma: Let the independent random variables $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ be complex Wishart distributed

$$\boldsymbol{X}_i \sim W_{\mathbb{C}}(p, n_i, \boldsymbol{\Sigma}), \qquad i = 1, 2. \tag{83}$$

Let

$$\boldsymbol{C} = (\boldsymbol{X}_1 + \boldsymbol{X}_2)^{-1/2} \tag{84}$$

be a matrix so that

$$\boldsymbol{C}(\boldsymbol{X}_1 + \boldsymbol{X}_2)\boldsymbol{C}^H = \boldsymbol{I}. \tag{85}$$

Then

$$\boldsymbol{X}_1 + \boldsymbol{X}_2 \tag{86}$$

and

$$\boldsymbol{C}\boldsymbol{X}_2\boldsymbol{C}^H = (\boldsymbol{X}_1 + \boldsymbol{X}_2)^{-1/2}\boldsymbol{X}_2\left\{(\boldsymbol{X}_1 + \boldsymbol{X}_2)^{-1/2}\right\}^H \tag{87}$$

are independent.

Proof: First, we consider random variables $\boldsymbol{U}_i, i = 1, 2$ that are Wishart distributed with $\boldsymbol{\Sigma} = \boldsymbol{I}$. In this case, the joint density is

$$f(\boldsymbol{u}_1, \boldsymbol{u}_2) = K\exp\left(-\mathrm{trace}(\boldsymbol{u}_1 + \boldsymbol{u}_2)\right)|\boldsymbol{u}_1|^{n_1-p}|\boldsymbol{u}_2|^{n_2-p}. \tag{88}$$

In this case, the theorem follows directly from the work in [39]. The general case now follows by considering the transformation

$$\boldsymbol{X}_i = \boldsymbol{\Sigma}^{1/2}\boldsymbol{U}_i(\boldsymbol{\Sigma}^{1/2})^H. \tag{89}$$

We then obtain

$$\boldsymbol{I} = \boldsymbol{C}(\boldsymbol{X}_1 + \boldsymbol{X}_2)\boldsymbol{C}^H \tag{90}$$

$$= \boldsymbol{C}\boldsymbol{\Sigma}^{1/2}(\boldsymbol{U}_1 + \boldsymbol{U}_2)(\boldsymbol{C}\boldsymbol{\Sigma}^{1/2})^H \tag{91}$$

$$\boldsymbol{C}\boldsymbol{X}_2\boldsymbol{C}^H = \boldsymbol{C}\boldsymbol{\Sigma}^{1/2}\boldsymbol{U}_2(\boldsymbol{C}\boldsymbol{\Sigma}^{1/2})^H. \tag{92}$$

Thus, the previous results apply to $\boldsymbol{U}_1, \boldsymbol{U}_2$, and the theorem follows. Q.E.D.

Returning to the likelihood ratio statistic Q, we can now prove the following.

Theorem 10: Given that

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_{j-1} \tag{93}$$

then the likelihood ratio test statistic R_j for testing the hypothesis

$$H_0 : \boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}_{j-1} \quad \text{against} \quad H_1 : \boldsymbol{\Sigma}_j \neq \boldsymbol{\Sigma}_{j-1} \tag{94}$$

is

$$R_j = \frac{j^{jpn}}{(j-1)^{(j-1)pn}}\frac{|\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_{j-1}|^{(j-1)n}|\boldsymbol{X}_j|^n}{|\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_j|^{jn}}. \tag{95}$$

Furthermore

$$Q = \prod_{j=2}^{k} R_j. \tag{96}$$

Proof: Straightforward. Q.E.D.

Theorem 11: If the hypothesis

$$H_0 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k \tag{97}$$

is true and $n \geq p$, then the quantities $R_2, \ldots, R_k$ are independent.

Proof: We find $\boldsymbol{C}_j$ so that

$$\boldsymbol{C}_j(\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_j)\boldsymbol{C}_j^H = \boldsymbol{I} \tag{98}$$

and put

$$\boldsymbol{E}_j = \boldsymbol{C}_j(\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_{j-1})\boldsymbol{C}_j^H \tag{99}$$

getting

$$\boldsymbol{I} - \boldsymbol{E}_j = \boldsymbol{C}_j\boldsymbol{X}_j\boldsymbol{C}_j^H. \tag{100}$$

From the aforementioned lemma, it follows that $\boldsymbol{E}_j$ is independent of $\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_j$ and is therefore also independent of $\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_j + \boldsymbol{X}_{j+1}$. This implies that

$$\boldsymbol{E}_j \text{ and } \boldsymbol{C}_{j+1}(\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_j)\boldsymbol{C}_{j+1}^H = \boldsymbol{E}_{j+1} \tag{101}$$

are independent. Arguing along those lines gives the theorem. Q.E.D.

Corollary: The quantities R_j and Q may be written as

$$R_j = \frac{j^{jpn}}{(j-1)^{(j-1)pn}} |\boldsymbol{E}_j|^{(j-1)n} |\boldsymbol{I} - \boldsymbol{E}_j|^n \tag{102}$$

$$Q = k^{kpn} \prod_{j=2}^{k} |\boldsymbol{E}_j|^{(j-1)n} |\boldsymbol{I} - \boldsymbol{E}_j|^n. \tag{103}$$

Proof: Straightforward. Q.E.D.

We conclude this section by stating a theorem on the distribution of the components R_j in the decomposition of the likelihood ratio criterion.

Theorem 12: Let the situation be as in Theorem 10. Letting

$$f = p^2 \tag{104}$$

$$\rho_j = 1 - \frac{2p^2 - 1}{6pn}\left(1 + \frac{1}{j(j-1)}\right) \tag{105}$$

$$\omega_{2j} = -\frac{p^2}{4}\left(1 - \frac{1}{\rho_j}\right)^2 + \frac{1}{24n^2}p^2(p^2-1) \times \left(1 + \frac{2j-1}{j^2(j-1)^2}\right)\frac{1}{\rho_j^2} \tag{106}$$

then

$$P\{-2\rho_j \ln R_j \leq z\} \simeq P\left\{\chi^2(f) \leq z\right\} + \omega_{2j}\left[P\left\{\chi^2(f+4) \leq z\right\} - P\left\{\chi^2(f) \leq z\right\}\right]. \tag{107}$$

Proof: Follows from the two-sample test for the equality of complex Wishart matrices. Q.E.D.

Coelho *et al.* [38] give other approximations to the distribution of the expression such as R_j.

REFERENCES

[1] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 6, pp. 764–773, Nov. 1988.

[2] E. J. Rignot and J. J. Van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 896–906, Jul. 1993.

[3] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874–887, Apr. 2005.

[4] G. Moser and S. B. Serpico, "Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2972–2982, Oct. 2006.

[5] Y. Ban and O. A. Yousif, "Multitemporal spaceborne SAR data for urban change detection in China," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1087–1094, Aug. 2012.

[6] H. Hu and Y. Ban, "Unsupervised change detection in multitemporal SAR images over large urban areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 8, pp. 3248–3261, Aug. 2014.

[7] C. Carincotte, S. Derrode, and S. Bourennane, "Unsupervised change detection on SAR images using fuzzy hidden Markov chains," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 432–441, Feb. 2006.

[8] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, Dec. 2005.

[9] T. Celik, "A Bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images," *Signal Process.*, vol. 90, no. 5, pp. 1471–1485, May 2010.

[10] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.

[11] F. Bovolo, C. Martin, and L. Bruzzone, "A hierarchical approach to change detection in very high resolution SAR images for surveillance applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2042–2054, Apr. 2013.

[12] P. Gamba, F. Dell'Acqua, and G. Lisini, "Change detection of multitemporal SAR data in urban areas combining feature-based and pixel-based techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2820–2827, Oct. 2006.

[13] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1445, May 2007.

[14] F. Chatelain, J.-Y. Tourneret, and J. Inglada, "Change detection in multisensory SAR images using bivariate Gamma distributions," *IEEE Trans. Image Process.*, vol. 17, no. 3, pp. 249–258, Mar. 2008.

[15] C. Pratola, F. Del Frate, G. Schiavon, and D. Solimini, "Toward fully automatic detection of changes in suburban areas from VHR SAR images by combining multiple neural-network modes," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2055–2066, Apr. 2013.

[16] S. Hachicha and F. Chaabane, "On the SAR change detection review and optimal decision," *Int. J. Remote Sens.*, vol. 35, no. 5, pp. 1693–1714, Feb. 2014.

[17] O. Yousif and Y. Ban, "Improving SAR-based urban change detection by combining MAP-MRF classifier and nonlocal means similarity weights," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 10, pp. 4288–4300, Oct. 2014.

[18] W. Dierking and H. Skriver, "Change detection for thematic mapping by means of airborne multitemporal polarimetric SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 618–636, Mar. 2002.

[19] M. Qong, "Polarization state conformation and its application to change detection in polarimetric SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 304–307, Oct. 2004.

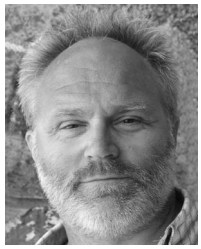
[20] G. Moser, S. B. Serpico, and G. Vernazza, "Unsupervised change detection from multichannel SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 278–282, Apr. 2007.

[21] G. Moser and S. B. Serpico, "Unsupervised change detection from multichannel SAR data by Markovian data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2114–2128, Jul. 2009.

[22] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 4–19, Jan. 2003.
[23] P. Formont, F. Pascal, G. Vasile, J. P. Ovarlez, and L. Ferro-Famil, "Statistical classification for heterogeneous polarimetric SAR images," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 567–576, Jun. 2011.
[24] A. Marino, S. Cloude, and J. Sanchez-Lopez, "A new polarimetric change detector in radar imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2986–3000, May 2013.
[25] V. Akbari, S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "The Hotelling–Lawley trace statistic for change detection in polarimetric SAR data under the complex Wishart distribution," in *Proc. IEEE IGARSS*, Melbourne, Vic., Australia, Jul. 21–26, 2013, pp. 4162–4165.
[26] M. Liu, H. Zhang, C. Wang, and F. Wu, "Change detection of multilook polarimetric SAR images using heterogeneous clutter models," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7483–7494, Dec. 2014.
[27] W. T. Coombs, J. Algina, and D. O. Oltman, "Univariate and multivariate omnibus hypothesis tests selected to control type I error rates when population variances are not necessarily equal," *Rev. Educ. Res.*, vol. 66, no. 2, pp. 137–179, Summer 1996.
[28] A. A. Nielsen, K. Conradsen, and H. Skriver, "Change detection in full and dual polarization, single—And multi-frequency SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 4041–4048, Aug. 2015.
[29] J. J. van Zyl and F. T. Ulaby, "Scattering matrix representation for simple targets," in *Radar Polarimetry for Geoscience Applications*, F. T. Ulaby and C. Elachi, Eds. Norwood, MA, USA: Artech House, 1990.
[30] S. N. Madsen, E. L. Christensen, N. Skou, and J. Dall, "The Danish SAR system: Design and initial tests," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 3, pp. 417–426, May 1991.
[31] E. L. Christensen *et al.*, "EMISAR: An absolutely calibrated polarimetric L—And C-band SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 6, pp. 1852–1865, Nov. 1998.
[32] R. G. Miller, Jr., *Simultaneous Statistical Inference*, 2nd ed. New York, NY, USA: Springer-Verlag, 1981.
[33] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. New York, NY, USA: Wiley, 2003.
[34] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 68–78, Jan. 1997.
[35] G. E. P. Box, "A general distribution theory for a class of likelihood criteria," *Biometrika*, vol. 36, no. 3/4, pp. 317–346, Dec. 1949.
[36] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*. New York, NY, USA: Springer-Verlag, 2013.
[37] S. N. Anfinsen and T. Eltoft, "Application of the matrix-variate Mellin transform to analysis of polarimetric radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2281–2295, Jun. 2011.
[38] C. A. Coelho, B. C. Arnold, and F. J. Marques, "The exact and near-exact distributions of the main likelihood ratio test statistics used in the complex multivariate normal setting," *TEST*, vol. 24, no. 2, pp. 386–416, Jun. 2015.
[39] A. K. Gupta and D. G. Kabe, "An n-variate characterization of the Gamma and the complex Wishart densities," *Appl. Math. Lett.*, vol. 10, no. 4, pp. 119–122, Jul. 1997.

Knut Conradsen received the Candidatus Scientiarum degree in mathematics from the University of Copenhagen, Copenhagen, Denmark, in 1970.

Since 1970, he has been with the Technical University of Denmark (DTU), Kongens Lyngby, Denmark, where he is currently a Professor of statistical image analysis with the Department of Applied Mathematics and Computer Science. From 1995 to 2010, he was a Provost (Deputy Rector) of DTU. His main research interests include the application of statistics and statistical models in primarily medical image analysis, remote sensing, and industrial applications. His work includes the analysis of multispectral/hyperspectral and multitemporal data, as well as optical radar sensors.



Allan Aasbjerg Nielsen received the M.Sc. and Ph.D. degrees from the Technical University of Denmark (DTU), Kongens Lyngby, Denmark, in 1978 and 1994, respectively.

From 1977 to 1978, he was with the Danish Defense Research Establishment. From 1978 to 1985, he worked on energy conservation in housing with the Thermal Insulation Laboratory, DTU. He was with the section for image analysis from 1985 to 2001 and with the section for geoinformatics from 2001 to 2006, both at the Department of Informatics and Mathematical Modeling, DTU. From 2007 to 2013, he was with the Danish National Space Center's section for geodesy. He is currently an Associate Professor with the Department of Applied Mathematics and Computer Science, DTU. Since 1985, he has worked on several national and international projects on the development, implementation, and application of statistical methods, and remote sensing in mineral exploration, mapping, geology, agriculture, environmental monitoring, oceanography, geodesy, and security funded by industry, the European Union, the Danish International Development Agency (Danida), and the Danish National Research Councils.

Henning Skriver (M'09) received the M.Sc. and Ph.D. degrees in electrical engineering from the Technical University of Denmark (DTU), Kongens Lyngby, Denmark, in 1983 and 1989, respectively.

Since 1983, he has been with DTU, where he is an Associate Professor, the Head of Microwaves and Remote Sensing, and the Deputy Head of Department with the National Space Institute. His work has been primarily concerned with topics related to the utilization of SAR data for different applications, such as sea ice parameter retrieval from SAR data, as well as different aspects of land applications of SAR data, such as forestry, agricultural, environmental, and topographic mapping applications using both satellite SAR data and data from polarimetric SARs, e.g., the Danish airborne polarimetric SAR, EMISAR, and the German experimental SAR ESAR. His research interests include methods for the processing of SAR data; SAR image simulation; SAR image filtering; speckle statistics; and texture analysis, segmentation, calibration, change detection, classification, and polarimetric analysis and processing.